



Survey paper

Identification of block-oriented nonlinear systems starting from linear approximations: A survey[☆]

Maarten Schoukens, Koen Tiels

Vrije Universiteit Brussel, Dept. ELEC, Pleinlaan 2, B-1050 Brussels, Belgium



ARTICLE INFO

Article history:

Received 4 July 2016

Received in revised form 9 June 2017

Accepted 27 June 2017

Available online 19 September 2017

Keywords:

System identification

Maximum likelihood

Nonlinear systems

Hammerstein

Wiener

Wiener–Hammerstein

Hammerstein–Wiener

Parallel cascade

Feedback

Best linear approximation

Linear fractional representation

ABSTRACT

Block-oriented nonlinear models are popular in nonlinear system identification because of their advantages of being simple to understand and easy to use. Many different identification approaches were developed over the years to estimate the parameters of a wide range of block-oriented nonlinear models. One class of these approaches uses linear approximations to initialize the identification algorithm. The best linear approximation framework and the ϵ -approximation framework, or equivalent frameworks, allow the user to extract important information about the system, guide the user in selecting good candidate model structures and orders, and prove to be a good starting point for nonlinear system identification algorithms. This paper gives an overview of the different block-oriented nonlinear models that can be identified using linear approximations, and of the identification algorithms that have been developed in the past. A non-exhaustive overview of the most important other block-oriented nonlinear system identification approaches is also provided throughout this paper.

© 2017 Elsevier Ltd. All rights reserved.

1. Introduction

Nonlinear models are often used to obtain a better insight into 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 (Paduart, 2008; Schön, Wills, & Ninness, 2011; Suykens, De Moor, & Vandewalle, 1995), NARX and NARMAX models (Billings, 2013; Leontaritis & Billings, 1985), and block-oriented nonlinear models (Giri & Bai, 2010). Due to the separation of the nonlinear dynamic behavior into linear time invariant (LTI) dynamics and the static nonlinearities (SNL), block-oriented nonlinear models are quite simple to understand and easy to use.

This paper provides a survey of the identification of block-oriented nonlinear models using input–output data only. The

[☆] This work was supported in part by the research council of the VUB (OZR bridging mandate), 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. The material in this paper was not presented at any conference. This paper was recommended for publication in revised form by Editor John Bailiul.

E-mail addresses: maarten.schoukens@vub.ac.be (M. Schoukens), ktiels@vub.ac.be (K. Tiels).

identification algorithms that are discussed in this paper are all initialized using linear approximation frameworks for nonlinear systems: the Best Linear Approximation (BLA) (Enqvist, 2005; Enqvist & Ljung, 2005; Pintelon & Schoukens, 2012) and the ϵ -approximation (Schoukens, Pintelon, Rolain, Schoukens, Tiels, Vanbeylen, Van Mulders, & Vandersteen, 2015a) framework. An overview on the practical use of the BLA is provided in Schoukens, Vaes, and Pintelon (2016a), however, that paper does not address the use of the BLA for block-oriented nonlinear modeling purposes. Many other block-oriented identification approaches, different from the linear approximation approaches, are described in the literature (e.g. overparametrization, iterative, or kernel-based methods). A non-exhaustive overview of the most important approaches for each model structure is given at the beginning of each section discussing that particular model structure. The presented identification approaches and the considered block-oriented structures go beyond the methods and structures that previous surveys, such as Billings (1980) Giri and Bai (2010) Hasiewicz, Mzyk, Śliwiński and Wachel (2012), present.

Block-oriented identification approaches starting from linear approximations have the advantage of addressing part of the model selection problem in an LTI-framework, which is better understood by many practitioners. These approaches obtain good

results in various benchmark problems (such as the Wiener–Hammerstein (Schoukens, Suykens, & Ljung, 2009b) and Silverbox benchmarks (Wigren & Schoukens, 2013)), as illustrated in Section 12 of this paper. Linear models of nonlinear systems can also be used to detect the underlying block-oriented structure of the system under test. Haber and Unbehauen (1990), Lauwers, Schoukens, Pintelon, and Enqvist (2008) and Schoukens et al. (2015a) discuss this structure detection problem using linear approximations, however, they do not present identification approaches there.

The different block-oriented structures that are studied throughout this paper are introduced in Section 2. Bussgang's theorem is very useful for determining the BLA of many block-oriented structures. This theorem is discussed in Section 3. The considered linear approximation frameworks are discussed next in Section 4. Section 5 introduces the cost function that is considered throughout this paper. The identification algorithms for the single branch models (Hammerstein, Wiener, Hammerstein–Wiener and Wiener–Hammerstein) are studied in Section 7. The parallel block-oriented nonlinear modeling approaches are discussed in Section 8. Section 9 considers the identification of different nonlinear feedback models. Section 10 provides an overview of the pros and the cons of the model structures and their identification based on the BLA. Some guidelines for the user are presented in Section 11. The good performance of the identification methods discussed in this paper is illustrated by some benchmark results in Section 12. Finally, some open research problems are highlighted in Section 13.

2. Block-oriented nonlinear models: Structures, representation and noise framework

Model structures

Block-oriented nonlinear models are constructed starting from two basic building blocks: a linear time-invariant (LTI) block and a static nonlinear block. 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 Fig. 1. These block-oriented nonlinear models are only a selection of the many different possibilities that one could think of. For instance, this paper does not consider the generalized Hammerstein–Wiener structure that Wills and Ninness (2012) discuss, nor the Hammerstein/nonlinear feedback model structure that Van Pelt and Bernstein (2001) discuss. A comparison of the model complexity and the identification procedures of the considered model structures is made in Section 10.

Model representation

The LTI blocks and the static nonlinear blocks can be represented in many different ways. The LTI block could for instance be modeled as a nonparametric frequency response function (Giri, Rochdi, Radouane, Broui, & Chaoui, 2013) or an impulse response (Lacy & Bernstein, 2003), or they can be parametrized using a state space (Verhaegen & Westwick, 1996), rational transfer function (Narendra & Gallman, 1966), or basis function expansion (Tiels & Schoukens, 2014). The static nonlinear block can again be represented in a nonparametric way using, for instance, kernel-based methods (Mzyk, 2014), or in a parametric way using, for instance, a linear in the parameters basis function expansion (polynomial, piecewise linear, radial basis function network, ...) (Schoukens, Pintelon, & Rolain, 2014a), neural networks (Schoukens, Marconato, Pintelon, Vandersteen, & Rolain, 2015b), or other dedicated parametrizations for static nonlinear functions.

The methods that are presented in this paper will typically use, but are not limited to, a rational transfer function representation for the LTI blocks and a linear in the parameters basis function expansion for the static nonlinearity.

Another issue of block-oriented nonlinear models is the uniqueness of the model parametrization. Gain exchanges, delay exchanges and equivalence transformations are present in many block-oriented structures (Schoukens, 2015; Schoukens et al., 2015b). This results in many different models with the same input–output behavior, but with a different parametrization. The Jacobian of the cost function can be rank deficient due to these indistinguishability issues that are common for block-oriented identification problems. The related degenerations in the Jacobian need to be taken into account during the optimization.

The indistinguishability issues do not only impact on the parametrization of the chosen model structure, they also affect the model structures themselves. The simple feedback structure (see Fig. 1) can both be represented using an LTI block in the forward path and a static nonlinearity in the feedback path, or the other way around, a static nonlinearity in the forward path and an LTI block in the feedback path as is depicted in Fig. 2. Both model structures lead to the same input–output representation, but are structurally different (Schoukens, Gommé, Van Moer, & Rolain, 2008). Indeed, we have that the outputs of the two possible simple feedback structures are given by:

$$\begin{cases} y(t) = G(q)[u(t) - g(y(t))], \\ \bar{y}(t) = \bar{g}(u(t) - \bar{G}(q)\bar{y}(t)). \end{cases} \quad (1)$$

Both representations should have the same input–output relationship, which results in the following equation:

$$G^{-1}(q)y(t) + g(y(t)) = \bar{g}^{-1}(\bar{y}(t)) + \bar{G}(q)\bar{y}(t). \quad (2)$$

It can be easily observed that this equation holds for $\bar{G}(q) = G^{-1}(q)$ and $\bar{g}(\cdot) = g^{-1}(\cdot)$. It is assumed here that both $G(q)$ and $g(y)$ are invertible. Note that the equivalence transformation can introduce an unstable (when $G(q)$ is not minimal-phase) and non-causal LTI block (when $G(q)$ contains a pure delay).

Noise framework

Assumption 1 (Noise Framework). 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:

$$y(t) = y_0(t) + n_y(t). \quad (3)$$

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.

The choice of the noise framework that is described in Assumption 1 is a simplified representation of reality. This simplification can lead to biased estimates when there are other noise sources present, located at other positions inside the system, e.g. process noise passing through a nonlinear subsystem (Hagenblad, Ljung, & Wills, 2008). A more realistic noise framework can be obtained by introducing multiple noise sources, or by placing the noise source at a different location in the considered system structure. For instance, Hagenblad et al. (2008), Lindsten, Schön, and Jordan (2013), Wahlberg, Welsh, and Ljung (2014) and Wills, Schön, Ljung, and Ninness (2013) consider multiple noise sources. This more realistic noise framework comes often at the cost of a more complex identification algorithm.

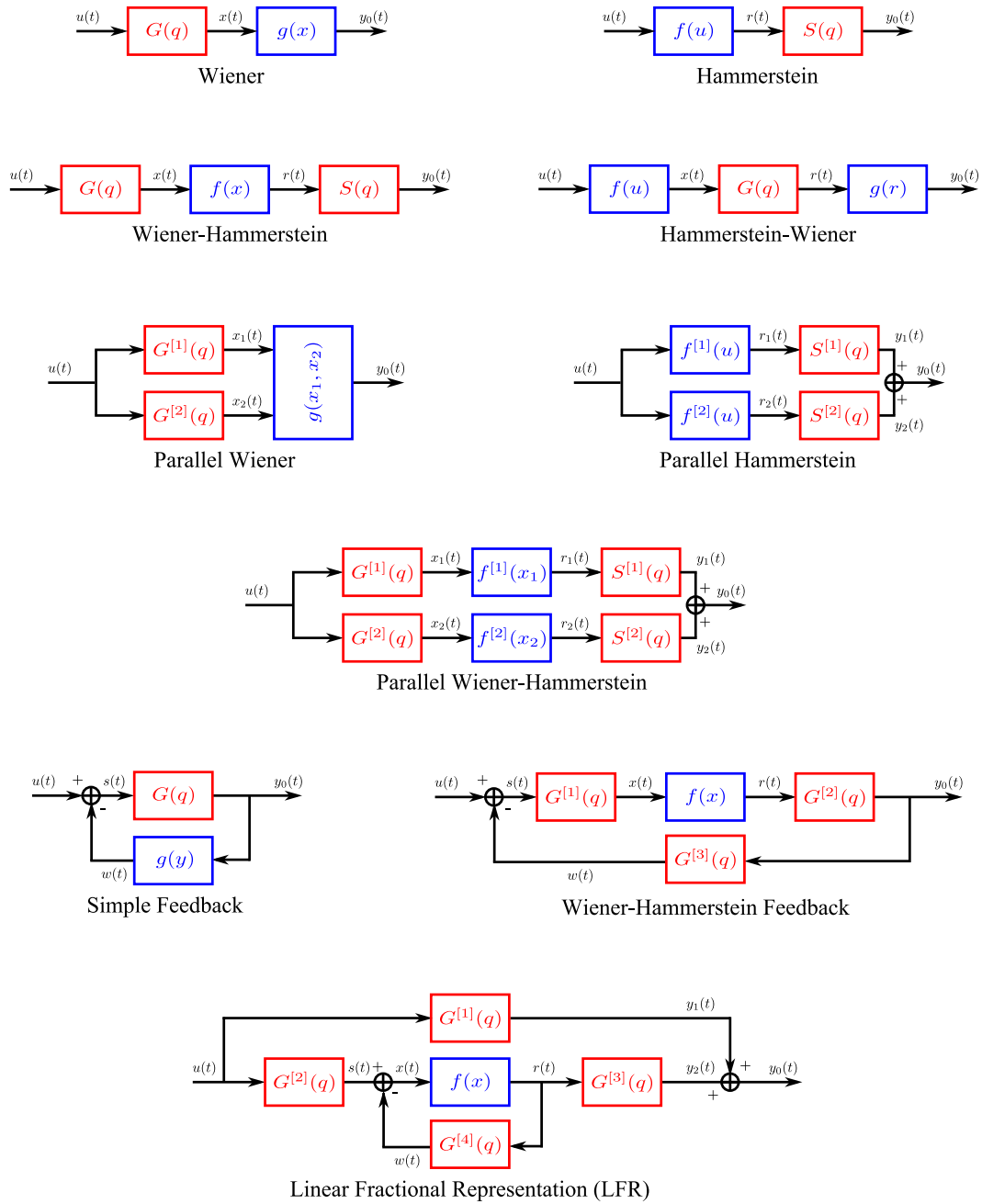


Fig. 1. The block-oriented structures that are considered in this paper. The different structures have been obtained by using series, parallel and feedback connections of LTI blocks $G(q)$ and $S(q)$ and static nonlinear blocks $f(\cdot)$ and $g(\cdot)$. There are three types of structure classes: single 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).

3. Bussgang's theorem and separable processes

3.1. Bussgang's theorem

Bussgang's theorem (Bussgang, 1952) is stated in Papoulis (1991) as follows:

Theorem 1. *If the input to a memoryless system $y = g(u)$ is a zero-mean stationary Gaussian input $u(t)$, the cross-correlation of $u(t)$ with the resulting output $y(t) = g(u(t))$ is proportional to $R_{uu}(\tau)$:*

$$R_{uy}(\tau) = \alpha R_{uu}(\tau) \quad \text{where} \quad \alpha = E\{g'(u(t))\} \quad (4)$$

Proof. See Bussgang (1952) and Papoulis (1991). \square

In the frequency domain this becomes:

$$S_{YU}(e^{j\omega T_s}) = \alpha S_{UU}(e^{j\omega T_s}), \quad (5)$$

where T_s is the sampling period, S_{YU} and S_{UU} denote the crosspower spectrum of $y(t)$ and $u(t)$ and autopower spectrum of $u(t)$, respectively.

Bussgang's theorem proves to be very valuable for the analysis of block-oriented structures using linear approximations as is shown in the remainder of this paper.

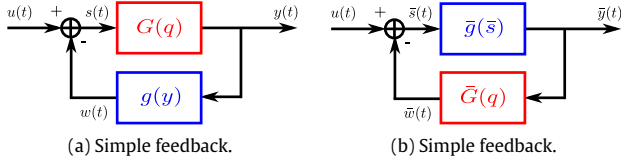


Fig. 2. Two simple feedback block-oriented nonlinear model structures: (a) LTI block $G(q)$ in the forward path and a static nonlinearity $g(\cdot)$ in the feedback path, (b) a static nonlinearity $\bar{g}(\cdot)$ in the forward path and an LTI block $\bar{G}(q)$ in the feedback path. Both structures share the same input–output behavior when $\bar{g}(\cdot) = g^{-1}(\cdot)$ and $\bar{G}(q) = G^{-1}(q)$.

3.2. Riemann equivalence class of asymptotically normally distributed excitation signals

Here we consider the Riemann equivalence class of asymptotically normally distributed excitation signals \mathbb{S}_U (Pintelon & Schoukens, 2012; Schoukens, Lataire, Pintelon, Vandersteen, & Dobrowiecki, 2009a). This signal class mainly contains Gaussian noise sequences, but considering the Riemann class allows one to extend the signal class to also contain periodic signal sets, while Bussgang's Theorem still applies (Pintelon & Schoukens, 2012).

Definition 1 (Riemann Equivalence Class of Asymptotically Normally Distributed Excitation Signals \mathbb{S}_U). Consider a signal $u(t)$ with a piecewise continuous power spectrum $S_{UU}(e^{j\omega T_s})$, with a finite number of discontinuities. A random signal $u(t)$ belongs to the Riemann equivalence class if it obeys any of the following statements:

- (1) $u(t)$ is a Gaussian noise signal with power spectrum $S_{UU}(e^{j\omega T_s})$.
- (2) $u(t)$ is a random multisine or random phase multisine (Pintelon & Schoukens, 2012) such that:

$$\frac{1}{N} \sum_{k=k_1}^{k_2} E \{ |U(k)|^2 \} = \frac{1}{2\pi} \int_{\omega_{k_1}}^{\omega_{k_2}} S_{UU}(e^{j\nu T_s}) d\nu + O(N^{-1}),$$

with $\omega_k = k \frac{2\pi f_s}{N}$, $k \in \mathbb{N}$, $0 < \omega_{k_1} < \omega_{k_2} < \pi f_s$, f_s is the sampling frequency, and $U(k)$ is the discrete Fourier spectrum of $u(t)$. $E \{ \cdot \}$ denotes the expected value operator.

A random phase multisine $u(t)$ is a periodic signal with period length $\frac{N}{f_s}$ defined in Pintelon and Schoukens (2012) as:

$$u(t) = N^{-1/2} \sum_{k=1}^{N/2-1} 2U_k \cos \left(2\pi k \frac{t}{N} f_s + \varphi_k \right). \quad (6)$$

The phases φ_k are random variables that are independent over the frequency and are a realization of a random process on $[0, 2\pi[$, such that $E\{e^{j\varphi_k}\} = 0$. For instance, the random phases can be uniformly distributed over the interval $[0, 2\pi[$. The (real-valued) amplitude U_k is set in a deterministic way by the user and is uniformly bounded by M_U ($0 \leq U_k \leq M_U < \infty$). Random phase multisines have the advantage of being periodic signals. This avoids the adverse effects of spectral leakage for a proper choice of the period length $\frac{N}{f_s}$. They also offer full control over the applied amplitude spectrum to the user. The Riemann equivalence ensures that a random phase multisine is asymptotically ($N \rightarrow \infty$) Gaussian distributed (Pintelon & Schoukens, 2012).

3.3. Separable processes

Bussgang's theorem has been extended to other classes of signals, besides the Gaussian class, using the concept of separable

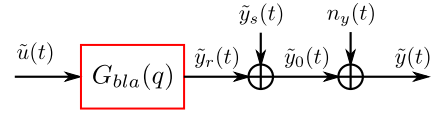


Fig. 3. The BLA of a nonlinear system for a given class of excitation signals \mathbb{U} consists of the resulting LTI model $G_{bla}(q)$, the unmodeled nonlinear contributions $\tilde{y}_s(t)$, and the additive noise source $n_y(t)$ that is assumed to be present at the output of the nonlinear system. The zero-mean input excitation $\tilde{u}(t)$ belongs to the signal class \mathbb{U} , $\tilde{u}(t)$ and $\tilde{y}(t)$ are defined in Eqs. (8) and (9).

processes that is introduced in Nuttall (1958). Gaussian processes, sine wave and phase modulated signals are shown to be separable processes in Nuttall (1958). Furthermore, McGraw and Wagner (1968) have shown that the signals belonging to the class of elliptically symmetric distributions are separable. Also random phase multisines with flat amplitude spectra are proven to be separable in Enqvist (2011). A more in-depth discussion of separable processes for nonlinear system identification can be found in Billings and Fakhouri (1978b), Enqvist (2010) and Enqvist and Ljung (2005).

4. Linear approximations of nonlinear systems

Most real-world systems do not behave completely linearly. Nevertheless, a linear model often explains a significant part of the behavior of a (weakly) nonlinear system. This approximate 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 under well-chosen system-specific boundary conditions. This paper uses the BLA framework (Enqvist, 2005; Enqvist & Ljung, 2005; Ljung, 2001; Pintelon & Schoukens, 2012; Schoukens et al., 2016a) and the ϵ -approximation (Schoukens et al., 2015a) to estimate a linear approximation of a nonlinear system. This section briefly introduces the theoretical framework of the BLA and the ϵ -approximation.

4.1. System class

Both the BLA and the ϵ -approximation consider the PISPO (periodic in, the same period out) system class. This class of nonlinear systems includes systems whose output can be approximated arbitrarily well in mean square sense by a uniformly bounded Volterra series (see Pintelon and Schoukens (2012), Schetzen (2006) and Schoukens, Dobrowiecki, and Pintelon (1998), for more details). Systems belonging to this class respond to a periodic input with a periodic steady-state output with the same period (Boyd, Chua, & Desoer, 1984), hence the acronym PISPO. This system class is linked with the bounded-input bounded-output (BIBO) stability criterion (Schoukens, 2015). Nonlinear phenomena generating sub-harmonics such as chaos and bifurcations are excluded. Saturation, clipping, and dead zone nonlinearities are, for example, still included in this system class.

4.2. Best linear approximation

The theoretical framework of the BLA results in a model whose structure is shown in Fig. 3. It includes four components: one boundary condition being an input excitation class \mathbb{U} , and three model constituents, namely a Linear Time Invariant (LTI) model labeled $G_{bla}(q)$, a perturbation noise source labeled $n_y(t)$, and a nonlinear distortion source labeled $\tilde{y}_s(t)$.

As mentioned above, the BLA of a system depends on the chosen signal class \mathbb{U} . Most identification methods that are discussed in this work consider \mathbb{U} to be the Riemann equivalence class of asymptotically normally distributed excitation signals \mathbb{S}_U . When this class of signals is considered to provide an excitation signal,

the BLA of many block-oriented nonlinear systems becomes, due to Bussgang's Theorem (see Section 3), a simple function of the linear dynamics that are present in that system. All signals are assumed to be stationary for the remainder of this work.

The BLA model of a nonlinear system is an LTI approximation of the behavior of that system. It is best in mean square sense for a fixed class of input signals \mathbb{U} only. The BLA is defined in Enqvist (2005), Enqvist and Ljung (2005) and Pintelon and Schoukens (2012) as:

$$G_{bla}(q) = \underset{G(q)}{\operatorname{argmin}} E \left\{ \left| \tilde{y}(t) - G(q)\tilde{u}(t) \right|^2 \right\}, \quad (7)$$

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

$$\tilde{u}(t) = u(t) - E\{u(t)\}, \quad (8)$$

$$\tilde{y}(t) = y(t) - E\{y(t)\}. \quad (9)$$

This definition of the BLA is equal to the definition of the linear time invariant second order equivalent model defined in Enqvist (2005), Enqvist and Ljung (2005) and Ljung (2001), when the stability and causality restrictions imposed there are omitted.

It is shown in Enqvist (2005), Enqvist and Ljung (2005) and Pintelon and Schoukens (2012) that Eq. (7) is minimized by:

$$G_{bla}(e^{j\omega T_s}) = \frac{S_{\tilde{y}\tilde{u}}(e^{j\omega T_s})}{S_{\tilde{u}\tilde{u}}(e^{j\omega T_s})}, \quad (10)$$

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

The output residuals $n_t(t)$ of the model are given by:

$$n_t(t) = \tilde{y}(t) - G_{bla}(q)\tilde{u}(t). \quad (11)$$

These residuals represent the total distortion that is present at the output of the system. The total distortion can be split into two contributions based on their nature as is depicted in Fig. 3. The nonlinear distortion $\tilde{y}_s(t)$ represents the unmodeled nonlinear contributions, while the noise distortion $n_y(t)$ is the additive noise that is assumed to be present at the system output. The nonlinear distortion and the noise distortion can be calculated separately as:

$$\tilde{y}_s(t) = \tilde{y}_0(t) - G_{bla}(q)\tilde{u}(t), \quad (12)$$

$$n_y(t) = \tilde{y}(t) - \tilde{y}_0(t), \quad (13)$$

where $\tilde{y}_0(t)$ is the unknown noiseless output. The nonlinear distortion $\tilde{y}_s(t)$ is uncorrelated with the input $\tilde{u}(t)$, however, $\tilde{y}_s(t)$ is not independent of the input $\tilde{u}(t)$ (Pintelon & Schoukens, 2012). The noise source $n_y(t)$, on the contrary, is assumed to be independent of the input $u(t)$.

The obtained model $G_{bla}(e^{j\omega T_s})$ does not only depend on the nonlinear system, but also on the class \mathbb{U} of input signals that is used. The class \mathbb{U} fixes both the probability density function (e.g. Gaussian inputs, uniform input distribution, etc.) and the power spectrum (or power spectral density) of the signals that are used to estimate it. Note that the BLA is equal to zero for zero-mean Gaussian signals applied to an even static nonlinear function (Bussgang, 1952).

The BLA can be estimated nonparametrically using the so-called robust and fast method (Pintelon & Schoukens, 2012; Schoukens, Pintelon, Dobrowiecki, & Rolain, 2005; Schoukens, Pintelon, & Rolain, 2012a). Afterwards, a parametric model can be estimated starting from the nonparametric one, by using the nonparametric disturbance model as a frequency weighting (Pintelon & Schoukens, 2012). One can also estimate a parametric model directly (Enqvist & Ljung, 2005; Ljung, 1999; Pintelon & Schoukens, 2012). This requires the user to make a careful selection of the

disturbance model that captures both the nonlinear and noise distortion contributions. Note that the estimation of a parametric model also results in a consistent estimate without the use of a disturbance model.

4.3. ϵ -approximation

The Riemann equivalence class of asymptotically normally distributed excitation signals \mathbb{S}_U is very useful to characterize a nonlinear system. However, the \mathbb{S}_U class has some drawbacks when it comes to structure detection and system analysis of some model structures. It is sometimes more convenient to consider infinitely small signals. For such excitation signals, the linearization of a cascade equals the cascade of the linearization.

Measuring infinitely or very small signals in a noisy setting is often not feasible. However, in quite some cases very high quality measurements can be obtained, e.g. during mechanical vibration tests or while measuring electronic circuits. In these cases very small signals can be applied to the system and can be measured with a sufficiently high signal-to-noise ratio. Deciding whether or not an input signal is small enough to act as an ϵ -excitation is mostly left to the user and the prior knowledge that is available on the system under test. Alternatively, the BLA framework can be used to make this decision (the estimated nonlinear distortion level should be below the estimated noise floor). Furthermore, the ϵ -excitation class allows formalizing some system properties, especially for the Hammerstein–Wiener and feedback structures considered in this paper.

The class of ϵ -excitations is defined in Schoukens et al. (2015a) as follows:

Definition 2 (Class of ϵ -excitations \mathbb{S}_ϵ). The signal $u_\epsilon(t)$, $t = 1, \dots, N$, belongs to the class of ϵ -excitations \mathbb{S}_ϵ , if it belongs to the class of Riemann equivalent excitation signals \mathbb{S}_U and

$$\sigma_u^2 = E\{\tilde{u}_\epsilon^2\} = \epsilon^2, \quad (14)$$

$$\tilde{u}_\epsilon(t) = u_\epsilon(t) - E\{u_\epsilon(t)\}. \quad (15)$$

ϵ -excitations are (asymptotically) normally distributed signals with a limited variance. In this paper, results are given for ϵ converging towards zero.

Definition 3 (ϵ -approximation). Consider the best linear approximation G_{bla} obtained for a random excitation $u = u_{DC} + u_\epsilon$, and $u_\epsilon \in \mathbb{S}_\epsilon$, where u_{DC} is a constant offset. Define:

$$\lim_{\epsilon \rightarrow 0} G_{bla}(e^{j\omega T_s})|_{u_\epsilon \in \mathbb{S}_\epsilon} = G_\epsilon(e^{j\omega T_s}). \quad (16)$$

The definition of G_ϵ can be extended to deterministic signals as discussed and formalized in Mäkilä and Partington (2003). The properties of the ϵ -approximation are discussed in more detail in Schoukens et al. (2015a).

5. Cost function

A least squares cost function is minimized to obtain the model parameters:

$$\hat{\theta} = \underset{\theta}{\operatorname{argmin}} \sum_{t=1}^N (y(t) - \hat{y}(t, \theta))^2, \quad (17)$$

where $\hat{y}(t, \theta)$ is the modeled output, depending on the parameters θ .

Note that the cost function used in Eq. (17) results in a maximum likelihood estimate if white Gaussian additive noise satisfying Assumption 1 is present at the output. A weighted version (using either a nonparametric noise model or a monic parametric

noise model) of the cost function needs to be used to obtain a (sample) maximum likelihood estimate in the case of additive colored Gaussian noise. The estimated sample variance of the noise can be used as a weighting function in a frequency domain cost function. A maximum likelihood estimator is asymptotically efficient (Cramér, 1946; Ljung, 1999; Pintelon & Schoukens, 2012). This means that the maximum likelihood estimator achieves the lowest asymptotic mean squared error possible with a consistent estimator, under the assumption that the system belongs to the model class, and using the noise framework specified in Assumption 1.

Unfortunately, this cost function is, in most of the cases considered in this paper, non-convex with respect to the parameters θ . A Levenberg–Marquardt algorithm (Fletcher, 1991; Levenberg, 1944; Marquardt, 1963; Pintelon & Schoukens, 2012) is used to minimize the cost function in a numerically stable way. This algorithm converges to a local minimum of the cost function. Hence, good initial values of the parameters are very important to ensure the convergence of the estimates to the global minimum. Such initial estimates can be obtained using the algorithms that are described in the following sections. Many of these initialization algorithms result in a consistent estimate, but they are not efficient.

6. Initialization methods: an overview

Besides the linear approximations-based methods that are described in the following sections, many other initialization methods for block-oriented nonlinear systems are available in the literature. Often two or more estimation methods are combined to identify the system under test. The most important approaches are highlighted here, a more detailed overview can be found in Giri and Bai (2010).

6.1. Overparametrization

The overparametrization approach is one of the most known methods for the identification of the Hammerstein and Hammerstein–Wiener type of structures (Bai, 1998; Chang & Luus, 1971; Hsia, 1976; Schoukens, Bai, & Rolain, 2012b).

In a first step, the number of parameters is artificially increased to obtain a problem that is linear in the parameters. In the next step, the estimated overparametrized parameter set is reduced, e.g. by using a singular value decomposition to retrieve the underlying low-rank structure of the overparametrized parameter set.

One of the main downsides of the overparametrization approach is the large number of parameters that needs to be estimated in the first step, resulting in a higher variance on the estimated parameters. Nuclear-norm regularization and kernel-based methods have been proposed to improve the classical approaches (Falck, Suykens, Schoukens, & De Moor, 2010; Risuleo, Bottegal, & Hjalmarsson, 2015), requiring less user interaction and reducing the variance on the estimated parameters.

6.2. Alternating least squares methods

Alternating least squares (ALS) is a powerful tool to solve problems that are multilinear in the parameters. ALS was one of the first successful identification methods for the identification of Hammerstein systems (Narendra & Gallman, 1966). Indeed, the Hammerstein identification problem can be written as a problem that is bilinear in the parameters. The convergence of this approach is analyzed in detail by Bai and Li (2004, 2010) and Stoica (1981). Similar techniques have been applied to Hammerstein–Wiener structures (Schoukens et al., 2012b; Zhu, 2002).

Alternatively, one can also use separable least squares approaches to estimate the parameters of block-oriented nonlinear models. This technique has been applied by Aljamaan, Bu bshait, and Westwick, 2011, Bai (2002) and Westwick and Kearney, 2001 for the identification of Hammerstein and Wiener models.

6.3. Inverse estimation

Some identification problems that seem to be hard to solve at first glance turn out to be much easier to solve when (part of) the system is inverted (if the inverse exists). Such an estimate can serve as a good initialization of the actual system identification problem at hand. Identifying a Wiener, or a Hammerstein–Wiener system is a well-known example of this approach.

The problem of identifying a Wiener system is nonlinear in the parameters in its most common formulation. However, the problem becomes linear in the parameters when the inverse of the static nonlinearity is identified instead of the static nonlinearity itself (Hagenblad & Ljung, 1998; Pajunen, 1992). A similar reasoning holds for the Hammerstein–Wiener case (Schoukens et al., 2012b; Zhu, 2002). Note that this approach results in biased estimates when there is noise present at the output of the system. Nevertheless, the estimates obtained by such an approach can still serve as initial values for a nonlinear optimization.

The inverse estimation approach can also be of interest for nonlinear feedback systems, such as the simple feedback and the Wiener–Hammerstein feedback structure. Indeed, by considering the inverse of the system structure the nonlinear feedback structure can be transformed in a parallel branch feedforward structure (see for example Section 4.3.4 in Markusson (2001)).

6.4. Other approaches

Besides the approaches discussed above many other block-oriented identification and initialization methods are described in the literature. Some of them are discussed shortly below, while others are included in the following sections when each specific block-oriented nonlinear structure is discussed in more detail.

Subspace methods have been successfully applied for the identification of MIMO (multiple input multiple output) Hammerstein, Wiener, Wiener–Hammerstein and Hammerstein–Wiener structures (Goethals, Pelckmans, Hoegaerts, Suykens, & De Moor, 2005; Katayama & Ase, 2016; Verhaegen & Westwick, 1996; Westwick & Verhaegen, 1996).

Frequency-domain approaches can be used to exploit the nonlinear behavior of the block-oriented nonlinear system, which is not visible in the time-domain, e.g. Giri, Radouane, Brouiri, & Chaoui (2014).

The structured nature of block-oriented nonlinear systems can be exploited by carefully designing the input signal used during the estimation, e.g. step signals, sine wave signals, sinesweeps or phase-coupled multisines (Castro-Garcia, Agudelo, Tiels, & Suykens, 2016; Giri et al., 2014; Rébillat, Ege, Gallo, & Antoni, 2016; Tiels, Schoukens, & Schoukens, 2015). Most of the methods described in the following sections make use of Gaussian input signals (Definition 1) to exploit Bussgang's Theorem (see Section 3).

The problem of identifying a block-oriented nonlinear system can sometimes be written as a convex optimization problem, or the problem can be relaxed such that it becomes convex. An example of such an approach is described in Cai and Bai (2011) and Han and de Callafon (2012).

7. Single branch models

The BLA of a single branch block-oriented nonlinear system (Hammerstein, Wiener, Wiener–Hammerstein) provides valuable information about the system. It is shown in the following sections that the BLA of these systems is equal to a scaled version of the linear dynamics that are present in the system. This allows the user to perform an important part of the model selection problem, i.e. the selection of the number of poles and zeros in the system dynamics, in the LTI framework. Although the case of the Hammerstein–Wiener system structure is a bit more involved, the BLA approach provides a good initial estimate for Hammerstein–Wiener identification.

7.1. Wiener, Hammerstein & Wiener–Hammerstein structure

Hammerstein structure

The Hammerstein structure is, together with the Wiener structure, one of the most simple block-oriented nonlinear model structures. A Hammerstein structure (Hammerstein, 1930) consists of a static nonlinear block $f(u)$ followed by an LTI block $S(q)$ (see Fig. 1). The Hammerstein structure is used to model nonlinear systems for which the static nonlinearity is present at the input of the system, such as a nonlinear actuator followed by a linear process, or some chemical processes and physiological systems (Giri & Bai, 2010).

Many different Hammerstein identification algorithms exist in the literature and are listed in Giri and Bai (2010). The first Hammerstein identification algorithm was introduced in Narendra and Gallman (1966), and further improved in Bai and Li (2004, 2010). Several other Hammerstein identification algorithms were developed in the last decades. A non-exhaustive list is given below where the methods are classified depending on their properties: kernel-based and mixed parametric–nonparametric identification algorithms (Hasiewicz et al., 2012; Mzyk, 2014; Risuleo et al., 2015), parametric approaches (Chang & Luus, 1971; Crama, Schoukens, & Pintelon, 2004; Schoukens, Widanage, Godfrey, & Pintelon, 2007), overparametrization approaches (Bai, 1998; Falck et al., 2010; Risuleo et al., 2015), blind identification algorithms (Bai, 2002; Vanbeylen, Pintelon, & Schoukens, 2008), a set-membership approach (Sznaier, 2009), and (continuous-time) instrumental variables approach (Laurain, Gilson, & Garnier, 2012). The feedback setting is considered in Giri, Ahmed-Ali, Brouni, Chaoui, and Radouane (2015) and the errors-in-variables setting in Mu, Han-Fu, Yi Wang, and Yin (2015). The MIMO Hammerstein case is considered in Goethals, Pelckmans, Suykens, and De Moor (2005) and Verhaegen and Westwick (1996). The works of Giri, Rochdi, Brouni and Chaoui (2011) Wang, Zhang, Mao and Zhou (2012) and Yong, Tan and Cham (2015) extend the Hammerstein structure to contain dynamic nonlinearities (e.g. backlash or hysteresis nonlinearities), and Guarín and Kearney (2015) and Sobhani Tehrani, Golkar, Guarín, Jalaaliddini and Kearney (2015) consider a time-varying Hammerstein structure.

Wiener structure

A Wiener structure (Wiener, 1958) consists of an LTI block $G(q)$ followed by a static nonlinear block $g(x)$ (see Fig. 1). A Wiener structure is often used to model systems where the nonlinear behavior is present at the output of the system. Some examples of such structures are systems with sensor nonlinearities, overflow valves and some physiological systems (Giri & Bai, 2010).

Giri and Bai (2010) list several Wiener system identification algorithms. A wide variety of Wiener identification algorithms have been developed in the last decades. Some are grouped here based on their properties. A non-exhaustive list of methods includes: nonparametric (or semi-parametric) identification algorithms (Hasiewicz et al., 2012; Lindsten et al., 2013; Mzyk, 2014), parametric approaches (Billings & Fakhouri, 1977; Crama & Schoukens, 2001; Hunter & Korenberg, 1986; Westwick & Kearney, 2003; Wigren, 1993), a minimal Lipschitz approach (Pelckmans, 2011), (orthogonal) basis function expansion approaches (Aljamaan, Westwick, & Foley, 2014; Lacy & Bernstein, 2003), blind identification algorithms (Vanbeylen, Pintelon, & Schoukens, 2009), a recursive approach (Greblicki, 2001), set-membership approaches (Cerone, Razza, & Regruto, 2015; Schnaier, 2009), and approaches for systems that contain dynamic nonlinearities (Giri et al., 2014; Yong et al., 2015). Janczak (2007) and Westwick and Verhaegen (1996) consider the MIMO Wiener system case. Most of the approaches consider that the noise source is present at the output of the system only, however some methods allow for process noise to be present in between the linear dynamics and the static nonlinearity (Hagenblad et al., 2008; Lindsten et al., 2013; Wahlberg et al., 2014; Wahlberg, Welsh, & Ljung, 2015).

Wiener–Hammerstein structure

A Wiener–Hammerstein model has a structure that is a bit more involved than the Wiener or the Hammerstein model structure. The static nonlinearity $f(x)$ is now sandwiched between two LTI blocks $G(q)$ and $S(q)$ (see Fig. 1). The presence of the two LTI blocks results in a problem that is harder to identify: the main challenge in identifying a Wiener–Hammerstein system lies in the separation of the dynamics over the front and the back LTI block.

The approaches described in Vandersteen, Rolain, and Schoukens (1997) and Tiels et al. (2015) estimate the two LTI blocks in a nonparametric way using carefully designed input signals. Weiss, Evans, and Rees (1998), Tan and Godfrey (2002), Kibangu and Favier (2006) and Ben Ahmed, Favier, and Derbel (2013) present Volterra (and Tensor decomposition) based approaches. Vörös (2007) discusses an iterative approach. Mu and Chen (2014) present a recursive EIV method, and Dewhurst, Simpson, Angarita, Allen, and Newland (2010) and Naitali and Giri (2016) present evolutionary-based approaches. Boutayeb and Darouach (1995) discuss the MISO (multiple input single output) Wiener–Hammerstein case, Katayama & Ase, (2016) the MIMO case. Many approaches use the BLA, or a similar correlation analysis, as a starting point for the algorithm e.g. Ase and Katayama (2015) Billings (1980) Billings and Fakhouri, (1978a) Crama and Schoukens (2005) Giordano and Sjöberg (2015) Katayama and Ase (2016) Korenberg and Hunter (1986) Lauwers (2011) Schoukens et al. (2014a) Sjöberg and Schoukens (2012) Tan, Wong and Godfrey (2012) Tiels, Schoukens and Schoukens (2014) Tiels et al. (2015) Vanbeylen (2014) Westwick and Kearney (2003) Westwick and Schoukens (2012).

Best linear approximation

The BLA of a Wiener–Hammerstein system is a scaled version of the LTI dynamics that are present in $G(q)S(q)$ when input signals belonging to \mathbb{S}_U are used.

Theorem 2. The BLA $G_{bla}(e^{j\omega T_s})$ of a Wiener–Hammerstein system excited by an input signal belonging to \mathbb{S}_U (Definition 1) is asymptotically given by (for the number of data growing to infinity):

$$G_{bla}(e^{j\omega T_s}) = \alpha G(e^{j\omega T_s})S(e^{j\omega T_s}), \quad \alpha \in \mathbb{R}. \quad (18)$$

Proof. See Enqvist (2010) and Schoukens (2015) \square

For the Hammerstein and the Wiener case, the BLA simplifies to:

$$G_{bla,H}(e^{j\omega T_s}) = \alpha S(e^{j\omega T_s}), \quad \alpha \in \mathbb{R}, \quad (19)$$

$$G_{bla,W}(e^{j\omega T_s}) = \alpha G(e^{j\omega T_s}), \quad \alpha \in \mathbb{R}. \quad (20)$$

Note that the gain α not only depends on the system characteristics, but also on the variance, the coloring, and the offset (DC value) of the input signal. In the case of the Hammerstein structure, the BLA does not depend on the coloring of the input signal since the input signal is applied directly to the nonlinearity of the system without prior filtering.

Hammerstein identification using the BLA

Many different approaches use the BLA, or similar correlation methods, to identify a Hammerstein system (Billings & Fakhouri, 1978a; Crama et al., 2004; Hunter & Korenberg, 1986; Westwick & Kearney, 2003). The identification of a Hammerstein system using the BLA is quite straightforward. The BLA approach decouples the identification of the LTI block and the static nonlinear block.

Firstly, the LTI dynamics are estimated using the BLA, this results in a consistent estimate of the LTI block $\hat{S}(q)$, up to an unknown gain exchange α (see Eq. (19)). Secondly, the static nonlinearity is estimated using a parametrization that is linear

in the parameters. This results in a problem that is linear in the parameters:

$$y(t) = S(q) [f(u(t))] \quad (21)$$

$$= \sum_{i=1}^{n_f} \gamma_i S(q) [f_i(u(t))], \quad (22)$$

where n_f denotes the number of basis functions $f_i(\cdot)$ that is used to describe the static nonlinearity. Note that Eq. (22) is linear in the parameters γ_i when $S(q)$ is replaced by its estimate $\hat{S}(q)$. Hence, the parameters γ_i can easily be obtained by a linear least squares estimation. This results in a consistent, but not efficient, estimator of the model parameters (Schoukens, 2015). An efficient estimate can be obtained with the nonlinear optimization described in Section 5.

Wiener identification using the BLA

The identification of Wiener systems using the BLA is very similar to the Hammerstein case. Billings and Fakhouri (1978a), Crama and Schoukens (2001), Hunter and Korenberg (1986) and Westwick and Kearney (2003) present BLA- or correlation-based identification methods for Wiener systems.

A two-step identification approach results in a consistent estimate of the model parameters. First, the LTI block is estimated using the BLA. Second, the static nonlinear block is identified starting from the estimated intermediate signal $\hat{x}(t) = \hat{G}(q)[u(t)]$ and the system output. One can use his/her favorite model structure (polynomial, neural network, ...) to describe the static nonlinearity. This results in a consistent, but not efficient, estimator of the model parameters (Schoukens, 2015). An efficient estimate can be obtained with the nonlinear optimization described in Section 5.

Wiener–Hammerstein identification using the BLA

There are many BLA-based Wiener–Hammerstein identification methods. In this section, we focus on the method that is presented in Sjöberg and Schoukens (2012). The challenge of separating the front and the back dynamics is reformulated as a pole and zero allocation problem: which poles and zeros belong to the front dynamic block and which poles and zeros belong to the back dynamic block?

The BLA is estimated and parametrized in a first step, and the poles and zeros of the BLA are computed. These poles and zeros are, due to Theorem 2 and under the assumption that the product $G(q)S(q)$ does not contain any pole-zero cancellations, the poles and zeros of the LTI blocks of the Wiener–Hammerstein system:

$$G_{bla} = \alpha G(q)S(q) = \alpha \frac{\prod_{i=1}^{n_b+n_d} (z_i - q^{-1})}{\prod_{i=1}^{n_a+n_c} (p_i - q^{-1})}, \quad (23)$$

where n_a and n_b are the number of poles and zeros in $G(q)$ respectively, while n_c and n_d are the number of poles and zeros in $S(q)$. The poles p_i and the zeros z_i of the parametrized BLA are the combined poles and zeros of the LTI blocks $G(q)$ and $S(q)$ of the Wiener–Hammerstein system.

These poles and zeros of the BLA need to be assigned to either $G(q)$ or $S(q)$:

$$G_{bla} = \alpha \hat{G}(q) \hat{S}(q), \quad (24)$$

$$\hat{G}(q) = \frac{\prod_{i=1}^{n_b+n_d} (z_i - q^{-1})^{\beta_{z_i}}}{\prod_{i=1}^{n_a+n_c} (p_i - q^{-1})^{\beta_{p_i}}}, \quad (25)$$

$$\hat{S}(q) = \frac{\prod_{i=1}^{n_b+n_d} (z_i - q^{-1})^{1-\beta_{z_i}}}{\prod_{i=1}^{n_a+n_c} (p_i - q^{-1})^{1-\beta_{p_i}}}, \quad (26)$$

β_{p_i} and β_{z_i} are binary parameters that assign the pole or zero to either $\hat{G}(q)$ ($\beta_{p_i} = 1$), or to $\hat{S}(q)$ ($\beta_{p_i} = 0$). The parameters

β_{p_i} and β_{z_i} are grouped in the binary parameter vector $\beta \in \{0, 1\}^{(n_a+n_b+n_c+n_d) \times 1}$.

There is a finite number of different realizations of the parameter vectors β . All possible realizations of the parameter vector are evaluated. To evaluate a realization of β , a static nonlinearity is estimated for every realization k . This problem is linear in the parameters if the static nonlinearity $\hat{f}_k(x(t))$ is described by a linear combination of basis functions. It is solved using linear least squares regression, resulting in the estimate \hat{f}_k . Finally, all the estimated models are ranked based on their mean squares error mse_k :

$$mse_k = \frac{1}{N} \sum_{t=1}^N (y(t) - \hat{y}_k(t))^2, \quad (27)$$

$$\hat{y}_k(t) = \hat{S}_k(q) [\hat{f}_k(\hat{G}_k(q)[u(t)])], \quad (28)$$

where N is the number of data points in the measured input–output record. The model with the lowest error is selected.

It is important to note that each complex conjugate pole or zero pair is assigned either to $S(q)$ or $G(q)$. The single elements of the pair are never assigned separately. This is required as the LTI blocks consist of transfer functions with real coefficients only. This also introduces a possible disadvantage of the approach: it can happen that two real poles (or zeros) are combined into a complex conjugate pole (or zero) pair during the parametrization of the BLA since the solution handles real poles and zeros as single elements. This pair cannot be correctly assigned if these poles (or zeros) originate from the two different LTI subsystems.

The total number of realizations of β depends on the number of poles and the number of zeros, n_p and n_z respectively. The total number is at least $2^{\frac{n_p+n_z}{2}}$ and at most $2^{n_p+n_z}$. The minimum number of realizations is obtained when all poles and zeros are part of a complex conjugate pair, the maximum number of realizations is obtained when all poles and zeros are real. This number increases very rapidly with the model order, which makes the brute-force scan approach in Sjöberg and Schoukens (2012) computationally expensive, since it equals the total number of least squares regressions that need to be performed to scan all possible pole-zero allocations.

A discrete optimization over the binary parameter vector β can be used to speed up the algorithm (Schoukens, Vandersteen, Rolain, & Ferranti, 2014b). The evaluation of all possible pole and zero allocations can be avoided by using a fractional approach (Gior-dano & Sjöberg, 2015; Vanbeylen, 2014), or by using a higher-order correlation, e.g. the so-called QBLA, to separate the dynamics over the front and the back LTI block (Billings & Fakhouri, 1978a; Schoukens et al., 2008, 2014a; Westwick & Schoukens, 2012). An equivalent subspace version of the method presented in Sjöberg and Schoukens (2012) is developed in Ase and Katayama (2015), and extended to the MIMO case in Katayama and Ase (2016).

Although Sjöberg and Schoukens (2012) show that this approach results in a consistent estimate of the system parameters, a nonlinear optimization of all the parameters simultaneously as described in Section 5, following the linear initialization step, can be used to reduce the variance of the estimates.

7.2. Hammerstein–Wiener structure

Another single branch model structure is the Hammerstein–Wiener model (static nonlinear block $f(u)$ – LTI block $G(q)$ – static nonlinear block $g(r)$, see Fig. 1). These models are used in a wide range of applications such as chemical processes (Giri & Bai, 2010), ionospheric dynamics (Palanthandalam-Madapusi, Ridley, & Bernstein, 2005), submarine detection (Abrahamsson, Kay, &

Stoica, 2007), and RF (radio frequency) power amplifier modeling (Taringou, Hammi, Srinivasan, Malhame, & Ghannouchi, 2010).

Different identification approaches are proposed to identify a Hammerstein–Wiener model or models that are very similar to Hammerstein–Wiener block structures. Amongst those are: iterative approaches (Vörös, 2004; Zhu, 2002), overparametrization methods (Bai, 1998; Schoukens, 2015; Schoukens et al., 2012b), frequency domain methods (Brouri, Amdouri, Chaoui, & Giri, 2014a; Crama & Schoukens, 2004), subspace methods (Goethals et al., 2005), stochastic algorithms (Wang & Ding, 2008), blind approaches (Bai, 2002), continuous-time instrumental variable approaches (Ni, Gilson, & Garnier, 2013), and maximum likelihood-based methods (Wills et al., 2013). The identification of Hammerstein–Wiener systems with a backlash input nonlinearity is discussed in Brouri, Amdouri, Chaoui, and Giri (2014b).

Best linear approximation

Contrary to Hammerstein and Wiener systems, the BLA of a Hammerstein–Wiener system cannot be simplified to a simple function of the system dynamics, even if input signals belonging to \mathbb{S}_U are used. The intermediate signal $r(t)$ (see Fig. 1) is not Gaussian due to the presence of the first static nonlinearity. Hence, Bussgang's theorem cannot be applied. However, Wong, Schoukens and Godfrey, (2012) shows that the BLA will not differ too much from a scaled version of the LTI block $G(q)$ in many practical cases (i.e. if the memory length of $G(q)$ is large w.r.t. the correlation length of the intermediate signal $x(t)$).

The theoretical framework of the ϵ -approximation can also be used to replace the BLA framework. The ϵ -approximation of a Hammerstein–Wiener system is a scaled version of the LTI dynamics $G(q)$ present in the system (Schoukens et al., 2015a).

Identification using the BLA

The BLA (or ϵ -approximation) is used as a starting point for the identification of a Hammerstein–Wiener system in Crama and Schoukens (2004). First, the BLA is estimated resulting in an estimate $\hat{G}(q)$ of the LTI block. Next, the front static nonlinearity and the inverse of the back static nonlinearity are identified by minimizing:

$$V_N = \sum_{t=1}^N \left(\sum_{i=1}^{n_f} \gamma_i \hat{G}(q) [f_i(u(t))] - \sum_{j=1}^{n_g} \delta_j g_j(y(t)) \right)^2 \quad (29)$$

where it is assumed that the static nonlinearity $f(u)$ and the inverse of the static nonlinearity $g(r)$ can be represented by a linear combination of n_f and n_g nonlinear basis functions respectively. This problem is linear in the parameters, it can be solved using a total least squares approach (Van Huffel & Lemmerling, 2002). In a next step the estimate of the inverse of the back static nonlinear block can be replaced by an estimate of the forward static nonlinearity $g(r)$. Note that the estimates obtained using the total least squares approach are not consistent due to the noise framework specified in Assumption 1. These estimates serve as an initialization for the nonlinear optimization described in Section 5, resulting in a consistent estimate of the model parameters (if all assumptions are satisfied).

8. Parallel branch model structures

Parallel block-oriented structures (parallel Hammerstein, parallel Wiener and parallel Wiener–Hammerstein, see Fig. 1) are a logical extension of the single branch block-oriented structures. The use of a parallel branch structure can be motivated by the presence of multiple signal paths inside the system under test.

Another possible reason for the use of parallel branch models is their increased modeling capacity. It is well-known that the parallel Wiener and the parallel Wiener–Hammerstein structure are universal approximators of fading memory systems (see Section 10 for more information).

In the case of parallel block-oriented structures the BLA will not only provide information about the LTI blocks that are present in the system, but also about the number of parallel branches in the system. By combining the BLAs obtained at different setpoints of the systems, the selection of the number of parallel branches and the selection of the number of poles and zeros that are present in the system can be performed (Schoukens, Pintelon, & Rolain, 2011; Schoukens & Rolain, 2012b; Schoukens et al., 2015b). This allows the user to perform the selection of the number of poles and zeros in the system dynamics in the LTI framework, and to select the number of parallel branches in the system at an early stage of the identification procedure.

The BLA can also be used to construct a set of orthonormal basis functions to represent the system dynamics (Tiels & Schoukens, 2014). These basis functions are based on a number of user-selected pole locations. A natural choice is to use the poles of the BLA as a starting point to construct the orthonormal basis functions.

8.1. Parallel Hammerstein and parallel Wiener Structures

Parallel Hammerstein and parallel Wiener models are generalized forms of the Hammerstein and Wiener structure. They consist of a parallel connection of more than one Hammerstein (or Wiener) system (see Fig. 1). Parallel Hammerstein models are sometimes called Uryson models, memory polynomial, or generalized Hammerstein models in the literature, see for example Billings (1980), Doyle, Pearson and Ogunnaike (2002), Gallman (1975) and Ghannouchi and Hammi (2009),

Parallel Hammerstein: A popular approach that is used to identify a parallel Hammerstein system is to feed the input signal to a parallel connection of some nonlinear basis functions that are each followed by a finite impulse response filter (FIR) to realize a branch. The corresponding LTI multiple-input-single-output identification problem is then solved using one's favorite FIR identification method in Gadringer, Silveira, and Magerl (2007). In earlier work (Gallman, 1975), Hermite polynomials were used as orthogonal basis functions to model the static nonlinearity for Gaussian input signals. These approaches result in a linear least squares identification problem, which is easy to solve. However, they suffer from two drawbacks: (i) the user gets no physical insight into the number of parallel paths in the device under test (DUT), as the number of parallel branches in the model is set by the degree of the nonlinearity, (ii) for systems with long memory effects, a large number of parameters is needed due to the FIR-nature of the dynamic model. The model is therefore not parsimonious. Rébillat et al. (2016) present a fully nonparametric approach, which requires almost no user interaction, based on an exponential sinesweep excitation.

Parallel Wiener: Previously proposed parallel Wiener estimation methods suffer from some disadvantages. Some methods rely on an estimate of the Volterra kernel of the system under test (Kibangu & Favier, 2009). This requires a very large amount of data for the identification. Other methods are limited to the use of finite impulse response models for the linear subsystems (Korenberg, 1991; Lyzell & Enqvist, 2012a; Westwick & Kearney, 1997). These approaches typically result in an unwanted high number of parallel branches.

Best linear approximation

Similar to the single branch case, the BLA of a parallel Hammerstein and parallel Wiener system is a simple function of the LTI dynamics that are present in the system.

Theorem 3. The BLA $G_{bla}(e^{j\omega T_s})$ of a parallel Hammerstein system excited by an input signal belonging to \mathbb{S}_U (Definition 1) is asymptotically given by (for the number of data growing to infinity):

$$G_{bla}(e^{j\omega T_s}) = \sum_{k=1}^{n_{br}} \alpha_k S^{[k]}(e^{j\omega T_s}), \quad \alpha_k \in \mathbb{R}, \quad (30)$$

where n_{br} is the number of parallel branches in the system.

Proof. The BLA of a Hammerstein system is given by Eq. (19). This holds for each separate branch in the parallel Hammerstein system. The output y is given by a summation of the outputs y_k of each parallel Hammerstein branch. Hence, the BLA of a parallel Hammerstein system is given by the sum of the BLAs of each parallel branch. \square

Theorem 4. The BLA $G_{bla}(e^{j\omega T_s})$ of a parallel Wiener system excited by an input signal belonging to \mathbb{S}_U is asymptotically given by (for the number of data growing to infinity):

$$G_{bla}(e^{j\omega T_s}) = \sum_{k=1}^{n_{br}} \alpha_k G^{[k]}(e^{j\omega T_s}), \quad \alpha_k \in \mathbb{R}, \quad (31)$$

under the assumption that the MISO static nonlinearity $g(\cdot)$ of the parallel Wiener system can be approximated arbitrarily well in least squares sense by a MISO polynomial.

Proof. See Schoukens (2015). \square

An important observation with respect to Eqs. (30) and (31) is that the input-dependent gain α_k only appears in the numerator. For example, the BLA in Eq. (30) is given by:

$$G_{bla}(q) = \sum_{k=1}^{n_{br}} \alpha_k S^{[k]}(q) = \sum_{k=1}^{n_{br}} \alpha_k \frac{B_s^{[k]}(q)}{A_s^{[k]}(q)}. \quad (32)$$

Introducing a common denominator results in the following BLA expression:

$$G_{bla}(q) = \frac{\sum_{k=1}^{n_{br}} \alpha_k B_s^{[k]}(q) \prod_{j=1, j \neq k}^{n_{br}} A_s^{[j]}(q)}{\prod_{k=1}^{n_{br}} A_s^{[k]}(q)}. \quad (33)$$

A similar result can be obtained for the parallel Wiener case.

This means that the poles of the identified BLA are also the poles of the LTI blocks that are present in the system. The zeros of the BLA of a parallel Hammerstein (or parallel Wiener) system may change when the variance or the offset (DC value) of the input signal changes. The gains α_k are also dependent on the coloring of the input signal in the case of a parallel Wiener system. Indeed, the denominator stays the same when the gains α_k change, but the numerator coefficients depend on the gains α_k . Schoukens (2015) and Schoukens et al. (2015a, 2015b) discuss this in further detail.

Parallel Wiener and parallel Hammerstein identification using the BLA–SVD approach

The main difficulty in identifying a parallel Wiener or parallel Hammerstein system is separating the dynamics over the different parallel branches. This problem can be made easier by allowing more parallel branches to be present in the estimated model (see e.g. (Billings, 1980; Doyle et al., 2002; Gallman, 1975; Korenberg, 1991; Lyzell & Enqvist, 2012a; Westwick & Kearney, 1997)). A singular value decomposition (SVD) based approach to identify a parallel Wiener or parallel Hammerstein system with a low

number of parallel branches is presented in Schoukens, Lyzell, and Enqvist (2013) and Schoukens et al. (2015b, 2011). This section explains the driving idea of this approach.

The approach proposed in Schoukens et al. (2013, 2015b) starts with an estimation of the BLA of the considered system for different operating conditions. The different operating conditions have been obtained using input signals with different power spectra. This includes the use of different magnitudes, different offsets, or different coloring of the power spectra.

Next, the measured BLAs are parametrized using a different LTI model for each operating condition. A common denominator model is used for all operating conditions simultaneously:

$$\hat{G}_{bla}^{[i_r]}(q, \hat{\theta}_{bla}) = \frac{\hat{d}_0^{[i_r]} + \hat{d}_1^{[i_r]} q^{-1} + \dots + \hat{d}_{n_d}^{[i_r]} q^{-n_d}}{\hat{c}_0 + \hat{c}_1 q^{-1} + \dots + \hat{c}_{n_c} q^{-n_c}}, \quad (34)$$

where superscript $[i_r]$ denotes the setpoint of the system. This is indeed possible, as Eq. (33) assures that the poles of the different measured BLAs are the same. A consistent estimate of the overall dynamics that are present in the nonlinear parallel Hammerstein or parallel Wiener system results. A matrix $\hat{\mathbf{D}}$ is constructed by stacking the estimated numerator coefficients of the BLAs at the R different operating conditions:

$$\hat{\mathbf{D}} = \begin{bmatrix} \hat{d}_0^{[1]} & \hat{d}_1^{[1]} & \dots & \hat{d}_{n_d}^{[1]} \\ \hat{d}_0^{[2]} & \hat{d}_1^{[2]} & \dots & \hat{d}_{n_d}^{[2]} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{d}_0^{[R]} & \hat{d}_1^{[R]} & \dots & \hat{d}_{n_d}^{[R]} \end{bmatrix}. \quad (35)$$

Starting from the parametrized BLAs, a decomposition of the overall dynamics at the different operating conditions is calculated. It uses the singular value decomposition (SVD) of the $\hat{\mathbf{D}}$ matrix:

$$\hat{\mathbf{D}} = \mathbf{U}_{bla} \boldsymbol{\Sigma}_{bla} \mathbf{V}_{bla}^T, \quad (36)$$

where the columns of \mathbf{V}_{bla} are the right singular vectors which act as an orthonormal basis for the row space of the $\hat{\mathbf{D}}$ -matrix, $\boldsymbol{\Sigma}_{bla}$ is a diagonal matrix containing the singular values, and the columns of \mathbf{U}_{bla} constitute an orthonormal basis for the column space.

The column vectors in \mathbf{V}_{bla} provide an estimate of the numerator coefficients for each branch k :

$$\hat{S}^{[k]}(q) = \frac{\hat{\delta}_0^{[k]} + \hat{\delta}_1^{[k]} q^{-1} + \dots + \hat{\delta}_{n_d}^{[k]} q^{-n_d}}{\hat{c}_0 + \hat{c}_1 q^{-1} + \dots + \hat{c}_{n_c} q^{-n_c}}, \quad (37)$$

where $\hat{\delta}_j^{[k]}$ is the element of the j th row and k th column of the matrix \mathbf{V}_{bla} .

The number of parallel branches in the system is obtained based on the estimated rank of the decomposed matrix (based on the singular values obtained in $\boldsymbol{\Sigma}_{bla}$). Note that the matrix $\hat{\mathbf{D}}$ needs to be whitened before the rank estimation can be performed (see Rolain, Schoukens, & Pintelon (1997) and Schoukens et al. (2015b) for more information).

Once an estimate of the LTI blocks of the parallel Hammerstein or parallel Wiener system has been obtained, the estimation of the static nonlinearities in the system boils down to a simple linear least squares problem as in the single branch Hammerstein or Wiener case.

Although Schoukens (2015) and Schoukens et al. (2015b) show that this approach results in a consistent estimate of the system parameters, a nonlinear optimization of all the parameters simultaneously as described in Section 5 can be used to reduce the variance of the estimates.

8.2. Wiener–Schetzen structure

A Wiener–Schetzen model (see Fig. 4) is a subset of parallel Wiener models where the dynamics are represented by

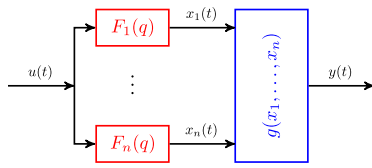


Fig. 4. A Wiener–Schetzen model ($F_1(q), \dots, F_n(q)$ are OBFs, $g(x_1, \dots, x_n)$ is a multivariate polynomial).

orthonormal basis functions (OBFs) (Heuberger, Van den Hof, & Wahlberg, 2005), and the nonlinearity is represented by a multivariate polynomial with the outputs of the OBFs as inputs. The parameters of this model are the polynomial coefficients. Nevertheless, the OBFs should be properly selected.

Wiener–Schetzen models appear in the literature under various names. In his original ideas, Wiener uses (continuous-time) Laguerre OBFs (Wiener, 1958) and the model is called a Laguerre system (see also (Boyd and Chua 1985)). Kibangou, Favier, and Hassani (2005b) use discrete-time Laguerre OBFs and use the name Laguerre-Volterra filters. When using generalized orthonormal basis functions (GOBFs), the names GOB-Volterra filters (Kibangou, Favier, & Hassani, 2005a) and Wiener/Volterra models (da Rosa, Campello, & Amaral, 2007) are used, amongst others.

Several methods have been proposed to identify Wiener–Schetzen models. Gómez and Baeyens (2004) propose a non-iterative method, where it is assumed that the nonlinearity is invertible and that this inverse can be described in terms of known basis functions. da Rosa et al. (2007) propose an optimal selection of one of the parameters in a truncated two-parameter Kautz OBF expansion. This optimal selection is done by minimizing an upper bound of the error made by approximating the system's Volterra kernels by the truncated OBF expansion. The Volterra kernels of the system are assumed to be known in this approach. A model structure very similar to a Wiener–Schetzen model is identified in Gómez and Baeyens (2012) and Tötterman and Toivonen (2009), where the nonlinearity is expanded in terms of support vector machines (SVMs).

Best linear approximation

Since a Wiener–Schetzen model is a parallel Wiener model with a MISO polynomial nonlinearity, its BLA is given by Theorem 4. In particular, the BLA of a Wiener–Schetzen model is a linear combination of the dynamics in the different branches. The poles of the BLA are thus also the poles of the LTI dynamics that are present in the system.

Identification using the BLA

OBFs are stable, discrete-time, proper, rational transfer functions $F_i(e^{j\omega})$ that are orthogonal with respect to the inner product

$$\langle F_1, F_2 \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} F_1(e^{j\omega}) F_2^*(e^{j\omega}) d\omega. \quad (38)$$

The outputs $x_1(t)$ and $x_2(t)$ of two OBFs F_1 and F_2 that share the same input are orthogonal if the input is white. This follows from the generalized expected value $\bar{E}\{x_1(t)x_2(t)\}$, which can be written as

$$\bar{E}\{x_1(t)x_2(t)\} = \frac{1}{2\pi} \int_{-\pi}^{\pi} F_1(e^{j\omega}) F_2^*(e^{j\omega}) |U(e^{j\omega})|^2 d\omega \quad (39)$$

using Parseval's theorem (Heuberger et al., 2005), and which reduces to the inner product of two orthonormal transfer functions if the input is white ($|U(e^{j\omega})|^2$ is constant). Although in general no claims can be made about the orthogonality for non-white inputs, the considered OBFs are quite robust with respect to the coloring of the OBFs (Heuberger et al., 2005, chap. 6).

Popular choices for the OBFs are Laguerre, Kautz, and Takenaka–Malmquist basis functions (Heuberger et al., 2005). The considered OBFs can be constructed from a set of poles $\{\xi_i\}$, and in its most general form they are the Takenaka–Malmquist basis functions:

$$F_i(q) = \frac{\sqrt{1 - |\xi_i|^2}}{q - \xi_i} \prod_{k=1}^{i-1} \frac{1 - \xi_k^* q}{q - \xi_k}, \quad (40)$$

for $i = 1, 2, \dots$; $\xi_i \in \mathbb{C}$; and $|\xi_i| < 1$. Choosing all poles ξ_i in the origin results in an FIR model. Choosing all poles identical and real results in a Laguerre basis. A two-parameter Kautz basis is obtained when choosing a pole structure $\{\xi_1, \xi_2, \xi_1, \xi_2, \dots\}$. Note that this pole structure can include multiple repetitions of ξ_1, ξ_2 , where ξ_1 and ξ_2 are either real or complex conjugated. GOBFs are obtained when a finite set of n_ξ poles is periodically repeated, i.e.

$$\xi_{i+(k-1)n_\xi} = \xi_i, \quad i = 1, 2, \dots, n_\xi, \quad k = 1, 2, \dots \quad (41)$$

A proper user choice of the pole locations is important in order to accurately represent the linear dynamics with a limited number of OBFs. Choosing the poles equal to the true poles of the underlying linear dynamics of the (parallel) Wiener system is optimal in this sense. The poles of the BLA are thus excellent candidates to be used in constructing the OBFs. Possible mismatches between the true poles and the estimates from the BLA can be compensated by periodically repeating the set of pole estimates. Fast convergence rates (convergence in probability of the modeled output to the system output) can be obtained (Tiels & Schoukens, 2014).

Once the OBFs are constructed, the intermediate signals $x_1(t), \dots, x_n(t)$ can be computed. Retrieving the polynomial coefficients then boils down to a linear regression problem. By choosing OBFs and Hermite polynomials (Schetzen, 2006), which are orthogonal for Gaussian inputs, this regression problem is optimally conditioned for white Gaussian inputs. Note that this is only true for infinitely long data records; for a finite data record, approximate orthogonality is obtained.

Compared to the BLA-SVD approach in the previous section, only one experiment suffices in this case. The model is also linear in the parameters, but this comes at the cost of a large number of parameters, since the number of polynomial coefficients increases combinatorially with the number of OBFs and the degree of the nonlinearity. Dimension reduction techniques, such as presented in Lyzell, Andersen, and Enqvist (2012), Lyzell and Enqvist (2012a, 2012b) and Schoukens et al. (2013), can be used to reduce the number of basis functions, and hence, the number of parameters.

8.3. Parallel Wiener–Hammerstein structure

Parallel Hammerstein and parallel Wiener model structures are restricted to systems with dominant output dynamics (Hammerstein) or input dynamics (Wiener). A further increase in flexibility of the model structure is obtained here by considering the parallel Wiener–Hammerstein structure (see Fig. 1). Previously published methods (Baumgartner & Rugh, 1975; Billings & Fakhouri, 1979, 1982; Wysocki & Rugh, 1976) studied a subclass of the parallel Wiener–Hammerstein structure. Identification methods based on repeated sine measurements (Baumgartner & Rugh, 1975; Wysocki & Rugh, 1976), or white Gaussian inputs (Billings & Fakhouri, 1979) are available for this model structure. Palm (1978, 1979) shows that a wide class of Volterra systems can be approximated arbitrarily well using a parallel Wiener–Hammerstein model structure. Dreesen, Westwick, Schoukens, and Ishteva (2017) present an identification method based on the decomposition of Volterra kernels.

The parallel Wiener–Hammerstein system class that is used here is a more general system class than the S_M system class that is used in Baumgartner and Rugh (1975), Billings and Fakhouri

(1979) and Wysocki and Rugh (1976) in the following sense. The S_M model has M parallel branches, and the m th branch contains a monomial nonlinearity that is fixed and equal to $(\cdot)^m$. This restricts the model to have a polynomial nonlinearity only, and to contain only one branch for each degree of this polynomial nonlinearity. Thus a parallel Wiener–Hammerstein system containing two parallel branches, each with different LTI subsystems, and with different polynomial nonlinearities can, in general, not be modeled by a S_M model. The methods that Schoukens et al. (2015b) and Schoukens (2015) present also make some extra assumptions on the parallel Wiener–Hammerstein system. However, even when these assumptions are met, the considered system class still allows for a much more complex system behavior than the S_M class does.

Best linear approximation

Similar to the parallel Hammerstein and parallel Wiener case, the BLA of a parallel Wiener–Hammerstein system is a simple function of the LTI dynamics that are present in the parallel branches of the system.

Theorem 5. The BLA $G_{bla}(e^{j\omega T_s})$ of a parallel Wiener–Hammerstein system excited by an input signal belonging to \mathbb{S}_U (Definition 1) is asymptotically given by (for the number of data growing to infinity):

$$G_{bla}(e^{j\omega T_s}) = \sum_{k=1}^{n_{br}} \alpha_k G^{[k]}(e^{j\omega T_s}) S^{[k]}(e^{j\omega T_s}), \quad \alpha_k \in \mathbb{R}. \quad (42)$$

Proof. The BLA of a Wiener–Hammerstein system is given by Eq. (18). This holds for each separate branch in the parallel Wiener–Hammerstein system. The output y is given by a summation of the outputs y_k of each parallel Wiener–Hammerstein branch. Hence, the BLA of a parallel Wiener–Hammerstein system is given by the sum of the BLA of each parallel branch. \square

Identification using the BLA–SVD approach

The identification algorithm for the parallel Wiener–Hammerstein case is a combination of the identification algorithms presented in Sections 7.1 and 8.1. Schoukens et al. (2015b) and Schoukens (2015) describe this in detail.

Firstly, the BLA decomposition approach is applied as described in Section 8.1 for the parallel Hammerstein and parallel Wiener case to split the BLA dynamics over the different parallel branches.

Next, the poles and zeros of each branch are allocated to the front or the back simultaneously using the brute-force pole-zero allocation scan described in Section 7.1. The static nonlinearity of the model is estimated during this step.

Although Schoukens (2015) and Schoukens et al. (2015b) show that this approach results in a consistent estimate of the system parameters, a nonlinear optimization of all the parameters simultaneously as described in Section 5 can be used to reduce the variance of the estimates.

Two disadvantages of this approach are (i) the high computational load due to the brute-force pole-zero allocation scan, and (ii) the introduction of a MIMO static nonlinearity in the final model although the original system is intrinsically decoupled. The appearance of the MIMO nonlinearity is due to equivalence transforms and the SVD approach to split the dynamics (Schoukens, 2015; Schoukens et al., 2015b). The brute-force pole-zero allocation scan can be made faster by allocating the poles and zeros of each branch separately instead of doing the allocation for all branches simultaneously (Schoukens, 2015). The coupled, MIMO nonlinearity can be decoupled into a number of SISO (single input single output) static nonlinearities in a second step using tensor decomposition methods as Dreesen, Schoukens, Tiels, and Schoukens (2015), Tiels and Schoukens (2013) and Schoukens and Rolain (2012a) show.

9. Feedback model structures

The use of a feedback structure can be motivated by the presence of nonlinear feedback in the system under test, e.g. many mechanical setups show resonant behavior in which the resonance frequency depends on the variance of the excitation signal. As discussed later in this section, this is a typical nonlinear feedback phenomenon. Feedback mechanisms are also often encountered in biological systems. A nonlinear feedback can also be included in the model structure to increase its modeling capacity.

The BLA or ϵ -approximation can be used in many cases for the identification of feedback block-oriented nonlinear model structures. Both frameworks allow the user to perform the selection of the number of poles and zeros in the system dynamics in the LTI framework, in the three cases considered in this section. Note that all feedback systems considered in this section are BIBO stable for the selected input class due to the assumed class of systems in Section 4. During the identification procedure, no special effort is made to obtain stable models. Multiple shooting techniques can be used to pass through unstable regions in the parameter space (Van Mulders, Schoukens, Volckaert, & Diehl, 2010). Since all the considered feedback structures can be restructured so that they include the simple nonlinear feedback structure (see Figs. 1 and 2), stability of the final model can be checked with e.g. the Popov criterion (Haddad & Bernstein, 1993, 1994; Khalil, 1996). This criterion provides a sufficient (not necessary) condition for the overall model to be absolutely stable. A typical assumption is that the linear subblock is observable and controllable, and that the nonlinearity fulfills a sector condition.

It is commonly assumed for nonlinear feedback systems to have at least one sample delay in either the forward or the backwards path of the feedback loop to avoid the presence of nonlinear algebraic loops during simulation. It is shown in Relan and Schoukens (2016) under band-limited assumptions that the approximation error made by introducing such a delay in the model can be kept arbitrary low by selecting a sufficiently high sampling frequency.

9.1. Simple feedback structure

The simple feedback structure consists of the feedback connection of an LTI block and a static nonlinear block (see Fig. 1). One can think of two possible simple feedback structures: LTI block in the feedforward and the nonlinearity in feedback or vice-versa. It can be shown that both combinations are equivalent from an input–output point of view (see Section 2 and (Schoukens et al., 2008)). Paduart (2008) and Paduart, Horváth, and Schoukens (2004) propose a BLA-based identification algorithm for the simple feedback structure. Van Pelt and Bernstein (2001) consider a slightly more advanced nonlinear feedback structure.

Best linear approximation

Contrary to most of the previous cases, the BLA of a simple feedback system cannot be simplified to an easy function of the system dynamics, even if input signals belonging to \mathbb{S}_U are used. The input signal of the nonlinearity $y(t)$ is not Gaussian due to the presence of the nonlinear feedback loop. Hence, Bussgang's theorem cannot be applied. However, in many practical cases it is fair to assume that the nonlinearity can be linearized to a gain α and thus that the BLA will be approximately given by:

$$G_{bla}(e^{j\omega T_s}) \approx \frac{G(e^{j\omega T_s})}{1 + \alpha G(e^{j\omega T_s})}, \quad \alpha \in \mathbb{R}. \quad (43)$$

The quality of this approximation depends on how non-Gaussian the input to the nonlinearity is. The approximation is relevant if the input to the nonlinearity is e.g. the output of a linear block with a sufficiently large memory length. It then follows from the

central limit theorem that the input to the nonlinearity resembles a Gaussian signal. The approximation in (43) is relevant if the memory length of $G(q)$ is sufficiently large.

The ϵ -approximation, contrary to the BLA, results in a simple and exact analytical expression of the linear approximation as a function of the system dynamics (Schoukens et al., 2015a). The expression for $G_\epsilon(e^{j\omega T_s})$ is as in (43) with α replaced by γ .

The gains α and γ do not only depend on the system characteristics, but they can also depend on the variance, power spectrum, and the offset (DC value) of the input signal.

Note that in the feedback case (also in the Wiener–Hammerstein feedback case in Sections 9.2), only the poles of the ϵ -approximation shift with a changing gain γ :

$$G_\epsilon(q) = \frac{B(q)}{A(q) + \gamma B(q)}, \quad (44)$$

where

$$G(q) = \frac{B(q)}{A(q)}. \quad (45)$$

Schoukens et al. (2015a) discuss this in further detail.

Identification using the BLA

The identification algorithm that is presented here follows the lines of the algorithm presented in Paduart (2008) and Paduart et al. (2004).

Firstly, it is important to note that an arbitrary gain β can be shifted from the static nonlinear block to the linear block in the model. Indeed we have that:

$$\begin{aligned} y(t) &= G(q)[u(t) - f(y(t))] \\ &= G(q)[u(t) - f(y(t)) + \beta y(t) - \beta y(t)] \\ (1 - \beta G(q))[y(t)] &= G(q)[u(t) - \tilde{f}(y(t))] \\ y(t) &= \tilde{G}(q)[u(t) - \tilde{f}(y(t))] \end{aligned} \quad (46)$$

where $\tilde{f}(y(t)) = f(y(t)) - \beta y(t)$ and $\tilde{G}(q) = \frac{G(q)}{1 + \beta G(q)}$.

This indicates that the BLA results in a good initial estimate of the linear subsystem up to the equivalence transforms present in the model structure:

$$\hat{G}(q) = G_{bla}(q) \approx \frac{G(q)}{1 + \alpha G(q)}. \quad (47)$$

Next, the static nonlinearity is estimated. The static nonlinearity is assumed to be represented as a linear combination of nonlinear basis functions. It is estimated by cutting the feedback loop:

$$\begin{aligned} \hat{G}(q)[u(t)] - y(t) &= \hat{G}(q) \left[\sum_{i=1}^{n_f} \gamma_i f_i(y(t)) \right], \\ &= \sum_{i=1}^{n_f} \gamma_i \hat{G}(q) [f_i(y(t))], \end{aligned} \quad (48)$$

where n_f denotes the number of basis functions that is used to describe the static nonlinearity. This equation is linear in the parameters γ_i . Hence, the parameters γ_i can easily be obtained by a linear least squares estimation.

The noisy output is used as a regressor during the initial identification of the static nonlinearity in Eq. (48). This results in a very simple initialization scheme which works well for high signal to noise ratios.

Finally, a nonlinear optimization (see Section 5) of all the parameters of the simple feedback structure together is applied such that a consistent estimate is obtained. The noise is assumed to be present at the output of the system (Assumption 1), i.e. outside the feedback loop. The feedback mechanism acts on the noise-free signal. The construction of the Jacobian of the simple feedback

structure is more involved and time consuming (compared to the single and parallel branch case) due to its recursive dependency on the output at previous time instances. This holds as well for the Wiener–Hammerstein feedback case and the LFR case in Sections 9.2 and 9.3.

9.2. Wiener–Hammerstein feedback structure

The Wiener–Hammerstein feedback structure is a generalized version of the simple feedback structure. It consists of a Wiener–Hammerstein system in the feedforward path combined with an LTI system in the feedback path (see Fig. 1). Both BLA-based approaches (Schoukens et al., 2008) and a convex relaxation-based approach (Sou, Megretski, & Daniel, 2008) have been developed for this system structure. It can be shown that this structure is equivalent to an LTI system in the forward path and a Wiener–Hammerstein system in the feedback path (Schoukens et al., 2008).

Best linear approximation

As for the simple feedback structure case, the BLA of a Wiener–Hammerstein feedback structure cannot be simplified to an easy function of the system dynamics if input signals belonging to \mathbb{S}_U are used. However, in many practical cases it is fair to assume that the BLA will be approximately of the form:

$$\begin{aligned} G_{bla}(e^{j\omega T_s}) &\approx \frac{\alpha G^{[1]}(e^{j\omega T_s}) G^{[2]}(e^{j\omega T_s})}{1 + \alpha G^{[1]}(e^{j\omega T_s}) G^{[2]}(e^{j\omega T_s}) G^{[3]}(e^{j\omega T_s})} \\ \alpha &\in \mathbb{R}. \end{aligned} \quad (49)$$

The approximation in (49) is relevant if the memory length of $G^{[2]}(q)G^{[3]}(q)G^{[1]}(q)$ is sufficiently large.

The ϵ -approximation, contrary to the BLA, results in a simple and exact analytical expression of the linear approximation in function of the system dynamics (Schoukens et al., 2015a). The expression for $G_\epsilon(e^{j\omega T_s})$ is as in (49) with α replaced by γ .

The gains α and γ depend not only on the system characteristics, but can also depend on the variance, power spectrum, and the offset (DC value) of the input signal.

Identification using the BLA

Schoukens et al. (2008) outline the BLA-based identification algorithm. The key observation is that the inverse of the BLA is approximately given by:

$$\frac{1}{G_{bla}(q)} \approx G^{[3]}(q) + \frac{\beta}{G^{[1]}(q)G^{[2]}(q)}, \quad (50)$$

where $\beta = \frac{1}{\alpha}$.

In a first step, the BLA of the system is identified at $R \geq 2$ different setpoints of the system:

$$\hat{G}_{bla}^{[j]}(q) \approx \frac{\alpha^{[j]} G^{[1]}(q) G^{[2]}(q)}{1 + \alpha^{[j]} G^{[1]}(q) G^{[2]}(q) G^{[3]}(q)}, \quad (51)$$

for $j = 1, \dots, R$.

A nonparametric initial estimate $\hat{G}^{[3]}(e^{j\omega T_s})$ of the linear feedback system is obtained as the mean of the inverses of the BLA:

$$\begin{aligned} \hat{G}^{[3]}(e^{j\omega_k T_s}) &= \frac{1}{R} \sum_{j=1}^R \frac{1}{\hat{G}_{bla}^{[j]}(e^{j\omega_k T_s})} \\ &\approx G^{[3]}(e^{j\omega_k T_s}) + \frac{\gamma}{G^{[1]}(e^{j\omega_k T_s}) G^{[2]}(e^{j\omega_k T_s})}, \end{aligned} \quad (52)$$

where $\gamma = \frac{1}{R} \sum_{j=1}^R \beta^{[j]}$.

A nonparametric estimate of a scaled version of the inverse of $S(e^{j\omega T_s}) = G^{[1]}(e^{j\omega T_s}) G^{[2]}(e^{j\omega T_s})$ is obtained at each setpoint:

$$E^{[j]}(e^{j\omega_k T_s}) = \frac{1}{\hat{G}_{bla}^{[j]}(e^{j\omega_k T_s})} - \hat{G}^{[3]}(e^{j\omega_k T_s}). \quad (53)$$

These estimates are collected for all frequencies and setpoints in a matrix as follows:

$$\mathbf{E} = \begin{bmatrix} E^{[1]}(e^{j\omega_1 T_s}) & \dots & E^{[R]}(e^{j\omega_1 T_s}) \\ \vdots & \ddots & \vdots \\ E^{[1]}(e^{j\omega_F T_s}) & \dots & E^{[R]}(e^{j\omega_F T_s}) \end{bmatrix}. \quad (54)$$

The SVD of this matrix results in:

$$\mathbf{E} = \mathbf{U} \Sigma \mathbf{V}^T. \quad (55)$$

The estimate $\hat{S}^{-1}(e^{j\omega_k T_s})$ is equal to the first left singular vector present in the matrix \mathbf{U} . Taking its inverse and ignoring the scaling factor (it will be accounted for in the estimate of the static nonlinear block), an estimate of the product $G^{[1]}(e^{j\omega T_s})G^{[2]}(e^{j\omega T_s})$ is obtained.

Splitting the dynamics $\hat{S}(e^{j\omega_k T_s})$ to the front and the back linear block of the Wiener–Hammerstein system can be done using a combination of the methods presented in Sections 7.1 and 9.1. This results in the estimates $\hat{G}^{[1]}(e^{j\omega_k T_s})$, $\hat{G}^{[2]}(e^{j\omega_k T_s})$ and $\hat{f}(\cdot)$.

Finally, a nonlinear optimization (see Section 5) of all the parameters of the closed loop Wiener–Hammerstein feedback structure together is applied such that a consistent estimate is obtained. Note from [Assumption 1](#) that the feedback mechanism acts on the noise-free signal.

9.3. Linear fractional representation

Vandersteen and Schoukens (1999), Hsu, Poolla, and Vincent (2008), Novara, Vincent, Hsu, Milanese, and Poolla (2011), Van Mulders, Schoukens, and Vanbeylen (2013) and Vanbeylen (2013) study the linear fractional representation (LFR) model structure. It is a combination of a Wiener–Hammerstein feedback system with a linear feedforward branch in parallel (see Fig. 1). It is the most general block-oriented nonlinear system that contains only one static nonlinear block. All other block-oriented nonlinear systems with only one static nonlinearity can be reformulated as an LFR structured system.

The approach that is presented in Vandersteen and Schoukens (1999) relies on a large set of two-tone measurements and Volterra theory to estimate the LFR model. The method presented by Van Mulders et al. (2013) starts from a polynomial nonlinear state space model and transforms it into a LFR model. The block-oriented, BLA-based method that is presented in Vanbeylen (2013) is studied in some more detail below.

Hsu et al. (2008) and Novara et al. (2011) present an extension to an LFR model with a MIMO static nonlinear block. This LFR structure with MIMO blocks includes both filtered output and process noise. While the methods presented in Vandersteen and Schoukens (1999), Van Mulders et al. (2013) and Vanbeylen (2013) start from input–output data only, the methods presented in Hsu et al. (2008) and Novara et al. (2011) assume that the inputs to the nonlinear block are measured (i.e. they can be obtained from the measured inputs and outputs and the known linear block). Furthermore, the static nonlinear block is identified under the assumptions that the MIMO LTI block is known and that the nonlinear block has a known block-diagonal structure with MISO blocks.

Best linear approximation

The BLA of a general LFR system cannot be simplified to an easy function of the system dynamics, even if input signals belonging to

\mathbb{S}_U are used. The input signal of the nonlinearity $y(t)$ is not Gaussian due to the presence of the nonlinear feedback loop. Hence, Busgang's theorem cannot be applied. However, in many practical cases it is fair to assume that the nonlinearity can be linearized to a gain α and thus that the BLA will be approximately of the form:

$$G_{bla}(e^{j\omega T_s}) \approx G^{[1]}(e^{j\omega T_s}) + \frac{\alpha G^{[2]}(e^{j\omega T_s})G^{[3]}(e^{j\omega T_s})}{1 + \alpha G^{[4]}(e^{j\omega T_s})}, \quad (56)$$

$\alpha \in \mathbb{R}.$

The approximation in (56) is relevant if the memory length of $G^{[4]}(q)$ is sufficiently large.

The ϵ -approximation, contrary to the BLA, results in a simple and exact analytical expression of the linear approximation in function of the system dynamics (Schoukens et al., 2015a). The expression for $G_\epsilon(e^{j\omega T_s})$ is as in (56) with α replaced by γ .

The gains α and γ do not only depend on the system characteristics, but they can also depend on the variance, power spectrum, and the offset (DC value) of the input signal.

Note that in the LFR case both the poles and zeros of the ϵ -approximation shift with a changing gain γ :

$$G_\epsilon(q) = G^{[1]}(q) + \frac{\gamma G^{[2]}(q)G^{[3]}(q)}{1 + \gamma G^{[4]}(q)}, \quad (57)$$

where

$$G^{[i]}(q) = \frac{B^{[i]}(q)}{A^{[i]}(q)}. \quad (58)$$

Introducing a common denominator results in an expression containing a dependency on γ in both numerator and denominator. Schoukens et al. (2015a) discuss this in further detail.

Identification using the BLA

The identification approach for the LFR model that is presented in Vanbeylen (2013) shares some of its aspects with the approaches that are explained above: it relies on at least two measured BLAs, and it needs to perform a scan over several possible solutions to allocate the dynamics over the different LTI-blocks in the model. The main aspects of the identification algorithm are described below.

First, the BLA is estimated at two different setpoints of the system. This BLA is converted to linear state space models.

Next, an equivalence transformation for each state space model is estimated by solving a nonsymmetric algebraic Riccati equation. The details of this equation can be found in Vanbeylen (2013). This results in different possible solutions.

Each solution is evaluated by constructing the corresponding LFR model. The model with the best fit is selected. This step is very similar to the pole-zero allocation scan during the Wiener–Hammerstein identification in Section 7.1.

A nonlinear optimization (see Section 5) of all the parameters of the closed loop simple feedback structure together is applied such that a consistent estimate is obtained. Note from [Assumption 1](#) that the feedback mechanism acts on the noise-free signal.

10. Pros and cons of the model structures

The model structures and identification methods that are explained in the previous sections all have their pros and cons. Some model structures have the ability to accurately model the input–output behavior of a large class of systems, but use a large number of parameters in return. Some identification methods can easily generate starting values from a single experiment, others require computationally expensive steps on multiple data sets. This section provides an overview of the model structures in terms of their descriptive power, number of parameters, and general difficulty in estimating the model. These criteria are explained below. Next, the

Table 1

Overview of the considered model structures and their identification using the BLA, as explained in Section 7–9. Hammerstein is abbreviated as H, Wiener as W, and Linear Fractional Representation as LFR.

Model structure	Section	Universal approximator	Number of parameters	Consistent initial values	General difficulties
Hammerstein (H)	7.1	No	Low	Yes	/
Wiener (W)	7.1	No	Low	Yes	/
W-H	7.1	No	Low	Yes	Mildly computationally expensive brute-force scan
H-W	7.2	No	Low	No ^a	Assumes an invertible output nonlinearity, iterative procedure
Parallel H	8.1	No	Low/medium	Yes	BLA needed at different setpoints
Parallel W	8.1	Yes/no ^b	Medium	Yes	BLA needed at different setpoints
Wiener–Schetzen	8.2	Yes	High ^c	Yes	Number of parameters strongly depends on the quality of the selected basis functions
Parallel W-H	8.3	Yes/no ^b	High ^d	Yes	BLA needed at different setpoints, computationally expensive brute-force scan
Simple feedback	9.1	No	Low	No	Stability nonlinear feedback hard to check
W-H feedback	9.2	No	Low	No	BLA needed at different setpoints, mildly computationally expensive brute-force scan, stability nonlinear feedback hard to check
LFR	9.3	No	Low/medium	No	BLA needed at different setpoints, mildly computationally expensive scan over solutions of a NARE ^e , stability nonlinear feedback hard to check

^a Only consistent in absence of noise and model errors.

^b Although the model structure is a universal approximator, the proposed identification methods cannot deal with the general class of Volterra systems; some limitations apply.

^c Multiple input single output polynomial.

^d Multiple input multiple output static nonlinearity.

^e Nonsymmetric algebraic Riccati equation.

different model structures are compared based on these criteria, and some general limitations of BLA-based block-oriented nonlinear modeling are discussed.

10.1. Criteria for comparison

In the context of this paper, a universal approximator is a model structure that can approximate arbitrarily well the input–output behavior of systems belonging to the PISPO system class (see Section 4.1) in mean square sense. Parallel Wiener and parallel Wiener–Hammerstein models are universal approximators without any constraint on the excitation class (Schetzen, 2006). Other approximation properties of these model structures have been shown in the past (Boyd & Chua, 1985; Korenberg, 1991; Palm, 1979).

A large number of independent parameters lead to a large variance on the estimated parameters (Söderström & Stoica, 1989) when a model is identified from noisy measurements. Therefore, a model is preferably parameter-parsimonious, as decreasing the variability on the estimated parameters requires larger data sets, thereby increasing the computational time of the optimization step in Section 5.

The optimization step in Section 5 converges to a local optimum of the (non-convex) cost function. Good initial estimates are thus required. A consistent initial estimate is preferable, as in that case a good initial estimate is guaranteed as more estimation data are used. As described in Section 5, the nonlinear estimator is consistent under a limited set of assumptions for each of the considered model structures.

Some identification approaches require multiple data sets, e.g. BLA estimates at different setpoints, which is not always possible due to time constraints.

10.2. Comparison of model structures

Table 1 provides an overview of the criteria for comparison as well as some general difficulties of the identification algorithms that are presented in the previous sections.

Only the parallel Wiener, Wiener–Schetzen, and parallel Wiener–Hammerstein model structures are universal approximators. Nevertheless, these feedforward structures cannot capture

chaotic behavior, sub-harmonics, or hysteresis, since this requires feedback. The Linear Fractional Representation (LFR) is the most general block-oriented structure that contains only one static nonlinear block. The Wiener, Hammerstein, Wiener–Hammerstein, simple feedback, and Wiener–Hammerstein feedback structures are thus all special cases of the LFR.

The single branch models of Section 7 all have a low number of parameters, since these structures consist of a limited series connection of SISO blocks. The parallel Hammerstein model structure consists of a number of parallel connections of Hammerstein models and thus still has a fairly low number of parameters. The parallel Wiener and Wiener–Schetzen model structures both have a MISO static nonlinear block, which typically results in a high number of parameters. For the parallel Wiener structure, however, the proposed identification method results in a low number of parallel branches, i.e. a low number of inputs to the nonlinear block, and thus a medium number of parameters. The parallel Wiener–Hammerstein structure has a MIMO static nonlinear block, resulting in a high number of parameters. The feedback structures of Section 9 consist of a limited (at most five) SISO blocks, and thus have a low number of parameters. The LFR structure, however, is typically parametrized in state space form (Vanbeylen, 2013), which introduces at least n^2 redundant parameters, where n is the model order. This is due to the existence of an $n \times n$ transformation matrix that leaves the input–output behavior unchanged. This transformation matrix should be nonsingular, but except for this constraint the elements can be freely chosen. The LFR structure can thus have a medium number of parameters, although the number of independent parameters is still low.

Most of the proposed identification methods result in consistent initial estimates. This is not the case for the Hammerstein–Wiener identification method and the methods for the feedback structures of Section 9, since these methods use the noise-corrupted output either in some of the regressors (Hammerstein–Wiener and simple feedback) or to reconstruct internal signals that constitute some of the regressors (Wiener–Hammerstein feedback and LFR). Moreover, an approximation to the BLA is made in the Hammerstein–Wiener structure and all feedback structures. This approximation can be avoided by

replacing the BLA-framework by the theoretical framework of the ϵ -approximation (Schoukens et al., 2015a).

The presented methods on model structures that include a Wiener–Hammerstein branch all involve a brute-force scan to separate the dynamics of the Wiener–Hammerstein branch over its two LTI blocks. A similar scan over the solutions of a non-symmetric algebraic Riccati equation is performed in the LFR identification method. These scans are mildly computationally expensive if only one Wiener–Hammerstein branch is involved (Wiener–Hammerstein, Wiener–Hammerstein feedback, and LFR structures). The scan is computationally expensive for the parallel Wiener–Hammerstein structure, since there, a scan needs to be performed for each branch and each time a MIMO nonlinearity needs to be estimated.

The BIBO stability of the feedforward models is easy to check, since if all subblocks are BIBO stable, the overall model is stable, and the stability of LTI and SNL subblocks can be easily checked. For the feedback models, the stability checks are more involved (see Section 9).

10.3. Limitations BLA-based block-oriented nonlinear modeling

From the comparison of the models in Table 1, it can be observed that the initialization can be cumbersome for the more complex model structures, while the flexibility of the simpler models can be limited.

The BLA-based block-oriented approaches presented in this paper are multistage approaches. In a first stage, the BLA is determined, and in one or more later stages, the full nonlinear model is estimated. Very noisy data, or strongly nonlinear systems (e.g. a hard saturation nonlinearity) can result in a poor BLA estimate. It is more likely that the nonlinear optimization will end up in a local minimum of the cost function in the noisy or very nonlinear case. However, we also know that more data will eventually result in better estimates since most of the initialization algorithms that are presented in this paper (and also the BLA estimator itself (Pintelon & Schoukens, 2012)) are consistent estimators. Therefore, we advice the user to collect more data, if possible, in case the measurements are (very) noisy or the system is strongly nonlinear. Using periodic excitations allows one to easily obtain an estimate of the level of the noise in the measurements, averaging over the periods allows one to decrease the influence of the noise on the measurements.

A particular case is the presence of an even nonlinearity. As pointed out in Section 4, the BLA is then zero. The quality of the estimate of the linear block(s), on which the presented methods rely, is typically poor in this case. An alternative for the BLA-based approach that overcomes this issue with an even nonlinearity is the weighted principal component analysis (PCA) method as reported by Zhang, Laurain, and Wang (2015) for the identification of Wiener systems.

11. Guidelines for the user

Two important steps in the system identification process are the input design and model structure selection step. This section provides some user guidelines for these two steps.

11.1. Model structure selection

The use of the BLA framework is not limited to obtaining a block-oriented nonlinear model of a nonlinear system. The BLA framework also offers a powerful analysis tool that allows the user, when combined with a well thought out input design strategy, to acquire a lot of insight into the nonlinear system at very low cost. The questions “How nonlinear is the system?”, “Do I need a nonlinear

model?”, and “What kind of nonlinear effects do I observe in the system?” can be (partly) answered within the BLA framework (Lauwers et al., 2008; Pintelon & Schoukens, 2012; Schoukens et al., 2012a, 2015a), an overview can be found in Schoukens et al. (2016a).

The BLA framework offers the user the possibility of obtaining both a nonparametric frequency domain estimate of the noise disturbance present at the output and of the system combined, and a nonparametric frequency domain estimate of the nonlinear disturbances generated by the system (Pintelon & Schoukens, 2012; Schoukens et al., 2012a). These can be used to quantify the level of the nonlinearities in the system and to determine whether or not the nonlinearities or the noise are the dominant disturbance source. Hence, it also allows the user to make a choice on whether or not a nonlinear model should be estimated, based on a quantitative measure. In some cases, depending on the final goal, a linear model could be accurate enough.

Exciting the nonlinear system under test at different setpoints of the system (e.g. different input variances or mean values of the input signal) can be very useful to detect what kind of nonlinear effects are present in the system. It is presented in Schoukens et al. (2015a) that a pole shift in the BLAs over the different setpoints indicates the presence of nonlinear feedback loops, while a zero shift indicates the presence of parallel nonlinear signal paths in the system. These pole and zero shifts are exploited by some of the presented identification algorithms to generate initial estimates of the linear dynamic blocks of the selected block-oriented structure. The BLA-based model structure selection approach can be combined with other model structure selection approaches, e.g. Haber and Unbehauen (1990) and Pearson (2003), to guide the user to a suitable model structure.

It is often not possible to perform an elaborate measurement campaign on the nonlinear system that needs to be modeled. A thorough system analysis and model structure selection step are not always possible under these circumstances. Fortunately, most of the presented modeling approaches are relatively simple and can be applied at a fairly low computational cost. This allows the user to try out multiple model structures and compare the quality of the resulting model on a validation data set.

11.2. Input design

A model is in many practical cases only an approximation of the system under study. Therefore, it is of great importance to use input signals that represent a realistic set of excitations to the system. The system should be excited within a realistic range of operation, both in amplitude and frequency. Also the amplitude distribution (e.g. the crest factor of the input) can play an important role. These considerations often exclude relatively simple excitation signals such as sine inputs and impulses.

The BLA framework prefers, but is not limited to, different realizations of periodic, random signals to perform a nonparametric analysis of the nonlinear system under test. An increase in the number of signal realizations decreases the variance on the BLA estimate due to both the nonlinear and noise distortion, while an increase of the number of periods only decreases the variance due to the noise distortion. A typical choice would be to use 7 different realizations and 2 (steady state) periods of a periodic, random signal (Pintelon & Schoukens, 2012; Schoukens et al., 2012a). Some of the presented identification algorithms require the BLA to be measured at different setpoints of the system. Esfahani, Schoukens, and Vanbeylen (2016) introduce an experimental algorithm to find a set of input signal rms-values and DC offsets such that the distortion on the BLA is minimized.

Multisine signals (and more specifically random phase multisines) are excellent candidate input signals since they are periodic

and offer a full control of the amplitude spectrum of the input signal (Pintelon & Schoukens, 2012). Small multisine signals can be superimposed on a more realistic input signal for the system under test to obtain a random input signal with a rich frequency content that can be used for a nonparametric analysis of the nonlinear system under test. An example of this can be found in Widanage, Stoev, Van Mulders, Schoukens, and Pinte (2011) and Marconato, Schoukens, Tiels, Widanage, Abu-Rmileh, and Schoukens (2014).

The identification algorithms that are presented in the previous sections make use of the Riemann equivalence class of asymptotically normally distributed excitation signals \mathbb{S}_U (including random phase multisine signals). This could be seen as a limitation of the identification algorithms. However, the presented algorithms are quite robust to deviations of this signal class due to the final nonlinear optimization step described in Section 5.

12. Benchmark and practical results

This section provides an overview of the benchmark setups and real-life systems on which the BLA-based block-oriented identification methods have been applied as well as the lessons learned from them. These results illustrate the good performance and the practical usefulness of the BLA-based block-oriented nonlinear modeling approaches.

The Wiener–Hammerstein identification techniques in Section 7.1 obtained excellent results on the Wiener–Hammerstein benchmark (Hjalmarsson, Rojas, & Rivera, 2012; Schoukens et al., 2009b; Sjöberg, Lauwers, & Schoukens, 2012). The simple feedback structure and the Wiener–Hammerstein feedback structure have been successfully applied on an electronic simulator of a hardening spring available as the Silverbox benchmark (Paduart, 2008; Wigren & Schoukens, 2013). The coupled electric drives benchmark in Wigren and Schoukens (2013) is an example of a system that is difficult to identify using the BLA-based approaches due to the presence of the purely even nonlinearity (see Section 4).

Other systems, not available as a benchmark, that have been successfully modeled by the block-oriented BLA-based approaches are: the insulin–glucose regulatory system (Wiener, Wiener–Schetzen and LFR models) (Marconato et al., 2014; Vanbeylen, Van Mulders, & Abu-Rmileh, 2014), a RF crystal detector (Wiener–Hammerstein feedback model) (Schoukens et al., 2008), a logarithmic amplifier (Wiener and parallel Wiener models) (Schoukens & Rolain, 2012b), a valve audio amplifier (Hammerstein and parallel Hammerstein models) (Schoukens et al., 2011) and a Doherty power amplifier (parallel Wiener–Hammerstein model) (Schoukens, 2015). Westwick and Kearney (2003) report applications in biomedical engineering and physiology using related correlation-based approaches.

From these and other practical results, it was found that including structural prior knowledge leads to a better trade-off between modeling performance and parameter parsimony (Marconato, Schoukens, Rolain, & Schoukens, 2013; Vanbeylen et al., 2014). Nevertheless, the selection of the model structure is still a user choice in which the intended application of the model plays a role. If the model will be used for control purposes for example, it makes sense to restrict the model class to invertible models (Gaasbeek & Mohan, 2016; Pajunen, 1992). If the model needs to be re-estimated often on new datasets, it makes sense to use a model structure that can be easily identified and that still has an acceptable error performance, even though it is known that the model is too simple (Rébillat et al., 2016).

13. Future research directions

This overview paper surveyed a quite broad range of block-oriented identification approaches, with the emphasis on methods

starting from linear approximations of the nonlinear system, for a wide range of block-oriented structures. It is also clear from the surveyed work that there are quite some open problems in the identification of block-oriented nonlinear systems. Some of these open problems are listed in this section. Quite some of these open problems are already addressed to some extent, but the full picture is not yet complete.

A first open research question is how to include dynamic nonlinearities in the block-oriented nonlinear modeling toolbox. Quite some research has already been performed on this topic, for backlash, backlash-inverse and hysteresis nonlinearities (Brouri et al., 2014b; Giri et al., 2014, 2011, 2013; Rochdi, Giri, Gning, & Chaoui, 2010; Wang et al., 2012). First steps to include dynamic nonlinearities in a BLA-based identification framework are made in Yong et al. (2015) for Hammerstein and Wiener model structures. Most of the methods that are developed until now consider Hammerstein and Wiener structures only, and are often limited by a specific input signal class. Further research on this topic is encouraged through the hysteretic nonlinear system identification benchmark presented in Noël and Schoukens (2016).

The extension of the available identification methods to more complex structures is another open research question. This includes for instance the development of flexible identification methods for complex, user-generated and problem-dependent block-oriented structures and the extension of the existing block-oriented structures towards a MIMO setting without losing the structured nature of a block-oriented nonlinear model. Some efforts have already been made to identify MIMO block-oriented structures, both for the simpler Hammerstein and Wiener block-oriented structures (Goethals et al., 2005; Janczak, 2007; Verhaegen & Westwick, 1996; Westwick & Verhaegen, 1996), as for the more involved Wiener–Hammerstein, parallel Hammerstein and model structures (Boutayeb & Darouach, 1995; Katayama & Ase, 2016; Schoukens & Rolain, 2011; Tiels, Heuberger, & Schoukens, 2012). These results could serve as a starting point for the development of other MIMO block-oriented identification algorithms.

Most block-oriented identification methods assume the presence of an output-error noise framework. This is often a simplified representation of reality, as discussed in Section 2. Developing identification methods that are able to deal with more complex noise frameworks is an ongoing research effort (see for instance Hagenblad et al., (2008) Lindsten et al. (2013) Mu and Chen (2014) Wahlberg et al., 2014 and Wills et al., 2013 and also the results on the benchmark presented in Schoukens and Noël (2016)).

A block-oriented nonlinear model is often an approximation of the true system behavior. How to deal with model errors in nonlinear system identification, and characterizing these model errors is an open research problem. The results of the BLA framework (Pintelon & Schoukens, 2012) could help as a guideline in how to develop a nonlinear system identification framework dealing with model errors.

In many practical cases, the user has some prior knowledge about the system behavior. This can be information about the system structure and the type of the static nonlinearity (saturation, discontinuity, smoothness), but also about the location of some poles and zeros and the stable nature of the system. Ideally, such prior knowledge should be taken into account during the identification of the system. The recently developed regularization-based identification methods allow one to include such prior knowledge (as is illustrated in Risuleo et al. (2015) for the Hammerstein model structure and (Birpoutsoukis & Schoukens, 2015) for Volterra kernel estimation). The work presented in Tiels and Schoukens (2014) proposes an approach to include prior pole-zero information during the identification for Wiener–Schetzen model structures.

The input of a nonlinear system is not always controlled by the user (e.g. when dealing with rainfall input), or it can be very costly

to collect a high number of data points. The identification of poorly excited systems, or the identification of systems with short data records is an open problem (see also the benchmark presented in Schoukens, Mattson, Wigren, and Noël (2016b)). Regularization-based identification methods are a possible solution to this problem (e.g. (Risuleo et al., 2015)). However, more research has to be performed to adapt these approaches for nonlinear system identification purposes.

Finally, an often neglected or underestimated problem in the development of a nonlinear system identification algorithm is its user friendliness. Most users are not experts in (nonlinear) system identification. Therefore, it is of great importance to develop identification algorithms for complex block-oriented structures that are robust with respect to the user choices, require as little user interaction as possible, and have little restrictions on the data used for identification. Such identification algorithms start to be available for the simpler block-oriented structures (e.g. Hammerstein and Wiener), but this is typically not yet the case for the more complex block-oriented structures (e.g. parallel Wiener–Hammerstein, Wiener–Hammerstein feedback, LFR). Developing algorithms that have less restrictions on the estimation data, and that are more user friendly in general could increase the applicability of such model structures significantly.

14. Conclusion

This paper surveys different block-oriented nonlinear systems that can be identified based on the best linear approximation, and it gives an overview of the identification algorithms that have been developed in the past. It turns out that a wide range of systems can be modeled using the best linear approximation approaches: single branch systems, parallel branch systems and nonlinear feedback systems. Furthermore, some future and ongoing research directions are discussed.

The best linear approximation framework allows the user to extract important information about the system (e.g. model order of the dynamics, system structure information), it guides the user in taking good modeling decisions, and it proves to be a good starting point for nonlinear system identification algorithms for a wide range of block-oriented nonlinear model structures. The best linear approximation provides in some cases a direct estimate of the linear dynamic block of the system. In other cases the best linear approximation can be combined with a pole-zero allocation scan, a singular value decomposition, orthogonal basis functions or a Riccati-equation-based approach to identify the nonlinear system under test.

References

- Abrahamsson, R., Kay, S., & Stoica, P. (2007). Estimation of the parameters of a bilinear model with applications to submarine detection and system identification. *Digital Signal Processing*, 17(4), 756–773.
- Aljamaan, I., Bu bshait, A., & Westwick, D. (2011). Separable least squares identification of Wiener Box–Jenkins models. In *19th world congress of the international federation of automatic control* (pp. 4434–4439).
- Aljamaan, I., Westwick, D., & Foley, M. (2014). Non-iterative identification of IIR Wiener systems using orthogonal polynomial. In *19th world congress of the international federation of automatic control* (pp. 487–492).
- Ase, H., & Katayama, T. (2015). A subspace-based identification of Wiener–Hammerstein benchmark model. *Control Engineering Practice*, 44, 126–137.
- Bai, E. (1998). An optimal two-stage identification algorithm for Hammerstein–Wiener nonlinear systems. *Automatica*, 34(3), 333–338.
- Bai, E. (2002). A blind approach to the Hammerstein–Wiener model identification. *Automatica*, 38(6), 967–979.
- Bai, E. (2002). Identification of linear systems with hard input nonlinearities of known structure. *Automatica*, 38(5), 853–860.
- Bai, E., & Li, D. (2004). Convergence of the iterative Hammerstein system identification algorithm. *IEEE Transactions on Automatic Control*, 49(11), 1929–1940.
- Bai, E., & Li, K. (2010). Convergence of the iterative algorithm for a general Hammerstein system identification. *Automatica*, 46(11), 1891–1896.
- Baumgartner, S., & Rugh, W. (1975). Complete identification of a class of nonlinear systems from steady state frequency response. *IEEE Transactions on Automatic Control*, 22(9), 753–759.
- Ben Ahmed, Z., Favier, G., & Derbel, N. (2013). Tensor-based methods for Wiener–Hammerstein system identification. In *10th international multi-conference on systems, signals & devices* (pp. 1–6).
- Billings, S. (1980). Identification of nonlinear systems: A survey. *IEE Proceedings D-Control Theory and Applications*, 127(6), 272–285.
- Billings, S. (2013). *Nonlinear system identification: NARMAX methods in the time, frequency, and spatio-temporal domains*. West Sussex, United Kingdom: John Wiley & Sons, Ltd..
- Billings, S., & Fakhouri, S. (1977). Identification of nonlinear systems using the Wiener model. *Electronics Letters*, 13(17), 502–504.
- Billings, S., & Fakhouri, S. (1978a). Identification of a class of nonlinear systems using correlation analysis. *Proceedings of the Institution of Electrical Engineers*, 125(7), 691–697.
- Billings, S., & Fakhouri, S. (1978b). Theory of separable processes with applications to the identification of nonlinear systems. *Proceedings of the Institution of Electrical Engineers*, 125(10), 1051–1058.
- Billings, S., & Fakhouri, S. (1979). Identification of non-linear S_m systems. *International Journal of Systems Science*, 10(10), 1401–1408.
- Billings, S., & Fakhouri, S. (1982). Identification of systems containing linear dynamic and static nonlinear elements. *Automatica*, 18(1), 15–26.
- Birpoutsoukis, G., & Schoukens, J. (2015). Nonparametric Volterra kernel estimation using regularization. In *IEEE international instrumentation and measurement technology conference* (pp. 222–227).
- Boutayeb, M., & Darouach, M. (1995). Recursive identification method for MISO Wiener–Hammerstein model. *IEEE Transactions on Automatic Control*, 40(2), 287–291.
- Boyd, S., & Chua, L. (1985). Fading memory and the problem of approximating non-linear operators with Volterra series. *IEEE Transactions on Circuits and Systems*, 32(11), 1150–1161.
- Boyd, S., Chua, L. O., & Desoer, C. A. (1984). Analytical foundations of Volterra series. *IMA Journal of Mathematical Control & Information*, 1, 243–282.
- Brouri, A., Amdouri, O., Chaoui, F., & Giri, F. (2014a). Frequency identification of Hammerstein–Wiener systems with piecewise affine input nonlinearity. In *19th world congress of the international federation of automatic control* (pp. 10030–10035).
- Brouri, A., Amdouri, O., Chaoui, F., & Giri, F. (2014b). Identification of Hammerstein–Wiener systems with backlash input nonlinearity bordered by straight lines. In *19th world congress of the international federation of automatic control* (pp. 475–480).
- Bussgang, J. (1952). Cross-correlation functions of amplitude-distorted Gaussian signals. In *Technical report 216*. MIT Laboratory of Electronics.
- Cai, Z., & Bai, E. (2011). Making parametric Hammerstein system identification a linear problem. *Automatica*, 47(8), 1806–1812.
- Castro-Garcia, R., Agudelo, O., Tiels, K., & Suykens, J. (2016). Hammerstein system identification using LS-SVM and steady state time response. In *2016 European conference on automatic control* (pp. 1063–1068).
- Cerone, V., Razza, V., & Regruto, D. (2015). One-shot set-membership identification of Wiener models with polynomial nonlinearities. In *17th IFAC symposium on system identification* (pp. 957–962).
- Chang, F., & Luus, R. (1971). A noniterative method for identification using Hammerstein model. *IEEE Transactions on Automatic Control*, 16(5), 464–468.
- Crama, P., & Schoukens, J. (2001). First estimates of Wiener and Hammerstein systems using multisine excitation. In *IEEE instrumentation and measurement technology conference*, vol. 2 (pp. 1365–1369).
- Crama, P., & Schoukens, J. (2004). Hammerstein–Wiener system estimator initialization. *Automatica*, 40(9), 1543–1550.
- Crama, P., & Schoukens, J. (2005). Computing an initial estimate of a Wiener–Hammerstein system with a random phase multisine excitation. *IEEE Transactions on Instrumentation and Measurement*, 54(1), 117–122.
- Crama, P., Schoukens, J., & Pintelon, R. (2004). Generation of enhanced initial estimates for Hammerstein systems. *Automatica*, 40(7), 1269–1273.
- Cramér, H. (1946). *Mathematical methods of statistics*. Princeton, N.J.: Princeton University Press.
- Dewhirst, O., Simpson, D., Angarita, N., Allen, R., & Newland, P. (2010). Wiener–Hammerstein parameter estimation using differential evolution: Application to limb reflex dynamics. In *International conference on bio-inspired systems and signal processing* (pp. 271–276).
- Doyle, F., Pearson, R., & Ogunnaike, B. (2002). *Identification and control using Volterra models*. Springer.
- Dreesen, P., Schoukens, M., Tiels, K., & Schoukens, J. (2015). Decoupling static nonlinearities in a parallel Wiener–Hammerstein system: A first-order approach. In *IEEE international instrumentation and measurement technology conference* (pp. 987–992).
- Dreesen, P., Westwick, D., Schoukens, J., & Ishteva, M. (2017). Modeling parallel Wiener–Hammerstein systems using tensor decomposition of Volterra kernels. In *The 13th international conference on latent variable analysis and signal separation* (pp. 16–25).

- Enqvist, M. (2005). *Linear models of nonlinear systems*. Sweden: Institute of technology, Linköping University.
- Enqvist, M. (2010). Identification of block-oriented systems using the invariance property. In F. Giri, & E. Bai (Eds.), *Lecture Notes in Control and Information Sciences: vol. 404. Block-oriented nonlinear system identification* (pp. 147–158). Berlin Heidelberg: Springer.
- Enqvist, M. (2011). Separability of scalar random multisine signals. *Automatica*, 47(9), 1860–1867.
- Enqvist, M., & Ljung, L. (2005). Linear approximations of nonlinear FIR systems for separable input processes. *Automatica*, 41(3), 459–473.
- Esfahani, A., Schoukens, F., & Vanbeylen, L. (2016). Using the best linear approximation with varying excitation signals for nonlinear system characterization. *IEEE Transactions on Instrumentation and Measurement*, 65(5), 1271–1280.
- Falck, T., Suykens, J., Schoukens, J., & De Moor, B. (2010). Nuclear norm regularization for overparametrized Hammerstein systems. In *49th IEEE conference on decision and control* (pp. 7202–7207).
- Fletcher, R. (1991). *Practical methods of optimization* (2nd ed.). Wiley & Sons.
- Gaasbeek, R., & Mohan, R. (2016). Control-focused identification of hysteric systems – selecting model structures? Think about the final use of the model!. In *Workshop on nonlinear system identification benchmarks* (p. 32).
- Gadringer, M., Silveira, D., & Magerl, G. (2007). Efficient power amplifier identification using modified parallel cascade Hammerstein models. In *IEEE radio and wireless symposium* (pp. 305–308).
- Gallman, P. (1975). Iterative method for identification of nonlinear-systems using a Uryson model. *IEEE Transactions on Automatic Control*, 20(6), 771–775.
- Ghannouchi, F., & Hammi, O. (2009). Behavioral modeling and predistortion. *IEEE Microwave Magazine*, 10(7), 52–64.
- Giordano, G., & Sjöberg, J. (2015). A time-domain fractional approach for Wiener-Hammerstein systems identification. In *17th IFAC symposium on system identification* (pp. 1232–1237).
- Giri, F., Ahmed-Ali, T., Brouri, A., Chaoui, F., & Radouane, A. (2015). Frequency identification of a feedback Hammerstein system. In *17th IFAC symposium on system identification* (pp. 945–950).
- Giri, F., & Bai, E. (Eds.). (2010). *Block-oriented nonlinear system identification*. In *Lecture notes in control and information sciences: vol. 404*. (p. 431). Berlin Heidelberg: Springer.
- Giri, F., Radouane, A., Brouri, A., & Chaoui, F. (2014). Combined frequency-prediction error identification approach for Wiener systems with backlash and backlash-inverse operators. *Automatica*, 50(3), 768–783.
- Giri, F., Rochdi, Y., Brouri, A., & Chaoui, F. (2011). Parameter identification of Hammerstein systems containing backlash operators with arbitrary-shape parametric borders. *Automatica*, 47(8), 1827–1833.
- Giri, F., Rochdi, Y., Radouane, A., Brouri, A., & Chaoui, F. (2013). Frequency identification of nonparametric Wiener systems containing backlash nonlinearities. *Automatica*, 49(1), 124–137.
- Goethals, I., Pelckmans, K., Hoegaerts, L., Suykens, J., & De Moor, B. (2005). Subspace intersection identification of Hammerstein-Wiener systems. In *44th IEEE conference on decision and control and european control conference* (pp. 7108–7113).
- Goethals, I., Pelckmans, K., Suykens, J., & De Moor, B. (2005). Identification of MIMO Hammerstein models using least squares support vector machines. *Automatica*, 41(7), 1263–1272.
- Gómez, J. C., & Baeyens, E. (2004). Identification of block-oriented nonlinear systems using orthonormal bases. *Journal of Process Control*, 14, 685–697.
- Gómez, J. C., & Baeyens, E. (2012). A combined SVM/OB-based Wiener model identification method. In *16th IFAC symposium on system identification*.
- Greblicki, W. (2001). Recursive identification of Wiener systems. *International Journal of Applied Mathematics and Computer Science*, 977, 977–991.
- Guarin, D., & Kearney, R. (2015). An instrumental variable approach for the identification of time-varying, Hammerstein systems. In *17th IFAC symposium on system identification* (pp. 196–201).
- Haber, R., & Unbehauen, H. (1990). Structure identification of nonlinear dynamic systems-A survey on input/output approaches. *Automatica*, 26(4), 651–677.
- Haddad, W., & Bernstein, D. (1993). Explicit construction of quadratic Lyapunov functions for the small gain, positivity, circle, and Popov theorems and their application to robust stability. Part I: continuous-time theory. *International Journal of Robust and Nonlinear Control*, 3, 313–339.
- Haddad, W., & Bernstein, D. (1994). Explicit construction of quadratic Lyapunov functions for the small gain, positivity, circle, and Popov theorems and their application to robust stability. Part II: discrete-time theory. *International Journal of Robust and Nonlinear Control*, 4, 249–265.
- Hagenblad, A., & Ljung, L. (1998). Maximum likelihood identification of Wiener models with a linear regression initialization. In *37th IEEE conference on decision and control*, vol. 1 (pp. 712–713).
- Hagenblad, A., Ljung, L., & Wills, A. (2008). Maximum likelihood identification of Wiener models. *Automatica*, 44(11), 2697–2705.
- Hammerstein, A. (1930). Nichtlineare Integralgleichungen nebst Anwendungen. *Acta Mathematica*, 54(1), 117–176.
- Han, Y., & de Callafon, R. (2012). Hammerstein system identification using nuclear norm minimization. *Automatica*, 48(9), 2189–2193.
- Hasiewicz, Z., Mzyk, G., Śliwiński, P., & Wachel, P. (2012). Mixed parametric-nonparametric identification of Hammerstein and Wiener systems-a survey. In *16th IFAC symposium on system identification* (pp. 464–469).
- Heuberger, P., Van den Hof, P., & Wahlberg, B. (Eds.). (2005). *Modelling and identification with rational orthogonal basis functions*. London: Springer.
- Hjalmarsson, H., Rojas, C., & Rivera, D. (2012). System identification: A Wiener-Hammerstein benchmark. *Control Engineering Practice*, 20(11), 1095–1096.
- Hsia, T. (1976). A multi-stage least squares method for identifying Hammerstein model nonlinear systems. In *1976 IEEE conference on decision and control including the 15th symposium on adaptive processes* (pp. 934–938).
- Hsu, K., Poola, K., & Vincent, T. L. (2008). Identification of structured nonlinear systems. *IEEE Transactions on Automatic Control*, 53(11), 2497–2513.
- Hunter, I., & Korenberg, M. (1986). The identification of nonlinear biological systems: Wiener and Hammerstein cascade models. *Biological Cybernetics*, 55(2–3), 135–144.
- Janczak, A. (2007). Instrumental variables approach to identification of a class of MIMO Wiener systems. *Nonlinear Dynamics*, 48(3), 275–284.
- Katayama, T., & Ase, H. (2016). Linear approximation and identification of MIMO Wiener-Hammerstein systems. *Automatica*, 71, 118–124.
- Khalil, H. (1996). *Nonlinear systems* (2nd). Prentice Hall.
- Kibangou, A., & Favier, G. (2006). Wiener-Hammerstein systems modeling using diagonal Volterra kernels coefficients. *IEEE Signal Processing Letters*, 13(6), 381–384.
- Kibangou, A., & Favier, G. (2009). Identification of parallel-cascade Wiener systems using joint diagonalization of third-order Volterra kernel slices. *IEEE Signal Processing Letters*, 16, 188–191.
- Kibangou, A., Favier, G., & Hassani, M. (2005a). Iterative optimization method of GOB-Volterra filters. In *16th triennial world congress* (pp. 773–778).
- Kibangou, A., Favier, G., & Hassani, M. (2005b). Laguerre-Volterra filters optimization based on Laguerre spectra. *EURASIP Journal on Applied Signal Processing*, 17, 2874–2887.
- Korenberg, M. (1991). Parallel cascade identification and kernel estimation for nonlinear systems. *Annals of Biomedical Engineering*, 19(4), 429–455.
- Korenberg, M. J., & Hunter, I. W. (1986). The identification of nonlinear biological systems: In cascade models. *Biological Cybernetics*, 55(2), 125–134.
- Lacy, S., & Bernstein, D. (2003). Identification of FIR Wiener systems with unknown, non-invertible, polynomial non-linearities. *International Journal of Control*, 76(15), 1500–1507.
- Laurain, V., Gilson, M., & Garnier, H. (2012). *System identification, environmental modelling, and control system design*. In L. Wang, & H. Garnier (Eds.), (pp. 27–47). London: Springer.
- Lauwers, L. (2011). *Some practical applications of the best linear approximation in nonlinear block-oriented modelling*. Vrije Universiteit Brussel.
- Lauwers, L., Schoukens, J., Pintelon, R., & Enqvist, M. (2008). A nonlinear block structure identification procedure using frequency response function measurements. *IEEE Transactions on Instrumentation and Measurement*, 57(10), 2257–2264.
- Leontaritis, I., & Billings, S. (1985). Input output parametric models for non-linear systems. Part II: Stochastic non-linear systems. *International Journal of Control*, 41(2), 329–344.
- Levenberg, K. (1944). A method for the solution of certain problems in least squares. *Quarterly of Applied Mathematics*, 2, 164–168.
- Lindsten, F., Schön, T., & Jordan, M. (2013). Bayesian semiparametric Wiener system identification. *Automatica*, 49(7), 2053–2063.
- Ljung, L. (1999). *System identification: Theory for the user* (2nd ed.). Upper Saddle River, New Jersey: Prentice Hall.
- Ljung, L. (2001). Estimating linear time-invariant models of nonlinear time-varying systems. *European Journal of Control*, 7(2–3), 203–219.
- Lyzell, C., Andersen, M., & Enqvist, M. (2012). A convex relaxation of a dimension reduction problem using the nuclear norm. In *51st IEEE conference on decision and control* (pp. 2852–2857).
- Lyzell, C., & Enqvist, M. (2012a). Inverse regression for the Wiener class of systems. In *16th IFAC symposium on system identification* (pp. 476–481).
- Lyzell, C., & Enqvist, M. (2012b). Sliced inverse regression for the identification of dynamical systems. In *16th IFAC symposium on system identification* (pp. 1575–1580).
- Mäkilä, P., & Partington, J. (2003). On linear models for nonlinear systems. *Automatica*, 39(1), 1–13.
- Marconato, A., Schoukens, M., Rolain, Y., & Schoukens, J. (2013). Study of the effective number of parameters in nonlinear identification benchmarks. In *52nd IEEE conference on decision and control* (pp. 4308–4313).
- Marconato, A., Schoukens, M., Tiels, K., Widanage, W., Abu-Rmileh, A., & Schoukens, J. (2014). Identification and comparison of several nonlinear models applied to a glucoregulatory system example. *IET Control Theory & Applications*, 8(17), 1921–1930.
- Markusson, O. (2001). *Model and system inversion with applications in nonlinear system identification and control*. Stockholm, Sweden: Royal Institute of Technology.
- Marquardt, D. (1963). An algorithm for least-squares estimation of nonlinear parameters. *SIAM Journal on Applied Mathematics*, 11(2), 431–441.

- McGraw, D., & Wagner, J. (1968). Elliptically symmetric distributions. *IEEE Transactions on Information Theory*, 14(1), 110–120.
- Mu, B., & Chen, H. (2014). Recursive identification of errors-in-variables Wiener-Hammerstein systems. *European Journal of Control*, 20(1), 14–23.
- Mu, B., Han-Fu, C., Yi Wang, L., & Yin, G. (2015). Recursive identification of Hammerstein with noisy observations. In *17th IFAC symposium on system identification* (pp. 1017–1022).
- Mzyk, G. (2014). Combined parametric-nonparametric identification of block-oriented systems. In *Lecture notes in control and information sciences: vol. 454*. Berlin Heidelberg: Springer.
- Naitali, A., & Giri, F. (2016). Wiener-Hammerstein system identification—an evolutionary approach. *International Journal of Systems Science*, 47(1), 45–61.
- Narendra, K., & Gallman, P. (1966). An iterative method for the identification of nonlinear systems using a Hammerstein model. *IEEE Transactions on Automatic Control*, 11, 546–550.
- Ni, B., Gilson, M., & Garnier, H. (2013). Refined instrumental variable method for Hammerstein-Wiener continuous-time model identification. *IET Control Theory Applications*, 7(9), 1276–1286.
- Noël, J., & Schoukens, M. (2016). Hysteretic benchmark with a dynamic nonlinearity. In *Workshop on nonlinear system identification benchmarks* (pp. 7–14) <http://homepages.vub.ac.be/mschouke/benchmarkBoucWen.html>.
- Novara, C., Vincent, T. L., Hsu, K., Milanesi, M., & Poolla, K. (2011). Parametric identification of structured nonlinear systems. *Automatica*, 47, 711–721.
- Nuttall, A. (1958). *Theory and application of the separable class of random processes*. MIT.
- Paduart, J. (2008). *Identification of nonlinear systems using polynomial nonlinear state space models*. Vrije Universiteit Brussel, Belgium.
- Paduart, J., Horváth, G., & Schoukens, J. (2004). Fast identification of systems with nonlinear feedback. In *6th IFAC symposium on nonlinear control systems* (pp. 381–385).
- Pajunen, G. (1992). Adaptive control of Wiener type nonlinear systems. *Automatica*, 28(4), 781–785.
- Palanthandalam-Madapusi, H., Ridley, A., & Bernstein, D. (2005). Identification and prediction of ionospheric dynamics using a Hammerstein-Wiener model with radial basis functions. In *American control conference* (pp. 5052–5057).
- Palm, G. (1978). On representation and approximation of nonlinear systems. *Biological Cybernetics*, 31, 119–124.
- Palm, G. (1979). On representation and approximation of nonlinear systems Part II: Discrete Time. *Biological Cybernetics*, 34, 49–52.
- Papoulis, A. (1991). *Probability, random variables, and stochastic processes*. McGraw-Hill.
- Pearson, R. (2003). Selecting nonlinear model structures for computer control. *Journal of Process Control*, 13(1), 1–26.
- Pelckmans, K. (2011). MINLIP for the identification of monotone Wiener systems. *Automatica*, 47(10), 2298–2305.
- Pintelon, R., & Schoukens, J. (2012). *System identification: A frequency domain approach* (2nd ed.). Hoboken, New Jersey: Wiley-IEEE Press.
- Rébillat, M., Ege, K., Gallo, M., & Antoni, J. (2016). Repeated exponential sine sweeps for the autonomous estimation of nonlinearities and bootstrap assessment of uncertainties. *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, SAGE Publications, 230(6), 1007–1018.
- Relan, R., & Schoukens, J. (2016). Recursive discrete-time models for continuous-time systems under band-limited assumptions. *IEEE Transactions on Instrumentation and Measurement*, 65(3), 713–723.
- Risuleo, R., Bottegal, G., & Hjalmarsson, H. (2015). A new kernel-based approach for overparameterized Hammerstein system identification. In *54th IEEE conference on decision and control* (pp. 115–120).
- Rochdi, Y., Giri, F., Gning, J., & Chaoui, F. (2010). Identification of block-oriented systems in the presence of nonparametric input nonlinearities of switch and backlash types. *Automatica*, 46(5), 864–877.
- Rolain, Y., Schoukens, J., & Pintelon, R. (1997). Order estimation for linear time-invariant systems using frequency domain identification methods. *IEEE Transactions on Automatic Control*, 42(10), 1408–1417.
- da Rosa, A., Campello, R., & Amaral, W. (2007). Choice of free parameters in expansions of discrete-time Volterra models using Kautz functions. *Automatica*, 43, 1084–1091.
- Schetzen, M. (2006). *The Volterra & Wiener theories of nonlinear systems*. Malabar, Florida: Krieger Publishing Company.
- Schön, T., Wills, A., & Ninness, B. (2011). System identification of nonlinear state-space models. *Automatica*, 47(1), 39–49.
- Schoukens, J., Dobrowiecki, T., & Pintelon, R. (1998). Parametric and non-parametric identification of linear systems in the presence of nonlinear distortions. A frequency domain approach. *IEEE Transactions on Automatic Control*, 43(2), 176–190.
- Schoukens, J., Gommé, L., Van Moer, W., & Rolain, Y. (2008). Identification of a block-structured nonlinear feedback system, applied to a microwave crystal detector. *IEEE Transactions on Instrumentation and Measurement*, 57(8), 1734–1740.
- Schoukens, J., Lataire, J., Pintelon, R., Vandersteen, G., & Dobrowiecki, T. (2009a). Robustness issues of the best linear approximation of a nonlinear system. *IEEE Transactions on Instrumentation and Measurement*, 58(5), 1737–1745.
- Schoukens, J., Pintelon, R., Dobrowiecki, T., & Rolain, Y. (2005). Identification of linear systems with nonlinear distortions. *Automatica*, 41(3), 491–504.
- Schoukens, J., Pintelon, R., & Rolain, Y. (2012a). *Mastering system identification in 100 exercises*. Hoboken, New Jersey: John Wiley & Sons.
- Schoukens, J., Pintelon, R., Rolain, Y., Schoukens, M., Tiels, K., Vanbeylen, L., et al. (2015a). Structure discrimination in block-oriented models using linear approximations: A theoretic framework. *Automatica*, 53, 225–234.
- Schoukens, J., Suykens, J., & Ljung, L. (2009b). Wiener-Hammerstein benchmark. In *15th IFAC symposium on system identification*.
- Schoukens, J., Vaes, M., & Pintelon, R. (2016a). Linear system identification in a nonlinear setting: nonparametric analysis of the nonlinear distortions and their impact on the best linear approximation. *IEEE Control Systems*, 36(3), 38–69.
- Schoukens, J., Widanage, W., Godfrey, K., & Pintelon, R. (2007). Initial estimates for the dynamics of a Hammerstein system. *Automatica*, 43(7), 1296–1301.
- Schoukens, M. (2015). *Identification of parallel block-oriented models starting from the best linear approximation*. Vrije Universiteit Brussel, Belgium.
- Schoukens, M., Bai, E., & Rolain, Y. (2012b). Identification of Hammerstein-Wiener Systems. In *16th IFAC symposium on system identification* (pp. 274–279).
- Schoukens, M., Lyzell, C., & Enqvist, M. (2013). Combining the best linear approximation and dimension reduction to identify the linear blocks of parallel Wiener systems. In *11th IFAC international workshop on adaptation and learning in control and signal processing* (pp. 372–377).
- Schoukens, M., Marconato, A., Pintelon, R., Vandersteen, G., & Rolain, Y. (2015b). Parametric identification of parallel Wiener-Hammerstein systems. *Automatica*, 51(1), 111–122.
- Schoukens, M., Mattson, P., Wigren, T., & Noël, J. (2016b). Cascaded tanks benchmark combining soft and hard nonlinearities. In *Workshop on nonlinear system identification benchmarks* (pp. 20–23) <http://homepages.vub.ac.be/mschouke/benchmarkCascadedTanks.html>.
- Schoukens, M., & Noël, J. (2016). Wiener-Hammerstein benchmark with process noise. In *Workshop on nonlinear system identification benchmarks* (pp. 15–19) <http://homepages.vub.ac.be/mschouke/benchmarkWienerHammerstein.html>.
- Schoukens, M., Pintelon, R., & Rolain, Y. (2011). Parametric identification of parallel Hammerstein systems. *IEEE Transactions on Instrumentation and Measurement*, 60(12), 3931–3938.
- Schoukens, M., Pintelon, R., & Rolain, Y. (2014a). Identification of Wiener-Hammerstein systems by a nonparametric separation of the best linear approximation. *Automatica*, 50(2), 628–634.
- Schoukens, M., & Rolain, Y. (2011). Parametric MIMO parallel Wiener identification. In *50th IEEE conference on decision and control and european control conference* (pp. 5100–5105).
- Schoukens, M., & Rolain, Y. (2012a). Crossterm elimination in parallel Wiener systems using a linear input transformation. *IEEE Transactions on Instrumentation and Measurement*, 61(3), 845–847.
- Schoukens, M., & Rolain, Y. (2012b). Parametric identification of parallel Wiener systems. *IEEE Transactions on Instrumentation and Measurement*, 61(10), 2825–2832.
- Schoukens, M., Vandersteen, G., Rolain, Y., & Ferranti, F. (2014b). Fast identification of Wiener-Hammerstein systems using discrete optimization. *IET Electronics Letters*, 50(25), 1942–1944.
- Sjöberg, J., Lauwers, L., & Schoukens, J. (2012). Identification of Wiener-Hammerstein models: Two algorithms based on the best split of a linear model applied to the SYSID'09 benchmark problem. *Control Engineering Practice*, 20(11), 1119–1125.
- Sjöberg, J., & Schoukens, J. (2012). Initializing Wiener-Hammerstein models based on partitioning of the best linear approximation. *Automatica*, 48(2), 353–359.
- Sobhani Tehrani, E., Golkar, M., Guarín, D., Jaleddini, K., & Kearney, R. (2015). Methods for the identification of time-varying Hammerstein systems with applications to the study of dynamic joint stiffness. In *17th IFAC symposium on system identification* (pp. 1023–1028).
- Söderström, T., & Stoica, P. (1989). *System identification*. Hemel Hempstead, Hertfordshire: Prentice Hall International (UK) Ltd.
- Sou, K., Megretski, A., & Daniel, L. (2008). Convex relaxation approach to the identification of the Wiener-Hammerstein model. In *47th IEEE conference on decision and control* (pp. 1375–1382).
- Stoica, P. (1981). On the convergence of an iterative algorithm used for Hammerstein system identification. *IEEE Transactions on Automatic Control*, 26(4), 967–969.
- Suykens, J., De Moor, B., & Vandewalle, J. (1995). Nonlinear system identification using neural state space models, applicable to robust control design. *International Journal of Control*, 62(1), 129–152.
- Sznaier, M. (2009). Computational complexity analysis of set membership identification of Hammerstein and Wiener systems. *Automatica*, 45(3), 701–705.
- Tan, A., & Godfrey, K. (2002). Identification of Wiener-Hammerstein models using linear interpolation in the frequency domain (LIFRED). *IEEE Transactions on Instrumentation and Measurement*, 51(3), 509–521.
- Tan, A., Wong, H., & Godfrey, K. (2012). Identification of a Wiener-Hammerstein system using an incremental nonlinear optimisation technique. *Control Engineering Practice*, 20(11), 1140–1148.

- Taringou, F., Hammi, O., Srinivasan, B., Malhame, R., & Ghannouchi, F. (2010). Behaviour modelling of wideband RF transmitters using Hammerstein-Wiener models. *IET Circuits Devices & Systems*, 4(4), 282–290.
- Tiels, K., Heuberger, P., & Schoukens, J. (2012). Parameter reduction of MISO Wiener-Schetzen models using the best linear approximation. In *IEEE international instrumentation and measurement technology conference* (pp. 2114–2118).
- Tiels, K., & Schoukens, J. (2013). From coupled to decoupled polynomial representations in parallel Wiener-Hammerstein models. In *52nd IEEE conference on decision and control* (pp. 4937–4942).
- Tiels, K., & Schoukens, J. (2014). Wiener system identification with generalized orthonormal basis functions. *Automatica*, 50(12), 3147–3154.
- Tiels, K., Schoukens, M., & Schoukens, J. (2014). Generation of initial estimates for Wiener-Hammerstein models via basis function expansions. In *19th world congress of the international federation of automatic control* (pp. 481–486).
- Tiels, K., Schoukens, M., & Schoukens, J. (2015). Initial estimates for Wiener-Hammerstein models using phase-coupled multisines. *Automatica*, 60, 201–209.
- Tötterman, S., & Toivonen, H. T. (2009). Support vector method for identification of Wiener models. *Journal of Process Control*, 19, 1174–1181.
- Van Huffel, S., & Lemmerling, P. (2002). *Total least squares and errors-in-variables modeling*. Kluwer Academic Publishers.
- Van Mulders, A., Schoukens, J., & Vanbeylen, L. (2013). Identification of systems with localised nonlinearity: From state-space to block-structured models. *Automatica*, 49(5), 1392–1396.
- Van Mulders, A., Schoukens, J., Volckaert, M., & Diehl, M. (2010). Two nonlinear optimization methods for black box identification compared. *Automatica*, 46(10), 1675–1681.
- Van Pelt, T., & Bernstein, D. (2001). Non-linear system identification using Hammerstein and non-linear feedback models with piecewise linear static maps. *International Journal of Control*, 74(18), 1807–1823.
- Vanbeylen, L. (2013). Nonlinear LFR block-oriented model: Potential benefits and improved, user-friendly identification method. *IEEE Transactions on Instrumentation and Measurement*, 62(12), 3374–3383.
- Vanbeylen, L. (2014). A fractional approach to identify Wiener-Hammerstein systems. *Automatica*, 50(3), 903–909.
- Vanbeylen, L., Pintelon, R., & Schoukens, J. (2008). Blind maximum likelihood identification of Hammerstein systems. *Automatica*, 44(12), 3139–3146.
- Vanbeylen, L., Pintelon, R., & Schoukens, J. (2009). Blind maximum-likelihood identification of Wiener systems. *IEEE Transactions on Signal Processing*, 57(8), 3017–3029.
- Vanbeylen, L., Van Mulders, A., & Abu-Rmieleh, A. (2014). Identification of a nonlinear model for a glucoregulatory benchmark problem. *Biomedical Signal Processing and Control*, 13, 168–173.
- Vandersteen, G., Rolain, Y., & Schoukens, J. (1997). Non-parametric estimation of the frequency-response functions of the linear blocks of a Wiener-Hammerstein model. *Automatica*, 33(7), 1351–1355.
- Vandersteen, G., & Schoukens, J. (1999). Measurement and identification of nonlinear systems consisting of linear dynamic blocks and one static nonlinearity. *IEEE Transactions on Automatic Control*, 44(6), 1266–1271.
- Verhaegen, M., & Westwick, D. (1996). Identifying MIMO Hammerstein systems in the context of subspace model identification methods. *International Journal of Control*, 63(2), 331–349.
- Vörös, J. (2004). An iterative method for Hammerstein-Wiener systems parameter identification. *Journal of Electrical Engineering*, 55(11–12), 328–331.
- Vörös, J. (2007). An iterative method for Wiener-Hammerstein systems parameter identification. *Journal of Electrical Engineering*, 58(2), 114–117.
- Wahlberg, B., Welsh, J., & Ljung, L. (2014). Identification of Wiener systems with process noise is a nonlinear errors-in-variables problem. In *53rd IEEE conference on decision and control* (pp. 3328–3333).
- Wahlberg, B., Welsh, J., & Ljung, L. (2015). Identification of stochastic Wiener systems using indirect inference. In *17th IFAC symposium on system identification* (pp. 620–625).
- Wang, D., & Ding, F. (2008). Extended stochastic gradient identification algorithms for Hammerstein-Wiener ARMAX systems. *Computers and Mathematics with Applications*, 56(12), 3157–3164.
- Wang, Z., Zhang, Z., Mao, J., & Zhou, K. (2012). A Hammerstein-based model for rate-dependent hysteresis in piezoelectric actuator. In *2012 24th chinese control and decision conference* (pp. 1391–1396).
- Weiss, M., Evans, C., & Rees, D. (1998). Identification of nonlinear cascade systems using paired multisine signals. *IEEE Transactions on Instrumentation and Measurement*, 47(1), 332–336.
- Westwick, D., & Kearney, R. (1997). Generalized eigenvector algorithm for nonlinear system identification with non-white inputs. *Annals of Biomedical Engineering*, 25, 802–814.
- Westwick, D., & Kearney, R. (2001). Separable least squares identification of nonlinear Hammerstein models: Application to stretch reflex dynamics. *Annals of Biomedical Engineering*, 29, 707–718.
- Westwick, D., & Kearney, R. (2003). *Identification of nonlinear physiological systems*. Hoboken, NJ: IEEE Press.
- Westwick, D., & Schoukens, J. (2012). Initial estimates of the linear subsystems of Wiener-Hammerstein models. *Automatica*, 48(1), 2931–2936.
- Westwick, D., & Verhaegen, M. (1996). Identifying MIMO Wiener systems using subspace model identification methods. *Signal Processing*, 52(2), 235–258.
- Widanage, W., Stoev, J., Van Mulders, A., Schoukens, J., & Pintel, G. (2011). Nonlinear system-identification of the filling phase of a wet-clutch system. *Control Engineering Practice*, 19(12), 1506–1516.
- Wiener, N. (1958). *Nonlinear problems in random theory*. Wiley.
- Wigren, T. (1993). Recursive prediction error identification using the nonlinear Wiener model. *Automatica*, 29(4), 1011–1029.
- Wigren, T., & Schoukens, J. (2013). Three free data sets for development and benchmarking in nonlinear system identification. In *European control conference* (pp. 2933–2938).
- Wills, A., & Ninness, B. (2012). Generalised Hammerstein-Wiener system estimation and a benchmark application. *Control Engineering Practice*, 20(11), 1097–1108.
- Wills, A., Schön, T., Ljung, L., & Ninness, B. (2013). Identification of Hammerstein-Wiener models. *Automatica*, 49(1), 70–81.
- Wong, H., Schoukens, J., & Godfrey, K. (2012). Analysis of best linear approximation of a Wiener-Hammerstein system for arbitrary amplitude distributions. *IEEE Transactions on Instrumentation and Measurement*, 61(3), 645–654.
- Wysocki, E., & Rugh, W. (1976). Further results on the identification problem for the class of nonlinear systems. *IEEE Transactions on Circuits and Systems*, 23(11), 664–670.
- Yong, A. Y., Tan, A. H., & Cham, C. L. (2015). Identification of block-oriented systems with rate saturation nonlinearity. In *17th IFAC symposium on system identification* (pp. 939–944).
- Zhang, Q., Laurain, V., & Wang, J. (2015). Weighted principal component analysis for Wiener system identification - regularization and non-Gaussian excitations. In *17th IFAC symposium on system identification* (pp. 602–607).
- Zhu, Y. (2002). Estimation of an N-L-N Hammerstein-Wiener model. *Automatica*, 38(9), 1607–1614.



Maarten Schoukens received the master's degree in electrical engineering: electronics and information technology and Ph.D. degree from the Vrije Universiteit Brussel (VUB), Brussels, Belgium, in 2010 and 2015, respectively. Since 2015, he has been a Post-Doctoral Researcher with the ELEC Department, VUB. His current interests include the measurement and identification of nonlinear systems.



Koen Tiels was born in Halle, Belgium, on April 11, 1987. He received the degree of master in Electromechanical Engineering in July 2010 and the degree of Doctor in Engineering (Ph.D.) in March 2015, both from the Vrije Universiteit Brussel (VUB), Brussels, Belgium. In September 2010, he joined the Electrical Engineering (ELEC) Department of the VUB where he is currently a post-doctoral researcher. His main interests are in the field of nonlinear system identification.