Preprints of the 19th World Congress
The International Federation of Automatic Control
Cape Town, South Africa. August 24-29, 2014

Generation of initial estimates for Wiener-Hammerstein models via basis function expansions ⋆

Koen Tiels * **Maarten Schoukens** ** **Johan Schoukens** ***

* *Department ELEC, Vrije Universiteit Brussel, 1050 Brussels, Belgium (e-mail: koen.tiels@vub.ac.be)*
** *Department ELEC, Vrije Universiteit Brussel, 1050 Brussels, Belgium (e-mail: maarten.schoukens@vub.ac.be)*
*** *Department ELEC, Vrije Universiteit Brussel, 1050 Brussels, Belgium (e-mail: johan.schoukens@vub.ac.be)*

Abstract: Block-oriented models are often used to model nonlinear systems. They consist of linear dynamic (L) and nonlinear static (N) sub-blocks. This paper proposes a method to generate initial values for a Wiener-Hammerstein model (LNL cascade). The method starts from the best linear approximation (BLA) of the system, which provides an estimate of the product of the transfer functions of the two linear dynamic sub-blocks. Next, the poles of the BLA are assigned to both linear dynamic sub-blocks. The linear dynamics are then parameterized in terms of rational orthonormal basis functions, while the nonlinear sub-block is parameterized by a polynomial. This allows to reformulate the model to the cascade of a parallel Wiener (with parallel LN structure) and a linear dynamic system, which is bilinear in its parameters. After a bilinear optimization, the parallel Wiener part is projected to a single-branch Wiener model. The approach is illustrated on a simulation example.

Keywords: Dynamic systems; Nonlinear systems; System identification; Wiener-Hammerstein model.

1. INTRODUCTION

Although nonlinear distortions are often present, many dynamical systems can be approximated by a linear model. When the nonlinear distortion level is too high, a linear approximation is insufficient, and a nonlinear model is needed.

One possibility is to use block-oriented models [Billings and Fakhouri, 1982, Giri and Bai, 2010], which are built up by linear dynamic and nonlinear static (memoryless) blocks. Due to this highly structured nature, block-oriented models offer insight about the system to the user. The simplest block-oriented models are the Wiener model (linear dynamic block followed by a nonlinear static block), and the Hammerstein model (linear dynamic block preceded by a nonlinear static block). They can be generalized to a Wiener-Hammerstein model (nonlinear static block sandwiched between two linear dynamic blocks, see Fig. 1).

Several identification methods have been proposed to identify single-branch Wiener-Hammerstein systems. Early work can be found in Billings and Fakhouri [1982] and Korenberg and Hunter [1986]. The maximum likelihood estimate is formulated in Chen and Fassois [1992]. The recursive identification of error-in-variables Wiener-Hammerstein systems is considered in Mu and Chen [2014]. Some other methods start from the best linear approximation (BLA) [Pintelon and Schoukens, 2012] of the Wiener-Hammerstein system [Sjöberg et al., 2012, Westwick and Schoukens, 2012]. These methods will be discussed in more detail in Section 3.

This paper presents a method to generate starting values for single-branch Wiener-Hammerstein systems. The method starts from the BLA of the system. Next, the poles of the BLA are used to construct generalized orthonormal basis functions (GOBFs) [Heuberger et al., 2005] that parameterize both the front and the back dynamics. Using a multivariate polynomial to describe the static nonlinearity, the model is reformulated to the cascade of a parallel Wiener and a linear dynamic system, which is bilinear in its parameters. After a bilinear optimization, the parallel Wiener part is projected to a single-branch Wiener model. This results in the initial estimate of the single-branch Wiener-Hammerstein system.

The rest of this paper is organized as follows. The basic setup is described in Section 2. Section 3 gives a brief overview of the BLA, and discusses three related identification methods. Section 4 presents the proposed approach, which is illustrated on a simulation example in Section 5. Finally, the conclusions are drawn in Section 6.

⋆ This work was supported in part by the Fund for Scientific Research (FWO-Vlaanderen), by the Flemish Government (Methusalem), the Belgian Government through the Interuniversity Attraction Poles (IAP VII) Programme, and by the ERC advanced grant SNLSID, under contract 320378.

2. PROBLEM STATEMENT

2.1 Setup

Consider the Wiener-Hammerstein system in Fig. 1, given by

$$\begin{aligned} x(t) &= R(q)u(t) \\ w(t) &= f(x(t)) \\ y(t) &= S(q)w(t) + v(t) \end{aligned} \quad , \tag{1}$$

where $R(q)$ and $S(q)$ are linear time-invariant (LTI) discrete-time transfer functions in the backward shift operator q^{-1} ($q^{-1}u(t) = u(t-1)$), i.e.

$$\begin{aligned} R(q) &= \frac{B_R(q)}{A_R(q)} = \frac{\sum_{l=0}^{n_R} b_{R,l} q^{-l}}{\sum_{l=0}^{m_R} a_{R,l} q^{-l}} \\ S(q) &= \frac{B_S(q)}{A_S(q)} = \frac{\sum_{l=0}^{n_S} b_{S,l} q^{-l}}{\sum_{l=0}^{m_S} a_{S,l} q^{-l}} \end{aligned} \quad , \tag{2}$$

where $f(x)$ is a static nonlinear function, and where $v(t)$ is additive output noise.

2.2 Assumptions

It is assumed that

(1) both $R(q)$ and $S(q)$ are proper, i.e. $n_R \leq m_R$, and $n_S \leq m_S$,
(2) there are no pole-zero cancellations in the product $R(q)S(q)$,
(3) $f(x)$ is non-even around the operating point,
(4) the input signal $u(t)$ has a Gaussian amplitude distribution, and
(5) the output noise $v(t)$ is a zero-mean filtered white noise that is independent of the input signal $u(t)$.

The reason for Assumption 4 is to obtain a good estimate of the product of the underlying dynamics $R(q)$ and $S(q)$ in (3). If Assumption 4 does not hold, a model error is made in (3) that drops rapidly with the length of the impulse response of $R(q)$ [Wong et al., 2012, Tiels and Schoukens, 2011].

2.3 Problem statement

The problem addressed in this paper is the following. Given a data sequence $\{u(t), y(t)\}$ for $t = 0, \ldots, N-1$, find initial estimates $\hat{R}(q)$, $\hat{f}(x)$, $\hat{S}(q)$ such that the simulated output $\hat{y}(t) = \hat{S}(q)\hat{f}\left(\hat{R}(q)u(t)\right)$ is close to $y(t)$ in mean-square sense.

Remark 1. From only input/output data, the linear dynamics and the static nonlinearity can only be estimated up to arbitrary non-zero scaling factors that can be exchanged between the linear dynamics and the static nonlinearity without affecting the input/output behavior, i.e. $\hat{S}(q)\hat{f}\left(\hat{R}(q)u(t)\right) = \left[\eta\hat{S}(q)\right] \frac{1}{\eta}\hat{f}\left(\frac{1}{\zeta}\left[\zeta\hat{R}(q)\right]u(t)\right)$.

3. THE BEST LINEAR APPROXIMATION OF A WIENER-HAMMERSTEIN SYSTEM

3.1 The best linear approximation

The BLA of a system is defined as the linear system whose output approximates the system's output best in mean-square sense [Pintelon and Schoukens, 2012]. Due to Bussgang's theorem [Bussgang, 1952], for a Gaussian excitation $u(t)$, the BLA of the considered Wiener-Hammerstein system is equal to

$$G_{BLA}(k) = cR(k)S(k) \quad , \tag{3}$$

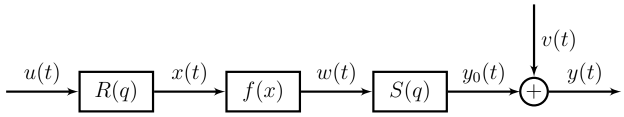

Fig. 1. A Wiener-Hammerstein system (R and S are linear dynamic systems and f is a nonlinear static system).

with c a constant depending on the static nonlinear function $f(x)$ and the power spectrum of the Gaussian excitation $u(t)$. This constant is non-zero under Assumption 3.

Under Assumption 2, it follows from (3) that the poles (zeros) of the BLA are equal to the poles (zeros) of both R and S. To obtain initial estimates for R and S, the poles and zeros of the BLA should be split over the individual transfer functions R and S.

3.2 Related initialization methods for Wiener-Hammerstein systems

Several methods have been proposed to make this split. Here we briefly discuss three of them, namely the brute-force and the advanced method in Sjöberg et al. [2012], and the QBLA method in Westwick and Schoukens [2012].

The brute-force method in Sjöberg et al. [2012] scans all possible splits. For each of these splits, the static nonlinearity is estimated via a linear least-squares regression. The obtained initial models are then tested on the data, and the best performing model is retained for further optimization. The drawback of this method is that the number of possible splits grows exponentially in the model order. This method can thus require a large computation time.

The advanced method in Sjöberg et al. [2012] uses a basis function expansion for R, based on the poles of the BLA, and a basis function expansion for the inverse of S, based on the zeros of the BLA. Like this, the poles of $\hat{R}$ and the zeros of $\hat{S}$ are fixed to those of the BLA. Hence, the model order of $\hat{R}$ and $\hat{S}$ is too large. By expressing the static nonlinearity in terms of two multivariate polynomials, the estimation of the remaining model parameters (the polynomial coefficients) is formulated linearly-in-the-parameters. Next, the model orders of $\hat{R}$ and $\hat{S}$ are reduced by performing several scans. In each scan, the effect of removing one basis function is verified, and the best performing model in terms of rms error is retained as an initial model. After each scan, one basis function is permanently removed.The initial models are then ranked with respect to their rms error. Typically, the rms error makes a strong jump when a necessary basis function was removed.

The method described in Westwick and Schoukens [2012] not only uses the BLA from the input $u(t)$ to the output $y(t)$, but also the so-called quadratic BLA (QBLA), which is a higher order BLA from the squared input $u^2(t)$ to the output residual $y_s(t) = y(t) - G_{BLA}(q)u(t)$. It is shown that the poles and zeros of the first linear dynamic system $R(q)$ shift in this QBLA, while the poles and zeros of the second linear dynamic system $S(q)$ remain invariant. This allows to spit the poles and zeros of the BLA over R and S. Due to the higher order nature of the QBLA, however, the estimation of the QBLA is difficult. A long measurement time is needed to obtain an accurate estimate.

4. THE PROPOSED METHOD

The proposed method is related to the advanced method in Sjöberg et al. [2012]. The main differences are the use of a basis function expansion for S, based on the poles of the BLA, rather than a basis function expansion for S^{-1}, based on the zeros of the BLA, and the use of only one MIMO (multiple input multiple output) polynomial to describe the static nonlinearity instead of two MISO (multiple input single output) polynomials. Like this, the proposed model structure is able to describe parallel Wiener-Hammerstein systems as well, but for now, the focus will be on single-branch Wiener-Hammerstein systems.

4.1 Basic idea

The basic idea is to use basis function expansions for R, f, and S

$$\hat{R}(q) = \sum_{j=1}^{n_\beta} \beta_{R,j} G_j(q) \quad , \tag{4}$$

$$\hat{f}(x) = \sum_{i=0}^{D} \gamma_i x^i \quad , \tag{5}$$

$$\hat{S}(q) = \sum_{j=1}^{n_\beta} \beta_{S,j} G_j(q) \quad , \tag{6}$$

where $\{G_1(q), \ldots, G_{n_\beta - 1}(q)\}$ are generalized orthonormal basis functions (GOBFs) [Heuberger et al., 2005] based on the poles of the BLA. One extra OBF, namely $G_{n_\beta} = 1$, is used that enables the estimation of a feed-through term and as such also enables the estimation of static systems [Tiels and Schoukens, 2011]. The basis function expansion for the static nonlinearity is here assumed polynomial for simplicity reasons. As a polynomial basis function expansion is used, the Wiener part described by (4) and (5) can be rewritten as a parallel Wiener system that is linear in its parameters (see e.g. Tiels and Schoukens [2011]). This results in the model structure shown in Fig. 2 that is described by

$$\begin{aligned}
p_j(t) &= G_j(q)u(t), & j &= 1, \ldots, n_\beta \\
z_j(t) &= g^{[j]}(p_1(t), \ldots, p_{n_\beta}(t)), & j &= 1, \ldots, n_\alpha \\
z(t) &= \sum_{j=1}^{n_\alpha} \alpha_j z_j(t) & & \\
\hat{y}(t) &= \sum_{j=1}^{n_\beta} \beta_j G_j(q) z(t) + \hat{y}_{DC} & &
\end{aligned} \quad , \tag{7}$$

where each $g^{[j]}(p_1, \ldots, p_{n_\beta})$ is a monomial of the multiple input multiple output (MIMO) polynomial $g(p_1, \ldots, p_{n_\beta})$ with corresponding coefficient α_j. The constant term that corresponds to γ_0 in (5) is not included in g, but is estimated at the output of the model ($\hat{y}_{DC}$ in Fig. 2). In principle, the MIMO static nonlinear mapping could be described by any basis function expansion that is linear-in-the-parameters, but this is out of the scope of this paper.

First, the polynomial coefficients α and the coefficients $\beta \triangleq \beta_S$ of the basis function expansion of $\hat{S}$ will be estimated. Next, the parallel Wiener part between $u(t)$ and $z(t)$ will be projected to a single-branch Wiener system.

4.2 Estimation of β and α

In the case that $S(q)$ can be exactly written as a linear combination of the basis functions $\{G_1(q), \ldots, G_{n_\beta}(q)\}$ (the true poles are known), the output spectrum at frequency k is given by

$$\begin{aligned}
Y(k) &= S(k)Z(k) \\
&= \begin{bmatrix} G_1(k) & \cdots & G_{n_\beta}(k) \end{bmatrix} \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_{n_\beta} \end{bmatrix} \begin{bmatrix} \alpha_1 & \cdots & \alpha_{n_\alpha} \end{bmatrix} \begin{bmatrix} Z_1(k) \\ \vdots \\ Z_{n_\alpha}(k) \end{bmatrix} \\
&= \mathcal{G}(k) \theta \mathcal{Z}^T(k)
\end{aligned} \tag{8}$$

Applying the "vec" operation on the last equation results in

$$\begin{aligned}
vec\,(Y(k)) &= Y(k) \\
&= (\mathcal{Z}(k) \otimes \mathcal{G}(k))\, vec(\theta) \\
&= \mathcal{K}(k) vec(\theta)
\end{aligned} \quad . \tag{9}$$

Note that $\mathcal{K}(k)$ is a row vector. Collecting (9) at all measured frequencies k results in

$$\mathcal{Y} = \mathcal{K} vec(\theta) \quad . \tag{10}$$

The matrix $\mathcal{K}$ can be easily obtained as

$$\mathcal{K} = (\mathcal{Z} \otimes J_{1,n_\beta}) \circ (J_{1,n_\alpha} \otimes \mathcal{G}) \quad , \tag{11}$$

where $J_{n,m}$ is an n by m matrix containing all ones, and $\circ$ is the Hadamard product (element-wise matrix product).

In a first step, $vec(\theta)$ in (10) is estimated via a linear least-squares approach. Next, the singular value decomposition (SVD) of the obtained parameter matrix θ_{LS} in (8) is taken, and is truncated to its first term to obtain estimates for α and β. This corresponds to the over-parametrization method in Bai and Liu [2005].

From numerical simulations, we have observed that the matrix $\mathcal{K} \in \mathbb{C}^{N \times n_\beta n_\alpha}$ has rank $n_1(n_\beta - 1) + n_\alpha$, where $n_1 = n_\alpha - \frac{(D+n_\beta-1)!}{D!(n_\beta-1)!} + 1$. The matrix $\mathcal{K}$ is thus rank-deficient as soon as $n_\beta > 1$. Therefore, the over-parameterization method is not convergent in this case, as the first step does not converge to the true $\theta = \beta\alpha^T$. Let $vec(\theta^{(i)})$ be in the null space of $\mathcal{K}$. Then θ_{LS} is just one of the infinitely many solutions $\theta_{all} = \theta_{LS} + \sum_{i=1}^{corank(\mathcal{K})} \lambda_i \theta^{(i)}$ to the least-squares problem in (10). To find the true θ, one would need to find the λ_i's such that θ_{all} is of minimal rank. This is known as the MinRank problem [Faugère et al., 2008], which is hard to solve. Nevertheless, the over-parameterization method can be used to obtain starting values for the coefficients α and β.

It should also be noted that the model in (7) is too complex, since both the parameterizations of $\hat{R}$ and $\hat{S}$ in (4) and (6) use the pole estimates of both R and S. To reduce the complexity of the model, the user can decide to start a simplified scanning procedure. In each scan, one

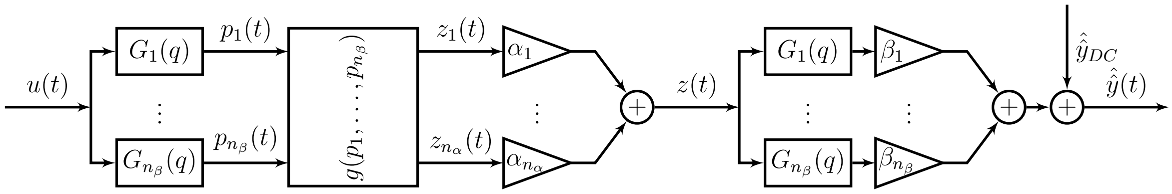

Fig. 2. Resulting intermediate model structure (G_1 and G_2 are basis function expansions, based on the poles of the BLA; $g(p_1,\ldots,p_{n_\beta})$ is a MIMO polynomial).

pole (or complex conjugate pole pair for complex poles) is removed from the full-complexity model. The least-squares estimate in (10) is calculated for the reduced-complexity model. If the rms error on the simulated output strongly deteriorates, the pole (or pole pair) is kept, otherwise, it is removed from the final model. After scanning, the over-parameterization method is applied to the final model. This scanning procedure requires at most $2n_\beta$ scans. Compared to the scanning procedures presented in Sjöberg et al. [2012], where the number of scans grows either exponentially (if no restrictions on properness of the subsystems, etc. are imposed) or combinatorially with the model order, this scanning procedure requires a number of scans that is proportional with the model order.

Next, the estimates of the coefficients α and β are optimized using a normalized iterative least-squares approach [Bai and Liu, 2005]. In each iteration, either the coefficients α or β are estimated via a linear least-squares approach, while the other set of coefficients remains constant. The norm of β is normalized to one in each iteration.

After this step, the estimate of $S(q)$ in (6) is available. To obtain estimates of $R(q)$ and $f(x)$, we need to project the parallel Wiener system between $u(t)$ and $z(t)$ to a single-branch Wiener system.

4.3 Returning to a single-branch model

First, the intermediate signals $p_j(t)$ and $z(t)$ in (7) are simulated using the measured input signal $u(t)$ and the estimated model. The simulated intermediate signal $z(t)$ is an estimate of $w(t)$ in the Wiener-Hammerstein system shown in Fig. 1. The BLA from $u(t)$ to $w(t)$ is equal to $R(q)$, up to an unknown scale factor (special case of (3)). Therefore, an estimate of $R(q)$ can be obtained by estimating the BLA from $u(t)$ to $z(t)$. The nonparametric and parametric estimation of this BLA are combined in one step as

$$\hat{\beta}_R = \underset{\beta_R}{\arg\min} \left\| z(t) - \sum_{j=1}^{n_\beta} \beta_{R,j} p_j(t) \right\|_2 \quad , \qquad (12)$$

where the parametrization in (4) is used for the parametric estimation of the BLA.

Approximate pole-zero cancellations in the estimates of $R(q)$ and $S(q)$ are removed.

Next, the polynomial coefficients in (5) are estimated using a linear least-squares approach.

Finally, the parameters of the single-branch Wiener-Hammerstein model (its transfer function and polynomial coefficients) can be further optimized using a Levenberg-Marquardt nonlinear optimization algorithm [Marquardt, 1963].

5. SIMULATION EXAMPLE

This section illustrates the approach on a simulation example.

5.1 Setup

The linear dynamic systems $R(q)$ and $S(q)$ are second-order Chebyshev filters, with a ripple of 10 and 20 dB, and a 3 dB bandwidth of $0.05f_s$ and $0.1f_s$ respectively. The sample frequency f_s is normalized to 1. The static nonlinearity is given by $f(x) = \mathrm{atan}(2x)$.

The system is excited with a random-phase multisine [Pintelon and Schoukens, 2012]

$$u(t) = A \sum_{k=1}^{N/6} cos(2\pi k \frac{f_s}{N} t + \phi_k) \qquad (13)$$

containing $N = 1024$ samples. The amplitude A is chosen such that $u(t)$ has rms value 1. The phases ϕ_k are independently uniformly distributed in the interval $[0, 2\pi[$. Seven phase realizations and three periods of the multisine are applied. The first period is removed to remove the influence of the transients. A zero-mean white Gaussian disturbance $v(t)$ is added to the output, with a signal-to-noise ratio $SNR_y = 60$ dB.

5.2 Model estimation and results

First, the BLA of the system is estimated nonparametrically

$$\hat{G}_{BLA}(k) = \frac{1}{M} \sum_{m=1}^{M} \frac{\frac{1}{P}\sum_{p=1}^{P} Y^{[m,p]}(k)}{U^{[m]}(k)} \quad , \qquad (14)$$

where $Y^{[m,p]}(k)$ is the DFT (discrete Fourier transform) spectrum of the output corresponding to the m^{th} realization and the p^{th} period of the input signal, and $U^{[m]}(k)$ is the DFT spectrum of the m^{th} realization of the input. Next, a parametric transfer function model

$$\frac{B_{BLA}(k)}{A_{BLA}(k)} = \frac{\sum_{l=0}^{n_{BLA}} b_{BLA,l} e^{-j\omega_k l}}{\sum_{l=0}^{m_{BLA}} a_{BLA,l} e^{-j\omega_k l}} \qquad (15)$$

is estimated on the nonparametric BLA estimate, with $n_{BLA} = m_{BLA} = 4$. The roots of $A_{BLA}(k)$ are calculated and used to construct the GOBFs. The nonlinearity is modeled using a low-degree ($D = 3$) MIMO polynomial g to limit the number of parameters α.

Starting values for the coefficients α and β are obtained using the over-parameterization method (see Section 4.2).

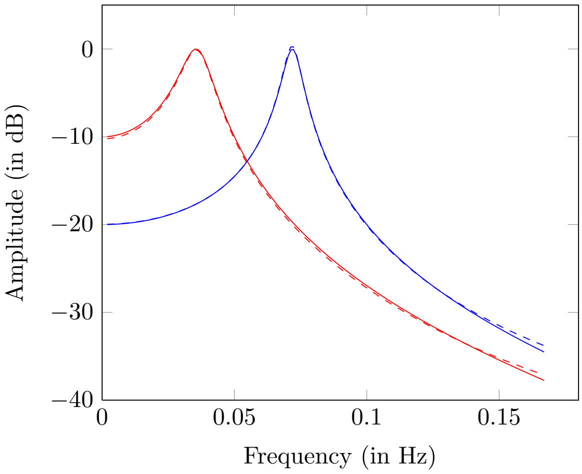

Fig. 3. True (full line) and estimated (dashed line) linear dynamics (R in red and S in blue).

These coefficients are then bilinearly optimized (see Section 4.2). Finally, the initial estimates for $R(q)$ and $f(x)$ are obtained as described in Section 4.3. In this step, the static nonlinearity is estimated with a higher-degree polynomial ($D = 5$), as the number of polynomial coefficients only increases proportionally with the degree as opposed to the combinatorial increase for the MIMO polynomial g.

Fig. 3 shows the true and estimated linear dynamics, while Fig. 4 shows the true and estimated static nonlinearity. The estimates are normalized, such that the estimated linear dynamic system $\hat{R}$ and the estimated static nonlinearity $\hat{f}$ match their true counterparts as well as possible in mean-square sense. The remaining normalization factor is taken into account in the estimate of $S(q)$ (see Remark 1). It can be observed that the estimates agree well with the true dynamics and nonlinearity.

To see the effect of the scanning procedure described in Section 4.2, the model estimation is repeated, this time also including the scanning procedure. Table 1 shows the results of the scanning procedure. It can be deduced that the estimated pole pair $0.8811 \pm 0.4275j$ should be assigned to the back dynamics, while the pole pair $0.9404 \pm 0.2163j$ should be assigned to the front dynamics. This is in good agreement with the true dynamics. Whereas $R(q)$ has a complex pole pair $0.9397 \pm 0.2162j$, $S(q)$ has a pole pair $0.8800 \pm 0.4275j$. Figs. 5 and 6 show the true and estimated linear dynamics and nonlinearity, respectively. Again, good initial estimates are obtained, but they are not as good as when the scanning procedure was not applied. Fig. 5 only shows amplitude information, but a phase error in the estimate of $S(q)$ is present as well. Although the initial estimates are not as good, the scanning procedure allowed us to reduce the complexity of the model before the bilinear optimization. Moreover, both initial estimates converge to similar estimates after a Levenberg-Marquardt optimization [Marquardt, 1963] (see Figs. 7 and 8).

6. CONCLUSION

An attempt was made to use generalized orthonormal basis functions for Wiener-Hammerstein system identification.

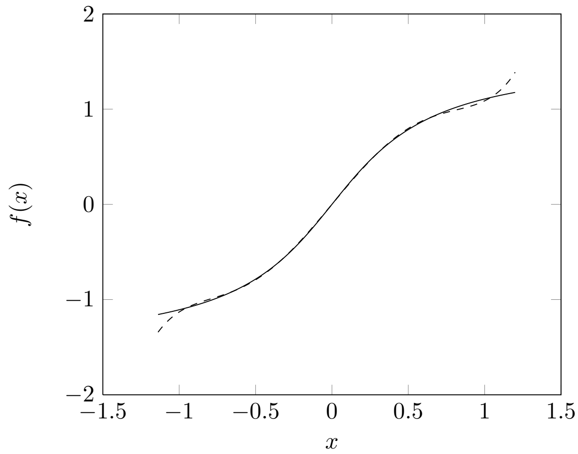

Fig. 4. True (full line) and estimated (dashed line) static nonlinearity.

Table 1. Normalized rms error (NRMSE) on the simulated output in dB when a complex conjugate pole pair is removed from the initial model in (10). The NRMSE for the initial model is equal to −24.9 dB.

pole estimates	removal from	
	front dynamics	back dynamics
0.8811±0.4275j	−24.4 dB	−14.9 dB
0.9404±0.2163j	−13.7 dB	−24.8 dB

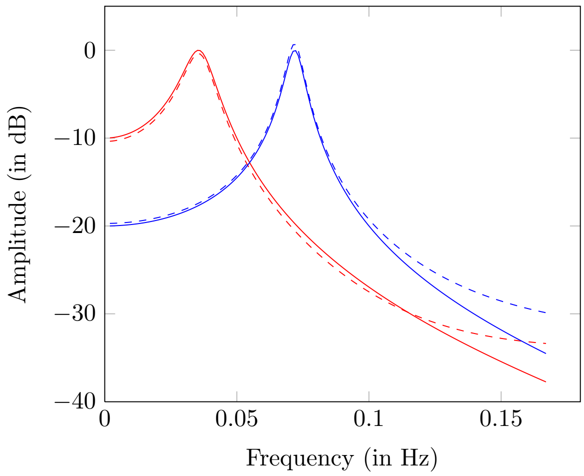

Fig. 5. True (full line) and estimated (dashed line) linear dynamics (R in red and S in blue when the scanning procedure is included).

The system was reformulated to the cascade of a parallel Wiener and linear dynamic system, which is bilinear in its parameters. After a bilinear optimization of the model parameters, the parallel Wiener part was projected to a single-branch Wiener system, resulting in an initial estimate of the Wiener-Hammerstein system.

The method, as presented in this paper, still has problems with high complexity systems. If the linear dynamics are high-order, the proposed method has problems to converge. The scanning procedure could however be simplified. Instead of an exponential or combinatorial increase

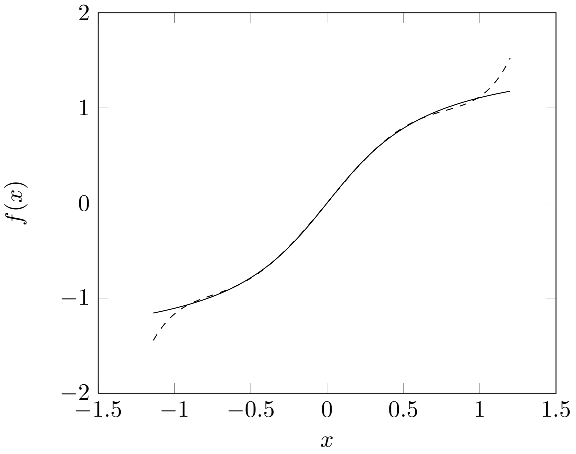

Fig. 6. True (full line) and estimated (dashed line) static nonlinearity when the scanning procedure is included.

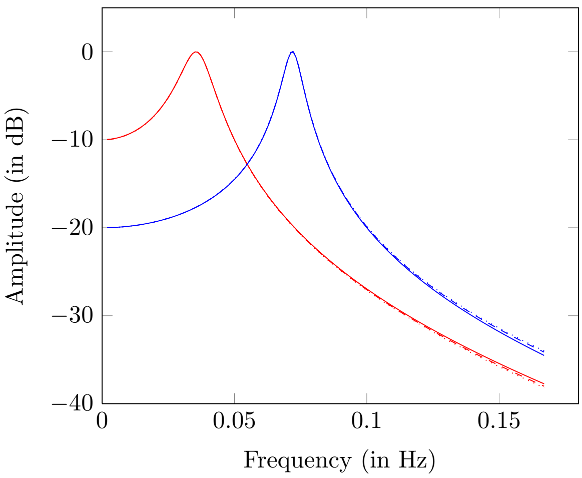

Fig. 7. True (full line) and estimated (dashed (no scanning) and dotted (scanning) line) linear dynamics (R in red and S in blue) after optimization.

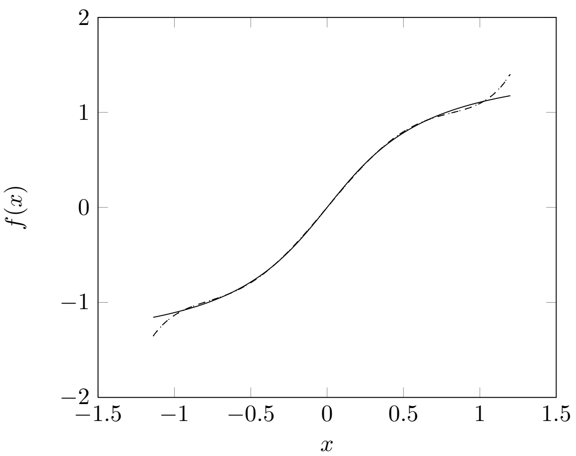

Fig. 8. True (full line) and estimated (dashed (no scanning) and dotted (scanning) line) static nonlinearity after optimization.

of the number of scans with the model order, the number of scans increases proportionally with the model order.

ACKNOWLEDGMENT

The authors would like to thank Mariya Ishteva for her helpful comments and suggestions.

REFERENCES

E.-W. Bai and Y. Liu. On the least squares solutions of a system of bilinear equations. In *Proceedings of the* 44^{th} *IEEE Conference on Decision and Control, and the European Control Conference 2005*, pages 1197–1202, 2005.

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

J. J. Bussgang. Crosscorrelation functions of amplitude-distorted Gaussian signals. Technical Report 216., MIT research laboratory of electronics, 1952.

C. H. Chen and S. D. Fassois. Maximum likelihood identification of stochastic Wiener-Hammerstein-type non-linear systems. *Mechanical Systems and Signal Processing*, 6(2):135–153, 1992.

J.-C. Faugère, F. Lévy-dit Vehel, and L. Perret. Cryptanalysis of MinRank. In *CRYPTO 2008 Proceedings of the* 28^{th} *Annual conference on Cryptology: Advances in Cryptology*, pages 280–296, 2008.

F. Giri and E.-W. Bai, editors. *Block-oriented Nonlinear System Identification*. Springer, 1 edition, 2010.

P. S. C. Heuberger, P. M. J. Van den Hof, and B. Wahlberg, editors. *Modelling and Identification with Rational Orthogonal Basis Functions*. Springer, London, 2005.

M. J. Korenberg and I. W. Hunter. The identification of nonlinear biological systems: LNL cascade models. *Biol. Cybern.*, 55:125–134, 1986.

D. Marquardt. An algorithm for least-squares estimation of nonlinear parameters. *SIAM Journal on Applied Mathematics*, 11(2):431–441, 1963.

B. Q. Mu and H.-F. Chen. Recursive identification of errors-in-variables Wiener-Hammerstein systems. *European Journal of Control*, 20:14–23, 2014.

R. Pintelon and J. Schoukens. *System Identification: A Frequency Domain Approach*. Wiley-IEEE Press, 2 edition, 2012.

J. Sjöberg, L. Lauwers, and J. Schoukens. 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: 1119–1125, 2012.

K. Tiels and J. Schoukens. Identifying a Wiener system using a variant of the Wiener G-Functionals. In 50^{th} *IEEE Conference on Decision and Control and European Control Conference (CDC-ECC11), Orlando, FL, USA*, December 2011.

D. T. Westwick and J. Schoukens. Initial estimates of the linear subsystems of Wiener-Hammerstein models. *Automatica*, 48:2931–2936, 2012.

H. K. Wong, J. Schoukens, and K. R. Godfrey. Analysis of best linear approximation of a Wiener-Hammerstein system for arbitrary amplitude distributions. *IEEE Trans. Instrum. Meas.*, 61(3):645–654, March 2012.