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

Generating initial estimates for Wiener-Hammerstein systems using phase coupled multisines☆

J. Schoukens[a], K. Tiels[a], M. Schoukens[a]

[a] *Vrije Universiteit Brussel, Department ELEC. Vrije Universiteit Brussel, Department ELEC.*

Abstract: Block oriented nonlinear models capture the dynamics of a nonlinear system with linear dynamic sub-systems (L), the nonlinear behavior is modelled using static nonlinear sub-blocks (N). In this paper we study the generation of initial estimates for the linear dynamic blocks of a Wiener-Hammerstein system that has a cascaded LNL structure. While it is very easy to identify the product of the transfer functions of the first and last dynamic block using linear system identification methods, it turns out to be very difficult to split the global dynamics over these individual blocks. In this paper a method is proposed that allows the poles of the best linear approximation to be assigned to the first or second linear block. Once this split is made, it is shown in the literature that the remaining initialization problem can be solved much easier than the original one. The first step of the method is the design of a special random phase multisine excitation, using pair-wise coupled random phases. Next, a modfied best linear approximation will be estimated on a shifted frequency grid. It will be shown that this procedure shifts the poles and zeros of the first linear sub-block with a known frequency offset, while those of the second sub-block are not changed. The shifted poles and zeros result in a transfer function with complex coefficients that can be identified using a modified frequency domain estimation method. This results in a simple initialization method, based on a linear system identification step.

Keywords: block oriented nonlinear system, Wiener-Hammerstein model, system identification

1. INTRODUCTION

Nonlinear system identification is much more involved than linear system identification. One of the major issues is the selection of a good model structure. Typical examples are nonlinear state space models or nonlinear ARX (NARX) and ARMAX (NARMAX) models that are well suited to capture the behavior of a dynamic nonlinear systems (Billings, 2013). Many successful applications are described. However, none of the above mentioned methods do perfectly match the needs of the design- and control engineers: typically a (very) large number of model parameters is used, and the models provide very little structural insight into the system behavior, all delayed inputs and outputs are nonlinearly combined. Moreover, the number of possible combinations of parameters grows very fast with the degree of the nonlinearity and the number of taps in the filters. Alternatively, block oriented nonlinear models like those shown in Figure 1 can be used (see also Billings and Fakhouri, 1982). These capture the dynamics of the system using linear dynamic sub-systems (L), while the nonlinear behavior is modelled using static nonlinear sub-blocks (N). This idea matches also with the observation that in many systems, the nonlinearity is localised at a few places in the system, embedded in the remaining linear dynamics. Although the identification of block oriented model structures is a hot topic, the actual state of the art is still struggling with very simple structures: most ($> 90\%$) of the recent publications on block oriented systems still deal with single branch structures consisting of sandwich systems like Wiener (LN), Hammerstein (NL), Wiener-Hammerstein (LNL), and Hammerstein-Wiener (NLN) as shown in Figure 1: a,b,e,f. In the recent edited book of Giri and Bai (2010), none of the 24 contributions was considering more complex systems, while it is known for a long time that structures with parallel branches of LNL systems (see for example Figure 1 h) are strongly needed to approximate a wide class of real-life nonlinear systems with a small(er) number of branches [Pal1979]. Some early attempts to identify parallel structures are reported in Billings and Fakhouri (1982), Hunter and Korenberg (1986), Korenberg (1991). Recently, the effort to identify parallel Hammerstein or Wiener systems (Figure 1 c) is strongly increased because these model structures are nowadays popular in the telecommunication field to linearise power amplifiers. Little or no information is available to identify parallel Hammerstein-Wiener or parallel

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

Email address: johan.schoukens@vub.ac.be (J. Schoukens)

Wiener-Hammerstein, and nonlinear feedback structures as shown in Figure 1 g, h, i) (Schoukens and Rolain, 2012) are hardly discussed in the literature.

The major difficulty in block-oriented identification is the generation of good starting values for the dynamics of the linear blocks, even for the single branch WH-model. Early attemps were published by (Vandersteen and Schoukens, 1999) using a series of very specific experiments. Also in (Haber and Keviczky, 1999), a number of methods is presented to separate the dynamics of the linear blocks, but in each of these methods, a set of nonlinear equations need to be solved. This raises again the problem of finding good initial values to start a numerical search procedure. Recently, it was shown that WH-systems could be modelled as a cascade of well selected Hammerstein-Wiener systems (Wills and Ninness, 2012). Other attempts started from the best linear approximation (BLA) of the nonlinear system, and next the poles and zeros are assigned to the first or second dynamic block of the system using, for example, a brute force scanning method by trying all possible combinations (Sjöberg and Schoukens, 2012).

An attempt to split the poles, using a more systematic procedure is given by Westwick and Schoukens (2012), using a higher order BLA based on the squared or cubed input. It is shown that the poles p_i of the first linear system will shift in this step to $2p_i$ or $3p_i$, while those of the second system remain invariant. This provides a tool to separate both sets. However, due to the higher order nature of the BLA, very long measurements are needed in order to get a sufficient precision. In this paper we will develop a similar approach, but using again the first order BLA in stead of the higher order BLA. Using a well designed excitation signal, we create again a shift of the sytem poles. Because we make no use of higher order BLA's, we can avoid the use of extremely long experiments.

We first will give a formal setup of the problem, followed by an analysis of the best linear approximation for a WH-system using a newly proposed class of excitation signals: the phase coupled multisines. Eventually, some simulation results are shown, followed by experimental results.

2. THE BLA OF A WH-SYSTEM USING RANDOM PHASE MULTISINES

In this section we give a brief introduction to the theory of the best linear approximation of a nonlinear system. We first define the class of systems, the class of excitation signals, and introduce formally the concept of the best linear approximation. Next we deliver explicit expressions for G_{BLA} for a WH-system.

2.1. System

In this paper we focus on a Wiener-Hammerstein single branch block-oriented system as given in Figure 2. It consists of a static nonlinear function f , that is acting

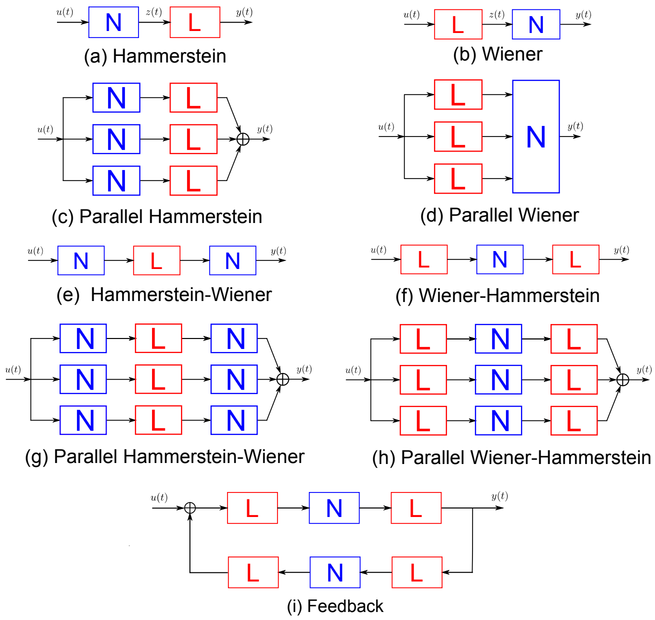

Figure 1: Examples of block-oriented nonlinear model structures.

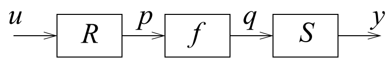

Figure 2: Wiener-Hammerstein system.

on the output of the linear dynamic system R. Its output is passed through the second linear dynamic system S.

In this paper, we consider, without loss of generality, discrete time systems. All results are also valid for continuous time systems. Starting from the measured input and output signal $u(t), y(t)$, with $t = 0, 1, \cdots, N-1$, we need to identify the linear dynamics R, S and the static nonlinearity f. The paper is completely focused on the generation of good initial estimates for R, S. For the moment we assume that there is no disturbing noise, all signals are exactly known. Adding disturbing noise to the output, that is independent of the input, will not change the conclusions of this paper since it is known that the classical least squares framework results in consistent estimates of the BLA under these conditions (Pintelon and Schoukens, 2012).

Define $Y(k), U(k)$ as the discrete Fourier transforms of $u(t), y(t)$, evaluated at the frequencies $k\frac{2\pi}{N}$. The analytic relation between Y, U for a Wiener-Hammerstein system is exactly known for polynomial nonlinearities, for example for a cubic system ($f(p) = p^3$), we have that (Pintelon and Schoukens, 2012):

$$Y(k) = \sum_{l_1=-N/2}^{N/2-1} \sum_{l_2=-N/2}^{N/2-1} \cdots \\ S(k)R(k-l_1-l_2)R(l_1)R(l_2)U(k-l_1-l_2)U(l_1)U(l_2) \tag{1}$$

In this expression we neglected the finite length effects (initial transient in the time domain, leakage in the frequency domain) without loss of generality. This will be done so in the rest of this paper. An alternative expression is

$$Y(k) = \sum_{l_1,l_2,l_3 s.t. l_1+l_2+l_3=k} \cdots \\ S(l_1+l_2+l_3)R(l_1)R(l_2)R(l_3)U(l_1)U(l_2)U(l_3) \quad (2)$$

This expression will be very convenient to find analytic expressions for the BLA.

2.2. The excitation signal

In this paper we consider Gaussian distributed excitation signals. This will lead to very simple expressions for the BLA in the next section. We restrict the class of excitations signals even more by focusing on random phase multisines:

$$u(t) = \sum_{k=-N/2+1}^{N/2-1} U_k e^{j2\pi kt/N} \quad (3)$$

where the Fourier coefficients U_k are either zero (the harmonic is not excited) or satisfy $|U_k| = \hat{U}(k/N)/\sqrt{N}$, with $S_{\hat{U}\hat{U}}(f) = \hat{U}^2(f)$, a uniformly bounded function on $[-1/2, 1/2]$ (the full unit circle). The phases $\varphi_k = \angle U_k = -\angle U_{-k}$ are i.i.d. with a uniform distribution on $[0, 2\pi[$.

2.3. Best linear approximation of a WH-system for random phase multisines

The method that is proposed to find initial estimates R, S relies strongly on the best linear approximation (BLA) of the nonlinear system. This is defined as:

$$g_{BLA} = \underset{g_{BLA}}{\arg\min}\ E_u(y - g_{BLA} * u)^2 \quad (4)$$

in the time domain (Enqvist and Ljung, 2005), or

$$G_{BLA} = \underset{G_{BLA}}{\arg\min}\ E_u|Y - G_{BLA}U|^2 \quad (5)$$

in the frequency domain (Pintelon and Schoukens, 2012). In these expressions, g_{BLA} is the impulse response of the best linear approximation, and G_{BLA} is the frequency response function (FRF) of the best linear approximation. The output of a nonlinear system can always be written as

$$Y(k) = G_{BLA}(k)U(k) + Y_S(k) \quad (6)$$

The first term describes that part of the output that is coherent with the input, the second part $Y_S(k)$ describes the non-coherent part. We will call it the nonlinear noise. If the nonlinear system is excited by Gaussian noise or random phase multisines, it is shown that $Y_S(l)$ is asymptotically independent of the coherent part $G_{BLA}(k)U(k)$, for all values of l, including $l = k$. The dependency converges to zero as an $O(N^{-1})$ (Pintelon and Schoukens, 2012). This results does not hold in the time domain, in that case there can be a strong dependency between $y_S(t)$ and $u(t)$ (the most simple example is the linear approximation of a static nonlinear system).

The coherent part describes the correlation between $Y(k)$ and $U(k)$. It is easy to show that

$$G_{BLA}(k) = S_{YU}(k)/S_{UU}(k) \quad (7)$$

where the expected value in the cross- and auto-correlation is calculated over the random input U. From this result, it can be easily understood that in (1) or (2), only those terms contribute to $G_{BLA}(k)$ where the product $U(l_1)U(l_2)U(l_3)$ has a phase $\varphi_k = \angle U(k)$. Terms that depend also on $\varphi_{l\neq k}$ will be eliminated in the expected value $E_u\{Y(k)\bar{U}(k)\}$, where the over-score denotes the complex conjugate. Hence for a third degree nonlinearity, the following result holds (Pintelon and Schoukens, 2012):

$$G_{BLA}(k) = 6S(k)R(k) \sum_{l=-N/2+1}^{N/2-1} |U(l)|^2 + O(N^{-1}) \quad (8)$$

The error term $O(N^{-1})$ is due to the fact that for $l = k$, only 3 different combinations can be made instead of 6.

This result shows that it is easy to measure the product $R(k)S(k)$ for a WH-system by measuring G_{BLA}. However, in order to get an initialization for the parametric modeling step, we should be able to split this product over the individual transfer functions $R(k), S(k)$. It is this split that turns the identification of WH-systems into a tough problem that is hard to solve. Recently, brute force search strategies were presented (Sjöberg and Schoukens, 2012). In a first step a linear transfer function is estimated for the measured G_{BLA}. Next, the poles and zeros are calculated, and finaly a coarse search is made by scanning over all the possible splits of these poles and zeros over R, S. For each of these combinations an initial estimate can be made for the static nonlinearity, and the final model can then be tested on the data. This allows the initial models to be ranked easily. The number of combinations to be tested grows very fast with the order of the systems. For that reason it would be much better if we can label the poles and zeros already in the preprocessing step. That will be discussed in the next section.

3. THE BLA OF A WH-SYSTEM USING COUPLED RANDOM PHASE MULTISINES

In the previous section, it was shown that it is possible to generate initial estimates for the linear dynamic blocks of a WH-system, at a cost of scanning a fast growing number of candidate models. In this section we will show that we can significantly reduce the number of possible pole/zero combinations by classifying the poles and zeros immediately after the G_{BLA} measurement. To do so, we introduce a more specific class of random phase multisines.

3.1. Coupled random phase multisines

In (3) we selected the phases at each frequency to be independent of each other. In a coupled random phase multisine, the same strategy will be used, but this time we assign a phase to well chosen couples of frequencies. We first set the ideas on a simple example, and next we introduce the general definition.

Simple example. Consider a random phase multisine that is exciting the following set of frequencies

$$\pm f \in \{(2,3),(6,7),\cdots,(2+4k,2+4k+1),\cdots\} \quad (9)$$

We assign an i.i.d. uniformly distributed random phase to each of these couples, and then we continue along the same lines as in (3). The presence of coupled frequencies will create additional contributions to G_{BLA}. As will be shown below, some of these contributions contain shifted terms like $G(\tilde{k}) = R(\tilde{k}-1)S(\tilde{k})$ at some frequencies $\tilde{k}$. This shift will also result in a shift of the poles/zeros of the first system. By generalizing the definition of the set of coupled frequencies, the user can control this shift, so that it will be easier to separate the shifted poles/zeros from those that do not move.

Coupled random phase multisine. Choose an even integer value $d \geq 4$ (e.g. $d = 4$). This value will set the frequency resolution of the coupled multisine. Choose a shift value $s = k_s d + 1$ that will set the shift of the poles/zeros of the first system. Then the frequency couples are given by

$$F_{shift} = \{(\frac{d}{2}+dk, \frac{d}{2}+dk+s), k=-k_{max},\cdots,k_{max}\} \quad (10)$$

The request that $d \geq 4$ is needed to create spectral lines in the output that will receive only shifted contributions to the BLA as will be explained in the next section. That will simplify the processing significantly.

For notational convenience, we define

$$m = \frac{d}{2} + dk \quad (11)$$

.

3.2. The best linear approximation

For simplicity, we again explain the idea on the third degree nonlinear example. The contributions to the BLA are those where only the input phase φ_k shows up in the contribution to $Y(k)$ in (2). For a coupled multisine, there are four possibilities that are listed below.

1) At frequency m

$$\begin{aligned} Y_1(m) \quad = \quad & S(m)R(m)U(m)\sum|R(l)|^2|U(l)|^2 \\ & + \\ & S(m)R(m+s)U(m)\sum R(l)\bar{R}(l+s)U(l)\bar{U}(l+s)+O(N^{-1}) \end{aligned} \quad (12)$$

2) At frequency $m+s$

$$\begin{aligned} Y_2(m+s) \quad = \quad & S(m+s)R(m+s)U(m+s)\sum|R(l)|^2|U(l)|^2 \\ & + \\ & S(m+s)R(m)U(m)\sum R(l+s)\bar{R}(l)U(l+s)\bar{U}(l)+O(N^{-1}) \end{aligned} \quad (13)$$

3) At frequency $m-s$

$$Y_3(m-s) \quad = \quad S(m-s)R(m)U(m)\sum R(l)\bar{R}(l+s)U(l)\bar{U}(l+s)+O(N^{-1}) \quad (14)$$

4) At frequency $m+2s$

$$\begin{aligned} Y_4(m+2s) \quad = \quad & S(m+2s)R(m+s)U(m+s)\sum R(l+s)\bar{R}(l)U(l+s)\bar{U}(l) \\ & +O(N^{-1}) \end{aligned} \quad (15)$$

In these expressions, the sum runs over all excited frequencies from $-N/2+1$ to $N/2-1$. From here on, we drop the $O(N^{-1})$ error term.

For our purpose, the contributions in group 3 and 4 are important. Since $Y_3(k) = \bar{Y_4}(-k)$, we can focus completely on one of both. Observe, for example, that Y_4 can be rewritten as:

$$Y_4(m+2s) = \alpha S(m+2s)R(m+s)U(m)$$

keeping in mind that $U(m+s) = U(m)$ by construction. The complex constant α equals in this case $\sum R(l+s)\bar{R}(l)U(l+s)\bar{U}(l)$. The term Y_4 contains only contributions where the dynamics of the two blocks are shifted over s with respect to each other. The resulting transfer function is $G_{BLA4}(k) = \alpha S(k)R(k-s)$. The frequency shift creates a shift of the poles and the zeros of R over the frequency s. These are no longer real, or paired in complex conjugated couples. Hence, the coefficients of the rational transfer function will be complex instead of being real. This will be used in the next section.

3.3. Parametric smoothing of the BLA, shifted poles and zeros

From the simulation results in the next section, it can be seen that in the nonparametric FRF measurements, the shifted resonance frequencies are clearly visible for lightly damped systems. However, by identifying a parametric model, it is possible to reduce the impact of the disturbing noise and the stochastic nonlinearities. At the same time, the user gets direct access to the poles of the system. Because the shifted poles result in a transfer function model with complex coefficients, an adapted frequency domain system will be used (Peeters *et al.*, 2001). Once the estimated transfer function is available, it is possible to search for shifted poles. This procedure is shortly discussed in this section. First the raw FRF data are selected, next a weighted least squares estimator is proposed.

Define the following FRF-vector by collecting the BLA-measurements $G = [G_L G_R]$, with m defined in (11):

$$\begin{aligned} G_L(f_i = -m+s) \quad &= \quad G_{BLA4}(-(m-s)) \\ &= \quad \alpha S(-m+s)R(-m) \end{aligned} \quad (16)$$

for $k = k_{max}, \cdots, 1$, and

$$\begin{aligned} G_R(f_i = m+2s) \quad &= \quad G_{BLA3}(m+2s) \\ &= \quad \alpha S(m+2s)R(m+s) \end{aligned} \quad (17)$$

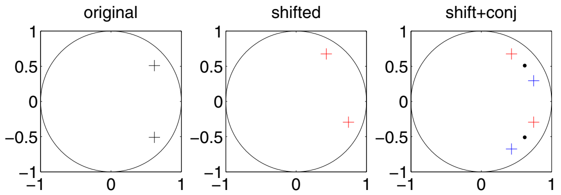

Figure 3: Shifted poles. Left: original poles; Middle: shifted poles; Right: shifted and conjugated poles. .

for $k = 1, \cdots, k_{max}$.

This creates an FRF $G(f_i)$ at the selected frequencies f_i

$$G(f_i) = \alpha S(f_i) R(f_i - s) \tag{18}$$

Observe that $G(f) \neq \bar{G}(-f)$. So, the frequency shift results in a transfer function with complex coefficients, and hence $G(f_i)$ should be modelled as a rational form with complex coefficients. The poles p_i (zeros z_i) of R will be no longer real nor complex conjugated. By comparing p_i and $\bar{p}_i$, a frequency shift $e^{j2\pi 2s}$ will be visible in the complex plane. The same operation on the poles (zeros) of S will not show this shift because the poles (zeros) of S are real, or complex conjugated.

In Figure 3, an example of the pole shifting is given. By splitting the poles (zeros) in shifted or non-shifted poles (zeros) we can assign in this step the poles and zeros to the first or second system in the WH-model. This solves eventually the initialization of the structure. It is clear that the stochastic nonlinearities or disturbing noise will disturb the estimated pole/zero positions, and this can blur a crisp view.

The transfer function coefficients are estimated by minimizing the following weighted least squares cost fucntion (Peeters *et al.*, 2001):

$$V(\theta) = \frac{1}{F} \sum_{k=-k_{max}}^{k_{max}} \frac{|G(f_k) - G(f_k, \theta)|^2}{\hat{\sigma}_G^2(k)} \tag{19}$$

with $G(f_k)$ the measured BLA at the selected set of frequencies, as explained in (16) and (17), and $\hat{\sigma}_G^2(k)$ the measured total variance (sum of the variance of the disturbing noise and the stochastic nonlinearities).

4. SIMULATION RESULTS

In this section, we illustrate the explained method on a simulation. The dynamic systems R, S were selected as Chebyshev filters of order 2, with a ripple of 10 and 20 dB, and a 3dB bandwidth of $0.05f_s$ and $0.1f_s$ respectively (see Figure 4). The sample frequency f_s is normalized to 1. The static nonlinearity $f(p)$ is shown in Figure 5 over the domain that is excited in the simulations.

The system is excited with a coupled multisine with $d = 4$, and a shift $s = 57$. The period length of the multisine is $N = 8192$. The WH-system is excited up to $f_s/6$. No external disturbing noise is added to the output.

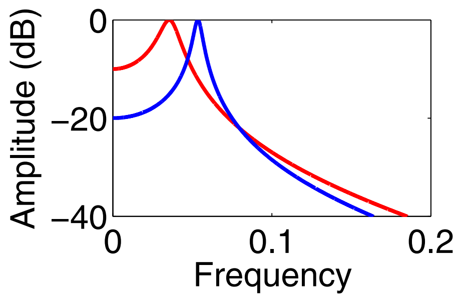

Figure 4: The transfer functions R (red) and S (blue) of the WH-system.

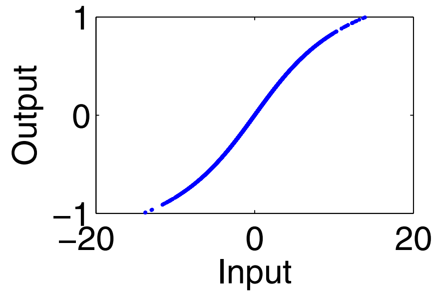

Figure 5: Static nonlinearity of the WH-system. .

In the simulation, we used only the steady state response of the system.

The FRF of $G_{BLA}$is measured by averaging over 10 realizations of the input. The averaged value is shown together with its standard deviation (due to the stochastic nonlinearities) in Figure 6. Observe that the FRF is not symmetric around the origin (complex coefficients in the transfer function!). It can also be seen that the signal-to-noise ratio is quite low, even in the absense of disturbing noise. This is due to the high level of the nonlinear distortions.

Using the weighted least squares method (19), a parametric transfer function with complex coefficients is estimated, and eventualy the estimated poles are calculated. These are plotted, together with their complex conjugates, in the z-domain. The first quadrant is shown in Figure 7. The poles of the S-system are shown in blue (almost on top of each other, and very close to the true value). Those of the R-system are shown in red. Here it can be clearly seen that they are nicely shifted with respect to the true value. This shows that in this case we could easily assign the poles to the first or second system. The results for the zeros were not useful in this case. All true zeros are in -1, and that part of the frequency band is not excited. For that reason, the uncertainty on the zero positions is very large, and this prohibits their classification.

5. CONCLUSIONS

In this paper we have first briefly introduced the best linear approximation (BLA). For a Wiener-Hammerstein

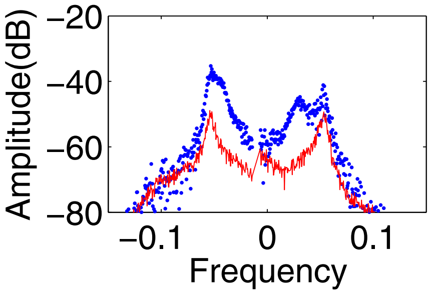

Figure 6: Measured FRF of the BLA (blue), and its standard deviation (red). Observe that the amplitude characteristic is not symmetric around the origin.

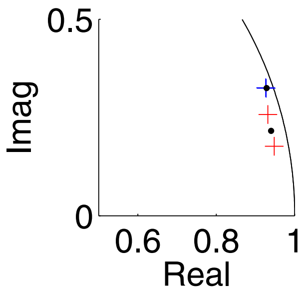

Figure 7: The estimated poles and their complex conjugates (only the first quadrant is shown). The black dots are the pole positions of the R, S-dynamic systems. The blue poles (S-system) did not shift, while the red poles (R-system) show a clear shift.

system that is excited by Gaussian noise or a random phase multisine, it is shown to be equal to the product of the transfer functions of the two dynamic blocks. By using more specialized coupled phase multisines, it was possible to create new contributions in the BLA that equals the product of frequency shifted transfer functions. On the basis of the nonparametric measurement of these components, it is shown to be possible to assign the identified poles and zeros to the first or second dynamic block of the WH-system, provided that they are properly excited.

REFERENCES

Billings, S. A. (2013). *Nonlinear system identification : NARMAX methods in the time, frequency, and spatio-temporal domains.* John Wiley & Sons Ltd.

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

Enqvist M. and L. Ljung. Linear approximations of nonlinear FIR systems for separable input processes. Automatica, 41(3):459-473, 2005.

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

Haber, R., and Keviczky, L. (1999). *Nonlinear System Identification: Input-Output Modeling Approach.* Dordrecht: Kluwer Academic Publishers.

Hunter I.W. and M.J. Korenberg (1986). The identification of nonlinear biological systems: Wiener and Hammerstein cascade models. *Biological Cybernetics*, 55, pp. 135-144.

M.J. Korenberg (1991). Parallel cascade identification and kernel estimation for nonlinear systems. Annals of Biomedical Engineering, 19:429–455.

Ljung, L (1999), *System Identification: Theory for the User (second edition).* Prentice Hall, Upper Saddle River, New Jersey.

Palm, G. (1979). On representation and approximation of nonlinear systems. Biological Cybern., 34, pp. 49-52, 1979.

Peeters F., R. Pintelon, J. Schoukens, Y. Rolain, E. S. Guiterrez and P. Guillaume (2001). Identification of Rotor-Bearing Systems in the Frequency Domain. Part I: Estimation of Frequency Response Functions, Mechanical Systems and Signal Processing, vol. 15, no. 4, pp. 759-773.

Pintelon, R. and J. Schoukens (2012). *System Identification. A Frequency Domain Approach.* 2nd edition. Wiley - IEEE-press, Piscataway.

Schoukens J., L. Gomme, W. Van Moer, et al. (2008). Identification of a block-structured nonlinear feedback system, applied to a microwave crystal detector. *IEEE Trans. Instrum. and Measurem.*, 57, pp. 1734-1740.

Schoukens M. and Y. Rolain. Parametric Identification of Parallel Wiener Systems (2012). *IEEE Trans. on Instrum. and Measurem.*, 61, pp. 2825-2832.

Sjöberg, J. and J. Schoukens (2012). Initializing Wiener-Hammerstein models based on partitioning of the best linear approximation. *Automatica*, 48, pp. 353-359

Söderström, T. and P. Stoica (1989). *System Identification.* Prentice-Hall, Englewood Cliffs 1989.

Vandersteen G. and J. Schoukens (1999). Measurement and identification of nonlinear systems consisting of linear dynamic blocks and one static nonlinearity. *IEEE Trans. on Automatic Control*, 44, pp. 1266-1271.

Westwick, D.T. and R.E. Kearney (2003). *Identification of Nonlinear Physiological Systems.* IEEE-Wiley.

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

Wills A. and B. Ninness (2012). Generalised Hammerstein-Wiener system estimation and a benchmark application. *Control Engineering Practice*, 20, pp. 1097-1108.