Organizing Committee:

Jean-Philippe Noël, University of Liège, Belgium

Maarten Schoukens, Eindhoven University of Technology, The Netherlands

Organized with the support of:
# PROGRAM

## WEDNESDAY 11/04/2018

<table>
<thead>
<tr>
<th>Time</th>
<th>Event</th>
<th>Speaker(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>08.30 – 09.00</td>
<td>Registration</td>
<td></td>
</tr>
<tr>
<td>09.00 – 09.30</td>
<td>Welcome &amp; Introduction</td>
<td>J.P. Noël &amp; M. Schoukens</td>
</tr>
<tr>
<td>09.30 – 10.30</td>
<td>Keynote 1</td>
<td>E. Chatzi – ETH Zürich, Switzerland</td>
</tr>
<tr>
<td></td>
<td>Data-Driven Assessment of Engineered Systems: Beyond LTI</td>
<td></td>
</tr>
<tr>
<td>10.30 – 11.00</td>
<td>Coffee</td>
<td></td>
</tr>
<tr>
<td>11.00 – 12.30</td>
<td>Session 1</td>
<td>T.O. Heinz, T. Münker and O. Nelles - Identification of Systems with Hysteretic Behavior Using NOBF Local Model Networks</td>
</tr>
<tr>
<td></td>
<td></td>
<td>K. Karami, D. Westwick and J. Schoukens - Identification of Decoupled Polynomial NARX Models using Simulation Error Minimization</td>
</tr>
<tr>
<td></td>
<td></td>
<td>J. Decuyper, A. Fakhrizadeh Esfahani, K. Tiels and J. Schoukens - Retrieving highly structured models starting from a black box state-space model: a case study on the Bouc-Wen hysteresis model</td>
</tr>
<tr>
<td>12.30 – 13.30</td>
<td>Lunch</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Nonlinear cortical responses in EEG evoked by continuous wrist manipulations</td>
</tr>
<tr>
<td>14.30 – 15.00</td>
<td>Session 2</td>
<td>B. Peeters and P.Z Csurcsia - The use of exotic multisines in MIMO structural dynamics and acoustic applications</td>
</tr>
<tr>
<td>15.00 – 15.30</td>
<td>Coffee</td>
<td></td>
</tr>
<tr>
<td>15.30 – 16.30</td>
<td>Session 3</td>
<td>M. Perne and M. Stepančič - Regressor selection using Lipschitz quotients on the F-16 aircraft benchmark</td>
</tr>
<tr>
<td></td>
<td></td>
<td>M. Mazzoleni, M. Scandella and F. Previdi - Kernel manifold regression for the coupled electric drives dataset</td>
</tr>
<tr>
<td>16.30 – 17.00</td>
<td>Discussion Session 1</td>
<td></td>
</tr>
<tr>
<td>19.00 – …</td>
<td>Workshop Dinner</td>
<td>Antidote restaurant, Rue Souverain Pont 4</td>
</tr>
<tr>
<td>Time</td>
<td>Session</td>
<td>Presenters</td>
</tr>
<tr>
<td>--------------</td>
<td>----------------</td>
<td>-----------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>09.00 – 10.00</td>
<td>Keynote 3</td>
<td>J.A.K. Suykens – KU Leuven, Belgium</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Function estimation, model representations and nonlinear system identification</td>
</tr>
<tr>
<td>10.00 – 10.30</td>
<td>Session 4</td>
<td>M.R.-H. Abdalmoaty and H. Hjalmarsson - Application of a Linear PEM estimator to a stochastic Wiener-Hammerstein Benchmark Problem</td>
</tr>
<tr>
<td>10.30 – 11.00</td>
<td>Coffee</td>
<td></td>
</tr>
<tr>
<td>11.00 – 12.30</td>
<td>Session 5</td>
<td>G. Birpoutsoukis and M. Schoukens - From the Volterra series to a Wiener-Hammerstein model</td>
</tr>
<tr>
<td></td>
<td></td>
<td>M. Stepančič, M. Perne and J. Kocijan - Regularised NFIR identification with Gaussian process model</td>
</tr>
<tr>
<td></td>
<td></td>
<td>D. Bouvier, T. Hélie and D. Roze - Phase-based homogeneous order separation for improving Volterra series identification</td>
</tr>
<tr>
<td>12.30 – 13.30</td>
<td>Lunch</td>
<td></td>
</tr>
<tr>
<td>13.30 – 14.30</td>
<td>Keynote 4</td>
<td>P. Young – Lancaster University, UK</td>
</tr>
<tr>
<td></td>
<td></td>
<td>State-Dependent Parameter (SDP) Nonlinear Models and a Hydrological Identification Benchmark</td>
</tr>
<tr>
<td>14.30 – 15.00</td>
<td>Session 6</td>
<td>N. Simidjievski, L. Todorovski, S. Džeroski, and J. Kocijan - Nonlinear System Identification with Equation Discovery</td>
</tr>
<tr>
<td>15.00 – 15.30</td>
<td>Coffee</td>
<td></td>
</tr>
<tr>
<td>15.30 – 16.30</td>
<td>Session 7</td>
<td>P.Z. Csurcsia, J. Schoukens and B. Peeters - Nonparametric Approximation of the Nonlinear SilverBox Data: a Linear Time-varying Approach</td>
</tr>
<tr>
<td></td>
<td></td>
<td>J. Schoukens - Simulation and Prediction Errors in the Presence of Model Errors: a Case Study on the Silverbox</td>
</tr>
<tr>
<td>16.30 – 17.00</td>
<td>Discussion Session 2</td>
<td></td>
</tr>
</tbody>
</table>
### FRIDAY 13/04/2018

<table>
<thead>
<tr>
<th>Time</th>
<th>Session</th>
<th>Presentation</th>
</tr>
</thead>
<tbody>
<tr>
<td>09.00 – 10.00</td>
<td>Keynote 5</td>
<td>J. Gonçalves – University of Luxembourg, Luxembourg&lt;br&gt;Challenges in system identification of biochemical systems</td>
</tr>
<tr>
<td>10.00 – 10.30</td>
<td>Session 8</td>
<td>R. Hostettler, F. Tronarp and S. Särkkä - Nonparametric Drift Model for Stochastic Differential Equations</td>
</tr>
<tr>
<td>10.30 – 11.00</td>
<td>Coffee</td>
<td></td>
</tr>
<tr>
<td>11.00 – 12.00</td>
<td>Session 9</td>
<td>M. Schoukens and R. Tóth - From Nonlinear Identification to Linear Parameter Varying Models: Benchmark Examples&lt;br&gt;A. Montazeri, M.M. Arefi and M. Kazemi - An Investigation of the Wiener Approach for Nonlinear System Identification Benchmarks</td>
</tr>
<tr>
<td>12.00 – 12.30</td>
<td>Discussion Session 3 &amp; Closing</td>
<td>J.P. Noël &amp; M. Schoukens</td>
</tr>
<tr>
<td>12.30 – 13.30</td>
<td>Lunch</td>
<td></td>
</tr>
</tbody>
</table>

The Workshop on Nonlinear System Identification Benchmarks is organized with the support of:

![Liège université](image1.png)  
![An event sponsored by Siemens](image2.png)
Benchmark Setups
The nonlinearbenchmark.org website collects high-quality and well-described datasets for the data-driven modelling of nonlinear systems. The datasets are hosted on the website, together with an overview of results that have been obtained in the past on these benchmarks.

Through this initiative, we aim to create a set of problems that are of interest for the overlapping communities dealing with data-driven modeling of nonlinear systems: e.g. the machine learning, systems and control and mechanical engineering communities. It is our hope that working on similar problems, and comparing the results obtained on these problems in detail, will lead to an increased interaction and understanding between these communities.

A total of 8 datasets have been collected over the past years:

1) Coupled Electric Drives (2017)
2) F-16 Ground Vibration Test (2017)
3) Cascaded Tanks System (2016)
5) Bouc-Wen System (2016)

Do you have a great nonlinear system identification dataset ready to feature as a benchmark? Feel free to contact us, maybe we can include it on this website!
Keynote Speakers
**Keynote Title:** Data-Driven Assessment of Engineered Systems: Beyond LTI

**Biography:** Eleni Chatzi is currently an Associate Professor, and the Chair of Structural Mechanics, at the Institute of Structural Engineering, of the Department of Civil, Environmental and Geomatic Engineering (DBÄUG), ETH Zürich. She has obtained her diploma (2004) and MSc (2006) in Civil Engineering, with honors, from the Department of Civil Engineering at the National Technical University of Athens (NTUA). In June 2010 she obtained her PhD Degree with distinction from the Department of Civil Engineering & Engineering Mechanics at Columbia University. In 2010 she was hired as the youngest Assistant Professor in ETH, and was promoted to an Associate Professor in 2017.

Prof. Chatzi's research couples novel simulation tools with state-of-the-art monitoring methodologies for smart infrastructure assessment, with the goal of providing actionable tools able to guide operators and engineers in the management of engineered systems. A key aspect of her research lies in extraction of quantifiable metrics that are indicative of structural performance across the component, system and network levels. Her research interests lie in the area of Structural Health Monitoring, with a strong focus on problems lying beyond the commonly adopted assumption of linear time invariant systems. Her research spans a broad range of topics, including applications on emerging sensor technologies and structural control, methods for curbing uncertainties in structural diagnostics and life-cycle assessment, as well as advanced schemes for nonlinear/nonstationary dynamics simulations.

Prof. Chatzi serves as editor for numerous peer-reviewed international journals with particular focus on system identification methods and topics relating to SHM. Since 2016, she is coordinating the joint ETH Zürich & University of Zurich PhD Programme in Computational Science. She is currently leading the ERC Starting Grant WINDMIL on the topic of "Smart Monitoring, Inspection and Life-Cycle Assessment of Wind Turbines", awarded by the European Research Council.

[Homepage](#)  
[Google Scholar Profile](#)
**ALFRED C. SCHOUTEN**

**Keynote Title:** Nonlinear cortical responses in EEG evoked by continuous wrist manipulations

**Biography:** Alfred Schouten is currently an Associate Professor at the Delft University of Technology (Faculty of Mechanical Engineering, Department of Biomechanical Engineering) and the University of Twente (Faculty of Engineering Technology, Department of Biomechanical Engineering). He obtained his MSc (1999) in Mechanical Engineering at the Delft University of Technology. In June 2004 he obtained his PhD Degree at the Man-Machine Systems group at the same institution.

Alfred Schouten's research focuses on methods and devices to assess and understand neuromuscular control in both healthy subjects and patients with neurological dysfunction. Key is the use of force-controlled robotic manipulators, neuromuscular modelling, and closed-loop system identification techniques to untangle the human motor control system.

Since his PhD he has developed system identification techniques that enable to quantify the contributions of different feedback pathways. His research resulted in a unique method that can distinguish and quantify the position, velocity and force feedback gains simultaneously. The key concept is the use of force-controlled robotic manipulators which allow for natural movement. Together with medical specialist from academic medical centers (LUMC, AMC, ErasmusMC) he applies the methodology to get insight in neurological dysfunction and working to use these manipulators as diagnostic tools.

In recent years he started to apply the methodology of perturbations and closed-loop system identification to understand the cortical contribution to human motor control. Very recently he demonstrated that the cortico-muscular coherence, a measure for the functional connectivity between the brain and muscles, increases with external perturbations.

[Homepage](#)

[Google Scholar Profile](#)
Keynote Title: Function estimation, model representations and nonlinear system identification

Biography: Johan A.K. Suykens was born in Willebroek Belgium, May 18 1966. He received the master degree in Electro-Mechanical Engineering and the PhD degree in Applied Sciences from the Katholieke Universiteit Leuven, in 1989 and 1995, respectively. In 1996 he has been a Visiting Postdoctoral Researcher at the University of California, Berkeley. He has been a Postdoctoral Researcher with the Fund for Scientific Research FWO Flanders and is currently a full Professor with KU Leuven.


He received an IEEE Signal Processing Society 1999 Best Paper Award and several Best Paper Awards at International Conferences. He is a recipient of the International Neural Networks Society INNS 2000 Young Investigator Award for significant contributions in the field of neural networks. He has served as a Director and Organizer of the NATO Advanced Study Institute on Learning Theory and Practice (Leuven 2002), as a program co-chair for the International Joint Conference on Neural Networks 2004 and the International Symposium on Nonlinear Theory and its Applications 2005, as an organizer of the International Symposium on Synchronization in Complex Networks 2007, a co-organizer of the NIPS 2010 workshop on Tensors, Kernels and Machine Learning, and chair of ROKS 2013. He has been awarded an ERC Advanced Grant 2011 and has been elevated IEEE Fellow 2015 for developing least squares support vector machines.

Homepage

Google Scholar Profile
Keynote Title: State-Dependent Parameter (SDP) Nonlinear Models and a Hydrological Identification Benchmark

Biography: Peter Young started his professional career in 1958 as an apprentice in the Aerospace Industry and he obtained a B.Tech. degree in Aeronautical Engineering from Loughborough University in 1962. He then completed an M.Sc. degree by research at Loughborough, before being awarded a Whitworth Fellowship and moving to Cambridge University, where he obtained his Doctoral degree for work of recursive estimation and adaptive control in 1970. Following two years as a civilian scientist, working for the U.S. Navy in California, he was appointed University Lecturer in Engineering and a Fellow of Clare Hall, Cambridge University, in 1970. As a result of his novel research on environmental modelling and forecasting, he was then invited to become Professorial Fellow and Head of the Systems Group in the newly formed Centre for Resource and Environmental Studies in the Institute of Advanced Studies at the Australian National University (ANU) in 1975. Finally, he moved back to U.K. in 1981, where he was Head of the Department of Environmental Science for seven years and then proposed the establishment of the Institute of Environmental & Biological Sciences, where he was Deputy Head for several years. He also established the Centre for Research on Environmental Systems and Statistics, where he was Director until his retirement in 2005. He is now Emeritus Professor at Lancaster and Adjunct Professor in the Fenner School of Environment and Society, ANU. Peter continues to work full-time on a number of research projects, as well as on-going research projects with colleagues in Spain (Departamento de Analisis Economico, Universidad Autonoma de Madrid), Australia (ANU and RMIT University, Melbourne), and France (University of Nancy).

Peter Young was one of the first scientists to recognise the importance of recursive estimation in the modelling and forecasting of stochastic dynamic systems. He is well known for his work in this area over the past 45 years, as well as research in other areas, such as environmental modelling and automatic control system design. He has evolved and promoted the inductive 'Data-Based Mechanistic' (DBM) approach to modelling uncertain dynamic systems and applied this in various areas of study, including the environment and climate, macro economics, and business applications. His most recent DBM modelling and forecasting research has been concerned with rainfall-flow modelling and real-time flood forecasting, where he worked on a project within the UK Flood Risk Management Research Consortium and published numerous papers, as listed in his publications. These include a recent paper which introduces 'Hypothetico-Inductive DBM' modelling and applies this to well known rainfall-flow data from the Leaf River in the USA. Most recently, he has worked on a unified approach to the optimal estimation of discrete and continuous time transfer function models from sampled data and the data-based mechanistic modelling and forecasting global climate data.

Peter Young is a Whitworth Fellow and a Chartered Engineer. He has received a variety of awards for his research contributions and has many publications in the open literature, including a monograph on recursive estimation, a jointly authored text on 'True Digital Control' and he has edited the 'Concise Encyclopedia of Environmental Systems'. His publishing duties include Departmental Editor of the Journal of Forecasting; Consultant Editor of the International Journal of Control; and member of the Board, Environmental Modelling and Software.
Jorge Goncalves

Keynote Title: Challenges in system identification of biochemical systems

Biography: Jorge Goncalves is a Professor at the Luxembourg Centre for Systems Biomedicine, University of Luxembourg and a Principal Research Associate at the Department of Engineering, University of Cambridge. He received his Licenciatura (5-year S.B.) degree from the University of Porto, Portugal, and the M.S. and Ph.D. degrees from the Massachusetts Institute of Technology, Cambridge, MA, all in Electrical Engineering and Computer Science, in 1993, 1995, and 2000, respectively. He then held two postdoctoral positions, first at the Massachusetts Institute of Technology for seven months, and from 2001 to 2004 at the California Institute of Technology with the Control and Dynamical Systems Division. At the Information Engineering Division of the Department of Engineering, University of Cambridge he was a Lecturer from 2004 until 2012, a Reader from 2012 until 2014, and since 2014 he is a Principal Research Associate. From 2005 until 2014 he was a Fellow of Pembroke College, University of Cambridge. From June to December 2010 and January to September 2011 he was a visiting Professor at the University of Luxembourg and California Institute of Technology, respectively. Since 2013 he is a Professor at the Luxembourg Centre for Systems Biomedicine, University of Luxembourg.

The Systems Control Group (SCG) headed by Jorge Goncalves at the University of Luxembourg aims to understand the source of diseases and find new therapies and cures. This is accomplished by building mathematical models and tools that pinpoint the exact location of diseases. The mathematical models are simple yet powerful, and capture the dynamics seen in biological systems. The predictive power of these models comes from mimicking the most important feature of biological systems: the fact that they are dynamic in nature, as molecular numbers change over time in response to internal or external stimulus. When pushed from their comfort zone, systems respond by activating a number of elements that try to restore their initial state. This transient behaviour is essential to understand how the components in the system interact with each other. Hence, time-series data offers a window into complex biochemical systems. Although time-series experiments are more expensive than steady-state, they provide an enormous amount of information, essential to understand biochemical pathways, organs and organisms. Theoretically, the SCG is moving ahead by developing the necessary tools to model and analyse time-series data, as we predict these data will become more and more frequent in the near future. Examples are tools to infer causal networks between measurements, to find differential expressed systems, and systematic engineering control tools for synthetic biology design. Applied, the group closely collaborates with experimental biologists to generate mathematical models, which offer new biological insights that can subsequently be tested experimentally, hence closing the loop between experiments and modelling. Examples range from the understanding the molecular details of circadian rhythms to learn about the sources of Parkinson's disease.

Homepage

Google Scholar Profile
Workshop Contributions
Identification of Systems with Hysteretic Behavior
Using NOBF Local Model Networks

Tim Oliver Heinz, Tobias Münker, and Oliver Nelles
University of Siegen, Mechanical Engineering, Automatic Control and Mechatronics
Paul-Bonatz-Str. 9-11, 57076 Siegen
tim.heinz@uni-siegen.de

1 Approach

In this contribution the Bouc-Wen system is identified using orthonormal basis functions (OBF) for the dynamics realization and local model networks for the approximation of the nonlinearity. Some of the results proposed here are described more precisely in [2]. Consider a given data set with \( N \) samples of input \( u(k) \) and output \( y(k) \) with \( k \) being the discrete time. The goal is to identify the relationship

\[
\hat{y}(k) = f(u(k), \ldots, u(k-n), y(k-1), \ldots, y(k-n))
\]

(1)

with the nonlinear function approximator \( f(\cdot) \) and the model order \( n \).

1.1 Local Model Networks

In this contribution local model networks (LMN) [3] are used, i.e., the approximator is described as

\[
\hat{y}(k) = \sum_{i=1}^{m} \phi_i(z(k))L_i(x(k))
\]

(2)

with the validity functions denoted by \( \phi_i(z(k)) \), the local linear models \( L_i(x(k)) \), and the number of local models \( m \). The variables \( x(k) \) and \( z(k) \) can both contain any delayed or filtered version of input and output values. The estimation of \( \phi_i(z(k)) \) is done by a tree-based identification method as described in [3].

1.2 Orthonormal Basis Functions

In [1] the universal approximation abilities of OBF in combination with a nonlinear readout map are derived. In the proposed approach the advantages of the OBF structure are combined with the beneficial property of LMNs to assign the inputs to the local models and to the nonlinear validity functions independently. Here the OBF filtered inputs are assigned to the local model inputs

\[
L_i(x(k)) = \sum_{j=1}^{n} g_{i,j}x_j(k) \quad \text{with} \quad x_j(k) = G_j(q)u(k)
\]

(3)

with the weights of the filtered inputs \( g_{i,j} \) and the Kautz filters \( G_j(q) \). To determine the nonlinear behavior an appropriate set of filtered inputs and outputs is assigned to the validity function \( \phi_i(z(k)) \). For the Bouc-Wen system it is a necessity to include delayed versions of \( y(k) \) in \( z \) to describe the hysteretic behavior of the system.

2 Results

The simulation error values (\( e_{\text{RMS}} \)) of the OBF LMN with \( n = 10 \) Kautz filters are depicted in Tab. 1. With delayed inputs and outputs in \( z \) the model is able to reproduce the hysteretic behavior of the Bouc-Wen system, see Fig. 1. In contrast, for \( x = z \) the model performs worse, see Tab. 1.

Table 1: \( e_{\text{RMS}} \) on the two provided test data sets.

<table>
<thead>
<tr>
<th>x-inputs</th>
<th>OBFinputs</th>
<th>ORBF</th>
</tr>
</thead>
<tbody>
<tr>
<td>multisine</td>
<td>1.2422e-4</td>
<td>1.3838e-4</td>
</tr>
<tr>
<td>sinesweep</td>
<td>1.1522e-4</td>
<td>1.4281e-4</td>
</tr>
</tbody>
</table>

References


Identification of Decoupled Polynomial NARX Models using Simulation Error Minimization

Kiana Karami  
University of Calgary  
kiana.karami@ucalgary.ca

David Westwick  
University of Calgary  
dwestwic@ucalgary.ca

Johan Schoukens  
Vrije Universiteit Brussel  
Johan.Schoukens@vub.ac.be

1 Introduction

The multivariate polynomial in a polynomial NARX model can often be replaced by a small number of univariate polynomials, each of which transforms a linear combination of the variables:

\[
y(t) = \sum_{r=1}^{R} g_r \left( \begin{bmatrix} u \\ y \end{bmatrix} \right) + e(t) \tag{1}\]

where the \( g_r(\cdot) \) are univariate polynomial functions, and \( R \) is the number of branches in the model.

NARX models are typically fitted using prediction error minimization (PEM) \([1]\), where the model has access to the past inputs and outputs from the system. Here we consider the simulation error minimization (SEM). In \([2]\), SEM was used to identify polynomial state-space models. Paduart showed that the Jacobian of the simulation error could be computed using nonlinear filtering operations. We applied this approach to the decoupled polynomial NARX model.

2 Results from the Bouc-Wen Benchmark

We generated identification data using the Bouc-Wen Benchmark \([3]\) simulation, but increased the standard deviation of the input multisine from the default of 50 N to 55 N, so that the identification data would cover a larger range than the validation sets. Previously, we had identified decoupled NARX models using PEM. The best such model had 5 branches each of which filtered 15 past inputs and 15 past outputs and transformed the result with a 9\(^{th}\) order polynomial nonlinearity. This model was used as the starting point for the SEM optimization.

Table 1: Validation results for models fitted using PEM and SEM.

<table>
<thead>
<tr>
<th>多于Sine Validation</th>
<th>PEM</th>
<th>SEM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simulation Error (dB)</td>
<td>-85.42</td>
<td>-84.54</td>
</tr>
<tr>
<td>Prediction Error (dB)</td>
<td>-100.26</td>
<td>-78.00</td>
</tr>
</tbody>
</table>

Table 1 presents the normalized RMS error, as specified in \([3]\) where the models were used both as 1-step ahead predictors and in simulation. When using the Multi-Sine validation signal, the model trained using SEM performed worse than the PEM model in both the prediction and simulation tests. The reason for the poor simulation performance can be seen in Fig. 1. The SEM trained model produced a large spike in the simulation error, followed by a transient decay, at about \( t = 4.5 \) seconds. These large errors dominate the RMS result. The swept-sine results are equally puzzling, in that the SEM model outperformed the PEM model in both prediction and simulation tasks.

3 Conclusion

The methods proposed in \([2]\) can be used in SEM of decoupled NARX models. The results from the Bouc-Wen benchmark are promising, but illustrate the problems with model stability and sensitivity.

References

Retrieving highly structured models starting from a black box state-space model: a case study on the Bouc-Wen hysteresis model.

J. Decuyper  
Pleinlaan 5, 1050 Brussel  
jan.decuyper@vub.be

A. Fakhrizadeh Esfahani  
Pleinlaan 2, 1050 Brussel  
Alireza.Fakhrizadeh.Esfahani@vub.be

K. Tiels  
PO Box 337, SE-75105 Uppsala  
koen.tiels@it.uu.se

J. Schoukens  
Pleinlaan 5, 1050 Brussel  
Johan.Schoukens@vub.be

1 Introduction

This work presents the results on modelling the Bouc-Wen hysteresis system given the provided benchmark data set. In particular we show that highly structured models can be retrieved starting from the models obtained using the black box nonlinear state-space approach, as was done in [1]. In [2] a tensor based decoupling algorithm was used to reduce the multivariate nonlinear part of the state-space model to univariate polynomial functions. We now show that an additional reduction, down to a single univariate polynomial function, can be accomplished, without suffering from a loss of accuracy.

2 Approach

2.1 Building a nonlinear state-space model of the Bouc-Wen system

The PNLSS model used as a starting point of this work is of third order (3 state variables) and includes the nonlinear degrees 2 and 3 in the state equation (details on the method can be consulted in [1]).

\[ x(k+1) = Ax(k) + Bu(k) + E\xi(k,x,u) \]
\[ y(k) = Cx(k) + Du(k), \]

with \( k \) indicating the time index and \( \xi \) including monomials of all possible cross products between \( x \) and \( u \) raised to the total powers 2 and 3.

2.2 Tensor based decoupling of multivariate polynomials

Using the first order information of the multivariate function \( E\xi(k,x,u) \) and the canonical polyadic tensor decomposition (CPD \(^1\)), the function can be decoupled in univariate polynomials,

\[ x(k+1) = Ax(k) + Bu(k) + W_1 g(V_1^Tu^+(k)) \]
\[ y(k) = Cx(k) + Du(k) + W_3 g(V^Tu^+(k)), \]

where \( u^+ \) is a vector containing the state variables and the input and \( g \) is a vector function of univariate polynomials. A decoupling into 3 univariate polynomial branches \((g_1 \text{ up to } g_3)\) of degree up to 5 was used.

\(^1\)Part of the Tensorlab v2.0 toolbox in MATLAB URL: http://www.tensorlab.net/.

3 Results

The number of parameters together with the corresponding rms errors on the validation data are presented in Table 1.

<table>
<thead>
<tr>
<th>Model</th>
<th>Num. para.</th>
<th>( \epsilon_{\text{RMSI}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>PNLSS model</td>
<td>106</td>
<td>2.88e-05</td>
</tr>
<tr>
<td>Decoupled model</td>
<td>52</td>
<td>3.06e-05</td>
</tr>
<tr>
<td>Decoupled unified model</td>
<td>42</td>
<td>2.99e-05</td>
</tr>
</tbody>
</table>

Table 1: Multisine validation results

2.3 Unifying resembling univariate polynomial branches

Studying the decoupled model obtained in Section 2.2 a strong similarity was observed amongst the 3 univariate polynomial branches. This suggests that even more structure is present in the nonlinear part of the model. This feature is exploited by unifying the branches into a single polynomial form. To convert the model to a single polynomial form an appropriate scaling on the internal inputs and outputs of the branches is applied by updating the linear transformation matrices \( V \) and \( W \). Finally all parameters are further optimised using the Levenberg-Marquardt algorithm.

4 Conclusion

A highly structured model containing a single univariate polynomial nonlinear part could be retrieved starting from a PNLSS model of the Bouc-Wen system. The number of parameters was reduced from 106 to 42 without any loss of accuracy.

5 Acknowledgement

This work was supported in part by the Fund for Scientific Research (FWO-Vlaanderen), and the ERC advanced grant SNLSID, under contract 320378.

References

Regressor selection using Lipschitz quotients on the F-16 aircraft benchmark

Matija Perne  
Jožef Stefan Institute  
matija.perne@ijs.si

Martin Stepančič  
Jožef Stefan Institute  
martin.stepancic@ijs.si

1 Introduction

When designing a nonlinear autoregressive model with exogenous input (NARX), it is necessary to select the regressors [1]. The more rigorous regressor selection methods can require a significant computational effort. We developed a computationally efficient approach based on Lipschitz quotients [2] and demonstrate it on the F-16 benchmark [3].

2 Method

He and Asada [2] introduce a method based on Lipschitz quotients for identifying system orders. They expect the output \( y \) to be a Lipschitz continuous function of the input \( x \),

\[
y = f(x) = f(x_1, x_2, \ldots , x_n). \tag{1}
\]

Lipschitz continuity means that the absolute value of a partial derivative \( |\partial f/\partial x_i| \) is bounded.

Consider input-output data pairs \((x_i, y_i), (x_j, y_j)\). The Lipschitz quotient is defined as

\[
q_{ij} = \frac{|y_i - y_j|}{\|x_i - x_j\|^2}, \quad i \neq j. \tag{2}
\]

Since the dependence of \( y \) on \( x \) is Lipschitz continuous, the value of the Lipschitz quotient is bounded as long as \( y \) depends only on the components of \( x \). However, if there is another input variable that influences \( y \) independently from the components of \( x \), then \( |y_i - y_j| \) may not be small when \( \|x_i - x_j\| \) is small and \( q_{ij} \) is not bounded. Skipping an important input variable thus significantly increases the Lipschitz quotients. Conversely, adding a component that has no influence on \( y \) to \( x \) only slightly decreases \( q_{ij} \).

He and Asada [2] construct the index

\[
q^{(n)} = \left( \prod_{k=1}^{p} \sqrt[n]{q^{(n)}(k)} \right)^{1/p} \tag{3}
\]

where \( q^{(n)}(k) \) is the \( k \)-th largest Lipschitz quotient, \( p \) is a positive number, usually between 0.01 and 0.02 times the number of input-output pairs \( N \), and \( n \) is the model order. The value of \( n \) beyond which \( q^{(n)} \) no longer decreases significantly with \( n \) is the system order.

We use the index for general regressor selection, not just to determine the system order. We thus observe the value of the index for various combinations of regressors.

3 Example

The Lipschitz index determination method is modified after the algorithm for system order determination from [4]. The value of \( p \) in (3) is 0.02 - \( N \). The method is used to select regressors for predicting the 2nd acceleration signal during multisinus excitation with a full frequency grid at level 7 of the F-16 aircraft benchmark [3] where the possible regressors are delayed values of the excitation force (0th to 19th delay) and the 2nd acceleration signal (1st to 20th delay). As investigating all the combinations would require too much computation, the MATLAB built-in function for backward elimination and forward selection methods is used until the solution stabilizes. On a computer running MATLAB R2015a, OS Ubuntu 16.04 LTS with Linux kernel 4.4.0-83-generic, 15.1 GiB of RAM and Intel® Core™ i7-3770K CPU @ 3.50GHz × 8, 10565 s are needed for selecting regressors based on 14742 regression vectors of 40 regressors, significantly reducing the number of regressors.

4 Acknowledgements

The authors acknowledge project “Method for the forecasting of local radiological pollution of atmosphere using Gaussian process models”, ID L2-8174, and research core funding No. P2-0001 were financially supported by the Slovenian Research Agency.

References


1 Brief presentation of the method

The aim of this work is to introduce the concept of manifold regularization to the identification of dynamic systems. The method has been tested on the coupled electric drives problem, using a purely black box approach instead of the parametric Wiener-Hammerstein model used in the technical report [3], the system is modeled using a generic NARX\((p,q)\) model:

\[
y(t+1) = f(\varphi_u(t), \varphi_y(t)) + e(t),
\]

where \(y(t)\) is the measured pulley velocity, \(u(t)\) is the input signal and \(e(t)\) is the measurement noise. The model in Equation (1) depends on past input \(\varphi_u(t) = [u(t), \ldots, u(t-p+1)]^T\) and past output \(\varphi_y(t) = [y(t), \ldots, y(t-q+1)]^T\) samples. The order of the exogenous part \(p\) and the order of the auto-regressive part \(q\) are considered not known a priori and will be properly estimated as hyperparameters.

Let’s assume now that \(f \in \mathcal{H}_k\), where \(\mathcal{H}_k\) is a RKHS with kernel \(k : \mathbb{R}^{p+q} \times \mathbb{R}^{p+q} \rightarrow \mathbb{R}\) and norm \(\|\cdot\|_{\mathcal{H}_k}\). The estimation of \(f\) is then carried out by solving the optimization problem in Equation (2), where \(\hat{f} = [f(\varphi_u(1), \varphi_y(1)), \ldots, f(\varphi_u(N), \varphi_y(N))]\) and \(L \in \mathbb{R}^{N \times N}\) is the Laplacian matrix computed on the complete graph whose nodes are the regressors \(\{\varphi_u(t), \varphi_y(t)\}_{i=1}^N\) :

\[
\hat{f} = \arg \min_{f \in \mathcal{H}_k} \left\{ \sum_{t} \left( f(\varphi_u(t), \varphi_y(t)) - y(t+1) \right)^2 + \lambda_T \|f\|^2_{\mathcal{H}_k} + \lambda_M \cdot f^T \cdot L \cdot f \right\}
\]

In Equation (2), we introduced the use of a proper manifold regularization term, in addition to the classical Tikhonov one. The effect of adding this term is to enforce additional local smoothness around the regressors of the dataset [1]. The kernel used is reported in Equation (3). It is composed by a Guassian kernel, that can detect the non-linearities of the system, and an affine linear kernel, that can explain the linear component of \(f\):

\[
k(a, b) = \lambda_{nl} \cdot \exp \left[ -\frac{\|a - b\|^2}{\sigma_k^2} \right] + \lambda_l \cdot a^T \cdot b + \lambda_c
\]

The hyperparameters of the method \((p, q, \lambda_T, \lambda_M, \lambda_{nl}, \sigma_k, \lambda_l, \lambda_c\), and the one needed to compute \(L\)) need to be identified from the data. In this work, we employed the Generalized Cross Validation (GCV) [2] scheme is employed.

### Table 1: Performance of the proposed method. The reported value are the RMSE on the test dataset.

<table>
<thead>
<tr>
<th>Method</th>
<th>Prediction</th>
<th>Simulation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tikhonov regularization</td>
<td>0.0609</td>
<td>0.2798</td>
</tr>
<tr>
<td>Manifold regularization</td>
<td>0.0371</td>
<td>0.1591</td>
</tr>
</tbody>
</table>

![Figure 1: Simulation output of the estimated model.](image-url)

2 Results

The method is applied on the Uniformly distributed input signal and validated on the other dataset of the same type. The performance are reported in the Table 1, where it is possible to observe that the proposed approach outperforms classical Tikhonov regression. A comparison between the simulation performance of the two regularization methods is presented in Figure 1. It is interesting to notice that the model obtained using the manifold regularization can simulate the spike around 1.5s very well. This is the major difference with the Tikhonov regularization and the results in the technical report [3].

### References


Application of a Linear PEM estimator to a stochastic Wiener-Hammerstein Benchmark Problem

Mohamed R.-H. Abdalmoaty  
abda@kth.se

Håkan Hjalmarsson  
hjalmars@kth.se

Automatic Control Department, School of Electrical Engineering and Computer Science  
KTH Royal Institute of Technology, Stockholm

1 Abstract

In this contribution, we apply a computationally attractive consistent PEM estimator [1] to a stochastic Wiener-Hammerstein benchmark problem using real-data [2]. The estimator is defined using a one-step ahead predictor, which is linear in the observed outputs but nonlinear in the input. This allows us to tackle several issues considered challenging from the perspective of the current smoothing-based approaches [3].

2 The Model

The estimated model is defined by the relations

\[ y_t = G_2(q; \theta) z_t + v_t, \quad t = 1, 2, 3 \ldots \]
\[ z_t = f(x_t; \theta) \]
\[ x_t = G_1(q; \theta) u_t + w_t, \]

in which \( q \) is the shift operator, \( G_1 \) and \( G_2 \) are third order, causal, and stable transfer operators (according to prior knowledge [2]). The nonlinearity is modeled by a sigmoid function of the form

\[ f(x; \theta) = L / (1 + \exp(-k x)) - L / 2. \]

It is assumed that: the input \( u \) is perfectly known, the process disturbance \( w \) is a stationary Gaussian process with zero mean and variance \( \lambda_w \), and the noise \( v \) is stationary with zero mean. Note that \( w \) may be a colored process.

3 Proposed Method

A consistent PEM estimator is defined as

\[ \hat{\theta} := \arg \min_{\theta, \lambda_w} \sum_{t=1}^{N} (y_t - \tilde{y}_{t-1}(\theta, \lambda_w))^2, \]

in which \( \tilde{y}_{t-1} = E[y_t | U_t; \theta, \lambda_w] \) is the OE-type predictor proposed in [1]. Instead of ignoring \( w \), the predictor averages the output over all possible values of \( w \). As indicated by the notations, it is a function of \( \lambda_w \), and the solution to (1) is sought jointly over \( \theta \) and \( \lambda_w \). Note that the used predictor can be efficiently computed for the assumed model.

4 Results

The parameters were initialized using the best split of the BLA model [4], which is estimated using the multisine data. To obtain a consistent split when \( w \) is present, the cost function in (1) is used as a splitting criterion.

Figure 1 compares the simulated output against the test data; the RMS of the residuals is reported in Table 1. The values in the third column is obtained when the 7th multisine experiment is used. This is the smallest RMS value over the 10 multisine experiments. The last column shows the RMSE when all the data in the 10 multisine experiments is used to solve (1). The estimation time is about 10 seconds for one experiment, and about 30 seconds when all the data is used.

Table 1: The residuals RMS of the estimated model using the multisine data (computed according to the instructions in [2])

<table>
<thead>
<tr>
<th></th>
<th>BLA</th>
<th>BLA+NL</th>
<th>OE-PEM (exp. 7)</th>
<th>OE-PEM (all data)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Swept-sine</td>
<td>0.0281</td>
<td>0.0127</td>
<td>0.0116</td>
<td>0.0091</td>
</tr>
<tr>
<td>Multisine</td>
<td>0.0339</td>
<td>0.0261</td>
<td>0.0171</td>
<td>0.0148</td>
</tr>
</tbody>
</table>

References


From the Volterra series to a Wiener-Hammerstein model

Georgios Birpoutsoukis
ICTEAM, Université Catholique de Louvain, B1348 Louvain la Neuve, Belgium
georgios.birpoutsoukis@uclouvain.be

Maarten Schoukens
Technische Universiteit Eindhoven, P.O. Box 513, 5600 MB Eindhoven, the Netherlands
m.schoukens@tue.nl

1 Introduction

The regularized Volterra series identification approach is applied to the coupled electric drive (CED) dataset [2]. This Volterra model is used in a second step to extract initial estimates of a Wiener and a Wiener-Hammerstein block-structured model. As a result, the error on validation data is decreased and a compact block-structured model for the nonlinear CED system is obtained.

2 Model structures and identification procedure

The data-generating system is initially modeled with the discrete time Volterra series:

\[ y(k) = h_0 + \sum_{m=1}^{M} \sum_{\tau_1=0}^{n_1-1} \cdots \sum_{\tau_m=0}^{n_m-1} h_m(\tau_1, \ldots, \tau_m) \prod_{\tau=\tau_1}^{\tau_m} u(k - \tau) + e(k) \]

where \( y(k) \) represents the measured system output at time instant \( k \) and \( u(k) \) is the input signal applied to the system. The \( m \)-th order Volterra kernel is denoted by \( h_m(\tau_1, \ldots, \tau_m) \) where \( \tau_i, i = 1, \ldots, m \), denote lag variables and \( e \) is i.i.d. noise (\( e \sim \mathcal{N}(0, \sigma_e^2) \)). The constant term \( h_0 \) in (1) represents the 0\textsuperscript{th} order Volterra kernel. The value \( n_m - 1 \) corresponds to the memory of \( h_m \). In this work, prior information about smoothness and exponential decay of the multi-dimensional kernels is used during the identification step [1] to tackle the high variance problem due to dimensionality issues. The estimated Volterra kernel is used to obtain an initial estimate for the linear block of a Wiener and further a Wiener-Hammerstein model structure. As such, similar to the linear identification case, a nonparametric model is used to extract information in order to obtain a compact model structure with a reduced number of parameters and validation error.

3 Results and conclusions

A characteristic second order Volterra kernel for the CED system is depicted in Fig. 1. For model estimation and validation, 300 and 200 data have been used, respectively. The relative rms error on the validation dataset for different measured datasets are given in Table 1. It can be verified that the Volterra series can be used to obtain reliable block structured models for the CED by providing initial estimates for the linear blocks.

4 Acknowledgments

This work was supported in part by the Fund for Scientific Research (FWO-Vlaanderen), by the Flemish Government (Methusalem), the Belgian Government through the Inter university poles of Attraction (IAP VII) Program, and by the ERC advanced grant SNLSID, under contract 320378.

References

Regularised NFIR identification with Gaussian process model

Martin Stepančič\textsuperscript{a,b}  
\textsuperscript{a}Jožef Stefan Postgraduate school  
martin.stepancic@ijs.si  
\textsuperscript{b}Jožef Stefan Institute  
matija.perne@ijs.si  

Juš Kocijan\textsuperscript{b,c}  
\textsuperscript{c}University of Nova Gorica  
jus.kocijan@ijs.si

1 Introduction

We focus on the identification problem of nonparametric and probabilistic nonlinear finite impulse response (NFIR) model in form of a Gaussian process with additional tuning parameters \([1]\). The NFIR structure may require a long digital delay line of the input signal. This leads to poor identification performance due to overfitting. The problem of high number of regressors is solved using the Bayesian or Akaike information criterion \([2]\). The criteria restrict the effective number of parameters in the nonparametric model, and hence the identification performance is improved. The performance of the proposed identification method is illustrated using the Bouc-Wen benchmark process from \([3]\).

2 The model structure and identification

Let the NFIR model structure be

\[
\hat{y}(k) = \phi(z(k), \hat{\theta})^T \hat{w} + E(k), \hat{w} \sim N(0, I), \tag{1}
\]

where \(z(k) = [z(k-1), \ldots, z(k-D)]^T\) is the regressors vector of delayed input signals at \(k\)th time instant, \(E(k) \sim N(0, \sigma^2_y)\) is the normally distributed noise with unknown variance \(\sigma^2_y\), \(\phi\) is the vector of basis functions, and \(\hat{\theta}\) is the vector of tuning parameters. The term \(\phi(z(k), \hat{\theta})^T \hat{w}\) is a Karhunen–Loève expansion of the Gaussian process with mean function \(m(z) = 0\) and covariance function \(K(z, z'; \hat{\theta}) = K_{zz'} = \phi(z, \hat{\theta})^T \phi(z', \hat{\theta})\) \([1]\). From now on, the covariance function argument \(\hat{\theta}\) is omitted for notational convenience.

The identification method is divided into two subsequent steps: (i) the maximum likelihood (ML) method for the estimation of \(\hat{\theta}, \sigma^2_y\) while assuming \(\hat{w} \sim N(0, I)\); (ii) the regularised least-squares (ReLS) method for estimating Gaussian process which depends on previously estimated \(\hat{\theta}, \sigma^2_y\). The ML method is based on the maximisation of the log probability density \(p(\hat{z}|y)\) where \(\hat{z} = [\hat{z}(k_1), \ldots, \hat{z}(k_N)]^T\) and \(y = [y(k_1), \ldots, y(k_N)]^T\) is the vector of measured output at time instants \(k_1, \ldots, k_N\). Once obtaining the ML estimates \(\hat{\theta}, \sigma^2_y\), the ReLS method modifies the Gaussian process estimate with new mean function \(m'(z) = K_{zz}(K_{zz} + \sigma^2_y \hat{I})^{-1}y\) and covariance function \(K_{zz'} = K_{zz}(K_{zz} + \sigma^2_y \hat{I})^{-1}K_{zz'}\), where \(Z = [z(k_1), \ldots, z(k_N)]\).

3 The regularization approach

The nonparametric model complexity is related with the effective number of parameters (EnP) \([2]\). The EnP of the Gaussian process model equals to \(d_f = K_{zz}(K_{zz} + \sigma^2_y \hat{I})^{-1}\) \([1]\). The regularization approach expands the ML criterion with a penalty term which results in the BIC criterion \(-2\log p(\hat{y}|y) - d_f \log N\) or the AIC criterion \(-2\log p(\hat{y}|y) - d_f\). The performance between the identification approaches in sections 2 and 3 is compared in Figure 1 using the Bouc-Wen benchmark example code \([3]\). The model with lowest RMSE is based on AIC at \(D = 160\) and its RMSE error on test multisine and sine-sweep dataset is 1.63 · 10^{-4} and 1.69 · 10^{-4}, respectively.

4 Conclusion

The results in Figure 1 show satisfactory identification performance of regularised approaches, namely BIC and AIC, up to \(D = 320\) and \(D = 680\), respectively. The ML approach is unsatisfactory at \(D > 100\) due to overfitting.

Acknowledgements

The authors acknowledge project “Method for the forecasting of local radiological pollution of atmosphere using Gaussian process models”, ID L2-8174, and research core funding No. P2-0001 were financially supported by the Slovenian Research Agency.

References


Phased-based homogeneous order separation for improving Volterra series identification

Damien Bouvier, Thomas Hélie and David Roze
Team Sound Signals and Systems: Audio/Acoustics, instruMents (S3AM), STMS (UMR9912), Sorbonne Université, 1, Place Igor Stravinsky, 75004, France
damien.bouvier@ircam.fr

1 Introduction and problem statement

The current work focuses on improving estimation of nonlinear systems that allows a polynomial-like input-output description; this is done by using an order separation method prior to identification. The results will be presented in the framework of Volterra series, but are applicable for polynomial block structure (Hammerstein, Wiener, ...) or even polynomial nonlinear state-space system.

For a system that admits a Volterra series representation, the output can be written

\[ y(t) = \sum_{n=1}^{\infty} y_n(t) = \sum_{n=1}^{\infty} V_n[u](t), \]

where \( y_n \) are the nonlinear homogeneous order of the system, and where each operator \( V_n \) is completely described by the Volterra kernel \( h_n \).

Most of Volterra identification methods exploit directly the raw output signal \( y \) in order to identify a set of kernels \( \{h_1, h_2, \ldots, h_N\} \) (as shown in Figure 1a). Due to the amplitude differences between the \( y_n \) (overall, it decreases as \( n \) increases), this simultaneous estimation will be less reliable for higher-order kernels. In order to circumvent this difficulty, it is possible to divide the identification into 2 steps (as shown in Figure 1b):

1. first we separate the nonlinear homogeneous orders \( y_n, n = 1, \ldots, N \) from a set of output signals;
2. then we identify separately the kernels \( h_n \) on each signal \( y_n \).

This two-part approach has the benefit of being modular: any existing identification methods can be used [1–3]. But the existing order separation method relies on amplitude differences between test signals [4], which leads to several drawbacks (large range of amplitudes needed for excitation, bad conditioning). Its use is in practice limited to the first few orders.

2 Methodology and applications

This contribution presents the design of a robust order separation method relying on phase deconstruction and reconstruction between tests signals. It extends the previous results of the authors that can be found in [5].

First a theoretical method using complex-valued signals is introduced. Then extension to real-valued signals is made, where multilinearity of the orders \( y_n \) has to be taken into account (instead of homogeneity). The proposed order separation method is evaluated and compared to the state-of-the-art one on simulated versions of the Wiener-Hammerstein and Silverbox benchmark systems.

References

Nonlinear System Identification with Equation Discovery

Nikola Simidjievski\textsuperscript{1}, Ljupčo Todorovski\textsuperscript{1,2}, Sašo Džeroski\textsuperscript{1}, and Juš Kocijan\textsuperscript{1,3}

\textsuperscript{1}Jožef Stefan Institute, Jamova Cesta 39, 1000 Ljubljana, Slovenia
\textsuperscript{2}Faculty of Administration, University of Ljubljana, Gosarjeva ulica 5, Ljubljana, Slovenia
\textsuperscript{3}University of Nova Gorica, Vipavska cesta 13, 5000 Nova Gorica, Slovenia
\*nikola.simidjievski@ijs.si, ljupco.todorovski@fu.uni-lj.si, saso.dzeroski@ijs.si, jus.kocijan@ijs.si

We present an Equation Discovery approach for system identification of nonlinear dynamic systems. Equation Discovery \cite{4} addresses the task of automated discovery of quantitative laws and models, expressed as equations, from measured data and domain-specific knowledge. Process-based modeling (PBM) \cite{1}, the most recent Equation Discovery approach, allows the modeler to formalize and use modeling knowledge specific to the domain of interest, which in turn, given measured data at input, is able to simultaneously tackle the tasks of structure identification and parameter estimation.

The PBM modeling-knowledge formalism is based on two general concepts of: entities and processes. Entities represent the static components of the observed system and include the variables and constants related to a specific component. On the other hand, each process specifies interaction among entities in terms of equations. More specifically, they provide the mathematical model of the process influence on the interacting entities, i.e. on the time derivatives of their variables. To this end, PBM allows for a higher-level representation of modeling knowledge, in terms of hierarchies of building blocks encoded, for the domain of interest, in a library of entities and processes templates. These templates represent generic modeling building blocks, that for a particular modeling scenario, are instantiated into specific components of a model of the observed system.

In this study we use the most recent implementation of PBM – the ProBMoT\textsuperscript{1} (Process-Based Modeling Tool) \cite{6,5}. More specifically, ProBMoT first instantiates the templates into a number of specific modeling components. Next, it performs a combinatorial optimization task to generate plausible candidate model structures by combining different model components. Finally, ProBMoT evaluates each candidate model structure by estimating the values of its constant parameters. The parameter estimation is tackled as a numerical optimization task where the objective function considered is the output error, i.e., the discrepancy between the simulated model response and the measurements. For model simulation, ProBMoT employs the CVode library \cite{2} for long-term simulating continuous-time models in a form of ordinary differential equations. For parameter estimation, ProBMoT employs the Differential Evolution algorithm, as implemented in the meta-heuristic optimization framework jMetal \cite{3}.

We test the utility of the process-based modeling approach on the cascaded water tanks benchmark \cite{7} from two aspects of: (1) structure identification and (2) parameter estimation. We first investigate PBM’s ability to discriminate among different instantiated candidate model structures. Next, we analyze its ability to correctly estimate the model’s parameter values. We perform these experiments both on synthetic and measured data.

The results show that the process-based modeling approach exhibits robust and accurate performance. Note that, while the empirical study is limited to one domain, i.e. the cascaded water tanks benchmark, the proposed process-based modeling approach is general enough and therefore can be readily applied to any other domain and to any other task of nonlinear systems identification.

References
\begin{thebibliography}{9}
\end{thebibliography}
Nonparametric Approximation of the Nonlinear SilverBox Data: a Linear Time-varying Approach

Péter Zoltán Csurcsia¹,², Johan Schoukens¹, Bart Peeters²
1) Vrije Universiteit Brussel, Pleinlaan 2, 1050 Elsene, Belgium
2) Siemens Industry Software NV., Interleuvenlaan 68, 3001 Leuven, Belgium
Email: peter.zoltan.csurcsia@vub.ac.be

1 Introduction and identification of LTV systems

In this abstract a methodology is presented to obtain a nonparametric two-dimensional impulse response function (IRF) estimate $\hat{h}[\tau, \tau]$ of linear time varying (LTV) systems using 2D regularization [1]. Unlike the linear time invariant systems where the IRF is unique, the time varying impulse response is not restricted to only one solution because the number of parameters is much higher than the number of data samples. The user can impose additional constraints to decrease this freedom. In the proposed case smoothness and exponential decaying are used over the system time $t$ (direction of the impulse responses, referring to the behavior) and smoothing constraint is used over the global time $\tau$ (referring to the system memory). The excess degrees of freedom can be removed by 2D regularization using these constrains [1][2].

A further issue is that a measurement of an LTV system cannot be usually repeated under the same conditions. Therefore it is needed to be able to estimate by using a single experiment with transient. A simple method is developed to avoid the above-mentioned problem.

2 The Silverbox data

The Silverbox system can be seen as an electronic implementation of the Duffing oscillator. It is build as a 2nd order linear time-invariant system with a 3rd degree polynomial static nonlinearity around it in feedback [3].

In this work a regularized LTV IRF modell is used to describe the nonlinear Silverbox benchmark data.

The time-varying behaviour of the data comes from the fact that the power of excitation increases over time in the experimantal data (see Fig 1.a). Silverbox - due to its nonlinearities - behaves at different power levels differently resulting in time-varying behaviour. Fig. 1. shows the input excitation, the measured output and the result of the proposed method [1][2].

3 Summary

In this work an LTV IRF approach is used to estimate a modell from the Silverbox data. The intended time domain estimation method is developed for smooth LTV systems. With respect to the system dynamics using the proposed method, it is possible to

1) decrease the effect of the disturbing noise,
2) eliminate the undesired transient term,
3) estimate from a single large experiment. This technique is illustrative, flexible and user friendly.

Despite the fact that the applied methodology is not able to extract details of nonlinear behaviour, the tracking of changes in system dynamics can be followed very well.

References

Simulation and Prediction Errors in the Presence of Model Errors: a Case Study on the Silverbox

Johan Schoukens
Vrije Universiteit Brussel, Dep. INDI
johan.schoukens@vub.be

1 Introduction

The behavior of simulation and prediction errors is illustrated on the Silverbox (forced Duffing oscillator) benchmark. Because the SNR of the measurements is very high (noise disturbances well below 1%), the simulation and prediction errors are completely dominated by model errors. The simulation and prediction errors will be shown for linear Box-Jenkins and ARX models, and for nonlinear NARX and NLSS models.

2 Simulation Errors and prediction errors

Consider the system

\[ y(t) = y_0(t) + v(t) = h(u(t), Z_{t-1}) + v(t). \] (1)

The term \( v(t) \) models the disturbances which are assumed to be additive and independent of the input. For nonlinear systems this is an oversimplification of the reality.

A simulation model \( \hat{y}_0(t) = h(u(t), Z_{t-1}, \theta) \) estimates the undisturbed output \( y_0(t) \), starting from a known input, without using measured output values. \( Z_{t-1} = [u(0), \ldots, u(t-1), y_0(t-1)] \) contains past inputs and past simulated outputs. This gives the possibility of experimenting on the model rather than on the actual, physical process to evaluate its behavior under various conditions.

A prediction model estimates a k-step ahead prediction \( \hat{y}(t+k) = g(u(t), Z_{t-1}, \theta) \) for the disturbed output \( y(t+k) \), starting from a known input \( u(t), t = 0, \ldots, u(t), \) and the measured past outputs \( y(t), t = 0, \ldots, y(t-1) \), in this case \( Z_{t-1} = [u(0), y(0), \ldots, y(t-1), y(t-1)] \).

Often the disturbance \( v(t) \) has a smooth behavior (it differs from white noise) so that its future values can be partly predicted from its past, which can be used to improve the quality of the prediction. The past values of \( v(t) \) are estimated as the difference between the past predicted and measured output values \( \hat{y}(t) = y(t) - \hat{y}(t) \). This is the intrinsic idea that is used in the development of optimal prediction models.

For white disturbances, the best prediction and simulation model become equal to each other because, by definition, it is impossible to predict white noise.

Whether a small prediction error is desirable or not depends on the nature of the disturbance \( v(t) \).

\( v(t) \) is dominated by measurement or sensor noise: sensor or measurement noise are not related at all with the process of interest. In that case the main goal is the elimination of this disturbance so that \( y_0(t) \) is the signal of interest, and hence a simulation model is the natural choice.

\( v(t) \) is dominated by process noise: process noise is an intrinsic part of the system output. It models that part of the system output that is due to inputs that are not known to the user. In control applications the goal is to impose a user defined behavior on the actual output \( y(t) = y_0(t) + v(t) \) so that the noise disturbance needs also to be included in the model. Moreover, the past outputs are available to decide on the next control action which turns the prediction model into a natural tool for these applications.

\( v(t) \) is dominated by model errors: If the model set is not rich enough to capture the true system, model errors appear which can also be represented as a ‘disturbance’ that is no longer independent of the input. The prediction method will decrease the impact of the model errors compared to the simulation method, using the correlation between neighboring errors. For that reason, it is more demanding to get a small simulation error than a small prediction error.

\[ y(t) = y_0(t) = h(u(t), Z_{t-1}, \theta) + h_{error}(u(t), Z_{t-1}) \] (2)

In this case \( v(t) \) is the model error that it depends on the input! If this error has some smoothness, it can again be partly predicted from its past values. However, for different classes of inputs this correlation will change, and the prediction errors can be come larger than the simulation errors.

3 Experimental illustration

The quality and the robustness of simulation and prediction models will be will be illustrated on the Silverbox benchmark. It will be shown, that in the case of model errors, all the results are strongly conditioned on the class of inputs.
Nonparametric Drift Model for Stochastic Differential Equations

Roland Hostettler, Filip Tronarp, Simo Särkkä
Department of Electrical Engineering and Automation, Aalto University, Finland
firstname.lastname@aalto.fi

1 Introduction

Stochastic differential equations (SDEs) are a powerful tool for modeling time series [1]. In this work, we consider nonparametric modeling of the nonlinear, time-varying drift function $f(x_t, u_t, t) = \left[f_1(x_t, u_t, t) \quad f_2(x_t, u_t, t) \quad \cdots \quad f_N(x_t, u_t, t)\right]^T$ in dynamic systems of the form

$$\begin{align*}
dx_t &= f(x_t, u_t, t)dt + dB_t, \quad (1a) \\
y_n &= Hx_{tn} + v_n, \quad (1b)
\end{align*}$$

where $t$ denotes continuous time, $x_t$ is the state vector, $u_t$ is a deterministic input, and $B_t$ is Brownian motion with diffusion matrix $Q$. Furthermore, $y_n$ is the measurement at time $t_n$, $H$ is the measurement matrix, and $v_n$ is a white noise sequence with covariance $R$. Given the SDE model (1), our aim is then to infer the drift function $f(x_t, u_t, t)$ and the diffusion matrix $Q$ from a set of system identification data $\{u_{1:N}, y_{1:N}\}$.

2 Model

Our approach is to model the drift function for the $t$th state variable as a zero-mean Gaussian process (GP) [2] with a covariance function that is separable in the input arguments $\{x_t, u_t\}$ and $t$ such that

$$f_t(x_t, u_t, t) \sim \mathcal{GP}(0, k_S(x_t, u_t, x'_t, u'_t)k_T(\tau)), \quad (2)$$

where $\tau = t - t'$. In order to reduce the computational complexity, we use eigenfunction decomposition of the covariance function $k_S(x_t, u_t, x'_t, u'_t)$ [3], together with spectral decomposition of the temporal covariance function $k_T(\tau)$ [4]. This yields the reduced rank GP drift model

$$\begin{align*}
f_t(x_t, u_t, t) &= \sum_{j=0}^{\infty} \alpha_j \psi_j(x_t, u_t), \quad (3a) \\
\dot{z}_{j,t} &= A_j \alpha_j dt + B_j \alpha_j dt, \quad (3b) \\
\alpha_j &= Cz_{j,t}, \quad (3c)
\end{align*}$$

where $\psi_j(x_t, u_t)$ is the $j$th eigenfunction of $k_S(x_t, u_t, x'_t, u'_t)$ and $\alpha_j$ is the corresponding time-varying coefficient. Furthermore, the dynamic system (3b)–(3c) results from the spectral decomposition of the covariance function $k_T(\tau)$ [4].

3 Results

We evaluate the proposed approach on the Bouc–Wen and the cascaded tanks benchmarks. In both examples, the hyperparameters of the GP priors are estimated by maximizing the marginal log-posterior of the training data using an extended Kalman filter and Euler discretization of the continuous-time model. Both benchmarks are evaluated in terms of the root mean squared error (RMSE) of the one step ahead prediction.

For the Bouc–Wen system, the validation RMSEs are $0.580 \times 10^{-5}$ and $0.096 \times 10^{-5}$ for the multisine and swept sine excitation test datasets, respectively. Furthermore, Fig. 1 shows the prediction error for the swept sine excitation validation dataset. For the cascaded tanks benchmark, the validation RMSE is $57.6 \times 10^{-3}$.

References

From Nonlinear Identification to Linear Parameter Varying Models: Benchmark Examples

Maarten Schoukens and Roland Tóth
Control Systems Group, Eindhoven University of Technology, Eindhoven, The Netherlands
Email: m.schoukens@tue.nl, r.toth@tue.nl

Introduction:
Linear parameter-varying (LPV) models form a powerful model class to analyze and control nonlinear systems [3]. Identifying an LPV model of a nonlinear system can be challenging due to the difficulty of selecting the scheduling variable(s) a priori, especially in case a first principles based understanding of the system is unavailable.


Approach:
A nonlinear system is identified first using a nonlinear block-oriented linear fractional representation (LFR) model [2], see Figure 1. This nonlinear LFR model class is embedded into the LPV model class by a specific factorization approach of the static nonlinear block present in the model. As a result of the factorization an LPV-LFR or an LPV state-space model with an affine dependency results. This approach also facilitates the selection of the scheduling variable from a data-driven perspective. The estimation is not affected by measurement noise on the scheduling variables, which is often left untreated by LPV model identification methods.

Results:
The result obtained on the two considered benchmark cases are shown in Figures 2 and 3. A large improvement can be observed by using the embedded LPV model compared to an LTI model, for both cases.

Acknowledgments:
Maarten Schoukens is supported the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement Nr 796627. Furthermore, 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).

References
An Investigation on the Wiener Approach for Nonlinear System Identification Benchmarks

Allahyar Montazeri  
Engineering Department, Lancaster University, Lancaster, LA14YW, UK  
a.montazeri@lancaster.ac.uk  

Mohamed Mehdi Arefi  
Department of Power and Control  
Shiraz University, Shiraz, Iran  
arefi@shirazu.ac.ir  

Mehdi Kazemi  
Department of Power and Control  
Shiraz University, Shiraz, Iran  
mhd.kazemi@gmail.com

1 INTRODUCTION

We evaluate the effectiveness of the Wiener model structure in modeling of the given benchmark problems. Two different approaches are proposed for parameter estimation. The results are compared for three problems, i.e. Silver Box, Wiener-Hammerstein, and Wiener-Hammerstein with noise. The aim is to evaluate the capability of the algorithms on the other benchmark problems in future works as well.

2 THE PROPOSED WIENER APPROACHES

In the following two different algorithms for parameter estimation of the Wiener model structure are proposed and they are applied to the benchmark problems.

A. Winere-Neural Nonlinear Identification [2]

State-space representation of a Wiener model can be stated as follows:

\[
\begin{align*}
x(k+1) &= Ax(k) + Bu(k) \\
z(k) &= Cx(k) + Du(k) \\
y(k) &= f(z(k)) + v(k)
\end{align*}
\]

where \( x(k) \) is \( n \times 1 \) the state vector at time, \( u(k) \) is \( m \times 1 \) the vector of control input, \( v(k) \) is the \( l \times 1 \) vector of measured output, and \( v(k) \) is a measurement noise assumed to be zero-mean and independent of \( u(k) \) for all \( k \). The system matrices \( A, B, C, D \) are real with proper dimensions and \( f(\cdot) \) is a nonlinear vector function defined on \( R^l \rightarrow R^n \). The first step is identification of linear part using state-space methods. So assuming the nonlinear mapping as an identity, the linear dynamics characterized by quadruple \( (A, B, C, D) \) will be identified. Then using the identified matrices \( (A, B, C, D) \), the output sequences of this LTI system \( \{z(k)\}_{k=1}^{\infty} \) will be computed. With this sequence a primary estimation of the nonlinear part of the Wiener model can be estimated. Here the static nonlinear term is identified using a single layer neural network.

B. Iterative Recursive Least Square Wiener [3]

The model structure of the Wiener model is shown in the figure below. It consists of an unknown linear transfer function followed by a parametrised polynomial nonlinearity. The noise model is also added to the output \( m(k) \). The parameters of the linear transfer function, static nonlinearity, and the noise model are estimated using an iterative recursive LS technique. Since in this formulation of the problem the intermediate signal \( x(n) \) and the noise signal \( e(n) \) are not already known, they are estimated in the iterative phase of the algorithm.

3 PRELIMINARY RESULTS

The results show that both techniques perform well to identify the SilverBox example [4]. The identification is performed against various scenarios for estimation and test signals and both techniques show robustness in this sense, i.e. the best fit is achieved for the multi-sine signal. Figure below shows the identification results for the Wiener-Hammerstein example [3]. The results confirm that the Wiener model structure captures the properties of the cascades Wiener-Hammerstein system when no process or measurement noise exist. Nevertheless, in case of noisy data the Wiener-Neural fails as it doesn’t have any measure to model the system noise while IR-Wiener shows some degrees of fit if the noisy data are used for estimation. Better tuning of the noise model may result in a better fit.

REFERENCES