

# On the Initialization of Nonlinear LFR Model Identification with the Best Linear Approximation

Maarten Schoukens\* Roland Tóth\*

\* Control Systems Group, Dept. of Electrical Eng., Eindhoven  
University of Technology, Postbus 513, 5600 MB Eindhoven, The  
Netherlands. (e-mail: {m.schoukens,r.toth}@tue.nl).

---

**Abstract:** Balancing the model complexity and the representation capability towards the process to be captured remains one of the main challenges in nonlinear system identification. One possibility to reduce model complexity is to impose structure on the model representation. To this end, this work considers the linear fractional representation framework. In a linear fractional representation the linear dynamics and the system nonlinearities are modeled by two separate blocks that are interconnected with one another. This results in a structured, yet flexible model structure. Estimating such a model directly from input-output data is not a trivial task as the involved optimization is nonlinear in nature. This paper proposes an initialization scheme for the model parameters based on the best linear approximation of the system and shows that this approach results in high quality models on a set of benchmark data sets.

*Keywords:* Nonlinear Identification, Neural Network, Linear Fractional Representation, Best Linear Approximation

---

## 1. INTRODUCTION

While the real world is nonlinear and time-varying, so far, we often treated it as linear and time invariant (LTI) when we model and control real-life systems. However, due to, for instance, increasing performance demands, introduction of more light-weight structures and increasing constraints on energy consumption of systems, nonlinear models and nonlinear control has become increasingly important (Schoukens and Ljung, 2019).

A wide range of nonlinear modelling approaches and frameworks is available to the user: a good overview of the variety of available methods is given by (Schoukens and Ljung, 2019). Including structure and prior knowledge has proven to be of key importance to obtain high-quality nonlinear models. While linear models are represented by a hyperplane in a high-dimensional regression space, nonlinear models are a manifold in this high-dimensional space. Including structure in the considered model class reduces the complexity of the modelling problem and makes it a tractable task.

Structure can be imposed or included in the considered model class in various ways. It can be included as a prior in kernel based regression (Pillonetto et al., 2011; Birpoutsoukis et al., 2017). Alternatively, the nonlinear dynamics can be represented as an interconnection of LTI

blocks and static nonlinear blocks as is done in nonlinear block-oriented modelling approaches (Giri and Bai, 2010; Schoukens and Tiels, 2017). Another possibility is to mine a nonlinear input-output relation out of the data, while trading of model complexity with model accuracy (Khandelwal et al., 2019). Structure can also be imposed as a second step in the identification algorithm as in (Fakhrizadeh Esfahani et al., 2018) where first a fully coupled nonlinear state-space model is estimated and the structure is only imposed in a second step by restricting the input dimension of the nonlinear state and output mapping using tensor decomposition techniques.

This paper considers models that are composed by the interconnection of a linear fractional representation (LFR) with a static nonlinearity, resulting in the NL-LFR model class (see Figure 1). This model class can be seen as a very general block-oriented structure. Just like block-oriented structures, the nonlinear dynamics are represented as an interconnection between an LTI block and a static nonlinear block. However the inner dynamics of the multiple-input multiple-output (MIMO) LTI block are not imposed. This allows for more flexibility compared to the more common Hammerstein and Wiener block-oriented structures (Schoukens and Tiels, 2017). The NL-LFR model class also has a direct link with robust control and linear parameter-varying control design approaches (Zhou et al., 1996; Tóth, 2010; Schoukens and Tóth, 2018).

The identification of NL-LFR structures has already been considered in previous publications. Some approaches derive the NL-LFR model starting from other model structures such as a Volterra series model (Vandersteen and Schoukens, 1999), or a nonlinear state-space model (Van

---

\* Maarten Schoukens is supported by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Fellowship (grant agreement nr. 798627). R. Tóth is supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement nr. 714663).

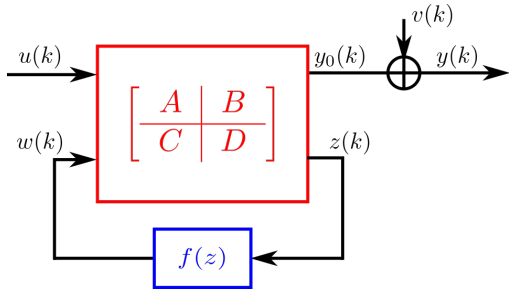


Fig. 1. The considered NL-LFR structure represented by MIMO LTI block interconnected with the MIMO static nonlinear function  $f(z)$ .

Mulders et al., 2013). Other approaches assume the knowledge of the LTI part of the NL-LFR model and focus on the identification of the static nonlinearity of the model (Hsu et al., 2008; Novara et al., 2011). Finally (Vanbeylen, 2013) starts the identification of the NL-LFR model from multiple linear approximations of the nonlinear systems. The main reason for the various restrictions of the considered identification problem, or on the required prior information (Volterra model, state-space model, multiple linear approximations) is due to the complexity of the involved parameter estimation problem.

This paper proposes an initialization approach of the nonlinear optimization scheme for the estimation of the NL-LFR structure that only requires a single linear approximation of the nonlinear system, which can be obtained using the best linear approximation (BLA) framework (Pintelon and Schoukens, 2012). This relaxes the assumed prior knowledge and data requirements of the identification algorithm significantly compared to the prior work available in the literature. The validity of the proposed approach is illustrated on two benchmark examples: the Bouc-Wen hysteretic system benchmark (Noël and Schoukens, 2016) and the parallel Wiener-Hammerstein benchmark system (Schoukens et al., 2015b).

The remainder of this paper first discusses the considered NL-LFR model structure in Section 2. The BLA is introduced in Section 3. Next, the proposed identification approach is discussed in Section 4. Finally, the results on the considered benchmark systems are discussed in Section 5.

## 2. NONLINEAR LFR MODEL STRUCTURE

### 2.1 Model Structure

The considered discrete-time nonlinear LFR model structure consists of a, possibly multiple input multiple output (MIMO), static nonlinearity interconnected with the linear fractional representation, as is shown in Figure 1. In this paper, the linear dynamics are represented using a state-space representation describing the dynamic relation between the LFR inputs  $u(k) \in \mathbb{R}^{n_u \times 1}$ ,  $w(k) \in \mathbb{R}^{n_w \times 1}$  and the outputs  $y(k) \in \mathbb{R}^{n_y \times 1}$ ,  $z(k) \in \mathbb{R}^{n_z \times 1}$ , where  $k$  denotes the sample index. The static nonlinearity is represented by a feedforward neural network with one hidden layer with  $n_n$  neurons using a nonlinear activation function  $\sigma(\cdot)$  and a linear output layer. This results in the following model equations:

$$\begin{aligned} x(k+1) &= Ax(k) + [B_u \ B_w] \begin{bmatrix} u(k) \\ w(k) \end{bmatrix} \\ \begin{bmatrix} z(k) \\ y_0(k) \end{bmatrix} &= \begin{bmatrix} C_z \\ C_y \end{bmatrix} x(k) + \begin{bmatrix} D_{zu} & 0 \\ D_{yu} & D_{yw} \end{bmatrix} \begin{bmatrix} u(k) \\ w(k) \end{bmatrix} \end{aligned} \quad (1)$$

and

$$w(k) = W_w \sigma(W_z z(k) + b_z) + b_w, \quad (2)$$

where  $\sigma(\cdot)$  is the nonlinear activation function which is often chosen as the hyperbolic tangent function or the radial basis function and  $W_z \in \mathbb{R}^{n_n \times n_z}$  and  $W_w \in \mathbb{R}^{n_n \times n_w}$  are the inner and outer weights respectively and  $b_z \in \mathbb{R}^{n_n \times 1}$  and  $b_w \in \mathbb{R}^{n_w \times 1}$  are the respective biases of the neural network. The states are represented by  $x(k) \in \mathbb{R}^{n_x \times 1}$ , while the matrices  $A \in \mathbb{R}^{n_x \times n_x}$ ,  $B_u \in \mathbb{R}^{n_x \times n_u}$ ,  $B_w \in \mathbb{R}^{n_x \times n_w}$ ,  $C_y \in \mathbb{R}^{n_y \times n_x}$ ,  $C_z \in \mathbb{R}^{n_z \times n_x}$ ,  $D_{zu} \in \mathbb{R}^{n_z \times n_u}$ ,  $D_{yu} \in \mathbb{R}^{n_y \times n_u}$  and  $D_{yw} \in \mathbb{R}^{n_y \times n_w}$  correspond to the parameters to be estimated together with the weights in eq. (2). Note that the term  $D_{zw}$  is not present in the considered state-space equation. This prevents the presence of algebraic equations in the model expression and hence avoiding well-posedness problems.

The nonlinear modelling capabilities of the NL-LFR structure can be tuned by increasing or decreasing the dimension of the  $z$  and  $w$  signals. This allows the model structure to go from very structured (only one SISO static nonlinearity present in the model) to rather unstructured (a high dimensional MIMO static nonlinearity). Observe as well that many of the commonly used block-oriented model structures such as the Wiener, Hammerstein and Hammerstein-Wiener model structures are subsets of the NL-LFR model structure. Furthermore by making use of a state-space representation of the dynamics, the model scales easily towards multiple inputs  $u$  and outputs  $y$ .

A zero-mean, possibly colored, additive noise source  $v(k)$  is assumed to present at the output  $y(k)$  only. The noise source is assumed to have a finite variance. In case the noise is white, this corresponds to the classical output-error noise framework.

### 2.2 Uniqueness of the Parametrization

The considered model representation is not unique. Beyond the well-known arbitrary state-space transformation that defines an equivalence class of models around (1)-(2), also the neural network representation of the static nonlinearity is not uniquely parametrized. A simple permutation of the inner and other weights and biases can result in the same input-output behavior of the neural network (2). Beyond the non-uniqueness of the LTI and static nonlinear blocks separately, a linear gain can also be exchanged between the static nonlinear and the LTI block (Schoukens and Tiels, 2017).

## 3. BEST LINEAR APPROXIMATION

The Best Linear Approximation (BLA) of a nonlinear system is an LTI approximation of the input-output map of the system, in a mean square sense (Pintelon and Schoukens, 2012; Enqvist, 2005). The BLA is obtained as:

$$\begin{aligned}
 G_{\text{bla}}(q) &= \arg \min_{G(q)} E_{u,v} \left\{ |\tilde{y}(k) - G(q)\tilde{u}(k)|^2 \right\}, \\
 \tilde{u}(k) &= u(k) - E_u \{u(k)\}, \\
 \tilde{y}(k) &= y(k) - E_{u,v} \{y(k)\},
 \end{aligned} \tag{3}$$

where  $E_{u,v} \{\cdot\}$  denotes the expected value operator taken w.r.t. the random variations due to the input  $u(k)$  and the output noise  $v(k)$ ,  $q^{-1}$  denotes the backwards shift operator. As can be observed in eq. (3), the BLA of a nonlinear system is dependent on the properties of the considered input class (Schoukens et al., 2015a).

#### 4. IDENTIFICATION OF A NONLINEAR LFR MODEL

##### 4.1 Parameter Estimation

The model parameters are obtained as the minimization of the mean squared simulation error:

$$V_N(\theta) = \frac{1}{N} \sum_{k=1}^N (y(k) - \hat{y}(k|\theta))^2, \tag{4}$$

$$\hat{\theta} = \arg \min_{\theta} V_N(\theta), \tag{5}$$

where  $\hat{y}(k|\theta)$  is the simulated output of the NL-LFR model given the parameter vector  $\theta$ . The parameter vector  $\theta$  contains all the model parameters: the state-space matrix entries and the weights and biases of the neural network representing the static nonlinearity.  $N$  represents the total number of samples over which the cost function is computed.

Since eq. (4) is typically nonlinear in the parameters and not convex, but its gradients can be efficiently computed, a gradient descent-type algorithm, like the Levenberg-Marquardt algorithm (Levenberg, 1944) is used to minimize the cost function. A 'data-driven coordinate frame' is used to get rid off equivalent gradient directions (resulting in a rank deficient Jacobian) due to the non-uniqueness of the model representation. In practice this is achieved using the singular value decomposition of the Jacobian (Wills and Ninness, 2008).

##### 4.2 Parameter Initialization

The Levenberg-Marquardt algorithm is not guaranteed to converge to the global minimum of the cost function, it converges to the 'closest' local minimum. The use of 'efficient' initial estimates to start the nonlinear optimization can help to guide the optimization algorithm to the global minimum.

The BLA estimate of a nonlinear system has proven to be a good initial point to start the identification of a nonlinear model (Paduart et al., 2010; Vanbeylen, 2013; Schoukens and Tiels, 2017). While the previous BLA-based approach (Vanbeylen, 2013) requires 2 BLA estimates at 2 different setpoints of the system, this work proposes an initialization procedure starting from only one state-space BLA estimate ( $A_{\text{bla}}, B_{\text{bla}}, C_{\text{bla}}, D_{\text{bla}}$ ).

First the BLA state-space matrices are transformed such that each of the states has a unit variance:

$$\begin{aligned}
 \bar{A}_{\text{bla}} &= TA_{\text{bla}}T^{-1}, & \bar{B}_{\text{bla}} &= TB_{\text{bla}} \\
 \bar{C}_{\text{bla}} &= C_{\text{bla}}T^{-1}, & \bar{D}_{\text{bla}} &= D_{\text{bla}}
 \end{aligned} \tag{6}$$

where  $T \in \mathcal{R}^{n_x \times n_x}$  is a diagonal matrix with the inverse of the standard deviation (taken over the samples) of  $x$  on its diagonal.

Next the BLA estimates are embedded into the NL-LFR model, while the other parameters are initialized as zero or as a random variable:

$$\begin{aligned}
 A &= \bar{A}_{\text{bla}} & B_u &= \bar{B}_{\text{bla}} \\
 C_y &= \bar{C}_{\text{bla}} & D_{yu} &= \bar{D}_{\text{bla}} \\
 B_w &= 0 & D_{yw} &= 0 \\
 W_w &\sim \mathcal{U}(-1, 1) & W_z &\sim \mathcal{U}(-1, 1) \\
 b_w &= 0 & b_z &\sim \mathcal{U}(-1, 1)
 \end{aligned} \tag{7}$$

where  $\mathcal{U}(a, b)$  denotes a uniformly distribution with a support from  $a$  to  $b$ . The uniform random initialization of the parameters is the common approach when training a neural network (Bishop, 1995). The initialization of  $C_z$  and  $D_{zu}$  is slightly different:

$$\begin{aligned}
 C_* &\sim \mathcal{U}(-1, 1) \\
 D_+ &\sim \mathcal{U}(-1, 1) & D_* &= D_+T_u^{-1} \\
 z_* &= C_*x + D_*u \\
 C_z &= T_{z_*}^{-1}C_* & D_{zu} &= T_{z_*}^{-1}D_*
 \end{aligned} \tag{8}$$

where  $T_{z_*} \in \mathcal{R}^{n_z \times n_z}$  and  $T_u \in \mathcal{R}^{n_u \times n_u}$  are a diagonal matrices with the inverse of the standard deviation (taken over the samples) of  $z_*$  and  $u$  respectively on its diagonal.

The transformations  $T, T_u, T_{z_*}$  ensure that the initial estimates of the  $x$  and  $z$  signals, which act as the input of the neural network, have a standard deviation equal to one and are zero-mean. This is generally recognized in the neural network literature to improve the estimation of the model parameters (Bishop, 1995). The initial model has the same performance as the BLA estimate since  $B_w$  and  $D_{yw}$  are initialized as zero matrices. It also ensures that, if the BLA estimate is stable, the initial estimate of the NL-LFR model is stable as well.

##### 4.3 BLA Identification

Many different approaches are available in the literature to estimate the BLA. For example, nonparametric and parametric frequency-domain methods (Pintelon and Schoukens, 2012; Paduart et al., 2010) or the prediction error method (Ljung, 1999) have been commonly used.

This paper directly estimates a state-space model of the BLA using the time-domain prediction error method as implemented in Matlab by the function `ssest` with the signals  $u(k)$  and  $y(k)$  as the input and output data respectively. This function initializes the parameter estimates using either a subspace approach or an iterative rational function estimation approach. The state-space matrices are refined subsequently using the prediction error minimization approach (Ljung, 1999).

When the system is strongly nonlinear, obtaining a sufficiently good estimate of the BLA might not be trivial. The lack of a good BLA estimate can result in sub-optimal NL-LFR model estimates. Worse, for specific setpoints of the system the BLA can be equal to zero (Schoukens and Tiels, 2017). The case when the BLA only captures part of the

system dynamics will again result in sub-optimal NL-LFR model estimates.

## 5. BENCHMARK RESULTS

Two benchmark datasets are considered: the Bouc-Wen Hysteretic system (Noël and Schoukens, 2016) and the parallel Wiener-Hammerstein datasets (Schoukens et al., 2015b).

The Bouc-Wen system is a hysteretic system featuring a dynamic nonlinearity: it is a mass-spring-damper system with a hysteretic restoring force. This hysteretic nonlinearity is governed by a differential equation containing hard nonlinearities such as the absolute value operator.

The parallel Wiener-Hammerstein system is obtained as the parallel cascade of two Wiener-Hammerstein systems (a linear input filter followed by a static nonlinearity, followed by a linear output filter). The system has 12th order dynamics ( $n_x = 12$ ), which is challenging for many black-box identification algorithms. The static nonlinearities are realized by a diode-resistor network resulting in a one-sided and a two-sided saturation nonlinearity.

### 5.1 Bouc-Wen Hysteretic System

*The Data:* The Bouc-Wen system is available as a simulation script. This allows the users to generate their own data for identification. We used one period of a random phase multisine input signal (Pintelon and Schoukens, 2012). The signal was  $N = 8192$  samples long and it excites the full frequency grid between 5 Hz and 150 Hz. The input signal amplitude is  $50 N_{\text{rms}}$ .

Two test datasets are available for the Bouc-Wen benchmark: a multisine and a sinesweep dataset. The multisine test output is obtained as the steady-state response of the system excited by a random phase multisine of  $N = 8192$  samples long, exciting the full frequency grid between 5 and 150 Hz, with a signal amplitude of  $50 N_{\text{rms}}$ . The sinesweep test output is obtained by exciting the system, starting from zero initial conditions, with a sinesweep signal with an amplitude of  $40 N_{\text{rms}}$ . The frequency band from 20 to 50 Hz is covered at a sweep rate of 10 Hz/min.

*Model Settings:* In line with previous finding the linear dynamics are described by a 3rd order state-space model ( $n_x = 3$ ) (Noël et al., 2017). From the system description (Noël and Schoukens, 2016) it can be concluded that the 2  $z$ -variables and one  $w$ -variable should result in a suitable model structure to capture the system dynamics. However, for illustrative purposes, three cases are considered here:  $\{n_z = 1, n_w = 1\}$ ,  $\{n_z = 2, n_w = 1\}$  and  $\{n_z = 2, n_w = 2\}$ . A total of 15 neurons ( $n_n = 15$ , tansig activation function) are used to represent the static nonlinearity.

*Results:* As a measure of model quality, the simulation RMSE (root mean squared error) on the multisine and sinesweep test dataset is reported:

$$e_{\text{RMSE}} = \sqrt{\frac{1}{N} \sum_{k=1}^N (y(k) - \hat{y}(k|\theta))^2}, \quad (9)$$

where  $y(k)$  is the observed test output and  $\hat{y}(k|\theta)$  is the simulated output using the estimated model. To make

sure the model output is in steady state for the multisine dataset two periods are simulated and the RMSE is calculated on the second period. The sinesweep data is not in steady state, hence, the first 2000 samples are ignored when calculating the RMSE to allow the transient to decay.

Table 1 shows a comparison of the results obtained using an LTI model (the model used to initialize the NL-LFR parameter optimization) and the various NL-LFR cases considered. It can be observed that the most simple NL-LFR configuration is not sufficiently rich to capture the complete nonlinear behaviour of the system, only a factor 3 reduction of the model error is obtained compared to the LTI case. The two cases where  $n_z = 2$  on the other hand succeed in reducing the model error with a factor 20 compared to the BLA model, indicating that this is a much more suited model structure for the Bouc-Wen benchmark system. This is indeed in line with the expectations as this matches with the model equations reported in (Noël and Schoukens, 2016). Both NL-LFR models with  $n_z = 2$  perform similar, which is an indication that the model structure with  $n_z = 2$  and  $n_w = 1$  is matching best with the underlying system structure.

The test results for the LTI model and the NL-LFR model with  $n_z = 2$  and  $n_w = 1$  are shown in Figures 2 and 3. It is apparent from the figures that the NL-LFR model significantly outperforms the LTI model. The signal and error behavior outside the excited frequency range indicate that the obtained model error is very close to the noise floor. The quality of the results are, to the authors knowledge among the best black-box identification results obtained on this dataset so far, the RMSE is 2-3 times lower than the one reported in (Fakhrizadeh Esfahani et al., 2018).

Table 1. Bouc-Wen benchmark results: simulation RMSE. The first column shows the results obtained with a 3rd order LTI model, the next three columns show the results obtained with the NL-LFR for different  $n_z$  and  $n_w$  values. The final column shows the results that are reported in (Fakhrizadeh Esfahani et al., 2018) using a decoupled polynomial nonlinear state-space approach. The rms errors (9) are reported in  $m_{\text{rms}}$ .

	LTI	$n_z = 1$ $n_w = 1$	$n_z = 2$ $n_w = 1$	$n_z = 2$ $n_w = 2$	decoupled PNLSS
Multisine	$15.8e^{-5}$	$5.31e^{-5}$	$0.72e^{-5}$	$0.74e^{-5}$	$1.34e^{-5}$
Sinesweep	$17.7e^{-5}$	$4.19e^{-5}$	$0.32e^{-5}$	$0.56e^{-5}$	$1.12e^{-5}$

### 5.2 Parallel Wiener-Hammerstein System

*The Data:* A detailed discussion of the estimation and test data for the parallel Wiener-Hammerstein system is given in (Schoukens et al., 2015b). This work used a subset of the available estimation data: 1 steady-state period of a random phase multisine signal of 16384 samples long at 5 different input amplitudes. The test data consists of another multisine realization at each of the 5 different amplitudes and a random Gaussian noise sequence with a linearly growing amplitude over time (denoted as the arrow signal later on).

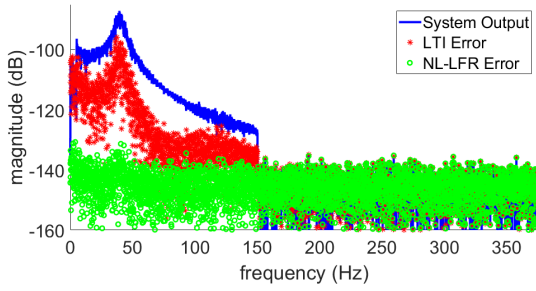


Fig. 2. Frequency-domain validation of the NL-LFR model with  $n_z = 2$  and  $n_w = 1$  using the Bouc-Wen multisine test data. The true system output is shown in blue, the residuals obtained with a 3rd order LTI model are shown in red and the residuals obtained with the NL-LFR model are shown in green.

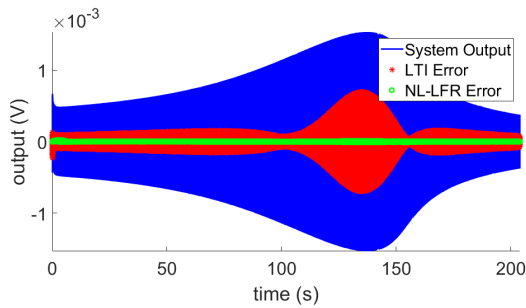


Fig. 3. Time-domain validation of the NL-LFR model with  $n_z = 2$  and  $n_w = 1$  using the Bouc-Wen sinesweep test data. The true system output is shown in blue, the residuals obtained with an LTI model are shown in red and the residuals obtained with the NL-LFR model are shown in green.

*Model Settings:* As described in (Schoukens et al., 2015b) the linear dynamics are described by a 12th order state-space model ( $n_x = 12$ ) and 2 parallel nonlinear branches are present in the system. This indicates 2  $z$ -variables and 2  $w$ -variable should result in the 'correct' model structure. A total of 15 neurons ( $n_n = 15$ , tanisig activation functions) are used to represent the static non-linearity.

*Results:* Table 2 shows a comparison of the results obtained using an LTI model and the obtained NL-LFR model for the 6 different test signals. Also the results obtained using the parallel Wiener-Hammerstein method described in (Schoukens et al., 2015b) are shown for comparison. Similar to the previous case study, to make sure the model output for both the multisine and arrow dataset is in steady state two periods are simulated, the RMSEs are calculated on the second period. It can be observed that the NL-LFR model outperforms the LTI models even though the LTI model has been specifically trained for each of the multisine input amplitudes separately.

The test results for the LTI model and the NL-LFR model on the largest multisine amplitude and on the arrow signal are shown in Figures 4 and 5. It is apparent from the figures that the NL-LFR model significantly outperforms the LTI model.

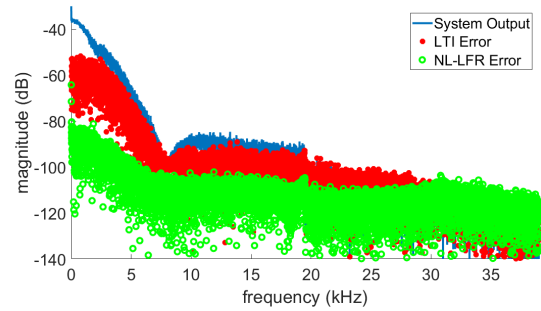


Fig. 4. Frequency-domain validation of the NL-LFR model using the parallel Wiener-Hammerstein multisine test data. The true system output is shown in blue, the residuals obtained with a 12th order LTI model are shown in red and the residuals obtained with the NL-LFR model are shown in green.

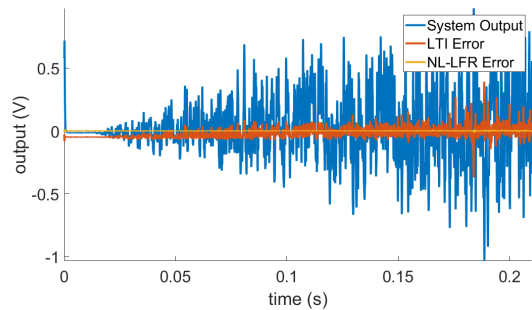


Fig. 5. Time-domain validation of the NL-LFR model using the parallel Wiener-Hammerstein arrow test data. The true system output is shown in blue, the residuals obtained with a 12th order LTI model are shown in red and the residuals obtained with the NL-LFR model are shown in green.

The obtained results are in line with the results reported in (Schoukens et al., 2015b) where a parallel Wiener-Hammerstein model has been fitted on the data. Even though the specific Wiener-Hammerstein nature of the system has not been imposed on the NL-LFR model, a similar model quality has been obtained.

Table 2. The parallel Wiener-Hammerstein benchmark results: simulation RMSE. The first column shows the results obtained with a 12th order LTI model specifically estimated for that amplitude. The LTI model obtained for the 1000 mV<sub>rms</sub> input signal also has been used for the arrow test signal. The second and third column shows the results obtained with the NL-LFR model (this paper) and a parallel Wiener-Hammerstein approach (Schoukens et al., 2015b). The rms errors (9) are reported in V<sub>rms</sub>.

	LTI	NL-LFR	pWH
Multisine 100 mV <sub>rms</sub>	$0.83e^{-3}$	$0.30e^{-3}$	$0.30e^{-3}$
Multisine 325 mV <sub>rms</sub>	$9.46e^{-3}$	$0.51e^{-3}$	$0.50e^{-3}$
Multisine 550 mV <sub>rms</sub>	$19.9e^{-3}$	$0.72e^{-3}$	$0.38e^{-3}$
Multisine 775 mV <sub>rms</sub>	$30.7e^{-3}$	$1.03e^{-3}$	$0.57e^{-3}$
Multisine 1000 mV <sub>rms</sub>	$37.7e^{-3}$	$1.37e^{-3}$	$1.10e^{-3}$
Arrow Signal	$45.5e^{-3}$	$1.42e^{-3}$	$2.66e^{-3}$

### 5.3 Discussion

The obtained results on the Bouc-Wen and parallel Wiener-Hammerstein benchmark datasets illustrate the versatile nature of the NL-LFR model structure and of the proposed identification algorithm. High-quality modelling results comparable with the state-of-the-art have been obtained even though the two systems that were considered here are significantly different in nature. Note as well that due to the structured nature of the NL-LFR, it has little problem in handling nonlinear systems with high-order dynamics such as the parallel Wiener-Hammerstein system if the dimension of the  $z$  and  $w$  signals can be kept under control.

## 6. CONCLUSION AND FUTURE WORK

The problem of identifying an NL-LFR model of a nonlinear system has been addressed in this paper. The NL-LFR model structure offers a flexible, yet structured model class which is a generalization of the popular block-oriented Hammerstein and Wiener model class. The proposed initialization of the nonlinear-in-the-parameters non-convex optimization associated with the identification problem, starting from a BLA estimate of the system, has proven to be successful on the two considered benchmark problems. A detailed theoretical analysis of the proposed identification approach is currently lacking and will be the subject of future work.

### REFERENCES

- Birpoutsoukis, G., Marconato, A., Lataire, J., and Schoukens, J. (2017). Regularized nonparametric volterra kernel estimation. *Automatica*, 82, 324–327.
- Bishop, C.M. (1995). *Neural Networks for Pattern Recognition*. Oxford University Press, Inc., New York, NY, USA.
- Enqvist, M. (2005). *Linear Models of Nonlinear systems*. Ph.D. thesis, Institute of technology, Linköping University, Sweden.
- Fakhrizadeh Esfahani, A., Dreesen, P., Tiels, K., Noël, J.P., and Schoukens, J. (2018). Parameter reduction in nonlinear state-space identification of hysteresis. *Mechanical Systems and Signal Processing*, 104, 884–895.
- Giri, F. and Bai, E. (eds.) (2010). *Block-oriented Nonlinear System Identification*, volume 404 of *Lecture Notes in Control and Information Sciences*. Springer-Verlag, London.
- Hsu, K., Poolla, K., and Vincent, T.L. (2008). Identification of structured nonlinear systems. *IEEE Transactions on Automatic Control*, 53(11), 2497–2513.
- Khandelwal, D., Schoukens, M., and Toth, R. (2019). Data-driven modelling of dynamical systems using tree adjoining grammar and genetic programming. In *2019 IEEE Congress on Evolutionary Computation, CEC 2019 - Proceedings*, 2673–2680. Wellington, New Zealand.
- Levenberg, K. (1944). A method for the solution of certain problems in least squares. *Quarterly of Applied Mathematics*, 2, 164–168.
- Ljung, L. (1999). *System Identification: Theory for the User (second edition)*. Prentice Hall, Upper Saddle River, New Jersey.
- Noël, J.P., Esfahani, A., Kerschen, G., and Schoukens, J. (2017). A nonlinear state-space approach to hysteresis identification. *Mechanical Systems and Signal Processing*, 84(B), 171–184.
- Noël, J.P. and Schoukens, M. (2016). Hysteretic benchmark with a dynamic nonlinearity. In *Workshop on Nonlinear System Identification Benchmarks*, 7–14. Brussels, Belgium.
- Novara, C., Vincent, T.L., Hsu, K., Milanese, M., and Poolla, K. (2011). Parametric identification of structured nonlinear systems. *Automatica*, 47, 711–721.
- Paduart, J., Lauwers, L., Swevers, J., Smolders, K., Schoukens, J., and Pintelon, R. (2010). Identification of nonlinear systems using polynomial nonlinear state space models. *Automatica*, 46(4), 647–656.
- Pillonetto, G., Quang, M.H., and Chiuso, A. (2011). A new kernel-based approach for nonlinear system identification. *IEEE Transactions on Automatic Control*, 56(12), 2825–2840.
- Pintelon, R. and Schoukens, J. (2012). *System Identification: A Frequency Domain Approach*. Wiley-IEEE Press, Hoboken, New Jersey, 2nd edition.
- Schoukens, J. and Ljung, L. (2019). Nonlinear system identification: A user-oriented road map. *IEEE Control Systems Magazine*, 39(6), 28–99.
- Schoukens, J., Pintelon, R., Rolain, Y., Schoukens, M., Tiels, K., Vanbeylen, L., Van Mulders, A., and Vandersteen, G. (2015a). Structure discrimination in block-oriented models using linear approximations: A theoretic framework. *Automatica*, 53, 225–234.
- Schoukens, M., Marconato, A., Pintelon, R., Vandersteen, G., and Rolain, Y. (2015b). Parametric identification of parallel Wiener-Hammerstein systems. *Automatica*, 51(1), 111–122.
- Schoukens, M. and Tiels, K. (2017). Identification of block-oriented nonlinear systems starting from linear approximations: A survey. *Automatica*, 85, 272–292.
- Schoukens, M. and Tóth, R. (2018). From nonlinear identification to linear parameter varying models: Benchmark examples. In *18th IFAC Symposium on system identification (SYSID)*. Stockholm, Sweden.
- Tóth, R. (2010). *Modeling and Identification of Linear Parameter-Varying Systems*, volume 403 of *Lecture Notes in Control and Information Sciences*. Springer-Verlag, Berlin Heidelberg.
- Van Mulders, A., Schoukens, J., and Vanbeylen, L. (2013). Identification of systems with localised nonlinearity: From state-space to block-structured models. *Automatica*, 49(5), 1392–1396.
- Vanbeylen, L. (2013). Nonlinear LFR Block-Oriented Model: Potential Benefits and Improved, User-Friendly Identification Method. *IEEE Transactions on Instrumentation and Measurement*, 62(12), 3374–3383.
- Vandersteen, G. and Schoukens, J. (1999). Measurement and identification of nonlinear systems consisting of linear dynamic blocks and one static nonlinearity. *IEEE Transactions on Automatic Control*, 44(6), 1266–1271.
- Wills, A. and Ninness, B. (2008). On gradient-based search for multivariable system estimates. *IEEE Transactions on Automatic Control*, 53(1), 298–306.
- Zhou, K., Doyle, J.C., and Glover, K. (1996). *Robust and Optimal Control*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA.