On the parametrization of APRBS and multisine test signals for the identification of nonlinear dynamic TS-models

Matthias Gringard
Department of Measurement and Control
University of Kassel
Kassel, Germany
Email: matthias.gringard@mrt.uni-kassel.de

Andreas Kroll
Department of Measurement and Control
University of Kassel
Kassel, Germany
Email: andreas.kroll@mrt.uni-kassel.de

***Abstract*—The quality of identified models is inherently linked to the amount of information contained in the data used for identification. After a short introduction into the topic and the presentation of the model family and identification process, general purpose test signals are discussed, some of which are manipulated in order to achieve suitable data for the identification of nonlinear dynamic TS-models. Before this contribution is concluded, a case study is presented.**

I. Introduction

Since the most powerful methods for simulation, prognosis and control design are model-based, high quality models are needed. Due to the ever rising complexity of technical systems as well as the need for nonlinear description of their components, theoretical modeling has become a difficult task that requires a lot of effort an expert knowledge. Therefore system identification is a more than valid alternative. However, identified models can in principle only describe the behavior that is present in the data used for identification and not be used for extrapolation like high quality physical models can. This directly leads to the task of generating suitable data for nonlinear system identification by exciting the system appropriately. Since Takagi-Sugeno-(TS-)Fuzzy models are universal approximators, they have been selected as the model class. The nonlinear behavior of locally affine TS-models is solely described by the partitioning, which emphasizes the importance of the structure identification.

The Design of Experiments (DoE) can be categorized in the following way: DoE can be done offline and in real time. If the design is done offline the complete system input will designed before the experiment is conducted. In an real time design the input will be adapted during the experiment. Since there are distinct computational difficulties with nonlinear model approaches, DoE approaches are distinguished between linear and nonlinear model approaches. In contrast to static systems where each design point can be chosen independently from each other, the system input of a dynamic systems is a time series and the system dynamics have to be considered as a side condition. Test signal design is a part of the DoE. The last major distinction is whether knowledge about the model and its structure is used in the process. If the design is done without considering the model structure, it is called model-free, otherwise it is model-based.

The model-based designs are typically based on the approach to optimize scalar measures on the Fisher Information Matrix (FIM). The FIM is constructed from the global model's first derivatives with respect to the model parameters. Exploiting the Cramér-Rao-Inequality the FIM is used to estimate the covariance of the parameters. Therefore the criteria can be interpreted as a way to minimize the model parameter uncertainty. The model structure often is assumed to be known [1]. The membership functions that determine the weighting of the locally-affine models mostly depend on the model prototypes nonlinearly. Since the partitioning is used to model system nonlinearities it is not sufficient to design the excitation signal without considering the prototypes. In contrast to FIM for models that are linear in their parameters, the FIM for models that are nonlinear in their parameters depends on the model parameters. They are however unknown at this stage of the experiment design. Experiments show that FIM-based design approaches are very sensitive to parameter mismatch [2], so a sequential robust design was proposed [3]. Therefore the current research investigates the application of manipulated general purpose test signals for the identification of nonlinear TS-models. This model-free approach was used to generate broadband test signals that were manipulated to achieve a specific amplitude distribution either in the time or frequency domain [4].

II. Model Family and Identification Process

As mentioned before, locally affine TS-models are used. TS-models are a superposition of local models weighted by their respective normalized membership functions. In case of (N)ARX-type models the i-th local model can be described like:

$$\hat{y}_i(k) = -\sum_{j=1}^{n_{yc}} a_{i,j} \cdot y(k-j) + \sum_{l=1}^{n_{uc}} b_{i,l} \cdot u(k-l) + a_{i,0} \quad (1)$$

$\hat{y}_i(k)$ is the output of the i-th local model, $a_{i,j}$ and $b_{i,l}$ are the coefficients of the inputs and outputs as well as their lags, n_{yc} and n_{uc} are the maximum lags, k is the discrete time and $a_{i,0}$ is the affine term. The difference equation (1) can be rewritten as:

$$\begin{aligned}\hat{y}_i(k) &= \underbrace{\begin{bmatrix}1 & \boldsymbol{\varphi}_y^{\mathrm{T}}(k-1) & \boldsymbol{\varphi}_u^{\mathrm{T}}(k-1)\end{bmatrix}}_{\boldsymbol{\varphi}^{\mathrm{T}}(k-1)} \underbrace{\begin{bmatrix}a_{i,0}\\ \boldsymbol{\Theta}_{i,y}\\ \boldsymbol{\Theta}_{i,u}\end{bmatrix}}_{\boldsymbol{\Theta}_i} \\ &= \boldsymbol{\varphi}^{\mathrm{T}}(k-1)\cdot\boldsymbol{\Theta}_i \end{aligned} \tag{2}$$

where $\boldsymbol{\varphi}(k-1)$ and $\boldsymbol{\Theta}_i$ are the regression vector and the i-th local model parameter vector. To get the global model, the local models (2) are superposed by their fuzzy basis functions functions $\phi_i(\mathbf{z}(k-1))$. To emphasize the dependance of the regression vector and the scheduling variable $\mathbf{z}(k-1)$ from past values of the in- and output, the argument $(k-1)$ is carried along. The fuzzy basis functions are:

$$\phi_i(\mathbf{z}(k-1)) = \frac{\mu_i\left(\mathbf{z}(k-1)\right)}{\sum_{j=1}^{c}\mu_j\left(\mathbf{z}(k-1)\right)} \tag{3}$$

with

$$\mu_i\left(\mathbf{z}\left(k-1\right)\right) = \left[\sum_{j=1}^{c}\left(\frac{||\mathbf{z}\left(k-1\right)-\mathbf{v}_i||_2}{||\mathbf{z}\left(k-1\right)-\mathbf{v}_j||_2}\right)^{\frac{2}{\nu-1}}\right]^{-1} \tag{4}$$

The membership functions μ_i are functions of the scheduling variable $\mathbf{z}(k-1)$. The scheduling variable is often a function of the regression vector but can be chosen to be entirely different if a priori knowledge is available. c is the number of local models, the $\mathbf{v}_i$ are the partition's prototypes and $\nu \in \mathbb{R}^{>1}$ is the fuzziness parameter. The membership functions have been chosen to be of the same type as in the Fuzzy c-means (FCM) clustering algorithm, because they fulfill the orthogonality condition (5). Besides the prototypes and the fuzziness parameter there is no need for further parametrization. When using Gaussian or triangular membership functions, more parameters have to be specified like the covariance matrices for the Gaussians. Since additional degrees of freedom are most helpful when used with additional knowledge and can be detrimental if not, the calculation of the membership functions in the simulation and optimization is done in the same way they are obtained in the first place.

$$\sum_{i=1}^{c}\mu_i\left(\mathbf{z}\left(k-1\right)\right) \equiv 1 \tag{5}$$

Then the fuzzy basis functions are equal to the membership functions:

$$\phi_i(\mathbf{z}(k-1)) = \mu_i(\mathbf{z}(k-1)) \tag{6}$$

Therefore the global model can be given as:

$$\hat{y}(k|k-1) = \sum_{i=1}^{c}\mu_i\left(\mathbf{z}(k-1)\right)\cdot\boldsymbol{\varphi}^{\mathrm{T}}(k-1)\cdot\boldsymbol{\Theta}_i \tag{7}$$

With the abbreviations $\mu_i(\mathbf{z}(k-1)) = \mu_{ik}$ and $\boldsymbol{\varphi}^{\mathrm{T}}(k-1) = \boldsymbol{\varphi}_k^{\mathrm{T}}$ follows from (7):

$$\hat{y}(k|k-1) = \begin{bmatrix}\mu_{1k}\boldsymbol{\varphi}_k^{\mathrm{T}} & \cdots & \mu_{ck}\boldsymbol{\varphi}_k^{\mathrm{T}}\end{bmatrix}\begin{bmatrix}\boldsymbol{\Theta}_1\\ \vdots\\ \boldsymbol{\Theta}_c\end{bmatrix} \tag{8}$$

With N observations (8) is evaluated $(N-n_{yc})$ times for $n_{yc} \geq n_{uc}$.

$$\hat{\mathbf{y}} = \boldsymbol{\Phi}_{\mathrm{E}}\boldsymbol{\Theta} \tag{9}$$

$\boldsymbol{\Phi}_{\mathrm{E}}$ is the extended regression matrix, $\boldsymbol{\Theta}$ the vector of all local model parameters and $\hat{\mathbf{y}}$ the regressand. With known $\mathbf{v}_i$ the estimation problem is linear in the parameters and the solution to the least squares method can be given as follows when using a quadratic cost function for identification:

$$\hat{\boldsymbol{\Theta}} = \left(\boldsymbol{\Phi}_{\mathrm{E}}^{\mathrm{T}}\boldsymbol{\Phi}_{\mathrm{E}}\right)^{-1}\boldsymbol{\Phi}_{\mathrm{E}}^{\mathrm{T}}\hat{\mathbf{y}} \tag{10}$$

The solution (10) is optimal with respect to the one-step-ahead-prediction (OSAP). To calculate it, c and $\mathbf{v}_i$ have to be known. Therefore structural decisions like the selection of the scheduling variable, number of local models and type of membership function have to be made before the parameter estimation. The FCM algorithm is used to determine the prototypes $\mathbf{v}_i$. The resulting prototypes are not optimal with respect to the prediction error. The estimated parameter vector $\hat{\boldsymbol{\Theta}}$ minimizes the squared sum of the deviations of the one-step-ahead-prediction. For simulation purposes it is more important to have a model with good predictions qualities when the model is evaluated recursively. This obviously is a harder criterion. Because of that the parameters are optimized with a nonlinear optimization algorithm with respect to the prediction error while the model is evaluated recursively. The local model parameters $\boldsymbol{\Theta}_i$ and prototypes $\mathbf{v}_i$ are aggregated to a parameter vector for the complete local model:

$$\boldsymbol{\Theta}_{\mathrm{agg},i}^{\mathrm{T}} = \begin{bmatrix}\boldsymbol{\Theta}_i^{\mathrm{T}} & \mathbf{v}_i^{\mathrm{T}}\end{bmatrix} \tag{11}$$

The local model parameters and the partition are now optimized simultaneously. With a quadratic cost function the following holds:

$$\hat{\boldsymbol{\Theta}}_{\mathrm{agg}} = \underset{\boldsymbol{\Theta}_{\mathrm{agg}}}{\operatorname{argmin}}\sum_{k=1}^{N}\left(y_k - \hat{y}_k\left(\left(\boldsymbol{\Theta}_{\mathrm{agg}}\right)\right)\right)^2 \tag{12}$$

with

$$\boldsymbol{\Theta}_{\mathrm{agg}}^{\mathrm{T}} = \begin{bmatrix}\boldsymbol{\Theta}_{\mathrm{agg},1}^{\mathrm{T}} & \cdots & \boldsymbol{\Theta}_{\mathrm{agg},c}^{\mathrm{T}}\end{bmatrix} \tag{13}$$

The optimization problem is solved with the `MATLAB`-function `lsqnonlin` which by default uses a Trust-Region-Reflective-Algorithm. The optimization is initialized with the local model parameters from (10) and the prototypes from the FCM clustering. The choice of ν is discussed in detail in [5].

III. Design of the excitation signals

This section is divided into two parts. In the first part standard test signals are discussed. The second part is about how some of these signals can be parametrized to obtain better identification data. The described methods of designing excitation signals in general can be used to identify the parameters of different model families but from the results of identifying TS-models there can be no generalization. The structure of TS-models is such that local model parameters need data near the center of the model and partition parameters need data near the borders which can be targeted with experimental design.

A. Standard test signals

Standard test signals are used to excite a system without any prior knowledge. These signals are generally obtained by using two basic signal types. On the one hand there are signals based on sinusoids like swept sines or multisines. On the other hand multi-level signals are used. Furthermore different kinds of signals based on random noise are used. Different standard test signals are compared in table I that bases on [6]. The crest factor [6] for a signal x_k is defined as:

$$c_{\mathrm{R}} := \frac{\max\limits_{1\leq k\leq N} |x_k|}{x_{\mathrm{RMS}}} \tag{14}$$

It is used to assess how well the available signal range is used. The main difference between a design for linear and nonlinear system identification is that one amplitude level is sufficient in the linear case since the choice of the amplitude levels does not yield any additional information. In the nonlinear case the amplitudes represent a degree of freedom that can be used to obtain a test signal that is better suited for identification.

B. Parametrization of test signals

As it was mentioned in the last section, multi-level and harmonic signals can be parametrized very comfortably, because the parameters can be interpreted easily. The signals that will be discussed here are multisine and APRBS (Amplitude-modulated Pseudo-random binary sequence) signals. The parameters of a multisine are the used frequencies, the corresponding phases as well as the amplitudes. While the used frequencies are determined by the frequency band of interest, the choice of the phases and amplitudes is not straightforward. When parametrizing a signal without special knowledge of the underlying system, goals can include to transfer as much energy into the system as possible. A uniform distribution in the scheduling space might yield better partitioning since there is no weight on certain regions. Due to the obvious lack of knowledge about the real system, these properties can only be approached indirectly. Without prior knowledge, the starting point for the design is to achieve a uniform amplitude distribution. Since a signal in the time domain always faces certain boundaries on maximum (and minimum) values, the phases are used to achieve such a distribution such that the amplitudes of the used frequencies can be scaled up to guarantee the maximal possible energy inserted into the system. A method for homogenization will be presented. For multi-level signals the impact of the parametrization is not that easy to interpret anymore. The values of the held amplitudes can be adjusted as well as the holding time. The amplitudes are chosen randomly.

1) Multisine Signals: The used homogenization method [7] iteratively adjusts the phases in the frequency domain until the signal in the time domain meets a criterion that is linked to the distribution. In this case a uniform distribution is used but the extension to arbitrary distributions is straightforward. As a measure of how well the signal distribution meets its target distribution a coverage index J_{CI} is introduced:

$$J_{\mathrm{CI}} = \sum_{i=1}^{n_{\mathrm{f}}} (n_i - n_{\mathrm{ideal}})^2 \tag{15}$$

n_{f} is the number of intervals of equal length of the range of the signal in the time domain, n_i is the number of data points within the i-th interval and n_{ideal} is the target number of data points per interval which with respect to a uniform distribution simply is:

$$n_{\mathrm{ideal}} = \frac{N}{n_{\mathrm{f}}} \tag{16}$$

where N is the total number of data points of the signal. The strategy of this method is to generate a timeseries of the same length of the signal within the upper and lower bounds of the signal. Since the sorted input signal can be viewed as the cumulative distribution function (cdf), the phases of the signal are replaced with the phases of the inversely (with respect to the sorted test signal) sorted ideally distributed signal until the distribution of the test signal reaches its target. Because the coverage index J_{CI} is not invariant for different N or n_{f} and therefore has no absolute meaning, a convergence oriented criterion is used. The simplest method is to check the convergence based on the difference between two consecutive iteration steps:

$$\Delta J_{\mathrm{CI}} = J_{\mathrm{CI},k+1} - J_{\mathrm{CI},k} \tag{17}$$

So the applied termination criterion is

$$\Delta J_{\mathrm{CI}} \leq \epsilon \tag{18}$$

where ϵ has to be chosen by the user. Alternatively or in addition cost-based criteria such as the number of iterations can be used. With the established termination criterion one iteration runs as follows:

1) Calculate J_{CI} for a signal $[u_k] = \mathbf{u}$ and check the criterion (18).
2) Sort the timeseries $\mathbf{u}$. The sorting can formally be expressed with the permutation matrix $\mathbf{T}$: $\mathbf{u}' = \mathbf{T}\mathbf{u}$
3) Define a timeseries $\mathbf{u}''$, which is a line.
4) Sort back the timeseries $\mathbf{u}''$ with $\mathbf{q} = \mathbf{T}^{-1}\mathbf{u}''$
5) Calculate the discrete fourier transforms (DFT) $\mathbf{Q} = \mathrm{DFT}(\mathbf{q})$ and $\mathbf{U} = \mathrm{DFT}(\mathbf{u})$
6) Generate new DFT $\mathbf{R} = [R_k]$, $R_k = |U_k| \cdot \exp\left(j \cdot \arg\left(Q_k\right)\right)$
7) Apply the inverse discrete fourier transform (IDFT) to $\mathbf{R}$, $\mathbf{r} = \mathrm{IDFT}(\mathbf{R})$

Signal	Definition	Crest Factor	Comment
Swept Sine/Chirp	$u(t) = A \cdot \sin((at+b)\,t)$ with $0 \leq t < T_0$ where T_0 is the period. With $f(t) = at + b$ and the constraints $f(0) = 2\pi k_1 f_0$ and $f(T_0) = 2\pi k_2 f_0$ the coefficients are calculated as $a = 2\pi\,(k_2 - k_1)\,f_0^2$ and $b = 2\pi k_1 f_0$	The crest factor is comparable to that of an ordinary sine wave. It is typically $c_{\mathrm{R}} = 1.45$.	The amplitude spectrum is not flat because a swept sine is not a superposition of single well defined sine waves but rather a smooth transition from the starting frequency to the end frequency. If the ground frequency f_0 is chosen to be $1/T_0$ and k_1 and k_2 are integers, this signal has no leakage
Schroeder Multisine	$u(t) = \sum_{k=1}^{F} A_k \cdot \sin(2\pi f_k t + \phi_k)$ where A_k are the Amplitudes of the excited frequencies f_k. ϕ_k are the phases of the corresponding sine waves. F is the number of excited frequencies. The Schroeder phases are defined as $\phi_k = -\frac{k(k-1)\pi}{F}$	The crest factor is typically $c_{\mathrm{R}} = 1.7$ when all amplitudes are chosen to be identical. In general the crest factor of a multisine depends on the choice of amplitudes and phases.	Since the amplitudes, phases and frequencies are independent of each other, many signal properties can be defined without any inherent constraints. Most of the time, the frequencies as well as the amplitudes are chosen with respect to the system in question. The phases then can be used to achieve a better crest factor.
(A)PRBS	A PRBS excitation signal is obtained from a pseudo random binary sequence. To obtain an APRBS signal the binary states are given individual amplitude levels.	For a PRBS signal the crest factor is equal to $c_{\mathrm{R}} = 1$ if all power is considered. For an APRBS signal the crest factor depends on the choice of the amplitude levels.	The choice of a multi-level signal versus a harmonic signal often depends on the plant friendliness. It can be bad for the actuators to be constantly excited so that a multisine signal cannot be used.
Random Noise	Filtered Noise Sequence	typically $c_{\mathrm{R}} = 2 \ldots 3$	

TABLE I
STANDARD TEST SIGNALS

8) Set $\mathbf{u} = \mathbf{r}$

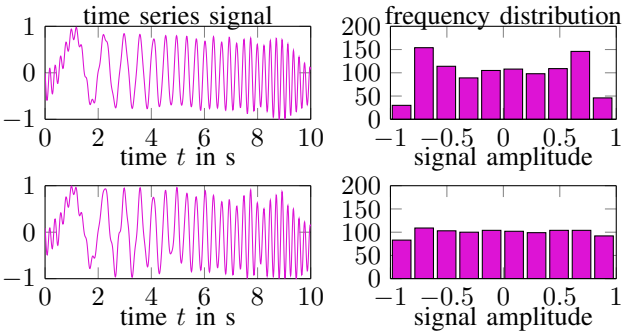

Fig. 1. Multisine signals with Schroeder phases (top) and after homogenization (bottom)

Figure 1 shows the impact of the homogenization of a full multisine with 50 frequencies between 0 Hz and 5 Hz and Schroeder phases [8]. It can be seen that both signals remain within the values of ± 1. Furthermore figure 2 shows the

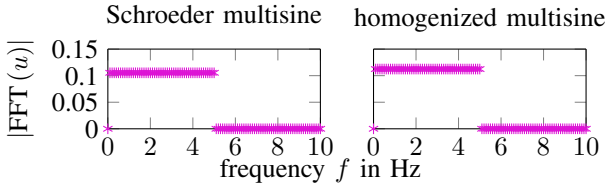

Fig. 2. Impact of the homogenization (frequency domain)

amplitude distribution of the multisines shown in figure 1. It can be seen that due to the homogenization the amplitudes can be scaled up while still satisfying the boundary conditions of $\max|u_k| < 1$. This allows a higher energy transfer while keeping the shape of the amplitude spectrum, the excited frequencies and the experiment time as it was in the initial signal. If Schroeder phases are used, the gain in amplitude level is very small, because Schroeder phases already have a good crest factor (Schroeder Phases: $c_{\mathrm{R}} = 1.89$, after Homogenization: $c_{\mathrm{R}} = 1.77$). Also the initialization with Schroeder phases leads to a homogenized signal that closely resembles the initial signal so that it can be assumed that the phase configuration is in the region of a local optimum.

2) APRBS signals: To generate an APRBS-Signal a shift register is used [9]. At first a Pseudo-Random Binary Sequence has to be generated. This sequence depends on the structure of the given shift register and the used seed. To obtain the multi-level signal from this sequence a change of the binary state is used to determine the change of the signal levels. The number of following equal states determines how long a level is held. By considering the holding time T_{H} and the sampling time T_{S} and the choice of the amplitude levels, the signal can be generated. An example of a pseudo-random binary sequence is given:

$$\mathbf{s}_{\mathrm{PRBS}} = \begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 1 & 0 \end{bmatrix} \tag{19}$$

The sequence $\mathbf{s}_{\mathrm{PRBS}}$ (19) is generated from a shift register with $n = 3$ states and therefore the repeating sequence has 2^{n-1} elements. Consecutive equal states are contracted to indicate a single held amplitude with a weight that will be proportional to the length of this amplitude. In this example, the first amplitude will be held for one T_{H}, the second one for two T_{H} and so on.

$$\mathbf{S} = \begin{bmatrix} 1 & 2 & 3 & 1 \end{bmatrix} \tag{20}$$

The length of $\mathbf{S}$ determines the number of amplitudes A_i needed to generate a real signal. With a holding time T_{H}, the sampling time T_{S} and the amplitudes A_i, the timeseries

is constructed by repeating the amplitudes A_i $S(i) \cdot \frac{T_\mathrm{H}}{T_\mathrm{S}}$ times. Figures 3 shows the difference between APRBS signals with identical amplitudes but different holding times, which has been exaggerated to show the effect in the frequency domain. If a shorter holding time is used, higher frequencies are

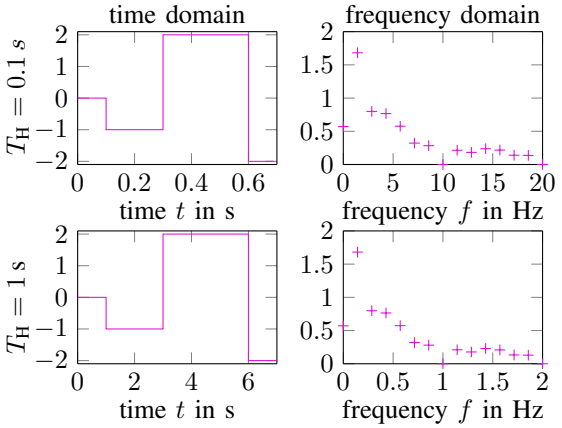

Fig. 3. Impact of the holding time T_H

excited. Since the amplitude levels and ratios between the lengths of held levels are identical for both inspected holding times, the amplitude spectra can be compared. Figure 3 shows that the shape of both spectra is identical. It is only stretched by the same factor that relates the holding times. A longer holding time results in longer signals and therefore in more data points given identical sampling times.

C. Assessment Criteria

The assessment criteria are divided into two groups. On the one hand test signal properties are analyzed. These include the coverage in the time domain, the crest factor and the signal energy. On the other hand it is of great interest whether the used signals can be used to obtain suitable identification data. Therefore two measures of the prediction error are used. The RMSE is used as an averaging measure as well as the MAE as a localized measure.

It is obvious that the crest factor is lower for a signal that is distributed more uniformly. A measure proportional to the energy of a signal $\{x_k\}$ is calculated in the frequency domain from the fourier transform $\{X_k\} = \mathrm{DFT}\,(\{x_k\})$:

$$E = \sum_{k=1}^{N} X_k^2 \tag{21}$$

For a measured signal y_k and its prediction $\hat{y}_k$ of length N the *root mean squared error* (RMSE) is calculated as follows:

$$J_\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{k=1}^{N}\left(y_k - \hat{y}_k\right)^2} \tag{22}$$

The *maximum absolute error* (MAE) is defined as:

$$J_\mathrm{MAE} = \max_{1\le k\le N} |y_k - \hat{y}_k| \tag{23}$$

The definitions of (22) and (23) are taken from [10].

IV. Case Study

As a case study an electro-mechanical throttle is used [11].

A. Test stand

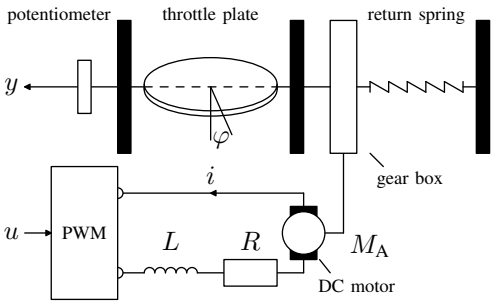

Fig. 4. Test stand schematic

Figure 4 shows a schematic of the test stand. The input signal u is the duty cycle for the pulse width modulator (PWM). The output signal y is the potentiometer signal. The throttle is actuated by a DC motor. Even though the setup is relatively simple, the modeling task is difficult because of the state-dependent friction as well as the nonlinear characteristic of the return spring. In addition physical parameters like moments of inertia are typically not known. The test stand is operated with a sampling time of $T_\mathrm{S} = 0.01\,\mathrm{s}$.

B. Excitation signals

Multisine and APRBS signals have been used to excite the system. For the multisine signals upper frequency bounds of $f_{\mathrm{u},1} = 5\,\mathrm{Hz}$, $f_{\mathrm{u},2} = 6\,\mathrm{Hz}$ and $f_{\mathrm{u},3} = 7.5\,\mathrm{Hz}$ have been used. They provide for an experiment duration of $T_{\mathrm{E},1} = 50\,\mathrm{s}$ and $T_{\mathrm{E},2} = 100\,\mathrm{s}$. The exited frequencies are 10 times the prime integers of the respective base frequencies, $f_{01} = 0.01\,\mathrm{Hz}$ for the longer experiment and $f_{02} = 0.02\,\mathrm{Hz}$ for the shorter experiment, up to the upper frequencies. For the longer experiment the frequency bound of $f_{\mathrm{u},2} = 6\,\mathrm{Hz}$ has been dropped. The amplitudes have been chosen to be identical but scaled up in a way to use the available range of the input signal as good as possible. Figure 5 shows an exemplary multisine signal and its amplitude spectrum.

For the APRBS signals different amplitude realizations (uniform and normal distribution as well as a sobol sequence) have been used with different holding times $T_{\mathrm{H},1} = 0.05\,\mathrm{s}$, $T_{\mathrm{H},1} = 0.1\,\mathrm{s}$ and $T_{\mathrm{H},1} = 0.2\,\mathrm{s}$. Figure 6 shows an exemplary APRBS signal and figure 7 shows the independent validation signal. To be able to compare the identification results, the prediction error on an independent validation set has been used. This set is used for validation in industrial applications and also has been introduced as a benchmark set in [11].

C. Identification approach and model evaluation

The number of local models $c = 8$ has been determined by cluster validation measures in [12]. All local models have the same structure. For the input signal the maximum lag is chosen

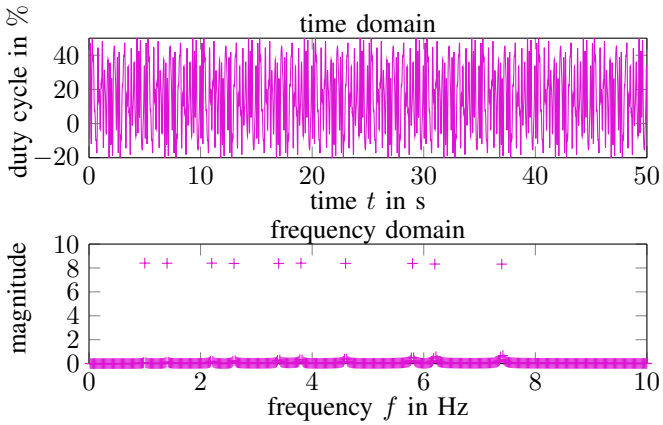

Fig. 5. Exemplary multisine, $f_o = 7.5\,\text{Hz}$

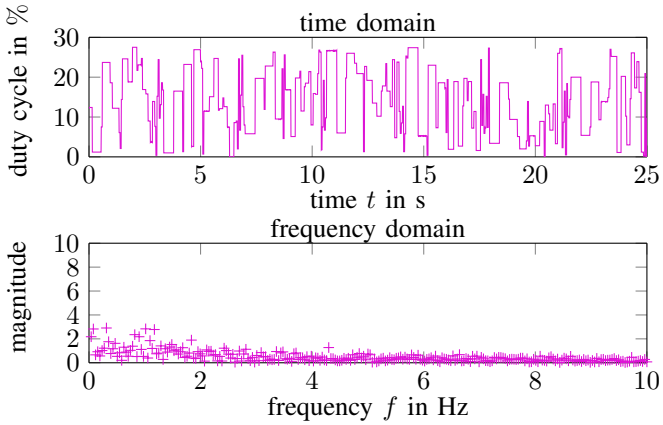

Fig. 6. Exemplary APRBS, $T_H = 0.1\,\text{s}$

to be $n_u = 1$, the maximum lag of the output signal is chosen to be $n_y = 2$, because the second order mechanical part is a lot faster than the electric part. The lags are identical for the antecedent and the consequent variables. The regression vector is:

$$\boldsymbol{\varphi}^{\mathrm{T}}(k-1) = \begin{bmatrix} u_{k-1} & y_{k-1} & y_{k-2} \end{bmatrix} \tag{24}$$

Since the friction and therefore a significant nonlinearity is velocity dependent, the scheduling variable is a linear combination of the regressor.

$$\mathbf{z}(k-1) = \begin{bmatrix} 1 & 0 \\ 0 & -1 \\ 0 & 1 \end{bmatrix} \cdot \boldsymbol{\varphi}^{\mathrm{T}} = \begin{bmatrix} u_{k-1} & y_{k-2} - y_{k-1} \end{bmatrix} \tag{25}$$

The fuzziness parameters for the clustering, estimation and model evaluation have been chosen to identically:

$$\nu_{\text{clust}} = \nu_{\text{est}} = \nu_{\text{eval}} = \nu = 1.1 \tag{26}$$

The models are evaluated in NOE-configuration. Since the maximum lag is $n_{\max} = 2$ the simulations are initialized with the first two values of the measured signal.

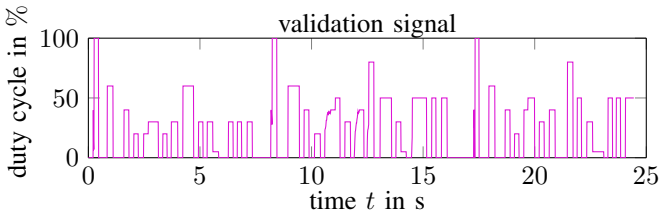

Fig. 7. Validation signal

D. Identification results

The homogenization method leads to a better coverage of the scheduling space as is shown in figure 8. In comparison to

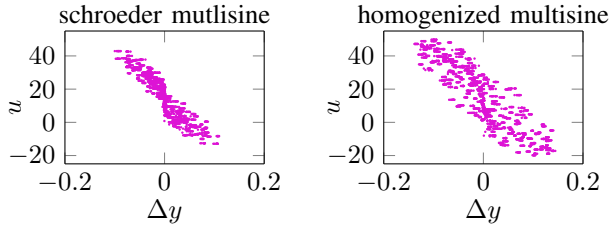

Fig. 8. Scheduling space comparison

the Schroeder multisines the homogenized signals in general yield better identification results. Figure 9 shows the measured system response to the multisine from figure 5 as well as the simulations with initially estimated parameters and the optimized parameters and the distribution of the residuals. Figure 10 shows the simulation results on the validation data.

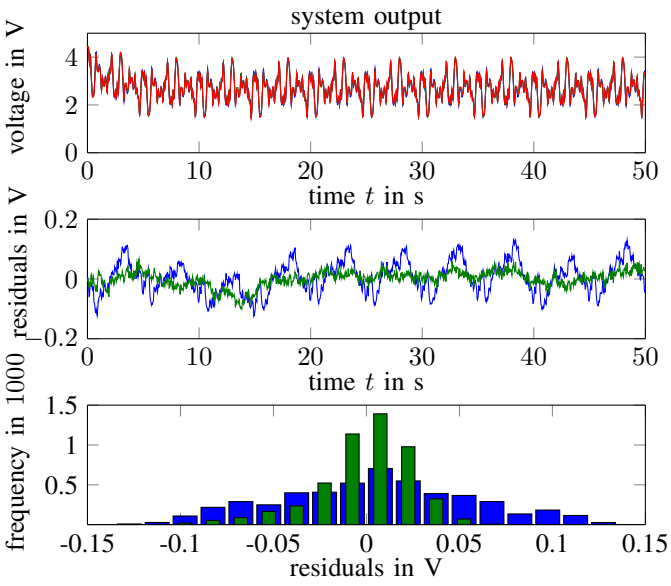

Fig. 9. Multisine identification signal, measured output (red) and predicted output for osap parameters (blue) and optimized parameters (green)

The residuals of the validation are larger on average than the

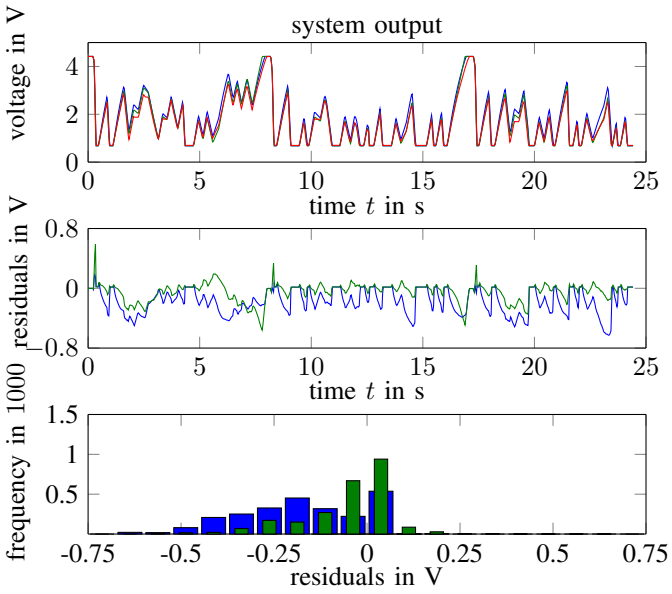

Fig. 10. Validation of the multisine identification, measured output (red) and predicted output for osap parameters (blue) and optimized parameters (green)

			J_{RMSE} in V			
R	duration	energy	id/osap	id/opt	val/osap	val/opt
S1	50 s	342.84	0.11	0.03	0.43	0.29
S2	50 s	342.60	0.07	0.02	0.23	0.28
S3	50 s	473.15	0.06	0.02	0.28	0.85
H1	50 s	342.85	0.07	0.02	0.17	0.14
H2	50 s	342.05	0.06	0.03	0.23	0.65
H3	50 s	494.65	0.05	0.03	0.23	0.13
S1	100 s	324.92	0.09	0.04	0.22	0.24
S3	100 s	364.29	0.14	0.04	0.43	0.29
H1	100 s	364.29	0.07	0.04	0.10	0.14
H3	100 s	325.01	0.06	0.04	0.18	0.11
A1	25 s	233.40	0.12	0.02	0.26	0.13
A1	50 s	234.32	0.12	0.06	0.27	0.21
A1	100 s	234.91	0.19	0.09	0.28	0.12
A2	25 s	209.07	0.18	0.03	0.12	0.14
A2	50 s	210.64	0.21	0.08	0.12	0.36
A2	100 s	211.32	0.25	0.07	0.16	0.29
A3	25 s	216.32	0.10	0.03	0.13	0.23
A3	50 s	217.54	0.11	0.04	0.17	0.31
A3	100 s	218.06	0.21	0.08	0.21	0.28

TABLE II
MULTISINE AND APRBS RESULTS

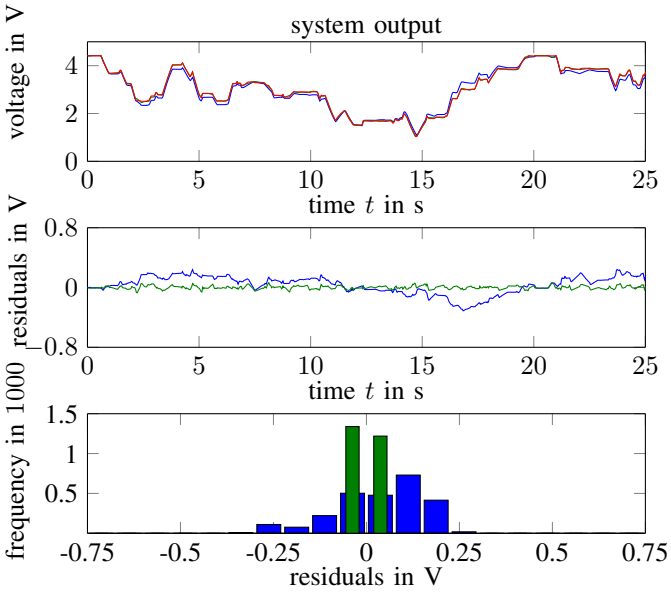

Fig. 11. APRBS identification, measured output (red) and predicted output for osap parameters (blue) and optimized parameters (green)

residuals on the identification data. The identified model is able to describe the validation data, even though the validation signal type is completely different from the identification signal type. The first half of table II shows an array of multisine identification results. The first column shows whether the used signal is a Schroeder multisine (S) or a homogenized one (H), where the number denotes the upper frequency. For the energy (21) the DC value has been excluded because it is identical for all signals. The following columns show the various RMSE values for the recursive model evaluation on the identification data as well as the validation data with respect to the OLS-estimated parameters and the parameters resulting from the nonlinear optimization of the summed squared prediction error in recursive model evaluation. The lower half of table II shows the APRBS results where the first column denotes the different amplitude initializations ($A1$: uniform distribution, $A2$: normal distribution, $A3$: Sobol-Sequence) Figure 11 shows the system response and simulations for the APRBS signal in figure 6. Figure 12 shows the simulation results for the validation data set. Since the holding time in this example is very small ($T_{\text{H}} = 0.05\,\text{s}$), the response doesn't look like the typical response to a multi-level signal. Because the APRBS is more similar to the validation signal, an optimization with respect to the identification data set is more likely to result in a better prediction as it can be seen in figure 12.

Table II shows the results of the APRBS identification with respect to different random amplitude realizations as well as the holding times T_{H}, because they directly correspond with the experiment duration. It can be seen that the amplitude realization has a great impact on the performance on the validation data. The optimization on the first amplitude realization results in a better average prediction error for every holding time. The performance of realizations 2 and 3 gets worse on the validation data after optimization. It also can be observed that with an exception for the optimized parameters based on the first realization on the validation data that a longer holding time in general results in a worse identification for this case study. Since the holding time is inversely proportional to the range of the the excited frequencies it is concluded that the electro-mechanical throttle is best excited with a short holding time. The signal energy does not change significantly with respect to the holding times because for equal amplitude levels the same work is done. Figure 13 shows that due to the longer holding time the amplitudes of the excited frequencies have a major drop off after 1 Hz whereas the amplitudes of the

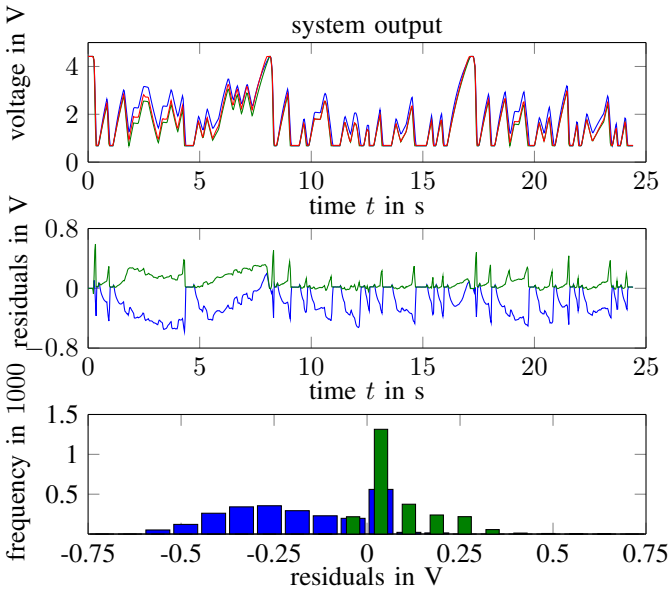

Fig. 12. Validation of the APRBS identification, measured output (red) and predicted output for osap parameters (blue) and optimized parameters (green)

excited frequencies of the signal with the very short holding time only drop off around the experimentally determined upper band limit of 5 Hz.

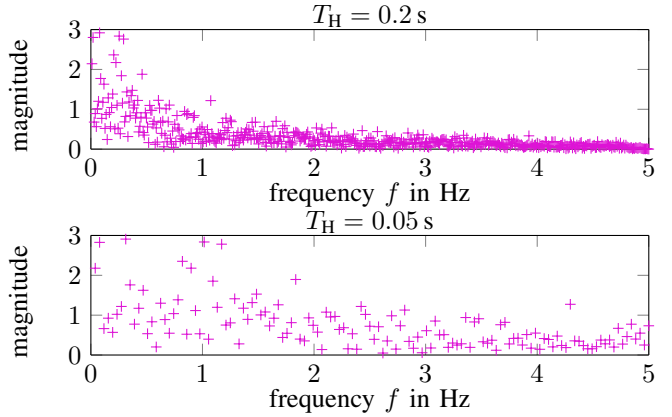

Fig. 13. Amplitude spectrum with respect to holding time

V. Conclusions & Outlook

In this paper parametrizations of the standard test signals, multisine and APRBS, have been conducted and analyzed. Often standard test signals are used with minimum adaptation to particular applications. The presented standard test signals have a defined structure. On the one hand this impedes the ability to adjust the signal properties because they are coupled with respect to the signal's parameters. On the other hand is it very difficult to design an input signal from scratch. In this contribution properties like the signal length and energy as well as coverage of scheduling space have been investigated. It can be concluded from the lower half of table II that the choice of the amplitude levels has a major impact on the excitation of a system with APRBS signals. The realizations of the amplitude levels of the APRBS signals are generated without considering any metrics beforehand. The design of the levels still has to be investigated. Since the inserted energy depends on the choice of amplitudes and not on the signal duration (when designed as proposed), the holding time can be used to target the frequency band of interest. The properties of a multisine are easily adjusted compared to an APRBS signal. The usage of a multisine or APRBS signal often is not a choice of the operator because it can be prohibited to use test signals so that the actuators are in constant motion. Therefore it is important to be able to tune the signal properties of an APRBS when multisines cannot be used. Furthermore the validation data is very similar to an APRBS. It can be seen in table II that the prediction error of the APRBS-based models is lower for the initially estimated parameters. When using standard test signals it is advised to choose a signal type close to the operational excitation.

To further validate the developed methods will be applied to other systems with different underlying physics as well as theoretical systems where the true parameters are known.

References

[1] C. Hametner, M. Stadlbauer, M. Deregnaucourt, M. Jakubek, T. Winsel, "Optimal experiment design based on local model networks and multiplayer perceptron networks," in *Engineering Applications of Artificial Intelligence*, vol. 26, no. 1, pp. 251–261, 2013.

[2] A. Kroll, A. Dürrbaum, "On joint optimal experiment design for identifying partition and local model parameters of Takagi-Sugeno models," in *Proc. 17th IFAC Symposium on System Identification (SysID)*, Beijing, China, pp. 1427–1432, 2015.

[3] A. Dürrbaum, A. Kroll, "On robust experiment design for identifying locally affine Takagi-Sugeno models," in *Proc. IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, Budapest, Hungary, 2016.

[4] M. Gringard, A. Kroll, "On the homogenization of broadband test signals for nonlinear system identification and the application to an electo-mechanical actuator [in German]," in *Proc. 25th Workshop Computational Intelligence*, Dortmund, Germany, pp. 145–162, 2015.

[5] A. Kroll, "On choosing the fuzziness parameter for identifying TS models with multidimensional membership functions," *Journal of Artificial Intelligence and Soft Computing Research*, vol. 1, no. 4, pp. 283–300, 2011.

[6] R. Pintelon, J. Schoukens, *System identification: A frequency domain approach,* 2nd edition. New York: Wiley-IEEE Press, 2012.

[7] S. Duym, J. Schoukens, "Design of excitation signals for the restoring force surface method," *Mechanical Systems and Signal Processing*, vol. 9, no. 2, pp. 139–158, 1995.

[8] M. Schroeder, "Synthesis of low-peak-factor signals and binary sequences with low autocorrelation," in *IEEE Transactions on Information Theory*, vol. 16, no. 1, pp. 85–89, 1970.

[9] K. Godfrey, *Perturbation Signals for System Identification,* 1st edition. Prentice Hall, 1993.

[10] A. Kroll, H. Schulte, "Benchmark problems for nonlinear system identification and control using Soft Computing methods: Need and overview," *Applied Soft Computing*, vol. 25, pp. 496–513, 2014.

[11] S. Zaidi, A. Kroll, "Electro-mechanical throttle as a benchmark problem for nonlinear system identification with friction," in *Proc. 24. Workshop Computational Intelligence*, Dortmund, Germany, pp. 173–186, KIT Scientific Publishing, 2014.

[12] Z. Ren, A. Kroll, F. Sofsky, F. Laubenstein, "On identification of piecewise-affine models for systems with friction and its application to electro-mechanical throttles," in *Proc. 16th IFAC Symposium on System Identification*, Brussels, pp. 1395–1400, 2012.