Block-Oriented Identification using the Best Linear Approximation: Benefits and Drawbacks

Maarten Schoukens
Dept. ELEC, Vrije Universiteit Brussel, Belgium
Email: maarten.schoukens@vub.ac.be

***Abstract*—Due to their simplicity and structured nature, block-oriented models are popular in nonlinear modeling applications. A wide range of block-oriented identification algorithms were developed over the years. One class of these approaches uses the so-called best linear approximation to initialize the identification algorithm. The best linear approximation framework allows the user to extract important information about the system, it guides the user in selecting good candidate model structures and orders, and it proves to be a good starting point for nonlinear system identification algorithms. This paper gives an overview of the benefits and drawbacks of using identification algorithms based on the best linear approximation.**

I. Introduction

Nonlinear models are often used these days to obtain a better insight in the behavior of the system under test, to compensate for a potential nonlinear behavior using predistortion techniques, or to improve plant control performance. Popular nonlinear model structures are, amongst others, nonlinear state-space models [1], NARMAX models [2], and block-oriented models [3]. This paper focuses on the block-oriented class of models.

Many different types of block-oriented identification algorithms exist [3], where the linear-approximation based algorithms are amongst the more popular. One particular method to obtain such a linear approximation is the Best Linear Approximation (BLA) framework. This paper discusses the benefits and drawbacks of using BLA-based block-oriented system identification algorithms.

II. Block-Oriented Systems

Block-oriented models are constructed starting from two basic building blocks: a linear time-invariant (LTI) block and a static nonlinear block. Due to the separation of the nonlinear dynamic behavior into linear time invariant dynamics and the static nonlinearities, block-oriented nonlinear models are quite simple to understand and easy to use. They can be combined in many different ways. Series, parallel and feedback connections are considered in this paper, resulting in a wide variety of block-oriented structures as is depicted in Figure 1. These block-oriented models are only a selection of the many different possibilities that one could think of. For instance the generalized Hammerstein-Wiener structure that is discussed in [4] is not considered in this paper.

The LTI blocks and the static nonlinear blocks can be represented in many different ways. The LTI blocks are most often represented as a rational transfer function. The model order selection of the order of the numerator and denominator of the different blocks is a challenging problem. The static nonlinear block can again be represented in a nonparametric way using, for instance, kernel-based methods, or in a parametric way using, for instance, a linear-in-the-parameters basis function expansion (polynomial, piecewise linear, radial basis function network, ...), neural networks, or other dedicated parametrizations for static nonlinear functions. Again, in the parametric case, the nonlinear function complexity (number of basis functions, neurons, ...) needs to be selected by the user.

Another issue of block-oriented models is the uniqueness of the model parametrization. Gain exchanges, delay exchanges and equivalence transformations are present in many block-oriented structures [5]. This results in many different models with the same input-output behavior, but with a different parametrization.

It is assumed throughout this paper that a Gaussian additive, colored zero-mean noise source $n_y(t)$ with a finite variance σ^2 is present at the output of the system only:

$$y(t) = y_0(t) + n_y(t). \tag{1}$$

This noise $n_y(t)$ is assumed to be independent of the known input $u(t)$. The signal $y(t)$ is the actual output signal and a subscript 0 denotes the exact (unknown) value.

III. Best Linear Approximation

A linear model often explains a significant part of the behavior of a (weakly) nonlinear system. This approximative linear model also provides the user with a better insight into the behavior of the system under test. It motivates the use of a framework that approximates the behavior of a nonlinear system by a linear time invariant model. This paper considers the Best Linear Approximation (BLA) framework [6], [7] to estimate a linear approximation of a nonlinear system..

The BLA is best in mean square sense for a fixed class of input signals $\mathbb{U}$ only, it is defined in [6], [7] as:

$$G_{bla}(q) \triangleq \underset{G(q)}{\arg\min}\, E\left\{\left|\tilde{y}(t) - G(q)\tilde{u}(t)\right|^2\right\}, \tag{2}$$

where $E\{.\}$ denotes the expected value operator. The expected value $E\{.\}$ is taken w.r.t. the random input $\tilde{u}(t)$. The zero-mean signals $\tilde{u}(t)$ and $\tilde{y}(t)$ are defined as:

$$\tilde{u}(t) \triangleq u(t) - E\{u(t)\}, \tag{3}$$

$$\tilde{y}(t) \triangleq y(t) - E\{y(t)\}. \tag{4}$$

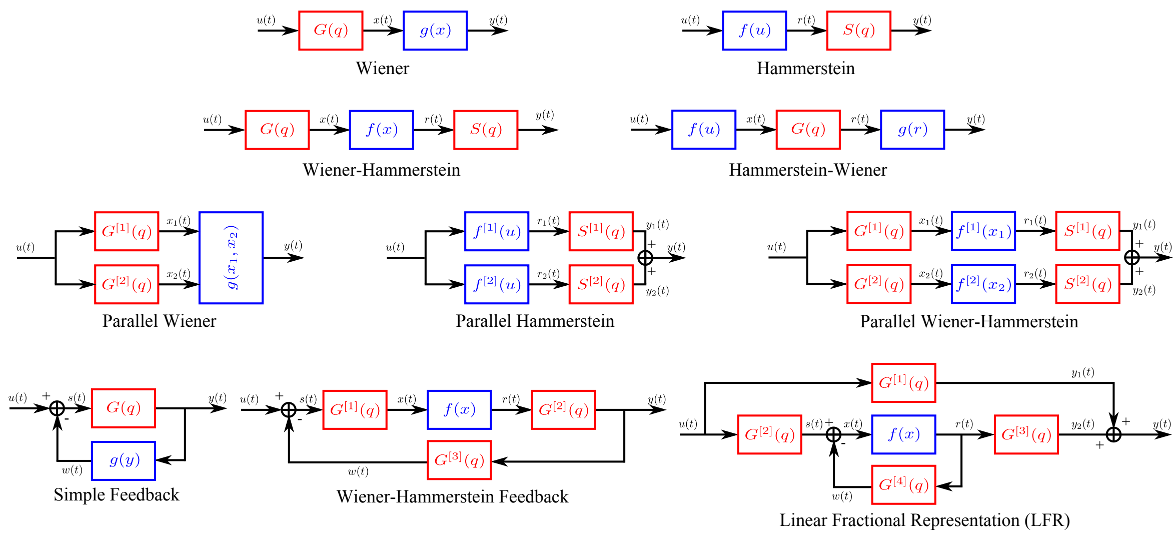

Fig. 1. An overview of possible block-oriented structures. The different structures are obtained by using series, parallel and feedback connections of LTI blocks ($G(q)$ and $S(q)$) and static nonlinear blocks ($f(\cdot)$ and $g(\cdot)$). There are three types of structure classes: single branch structures (Wiener, Hammerstein, Wiener-Hammerstein and Hammerstein-Wiener), parallel branch structures (parallel Wiener, parallel Hammerstein, parallel Wiener-Hammerstein) and feedback structures (simple feedback structure, Wiener-Hammerstein feedback and LFR).

The error in eq. (2) is minimized by [6], [7]:

$$G_{bla}(e^{j\omega T_s}) = \frac{S_{\tilde{Y}\tilde{U}}(e^{j\omega T_s})}{S_{\tilde{U}\tilde{U}}(e^{j\omega T_s})}, \tag{5}$$

where $S_{\tilde{Y}\tilde{U}}$ and $S_{\tilde{U}\tilde{U}}$ denote the crosspower of $\tilde{y}(t)$ and $\tilde{u}(t)$ and autopower of $\tilde{u}(t)$ respectively.

The BLA of a system depends on the signal class $\mathbb{U}$ that is used. This work considers $\mathbb{U}$ to be the Riemann equivalence class of asymptotically normally distributed excitation signals. The signal class $\mathbb{U}$ contains Gaussian noise sequences, but contains also periodic signal sets known as random phase multisines [7]. This class of signals will be referred to as Gaussian signals in the remainder of this paper. When stationary Gaussian excitation signals are used, the BLA of many block-oriented systems becomes a simple function of the linear dynamics that are present in that system.

A. BLA of Single Branch Structures

The BLA of Wiener, Hammerstein and Wiener-Hammerstein structures are a very simple expression of the LTI-blocks present in the block-oriented system when Gaussian excitation signals are used. Due to Bussgang's Theorem [8] one obtains the following expression for the BLA of a Wiener-Hammerstein structure [9]:

$$G_{bla}(q) = \lambda G(q) S(q), \tag{6}$$

where q^{-1} is the backwards shift operator, and λ is a gain depending on the system and considered class of input (input power, offset and power spectrum). Note that the poles and zeros of the BLA are the poles and zeros of the LTI blocks present in the system [9].

B. BLA of Parallel Branch Structures

The output of a parallel branch structures is the summation of multiple single branch system, hence, the BLA of a parallel Wiener-Hammerstein system is given by:

$$G_{bla}(q) = \sum_{i=1}^{n_{br}} \lambda_i G^{[i]}(q) S^{[i]}(q). \tag{7}$$

Note that zeros of the BLA depend on the gains λ_i, while the poles of the BLA are the poles of the LTI blocks of the parallel Wiener-Hammerstein system [10].

C. BLA of structures containing feedback

Bussgangs Theorem cannot be used anymore in the case of nonlinear feedback structures since the input of the static nonlinear block is not Gaussian anymore. Therefore, only approximate expressions of the BLA are given here. The BLA of a simple feedback structure is approximately given by [1]:

$$G_{bla}(q) \approx \frac{G(q)}{1 + \lambda G(q)}. \tag{8}$$

The case of the LFR structure is more involved [11]:

$$G_{bla}(q) \approx G^{[1]}(q) + \frac{G^{[2]}(q) G^{[3]}(q)}{1 + \lambda G^{[4]}(q)}. \tag{9}$$

It can be observed that the poles of the BLA of a simple feedback structure depend on the gain $lambda$. In the case of the LFR structure both the poles and the zeros depend on the gain $lambda$.

IV. Detection of Nonlinearity

Although one might know beforehand that a given system is nonlinear, it can very well turn out that, for the class of signals that will realistically act on the system, and for the frequency region of interest, and application on hand, no significant nonlinear behavior is observed. In such a case, much modeling effort can be spared by simply estimating the BLA of the system, and using it later on for its intended task (control design, simulation, system analysis, ...).

The BLA framework allows a user to detect and quantify the level of the nonlinear distortions. Based on this analysis one can see, for a chosen class of input signals, if the effects of the nonlinearity are dominant or not, in which frequency region the nonlinearity is active, and how much can be gained by estimating a nonlinear model [7].

V. Model Order and Model Structure Selection

The model structure and model order selection problem is a though challenge in many nonlinear system identification problems. Given an input/output dataset the user has to decide what type of nonlinear model will be used (e.g. Hammerstein, Wiener, nonlinear feedback, ...) and which model orders (e.g. orders of the dynamics and degree of the static nonlinearity) are to be selected.

A. Model Order Selection

The model order selection problem in block-oriented modeling problems is much harder than the one in the LTI framework. One needs to select the model order of each LTI block separately and on top of this, also the complexity of the static nonlinearity needs to be decided on. Using the BLA framework to start the modeling of a block-oriented system allows one to extract the model orders of the combined linear dynamics present in the block-oriented structure. Indeed, the BLA is in many cases a simple function of the underlying linear blocks of the block-oriented system under test (see Section III).

An important part of the model order selection problem, the selection of the order of the dynamics of the system, is now taking place in a linear framework on the BLA, separate from the nonlinearity selection problem. This results in a problem which is much more simple and better understood by many researchers and practitioners.

The selected model orders of the BLA can be used directly in the nonlinear modeling step (Hammerstein, Wiener, Simple Feedback structure), or they are translated automatically in a second step into the model orders of each LTI block separately using either pole-zero allocation algorithms [9], [10] to split the dynamics over the front and the back (Wiener-Hammerstein, parallel Wiener-Hammerstein and Wiener-Hammerstein Feedback structure), singular value decomposition approaches [10], [12] to split the dynamics over the parallel branches (parallel Hammerstein, parallel Wiener and parallel Wiener-Hammerstein), or by solving a Riccati equation in the case of the LFR structure [11].

B. Model Structure Selection

The model structure selection problem can also be tackled in part by taking a closer look at the expression of the BLA for the different block-oriented model structures. It is discussed in [13] how the BLA behaves when it is estimated at different setpoints of the system (different input amplitudes, constant offsets or power spectra).

Single-branch system structures such as the Wiener-Hammerstein structure will only exhibit a varying gain over the different BLA setpoints, while parallel branch systems exhibit varying zeros and feedback structures exhibit varying poles over the different BLA setpoints (see Table I). Note that the LFR structure is both a parallel branch and a nonlinear feedback structure. Hence, both poles and zeros of the BLA will depend on the input power, constant offset and power spectrum. This analysis demonstrates how one can quickly detect some important structural features of the nonlinear system under test using only linear approximations of that system.

VI. Drawbacks of BLA-Based Modeling

Of course, obtaining a sufficiently high-quality estimate of the BLA (sufficiently low variance on the BLA) comes at a cost. The variance on the estimated BLA depends on how nonlinear the system under test is. If the system is very nonlinear, a significant error is introduced when the least-squares linearization is performed. This results in a high variance on the estimate. The classical approach to lower the variance on the estimated BLA is to use more input-output data. Hence, it can be the case that to model a strongly nonlinear system using the BLA, a larger dataset is required compared to some of the approaches that take the nonlinearity directly into account.

Another issue can be the presence of nonlinearities which give rise to a BLA equal to zero over all frequencies. Of course this is input dependent: this problem can be circumvented by doing measurements at different constant offset of the input signal. For example, the BLA of an even nonlinearity using a zero-mean Gaussian input is equal to zero [8]. However, when an constant offset is added to this input signal a non-zero BLA is obtained.

A last remark concerns the systems with nonlinear feedback (simple feedback structure, Wiener-Hammerstein feedback and LFR). The BLA expressions given in this paper for these systems are not exact. Although they are a good approximation of reality, more involved effects come into play due to the non-Gaussianity of the signal at the input of the static nonlinearity. However, it is observed in many practical applications that the simplification used in this paper does lead to good model estimates, e.g. [1].

Note that, aside the drawbacks listed above, many other challenges exists. The selection of the nonlinearity present in the model, the validation of the system structure over a wide range of use, dealing with model errors in a proper manner, using more involved noise frameworks, and many more are all open problems in nonlinear system identification.

VII. Example: Silverbox

The Silverbox system (an electronic realization of the duffing oscillator) is studied (see for instance [1]) here as a simple illustration of the theory explained in the previous sections.

TABLE I

MODEL STRUCTURE SELECTION USING THE BLA BY OBSERVING THE GAIN, POLES AND ZEROS OF THE BLA ESTIMATED AT MULTIPLE SETPOINTS OF THE SYSTEM (W-H STANDS FOR WIENER-HAMMERSTEIN).

	LTI	Wiener	Hammerstein	W-H	Parallel W-H	Simple Feedback	W-H Feedback	LFR
Gain	fixed	varying	varying	varying	varying	varying	varying	varying
Poles	fixed	fixed	fixed	fixed	fixed	varying	varying	varying
Zeros	fixed	fixed	fixed	fixed	varying	fixed	fixed	varying

As a first step the nonparametric BLA is estimated. The estimated FRF and the total (noise + nonlinearities) and noise distortion variance are shown in Figure 2 for two different amplitudes of the input excitation. Based on this figure, the user can observe that a nonlinear model would not offer much improvement in the low input level case. On the other hand, the nonlinear contribution are almost as large as the linear one for the high input level case.

A clear shift in the resonance frequency can be observed. This is a strong indication for a shifting pole, and hence, the presence of a nonlinear feedback in the system. The system dynamics are also clearly visible, and can easily be determined using the linear model order selection techniques, 2nd order dynamics are clearly present.

To conclude, a nonlinear feedback model structure with 2nd order linear dynamics should be a good candidate to model the behavior of the Silverbox system when it is excited by the high amplitude input level. This corresponds with the known underlying structure of the Silverbox system, it is a simple feedback structure. A linear model will be qualitative enough in the low input case.

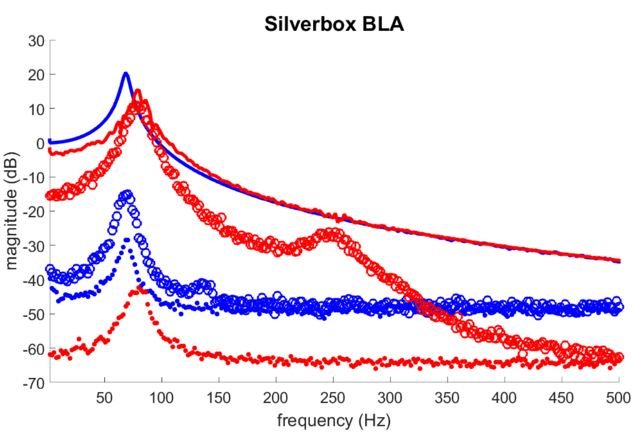

Fig. 2. The BLA of the Silverbox system for low (blue) and high (red) level of excitation. The FRF (full line), total distortion levels (circles) and noise distortion levels (dots) are shown. The size of the gap between the noise and total distortion level indicate how nonlinear the system behaves.

VIII. CONCLUSION

The paper has presented an overview on how the complexity of the (block-oriented) nonlinear modeling process can be reduced significantly using the Best Linear Approximation. The BLA framework offers answers to the questions: "Should I use a nonlinear model for the application at hand?", "What model structure should I select?", and "How can I select the model orders in a simple but efficient way?".

The main disadvantage of using the BLA framework is that it possibly requires more data than some of the other nonlinear modeling approaches which are available. Furthermore, one has to be aware that the BLA can be equal to zero in the presence of nonlinearities which are even around the setpoint of the input signal. This can, of course, easily be solved by changing the constant offset of the input signal used.

ACKNOWLEDGMENT

This work was partly funded by the Belgian Government through the Inter university Poles of Attraction IAP VII/19 DYSCO program, and the ERC advanced grant SNLSID, under contract 320378.

REFERENCES

[1] J. Paduart, "Identification of Nonlinear Systems using Polynomial Nonlinear State Space Models," Ph.D. dissertation, Vrije Universiteit Brussel, Belgium, 2008.

[2] S. Billings, *Nonlinear System Identification: NARMAX Methods in the Time, Frequency, and Spatio-Temporal Domains*. West Sussex, United Kingdom: John Wiley & Sons, Ltd., 2013.

[3] F. Giri and E. Bai, Eds., *Block-oriented Nonlinear System Identification*, ser. Lecture Notes in Control and Information Sciences. Berlin Heidelberg: Springer, 2010, vol. 404.

[4] A. Wills and B. Ninness, "Generalised Hammerstein-Wiener system estimation and a benchmark application," *Control Engineering Practice*, vol. 20, no. 11, pp. 1097–1108, 2012.

[5] M. Schoukens, R. Pintelon, and Y. Rolain, "Identification of Wiener-Hammerstein systems by a nonparametric separation of the best linear approximation," *Automatica*, vol. 50, no. 2, pp. 628–634, 2014.

[6] M. Enqvist and L. Ljung, "Linear approximations of nonlinear FIR systems for separable input processes," *Automatica*, vol. 41, no. 3, pp. 459–473, 2005.

[7] R. Pintelon and J. Schoukens, *System Identification: A Frequency Domain Approach*, 2nd ed. Hoboken, New Jersey: Wiley-IEEE Press, 2012.

[8] J. Bussgang, "Cross-correlation functions of amplitude-distorted Gaussian signals," MIT Laboratory of Electronics, Tech. Rep. 216, 1952.

[9] J. Sjöberg and J. Schoukens, "Initializing Wiener-Hammerstein models based on partitioning of the best linear approximation," *Automatica*, vol. 48, no. 2, pp. 353–359, 2012.

[10] M. Schoukens, A. Marconato, R. Pintelon, G. Vandersteen, and Y. Rolain, "Parametric identification of parallel Wiener-Hammerstein systems," *Automatica*, vol. 51, no. 1, pp. 111 – 122, 2015.

[11] L. Vanbeylen, "Nonlinear LFR Block-Oriented Model: Potential Benefits and Improved, User-Friendly Identification Method," *IEEE Transactions on Instrumentation and Measurement*, vol. 62, no. 12, pp. 3374–3383, 2013.

[12] M. Schoukens and Y. Rolain, "Parametric Identification of Parallel Wiener Systems," *IEEE Transactions on Instrumentation and Measurement*, vol. 61, no. 10, pp. 2825–2832, 2012.

[13] J. Schoukens, R. Pintelon, Y. Rolain, M. Schoukens, K. Tiels, L. Vanbeylen, A. Van Mulders, and G. Vandersteen, "Structure discrimination in block-oriented models using linear approximations: A theoretic framework," *Automatica*, vol. 53, pp. 225–234, 2015.