

# Identification of the Silverbox Benchmark Using Nonlinear State-Space Models

Anna Marconato\* Jonas Sjöberg\*\* Johan Suykens\*\*\*  
Johan Schoukens\*

\* Dept. ELEC, Vrije Universiteit Brussel,  
Pleinlaan 2, 1050 Brussels, Belgium.  
(e-mail: anna.marconato@vub.ac.be)

\*\* Dept. of Signals and Systems, Chalmers University of Technology,  
SE412 96 Gothenburg, Sweden.

\*\*\* Dept. ESAT-SCD-SISTA, Katholieke Universiteit Leuven,  
Kasteelpark Arenberg 10, 3001 Leuven (Heverlee), Belgium.

---

**Abstract:** This work presents the application of an initialization scheme for nonlinear state-space models on a real data benchmark example: the Silverbox problem. The goal of the proposed approach is to transform the identification of a nonlinear dynamic system into an approximate static problem, so that system dynamics and nonlinear terms are identified separately. Classic identification techniques are used to handle dynamics, while regression methods from the statistical learning community are introduced to estimate the nonlinearities in the model. Results obtained on the Silverbox problem are discussed and compared with the performance of other related methods.

Keywords: State-space models; benchmark examples; system identification; nonlinear systems; dynamic systems; neural networks.

---

## 1. INTRODUCTION

The field of nonlinear system identification is a very active research area and presents still many challenges and open problems (Ljung [2010]). Various methods have been developed to deal with the difficult task of modeling nonlinear (dynamic) systems, including a whole set of grey shades between white-box and black-box models (Ljung [2010], Sjöberg et al. [1995]), nonparametric and parametric techniques (Rugh [1981], Billings and Fakhouri [1982]), block-oriented approaches (Giri and Bai [2010]), just to mention a few examples.

The challenge becomes particularly interesting when one wants to model a real system, for which knowledge about the problem might not be entirely available, based on a set of measurement data that are typically corrupted by noise. In this framework, a good opportunity to learn about advantages and disadvantages of several identification methods is given by considering real-life benchmark problems, on which the different approaches can be tested. Traditionally, a number of special sessions in identification and control conferences have been organized around challenging benchmark examples (see e.g. Schoukens et al. [2009b], Cham et al. [2010] for recent examples).

This paper discusses the application of an identification technique for nonlinear state-space models on a benchmark problem of a special session organized at the IFAC Symposium on Nonlinear Control Systems (NOLCOS) in 2004: the Silverbox.

The Silverbox is an electrical circuit describing the behavior of a mass-spring-damper system. It is an example of

nonlinear dynamic system with feedback, where the linear contributions are dominant (Schoukens et al. [2003]).

Six papers were presented in the Silverbox special session, describing different types of approaches, including white-box (Hjalmarsson and Schoukens [2004], Paduart et al. [2004]), grey-box (Ljung et al. [2004]), and black-box methods (Espinoza et al. [2004], Sragner et al. [2004], Verdult [2004], Ljung et al. [2004]), and since then other techniques have been applied to solve the Silverbox problem (e.g. Paduart [2008], Van Mulders et al. [2011]). A discussion on these methods and their performance on the Silverbox case study is provided at the end of this paper.

Among the many possible examples in the class of black-box approaches, the identification of state-space models can lead to some advantages. For example, this general model structure describes naturally system dynamics and is proven to be equivalent to several classic block-oriented models (including models with feedback) (Paduart [2008]). However, identifying nonlinear state-space models is still considered a difficult task.

In this work a method to efficiently obtain good initial estimates for nonlinear state-space models is presented, combining system identification techniques (Pintelon and Schoukens [2001]) with regression methods from the statistical learning community (Hastie et al. [2009]). The proposed approach targets systems that are weakly nonlinear and systems that are assumed to have only one equilibrium point.

This paper is organized as follows. Section 2 presents more in details the problem to be solved and describes

the proposed technique for the identification of nonlinear state-space models, with focus on a three-steps procedure for the initialization of the model. The application of the method on the Silverbox identification problem is discussed in Section 3. The obtained results are compared in Section 4 with the ones provided by other related methods that have been applied to the same benchmark data. Some final remarks conclude the paper in Section 5.

## 2. IDENTIFICATION OF NONLINEAR STATE-SPACE MODELS

### 2.1 Considered problem

In this work we deal with the problem of identifying nonlinear state-space models of the form:

$$x(t+1) = f(x(t), u(t)) \quad (1)$$

$$y(t) = g(x(t), u(t)) \quad (2)$$

where  $u(t) \in \mathbb{R}^{n_u}$  and  $y(t) \in \mathbb{R}^{n_y}$  are the given input and output signals at time  $t$ ,  $x(t) \in \mathbb{R}^{n_x}$  is the unknown state of the system, and  $f(\cdot)$  and  $g(\cdot)$  are the nonlinear functions that one wants to estimate with the identification procedure.

To find a good model, a cost function (e.g. the root mean square error between the modeled output and the true output signal) is typically minimized with respect to the parameters of the model. When the parameters appear nonlinearly in the cost function, the optimization problem is solved iteratively by means of algorithms such as the Levenberg-Marquardt technique (Ljung [1999], Pintelon and Schoukens [2001]). Therefore, an important issue is how to generate good starting values for the model parameters, so that the search for the cost function minimum can converge fast to a good solution.

In this paper, an initialization scheme for the identification of nonlinear state-space models is proposed, based on the idea of separating the estimation of the system dynamics and the modeling of the nonlinearities in Eqs. (1-2) (see also Marconato et al. [2012]). To achieve this, one needs to cut the recursion in the state equation (1), in order to obtain a static version of the nonlinear dynamic problem, that can be solved more efficiently.

More in details, assuming that the nonlinear dynamics in (1-2) are modeled as:

$$\begin{aligned} x(t+1) &= f(x(t), u(t)) = \\ &= Ax(t) + Bu(t) + f_{NL}(x(t), u(t)) \end{aligned} \quad (3)$$

$$\begin{aligned} y(t) &= g(x(t), u(t)) = \\ &= Cx(t) + Du(t) + g_{NL}(x(t), u(t)) \end{aligned} \quad (4)$$

where  $A \in \mathbb{R}^{n_x \times n_x}$ ,  $B \in \mathbb{R}^{n_x \times n_u}$ ,  $C \in \mathbb{R}^{n_y \times n_x}$ ,  $D \in \mathbb{R}^{n_y \times n_u}$  and nonlinear terms  $f_{NL}(\cdot)$  and  $g_{NL}(\cdot)$  have  $n_x$  and  $n_y$  outputs respectively, an approximate static version of the same identification problem can be formulated as follows, starting from an estimate  $\hat{x}(t)$  of the state  $x(t)$ :

$$\begin{aligned} \hat{x}(t+1) &= f(\hat{x}(t), u(t)) + r(t) = \\ &= A\hat{x}(t) + Bu(t) + f_{NL}(\hat{x}(t), u(t)) + r(t) \end{aligned} \quad (5)$$

$$\begin{aligned} y(t) &= g(\hat{x}(t), u(t)) + e(t) = \\ &= C\hat{x}(t) + Du(t) + g_{NL}(\hat{x}(t), u(t)) + e(t) \end{aligned} \quad (6)$$

Since  $\hat{x}(t)$  differs from the true state  $x(t)$ , error terms  $r(t)$  and  $e(t)$  are included in the state and in the output equation respectively.

At this point, Eqs. (5-6) can be regarded as two different static regression problems, that can be addressed independently, and can therefore be solved much more efficiently. In this work, regression methods from the statistical learning community will be employed to model nonlinear terms  $f_{NL}$  and  $g_{NL}$ . Several examples of nonlinearities can be easily applied to solve this task (Hastie et al. [2009], Suykens et al. [1995, 1996]); here multilayer perceptrons (MLPs) will be used to illustrate the method, and to solve the Silverbox benchmark problem.

A nonlinear search algorithm is finally applied to optimize all model parameters.

In the following, the complete algorithm for the identification of nonlinear state-space models, consisting of an initialization phase and an optimization phase, is presented in details.

### 2.2 Initialization scheme

First of all, an initialization procedure is proposed, that exploits the separation between linear dynamics and nonlinearities resulting from the transformation of problem (3-4) into problem (5-6).

It consist of three main steps:

*Estimation of a linear model.* As a first step, estimates  $\hat{A}$ ,  $\hat{B}$ ,  $\hat{C}$ ,  $\hat{D}$  of the linear part of the model are obtained, to capture system dynamics and linear behavior. In particular, the Best Linear Approximation (BLA), which is optimal in least square sense, of the nonlinear system is estimated and transformed into state-space form (Pintelon and Schoukens [2001]).

*Approximation of the nonlinear state.* Based on the BLA and on the set of available input/output data  $\{u(t), y(t)\}_{t=1}^N$ , an approximation of the unknown nonlinear state  $x(t)$  can be determined, so that it will be possible to cut the recursion in the state equation, as shown in Section 2.1.

The approximate state  $\hat{x}(t)$  is estimated by solving the following Least Squares problem:

$$\begin{aligned} \hat{x}(t) &= \arg \min_{\{x(t)\}} \sum_t (y(t) - \hat{C}x(t) - \hat{D}u(t))^2 \\ &\quad + \lambda \sum_t (x(t+1) - \hat{A}x(t) - \hat{B}u(t))^2 \end{aligned} \quad (7)$$

where the first term of the cost function represents the fit to data, and the second term represents the fit to the linear model; the balance between the two criteria can be set by changing the value of the trade-off parameter  $\lambda$ .

One could of course obtain the linear state directly from the BLA, but here the idea is to allow a deviation from the linear state (by changing the value of  $\lambda$ ), to take into account the nonlinear terms in Eqs. (3-4).

Kalman filtering could also be considered as an alternative to approximate the state.

Notice that this state approximation is used only to initialize the nonlinear model; a nonlinear optimization of all model parameters is performed at a later stage.

*Modeling of the nonlinearities.* Once the BLA and the approximate state are determined, the identification task can be solved by considering the static version of the problem in Eqs. (5-6). Any regression method can be used at this point to estimate nonlinear terms  $f_{NL}$  and  $g_{NL}$ . The statistical learning community offers many examples of nonlinearities that can be simply plugged in the considered problem.

In particular here MLPs with one hidden layer and  $\tanh(\cdot)$  as activation function are considered to illustrate the main idea, but the method could be extended to include other choices of nonlinear functions.

### 2.3 Nonlinear optimization of the model parameters

Once the previous three steps are performed, and therefore initial values for the parameters of both linear and nonlinear parts of the model are determined, the initialized nonlinear state-space model can be simulated, reintroducing the recursion in the state equation, i.e. going back to the nonlinear dynamic problem in Eqs. (3-4).

Finally, after initialization, all model parameters need to be optimized applying a nonlinear optimization algorithm, e.g. a Levenberg-Marquardt technique.

## 3. THE SILVERBOX EXAMPLE

### 3.1 Description of the data

The Silverbox is an electrical circuit simulating a mass-spring-damper system. The system's behavior can be approximately described by the following equation:

$$m\ddot{y}(t) + d\dot{y}(t) + k_1y(t) + k_3y^3(t) = u(t) \quad (8)$$

where  $u(t)$  represents the input force applied to the mass  $m$  and output  $y(t)$  is the mass displacement. Parameters  $k_1$  and  $k_3$  describe the (nonlinear) behavior of the spring, and  $d$  is the damping of the system (Pintelon and Schoukens [2001]).

The dataset was generated by exciting the system with the input signal shown in Figure 1.

The first part of the excitation signal is a filtered Gaussian noise sequence with increasing RMS value. This sequence is made of 40000 samples and is used for the final testing of the models. The second part of the excitation consists of ten different realizations of an odd random phase multisine (Schoukens et al. [2009a]), each realization consisting of 8192 samples (plus 500 transient samples). This portion of the data is used to estimate the models. For both the estimation and the test data, the excitation signal has a bandwidth of 200 Hz, and an RMS value equal to 22.3 mV (notice that for the test data the signal is non-stationary, so this value indicates the average RMS value). The considered sampling frequency is 610.35 Hz. Notice that since the amplitude of the last part of the test data is bigger than the amplitude of the data portion used for estimation, extrapolation issues will be present when testing the models on this part of the data.

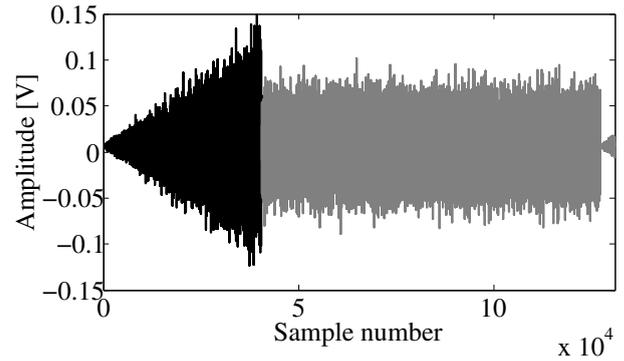


Fig. 1. Excitation signal for the Silverbox problem. The first part (in black) is a filtered Gaussian noise signal used as test set. The second part (in grey) is an odd random phase multisine signal used to estimate the models.

### 3.2 Initialization of the nonlinear state-space model

The initialization scheme proposed in Section 2.2 is now applied to the benchmark data, to obtain initial estimates for the nonlinear state-space model parameters.

*Estimation of the BLA.* First of all, a linear model is determined to describe the dynamic behavior of the system. The BLA is computed by exploiting the fact that 10 realizations of the multisine signal are available. More in details, the FRF of the different realizations is first computed and then the average over the realizations is taken (Pintelon and Schoukens [2001]). Transient terms are removed by means of a local polynomial approximation (Pintelon et al. [2010]). Starting from this nonparametric BLA, a parametric second order linear model is estimated, and transformed into state-space form, so that  $\hat{A}$  is a  $2 \times 2$  matrix,  $\hat{B}$  is a  $2 \times 1$  vector,  $\hat{C}$  is a  $1 \times 2$  vector, and  $\hat{D}$  is a scalar value (see Eqs. (3-4)).

The error of the BLA model on the final test data is shown in Figure 2, together with the test output. The obtained linear model gives a RMSE equal to 7.2 mV on the estimation data and 13.5 mV on the test data (notice the difference in the error values, due to the presence of an extrapolation region in the test data, where large spikes of the error signal are located).

*Note.* In the following, since the algorithm will not make explicit use of the multisine properties of the excitation signal, this splitting of the data is considered:

- *estimation data:* the first portion of the multisine data (40000 samples) will be used to build the models;
- *validation data:* the second portion of the multisine data (45000 samples) will be used to validate the models during model selection;
- *test data:* the Gaussian noise part of the data (40000 samples) will be used to assess the performance of the models.

*Approximation of the nonlinear state.* To approximate the unknown nonlinear state  $x(t)$ , the Least Squares problem in (7) was solved. The trade-off parameter  $\lambda$  was tuned choosing from a grid of values, by observing the trend of the RMSE (on the validation data) of the obtained

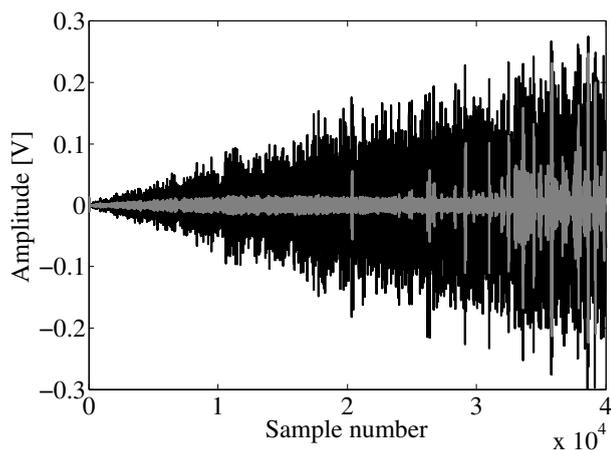


Fig. 2. Error signal given by the BLA model on the test data (grey line), and test output signal (black line).

Model	$n_f$	$n_g$	Number of parameters	Est. RMSE	Val. RMSE
A	3	2	40	2.5	2.3
B	2	2	34	4.5	4.1
C	1	1	23	6.5	6.1
BLA	-	-	9	7.2	6.9

Table 1. Initialized models. Number of parameters and RMSE values (in mV) on estimation and validation data are shown, for different choices of the number of neurons for  $f_{NL}$  ( $n_f$ ) and  $g_{NL}$  ( $n_g$ ). As a reference, the BLA result is also shown.

initialized nonlinear models. The lowest RMSE values were obtained by fixing  $\lambda = 0.01$ .

*Modeling of the nonlinearities.* Using the estimated  $\hat{x}(t)$ , the identification problem is transformed into the static form (5-6). One-hidden-layer MLPs with  $\tanh(\cdot)$  as activation function are used to model the nonlinear terms  $f_{NL}$  and  $g_{NL}$ .

*Model selection.* In order to reduce the RMSE value of the initialized model on the validation data, and at the same time keeping the number of parameters in the model low, several numbers of neurons have been taken into account during model selection. For each choice of number of neurons, 10 different initializations of the MLPs were considered, and the best resulting initialized model in terms of RMSE on the validation data was selected.

In Table 1 the results obtained with the initialized models on the different data portions are summarized. The total amount of parameters resulting from the different choices for the number of neurons is also shown.

Although the initialized model C with only one neuron in the hidden layer for both  $f_{NL}$  and  $g_{NL}$  gives a poor result in terms of RMSE (it does not really improve much the performance of the BLA estimate), it was decided to optimize the model parameters starting from all the three initialized models A, B and C, to try to obtain a good final fitted model characterized by a small number of parameters.

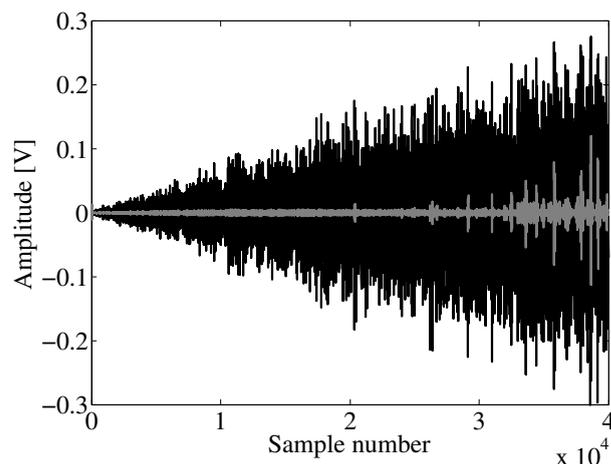


Fig. 3. Error signal given by initialized model A on the test data (grey line), and test output signal (black line).

Model	Number of parameters	Est. RMSE	Val. RMSE	Test RMSE
A	40	0.22	0.23	0.50
B	34	0.25	0.25	0.49
C	23	0.24	0.24	0.34

Table 2. Final fitted models. Number of parameters and RMSE values (in mV) on estimation, validation and test data are shown.

In Figure 3 the error on the test data of the initialized model A (the best among the initialized models) is plotted, together with the test output. Notice again the spikes in the extrapolation region.

### 3.3 Optimization of all model parameters

Starting from the three initialized models, a Levenberg-Marquardt algorithm was applied to optimize all model parameters. The optimization algorithm was let run for 1000 iterations, although in most cases only approximately 100 iterations were needed for convergence.

The results obtained in terms of RMSE by the final fitted models are summarized in Table 2.

All the three fitted models give very low values of the RMSE on the validation data, including model C for which the initial parameter values gave poor results. Since the goal was to obtain a model with a good trade-off between low RMSE on the validation data and reduced number of parameters, model C is preferred to model A, given that its RMSE is slightly higher, but the number of parameters is almost 50% smaller.

The higher RMSE on the test data is due only to the difficulty of the models to extrapolate on the last portion of the test data. If results on the first 30000 samples (data for which no extrapolation occurs) are considered, it can be observed that for the three models the RMSE drops to the same value as for the validation data. However, observe that model C achieves good results also when the RMSE is computed on the whole test data set. The error plot for model C on the test data is shown in Figure 4.

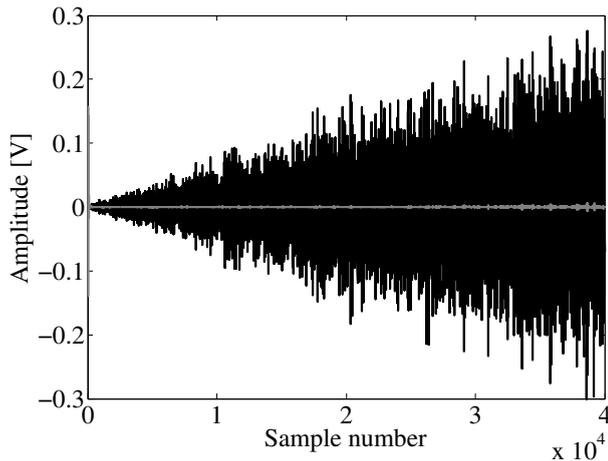


Fig. 4. Error signal given by model C (after optimization) on the test data (grey line), and test output signal (black line).

## 4. DISCUSSION

### 4.1 Obtained results

As seen in the previous section, the application of the proposed initialization scheme for nonlinear state-space models gives satisfactory results on the Silverbox problem. In particular, low values of the RMSE were obtained for the final fitted models, starting from initialized models that improved the performance of the linear model. Moreover, it was possible to significantly reduce the number of parameters by considering less neurons in the hidden layer, obtaining the best result in terms of RMSE on the test data, with an optimized model characterized by only 23 parameters.

A note on extrapolation issues: the considered nonlinearity did not always allow to generalize very well in the extrapolation region. This is due to the fact that there is a mismatch with the intrinsic nonlinear structure of the underlying system (i.e. a cubic nonlinearity of the output in the feedback loop, as shown in Eq. (8)). The advantage of including in the model the same kind of nonlinear terms will be clear from the discussion of the performance of other methods applied to the Silverbox. It should be underlined that, although this was one of the features of the benchmark problem, it is not completely fair to compare the methods on test data that include an extrapolation part, since achieving good extrapolation is actually a matter of luck.

### 4.2 Related methods

In the special session at NOLCOS'04 different methods were applied to the Silverbox benchmark. A brief overview on those techniques is provided here, that will also help to get a better understanding of the approach proposed in this paper.

Two white-box approaches, exploiting information about the physical structure of the device, were presented. In (Hjalmarsson and Schoukens [2004]) the internal model structure of the Silverbox is transformed into a MISO Hammerstein system, to which a relaxation algorithm is

applied to estimate the model parameters. The obtained model is characterized by only 5 parameters and results in a RMSE value of 0.95 mV.

In (Paduart et al. [2004]) a LTI system with a cubic nonlinearity in the feedback loop is identified, obtaining a RMSE of 0.38 mV on the test data, with a model characterized by only 10 parameters.

The work in (Ljung et al. [2004]) describes a number of grey-box and black-box methods applied to the Silverbox. Different types of nonlinearities are considered, including a one-hidden-layer sigmoidal neural network. The results are significantly improved if a cubic regressor is embedded in the nonlinear models. The best result in terms of RMSE (0.30 mV) is achieved with a neural network model containing more than 700 parameters.

Two other contributions make use of nonlinearities from the statistical learning community. In (Espinoza et al. [2004]) several examples of Least Squares Support Vector Machines (LS-SVMs) are compared on the benchmark problem. Using standard LS-SVMs solved in the dual space requires subsampling of the estimation data, since this method can not deal with very large data sets. Fixed Size LS-SVMs solved in the primal space obtain the best performance on the test set (0.32 mV), with a fixed number of 500 support vectors. In all cases the use of polynomial kernels gives better results than radial basis functions, given the dominant linear behavior of the device.

Several examples of neural networks are considered in (Sragner et al. [2004]), resulting in large RMSE values on the test data (the best model achieves 7.8 mV), and high number of parameters (around 600).

In (Verdult [2004]) a weighted combination of local linear state-space models is used to approximate the nonlinear behavior of the system, considering radial basis functions as weighing vectors. A limited portion of the data is considered due to computational complexity constraints. The best obtained model, characterized by 16 parameters, achieves a RMSE of 1.3 mV on the test data.

More recently, two nonlinear state-space methods have been applied to the Silverbox problem. In (Paduart [2008]), a Polynomial Nonlinear State-Space (PNLSS) model with a cubic nonlinearity is proposed, obtaining a RMSE equal to 0.26 mV, with 37 parameters. The work in (Van Mulders et al. [2011]) aims at reducing the number of parameters of a PNLSS model, by transforming it into a general block-structured model. Using that approach, it is possible to decrease the number of parameters to 12, with a RMSE on the test data equal to 0.35 mV.

Table 3 summarizes some of the best results obtained on the Silverbox problem with the methods discussed above.

The method proposed in this work performs comparably with the other approaches, showing a good trade-off between low RMSE and low number of parameters. Notice that the best model in this sense (model C) is characterized by a significantly lower number of parameters if compared with the other methods that make use of nonlinear functions from the statistical learning community. Moreover, it can be observed that, among the different approaches that obtain a RMSE lower than 1 mV on the test data,

Approach	Number of parameters	Test RMSE
Hjalmarsson and Schoukens [2004]	5	0.95
Paduart et al. [2004]	10	0.38
Ljung et al. [2004]	>700	0.30
Espinoza et al. [2004]	500 (SVs)	0.32
Verdult [2004]	16	1.3
Paduart [2008]	37	0.26
Van Mulders et al. [2011]	12	0.35
This paper	23	0.34

Table 3. Comparison of different methods applied to the Silverbox benchmark. The RMSE values are expressed in mV.

the proposed method is the only one that does not include explicitly a polynomial representation of the nonlinearity to solve the problem, meaning that it was possible to achieve a low RMSE on the test data (including the extrapolation region) without using any *a priori* knowledge on the internal structure of the device.

## 5. CONCLUSION

This paper has presented a method based on ideas from both system identification and statistical learning for the initialization of nonlinear state-space models. To be able to exploit nonlinear regression methods such as MLPs, the identification of nonlinear dynamic systems is transformed into a static problem by cutting the recursion in the state equation. The proposed approach has been applied to the Silverbox identification problem, obtaining good results both in terms of RMSE and of number of parameters.

## ACKNOWLEDGEMENTS

This work is sponsored by the Fund for Scientific Research (FWOVlaanderen), the Flemish Government (Methusalem Fund, METH1) and the Belgian Federal Government (IAP VI/4). J. Suykens acknowledges support from K.U.Leuven, the Flemish Government, FWO and the Belgian federal science policy office (CoE EF/05/006, GOA MANET, IUAP DYSCO).

## REFERENCES

S.A. Billings and S.Y. Fakhouri. Identification of systems containing linear dynamic and static nonlinear elements. *Automatica*, 18(1):15–26, 1982.

C.L. Cham, A.H. Tan, and W.H. Tan. Hyperfast switching Peltier cooling system benchmark. In *UKACC International Conference on Control*, Coventry, UK, 2010.

M. Espinoza, K. Pelckmans, L. Hoegaerts, J.A.K. Suykens, and B. De Moor. A comparative study of LS-SVM's applied to the Silver box identification problem. In *IFAC Symposium on Nonlinear Control Systems (NOLCOS2004)*, Stuttgart, Germany, 2004.

F. Giri and E.W. Bai, editors. *Block-oriented Nonlinear System Identification*. Springer, 2010.

T. Hastie, R. Tibshirani, and J. Friedman. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer-Verlag, 2009.

H. Hjalmarsson and J. Schoukens. On direct identification of physical parameters in non-linear models. In *IFAC Symposium on Nonlinear Control Systems (NOLCOS2004)*, Stuttgart, Germany, 2004.

L. Ljung. Perspectives on system identification. *Annual Reviews in Control*, 34:1–12, 2010.

L. Ljung. *System Identification: Theory for the User (2nd ed.)*. Prentice Hall, New Jersey, 1999.

L. Ljung, Q. Zhang, P. Lindskog, and A. Juditski. Estimation of grey box and black box models for non-linear circuit data. In *IFAC Symposium on Nonlinear Control Systems (NOLCOS2004)*, Stuttgart, Germany, 2004.

A. Marconato, J. Sjöberg, J.A.K. Suykens, and J. Schoukens. Separate initialization of dynamics and nonlinearities in nonlinear state-space models. In *IEEE International Instrumentation and Measurement Technology Conference*, Graz, Austria, 2012.

J. Paduart. *Identification of Nonlinear Systems using Polynomial Nonlinear State Space Models*. PhD Thesis, Vrije Universiteit Brussel, 2008.

J. Paduart, G. Horvath, and J. Schoukens. Fast identification of systems with nonlinear feedback. In *IFAC Symposium on Nonlinear Control Systems (NOLCOS2004)*, Stuttgart, Germany, 2004.

R. Pintelon and J. Schoukens. *System Identification. A Frequency Domain Approach*. IEEE Press, 2001.

R. Pintelon, J. Schoukens, G. Vandersteen, and K. Barbé. Estimation of nonparametric noise and frf models for multivariable systems - part i: Theory. *Mechanical Systems and Signal Processing*, 24(3):573–595, 2010.

W. Rugh. *Nonlinear System Theory - The Volterra/Wiener Approach*. The Johns Hopkins University Press, Baltimore, 1981.

J. Schoukens, G. Nemeth, P. Crama, Y. Rolain, and R. Pintelon. Fast approximate identification of nonlinear systems. *Automatica*, 39(7):1267–1274, 2003.

J. Schoukens, J. Lataire, R. Pintelon, G. Vandersteen, and T. Dobrowiecki. Robustness issues of the best linear approximation of a nonlinear system. *IEEE Trans. Instrumentation and Measurement*, 58(5):1737–1745, 2009a.

J. Schoukens, J.A.K. Suykens, and L. Ljung. Wiener-Hammerstein benchmark. In *15th IFAC Symposium on System Identification*, Saint-Malo, France, 2009b.

J. Sjöberg, Q. Zhang, L. Ljung, A. Benveniste, B. Delyon, P.-Y. Glorennec, H. Hjalmarsson, and A. Juditski. Nonlinear black-box modeling in system identification: a unified overview. *Automatica*, 31(12):1691–1724, 1995.

L. Sragner, J. Schoukens, and G. Horvath. Modelling of a slightly nonlinear system: a neural network approach. In *IFAC Symposium on Nonlinear Control Systems (NOLCOS2004)*, Stuttgart, Germany, 2004.

J.A.K. Suykens, B. De Moor, and J. Vandewalle. Nonlinear system identification using neural state space models, applicable to robust control design. *International Journal of Control*, 62(1):129–152, 1995.

J.A.K. Suykens, J. Vandewalle, and B. De Moor. *Artificial Neural Networks for Modeling and Control of Non-Linear Systems*. Springer, 1996.

A. Van Mulders, L. Vanbeylen, and J. Schoukens. Identification of systems with localised nonlinearity: from state-space to block-structured models. (*submitted to Automatica*), 2011.

V. Verdult. Identification of local linear state space models: the Silver box case study. In *IFAC Symposium on Nonlinear Control Systems (NOLCOS2004)*, Stuttgart, Germany, 2004.