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

Identification of parallel Wiener-Hammerstein systems with a decoupled static nonlinearity ⋆

M. Schoukens * **K. Tiels** * **M. Ishteva** * **J. Schoukens** *

* *Vrije Universiteit Brussel, Brussels, Belgium (e-mail: maarten.schoukens@vub.ac.be)*

Abstract: Block-oriented models are often used to model a nonlinear system. This paper presents an identification method for parallel Wiener-Hammerstein systems, where the obtained model has a decoupled static nonlinear block. This decoupled nature makes the interpretation of the obtained model more easy. First a coupled parallel Wiener-Hammerstein model is estimated. Next, the static nonlinearity is decoupled using a tensor decomposition approach. Finally, the method is validated on real-world measurements using a custom built parallel Wiener-Hammerstein test system.

1. INTRODUCTION

Block-oriented models are often used to model nonlinear systems. A block-oriented model consists of two types of blocks: linear-time invariant (LTI) blocks and static nonlinear (SNL) building blocks. They offer insight about the system to the user due to this highly structured nature. There are many different types of block-oriented models, as is discussed in Giri and Bai [2010]. The simplest ones are Hammerstein (SNL-LTI) and Wiener (LTI-SNL) models. These two basic block-oriented models can be extended to Wiener-Hammerstein (LTI-SNL-LTI), and Hammerstein-Wiener (SNL-LTI-SNL) models by adding blocks in series. Another extension can be made to parallel Hammerstein and parallel Wiener models by connecting a number of Hammerstein or Wiener models in parallel (see Schoukens et al. [2011] and Schoukens and Rolain [2012b]). This paper presents an identification method for parallel Wiener-Hammerstein systems.

A parallel Wiener-Hammerstein system consists of a number of Wiener-Hammerstein systems placed in parallel, as is shown in Figure 1. All the Wiener-Hammerstein subsystems share the same input, and the output of the parallel Wiener-Hammerstein system is obtained by summing up the outputs of the different subsystems. It is shown in Palm [1979] that a wide class of discrete time Volterra systems can be approximated arbitrary well using a parallel Wiener-Hammerstein model structure. However, no method is presented there to identify such a model.

A parallel Wiener-Hammerstein system identification approach is presented in Schoukens et al. [2013], but the resulting model has one multiple-input-multiple-output (MIMO) static nonlinearity, rather then a single-input-single-output (SISO) static nonlinearity for each parallel branch. Other parallel Wiener-Hammerstein system identification methods are presented in Baumgartner and Rugh [1975], Wysocki and Rugh [1976] and Billings and Fakhouri [1979]. However these methods are restricted to the S_M system class. A system belonging to the S_M system class has M parallel branches, where the static nonlinearity of branch m is given by the monomial $(.)^m$. Only one branch for each degree of the nonlinearity is allowed. This reduces the model flexibility significantly.

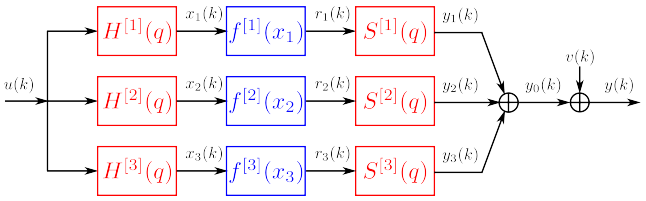

Fig. 1. A 3-branch parallel Wiener-Hammerstein system: different Wiener-Hammerstein systems placed in parallel. The static nonlinear block $f^{[i]}$ of the i-th branch sandwiched in between the LTI blocks $H^{[i]}(q)$ and $S^{[i]}(q)$. The noise source $v(k)$ is additive colored noise.

This paper presents a method to identify a parallel Wiener-Hammerstein model with a SISO static nonlinearity for each branch, rather than one big static nonlinearity for all the branches together. First a parallel Wiener-Hammerstein model containing one MIMO static nonlinearity is estimated with the parallel Wiener-Hammerstein identification method proposed in Schoukens et al. [2013]. The MIMO static nonlinearity is modeled by a set of multivariate polynomials. Next, the MIMO static nonlinearity is decoupled using a tensor decomposition approach, imposing common factors. The decoupling algorithm is an extended version of the algorithm that is presented in Tiels and Schoukens [2013].

A decoupled nonlinearity has different advantages. Firstly, it is more easy to increase the model complexity. The number of model parameters tends to grow very fast for an increasing model complexity of a MIMO static nonlinearity, for sure when for instance a multivariate

⋆ This work was funded by the Fund for Scientific Research (FWO), the Methusalem grant of the Flemish Government (METH-1), the IAP VII/19 DYSCO program, and the ERC advanced grant SNLSID, under contract 320378. Mariya Ishteva is an FWO Pegasus Marie Curie Fellow

polynomial is used. A linear dependency of the degree is achieved when decoupled polynomials are used. Secondly, the ability to interpret the model is increased by using different SISO static nonlinearities instead of one MIMO static nonlinearity.

The contribution of this paper is twofold. First, this paper improves the decoupling method presented in Tiels and Schoukens [2013]. Second, the decoupling approach is integrated with the parallel Wiener-Hammerstein identification approach that is presented in Schoukens et al. [2013], and applied on a custom built parallel Wiener-Hammerstein test system.

in Section 2, the system, signals and stochastic framework are introduced. Next, the proposed identification approach is explained in Section 3. Finally, the proposed method is validated on a measurement example in Section 4.

2. SYSTEM, SIGNALS AND STOCHASTIC FRAMEWORK

This section describes the considered class of systems, introduces the noise framework and defines the signal class that is considered in this paper.

2.1 The system class

We consider parallel Wiener-Hammerstein systems. These systems consists of different Wiener-Hammerstein systems that share the same input signal (see Figure 1). The output of the total system is obtained as the sum of the outputs of the different branches.

All the LTI blocks are considered to be infinite impulse response (IIR) filters, parametrized by a rational polynomial in the backward shift operator q^{-1}:

$$H^{[i]}(q) = \frac{B_h^{[i]}(q)}{A_h^{[i]}(q)} = \frac{b_{h,0}^{[i]} + \ldots + b_{h,n_{b_{h,i}}}^{[i]} q^{-n_{b_{h,i}}}}{a_{h,0}^{[i]} + \ldots + a_{h,n_{a_{h,i}}}^{[i]} q^{-n_{a_{h,i}}}},$$

$$S^{[i]}(q) = \frac{B_s^{[i]}(q)}{A_s^{[i]}(q)} = \frac{b_{s,0}^{[i]} + \ldots + b_{s,n_{b_{s,i}}}^{[i]} q^{-n_{b_{s,i}}}}{a_{s,0}^{[i]} + \ldots + a_{s,n_{a_{s,i}}}^{[i]} q^{-n_{a_{s,i}}}},$$

where $n_{b_h,i}$ and $n_{a_h,i}$ are the orders of the numerator and denominator of the front dynamics of the i-th parallel branch respectively, $n_{b_s,i}$ and $n_{a_s,i}$ are the orders of the numerator and denominator of the back dynamics of the i-th parallel branch.

The noiseless output $y_0(k)$ of a parallel Wiener-Hammerstein system is given by:

$$x_i(k) = H^{[i]}(q)u(k), \qquad i = 1, \ldots, n_{br} \quad (1)$$

$$r_i(k) = f^{[i]}(x_i(k)), \qquad i = 1, \ldots, n_{br} \quad (2)$$

$$y_0(k) = \sum_{i=1}^{n_{br}} S^{[i]}(q) r_i(k), \quad (3)$$

where n_{br} is the number of parallel branches in the system, $f^{[i]}(x_i)$ is the static nonlinearity of branch i and the signals are defined in Figure 1.

2.2 Signals and noise

The excitation signal $u(k)$ belongs to the Riemann equivalence class of asymptotically normally distributed excitation signals as defined in Pintelon and Schoukens [2012].

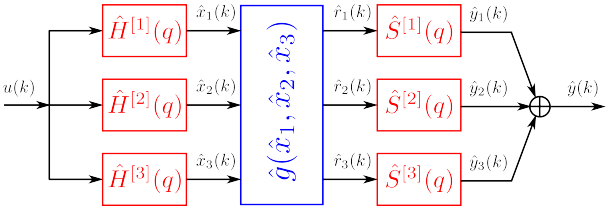

Fig. 2. A 3-branch parallel Wiener-Hammerstein model with a MIMO static nonlinearity. A MIMO static nonlinear block $\hat{g}$ sandwiched in between the LTI blocks $\hat{H}^{[i]}(q)$ and $\hat{S}^{[i]}(q)$.

This signal class includes random Gaussian noise signals, but also periodic Gaussian noise signals and random phase multisines.

The output is disturbed by an additive, colored Gaussian noise disturbance $v(k)$:

$$y(k) = y_0(k) + v(k). \quad (4)$$

2.3 Identifiability

A parallel Wiener-Hammerstein system suffers from some identifiability issues. A finite gain and an arbitrary delay can be exchanged between the blocks that are connected in series. On top of this, a full rank linear transform can be introduced between the LTI blocks and the static nonlinear blocks. This transforms the parallel SISO nonlinearity in a MIMO static nonlinearity. This issue is discussed in more detail in Schoukens et al. [2013].

3. IDENTIFICATION APPROACH

A two step identification algorithm is proposed. First, a parallel Wiener-Hammerstein model with a MIMO static nonlinear block (see Figure 2) is estimated using the algorithm that is presented in Schoukens et al. [2013], and which is repeated briefly in this paper. Next, the MIMO static nonlinearity is decoupled into a single SISO static nonlinearity for each branch in the parallel Wiener-Hammerstein model. This last step can increase the number of parallel branches in the model.

3.1 Identification of a coupled model

The parallel Wiener-Hammerstein identification method presented in Schoukens et al. [2013] starts by estimating the best linear approximation (BLA) of the system under test at different operating points. The BLA G_{bla} of a system for any type of input signal, best in mean squares sense, is defined in Pintelon and Schoukens [2012]:

$$G_{bla}(j\omega_k) = \underset{G(j\omega_k)}{\arg\min}\, E_U \left\{ \| Y(j\omega_k) - G(j\omega_k) U(j\omega_k) \|^2 \right\}, \quad (5)$$

where $Y(j\omega_k)$ is the discrete Fourier transform (DFT) of the output y evaluated at frequency $j\omega_k$, and where $U(j\omega_k)$ is the DFT of the input u evaluated at frequency $j\omega_k$. $G_{bla}(j\omega_k)$ is the value of the frequency response function (FRF) of the BLA at $j\omega_k$. The expectation is taken with respect to the random input U. Each BLA is obtained using excitation signals that have a different

power spectrum. This changes the operating point of the system under test slightly for each BLA.

When input signals belonging to the Riemann equivalence class of asymptotically normally distributed excitation signals are used, the BLA of a parallel Wiener-Hammerstein system is given by:

$$G_{bla}(j\omega_k) = \sum_{i=1}^{n_{br}} \alpha_i H^{[i]}(j\omega_k) S^{[i]}(j\omega_k). \tag{6}$$

The scaling factors α_i depend on the input power spectrum, and on the nonlinearities that are present in the system. It follows from eq. (6) that the poles of the BLA are constant and independent of the applied excitation signal, while the zeros of the BLA shift depending on the applied excitation signal.

The BLAs of the different operating points i_r are parametrized using a common denominator approach:

$$\hat{G}_{bla}^{[i_r]}\left(q, \hat{\boldsymbol{\theta}}_{bla}\right) = \frac{\hat{d}_0^{[i_r]} + \hat{d}_1^{[i_r]} q^{-1} + \ldots + \hat{d}_{n_d}^{[i_r]} q^{-n_d}}{\hat{c}_0 + \hat{c}_1 q^{-1} + \ldots + \hat{c}_{n_c} q^{-n_c}}, \tag{7}$$

where each BLA has different numerator coefficients $\hat{d}_i^{[i_r]}$, but they all share the same denominator coefficients $\hat{c}_i$.

Next, the dynamics present in the BLAs are decomposed over the different parallel branches of the model. To do so, the singular value decomposition (SVD) is taken of the matrix composed of the numerator coefficients:

$$\boldsymbol{D} = \begin{bmatrix} \hat{d}_0^{[1]} & \hat{d}_1^{[1]} & \ldots & \hat{d}_{n_d}^{[1]} \\ \hat{d}_0^{[2]} & \hat{d}_1^{[2]} & \ldots & \hat{d}_{n_d}^{[2]} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{d}_0^{[R]} & \hat{d}_1^{[R]} & \ldots & \hat{d}_{n_d}^{[R]} \end{bmatrix}^T, \tag{8}$$

where R is the number of estimated BLAs. The SVD of $\boldsymbol{D}$ yields an orthonormal basis for the space spanned by the $\boldsymbol{D}$-matrix:

$$\boldsymbol{D} = \boldsymbol{\Delta}_{bla} \boldsymbol{\Sigma}_{bla} \boldsymbol{V}_{bla}^T, \tag{9}$$

The matrix $\boldsymbol{\Delta}_{bla}$ contains an estimate of the numerator coefficients for each branch:

$$\hat{G}_i(q) = \frac{\hat{\delta}_0^{[i]} + \hat{\delta}_1^{[i]} q^{-1} + \ldots + \hat{\delta}_{n_d}^{[i]} q^{-n_d}}{\hat{c}_0 + \hat{c}_1 q^{-1} + \ldots + \hat{c}_{n_c} q^{-n_c}}, \tag{10}$$

where $\hat{\delta}_j^{[i]}$ is the element of the j-th row and i-th column of the matrix $\boldsymbol{\Delta}_{bla}$. An estimate of the number of parallel branches n_{br} present in the system is obtained by looking at the spectrum of the singular values of $\boldsymbol{D}$. $\hat{G}_i(q)$ is an estimate of the dynamics that are present in branch i of the parallel Wiener-Hammerstein model.

Finally, the estimated dynamics of each branch of the parallel Wiener-Hammerstein model need to be allocated to the front and back LTI blocks of the model. To do so, all combinations of poles and zeros in the different blocks are scanned, and a MIMO static nonlinear block is estimated for each combination with a multivariate polynomial which is linear in the parameters. Finally, the model with the smallest simulation error is selected. The parameters of the selected model are optimized further using a Levenberg-Marquardt nonlinear optimization algorithm to refine the model estimate. A more detailed description of this step in the algorithm can be found in Schoukens et al. [2013].

In this paper, a low order multivariate polynomial is used to model the MIMO static nonlinearity to limit the number of coefficients. The next section presents a method to decouple the MIMO static nonlinearity to one SISO static nonlinearity for each branch. The complexity of the models that describe the decoupled static nonlinearities is increased in a later step. This decreases the model error, while keeping the number of parameters in the model relatively low (compared to a high order MIMO static nonlinearity).

3.2 Decoupling the static nonlinearity

The decoupling method generates starting values for the parameters of the decoupled Wiener-Hammerstein model, described by:

$$\tilde{x}_i(k) = \tilde{H}^{[i]}(q)u(k), \qquad i = 1, \ldots, n_r \tag{11}$$

$$\tilde{r}_i(k) = \tilde{f}^{[i]}(\tilde{x}_i(k)), \qquad i = 1, \ldots, n_r \tag{12}$$

$$\tilde{y}(k) = \sum_{i=1}^{n_r} \tilde{S}^{[i]}(q)\tilde{r}_i(k), \tag{13}$$

starting from its coupled polynomial representation:

$$\hat{x}_i(k) = \hat{H}^{[i]}(q)u(k), \qquad i = 1, \ldots, n_{br} \tag{14}$$

$$\hat{r}_i(k) = \hat{g}^{[i]}(\hat{\boldsymbol{x}}(k)), \qquad i = 1, \ldots, n_{br} \tag{15}$$

$$\hat{y}(k) = \sum_{i=1}^{n_{br}} \hat{S}^{[i]}(q)\hat{r}_i(k), \tag{16}$$

where $\hat{\boldsymbol{x}}(k) = [\hat{x}_1(k), \cdots, \hat{x}_{n_{br}}(k)]^T$. As already pointed out, the decoupling step can increase the number of branches (n_r can be larger than n_{br}), but the goal is to keep the number of branches n_r small.

Some methods already exist to decouple the polynomial representations in Volterra models (see Favier and Bouilloc [2009]), and parallel Wiener models (see Schoukens and Rolain [2012a]). For example, the method in Schoukens and Rolain [2012a] splits the polynomial in a sum of homogeneous polynomials, and uses tensor decomposition methods to eliminate the cross-terms in each homogeneous polynomial separately. This technique can be directly applied to each polynomial $\hat{g}^{[i]}(\hat{\boldsymbol{x}}(k))$, and is in Tiels and Schoukens [2013] referred to as "separate decoupling". This results in a sum of optimally decoupled polynomial representations (optimal in the sense of the smallest number of branches), but the total number of branches is not necessarily optimal. Two other methods are presented in Tiels and Schoukens [2013], referred to as "simultaneous homogeneous" and "simultaneous all", that can result in a smaller total number of branches. This is realized by imposing common factors in the tensor decompositions. The main idea behind these methods is to impose that the resulting SISO polynomials share the same input signals, thus keeping the number of input filters $H^{[i]}(q)$ small. The "simultaneous all" approach suffers from the drawback that, although the same input dynamics are shared for all polynomials, the output dynamics in general differ for different degrees of nonlinearity. Another drawback of the method is that the coefficients of the resulting SISO polynomials are all equal to one.

Here we present an improved version of the "simultaneous all" method. Without loss of generality, the focus is

on quadratic and cubic polynomials. Compared to the "simultaneous all" approach in Tiels and Schoukens [2013], in this paper, branches that share the same input dynamics are imposed to share the same output dynamics as well. Moreover, the flexibility of each branch in the decoupled structure is increased by allowing arbitrary coefficients β for the SISO polynomials. Although this increases the complexity of the optimization problem, the increased flexibility of each branch allows for a smaller total number of branches.

Assume that $\hat{g}^{[i]}(\hat{\boldsymbol{x}}(k))$ is the sum of a quadratic and a cubic homogeneous polynomial:

$$\begin{aligned}\hat{g}^{[i]}(\hat{\boldsymbol{x}}(k)) = &\sum_{j_1,j_2=1}^{n_{br}} \gamma_{j_1 j_2}^{[i]} \hat{x}_{j_1}(k)\hat{x}_{j_2}(k) \\ &+ \sum_{j_1,j_2,j_3=1}^{n_{br}} w_{j_1 j_2 j_3}^{[i]} \hat{x}_{j_1}(k)\hat{x}_{j_2}(k)\hat{x}_{j_3}(k),\end{aligned} \tag{17}$$

with $\boldsymbol{\Gamma}^{[i]}$ the symmetric matrix of polynomial coefficients $\gamma_{j_1 j_2}^{[i]}$, and $\boldsymbol{\mathcal{W}}^{[i]}$ the symmetric tensor of polynomial coefficients $w_{j_1 j_2 j_3}^{[i]}$. The main tool to decouple these matrices and tensors will be the canonical polyadic decomposition (CPD) (see Carroll and Chang [1970], Harshman [1970], Kolda and Bader [2009]). The CPD approximates a tensor - in least squares sense - by a sum of rank-one tensors. Say, for example, that $\boldsymbol{\mathcal{W}}^{[i]}$ has a CPD

$$\boldsymbol{\mathcal{W}}^{[i]} \approx \sum_{r=1}^{n_r} \psi_r^{[i]} \boldsymbol{p}_r \circ \boldsymbol{p}_r \circ \boldsymbol{p}_r, \tag{18}$$

where $\circ$ denotes the tensor product. This means that element-wise:

$$w_{j_1 j_2 j_3}^{[i]} \approx \sum_{r=1}^{n_r} \psi_r^{[i]} p_{j_1 r}^{[i]} p_{j_2 r}^{[i]} p_{j_3 r}^{[i]}. \tag{19}$$

The cubic multivariate homogeneous polynomial described by $\boldsymbol{\mathcal{W}}^{[i]}$ is thus transformed into a sum of n_r univariate homogeneous polynomials $\psi_r^{[i]} \tilde{x}_r^3(k)$, with $\tilde{x}_r = \boldsymbol{p}_r^T \hat{\boldsymbol{x}}(k)$. The CPD is often calculated via an alternating least-squares (ALS) approach (see e.g. Kolda and Bader [2009]). Recently, other algorithms to calculate the CPD of a tensor were proposed in Sorber et al. [2013a] that obtain a better overall performance than ALS.

To impose that the univariate polynomials, obtained from decoupling all the matrices $\boldsymbol{\Gamma}^{[i]}$ and all the tensors $\boldsymbol{\mathcal{W}}^{[i]}$, share the same input signals $\tilde{x}_r$, these matrices and tensors are stacked in a partially symmetric tensor $\boldsymbol{\mathcal{T}} \in \mathbb{R}^{(n_{br}+1)\times(n_{br}+1)\times(n_{br}+1)\times 2\times n_{br}}$, such that:

$$\boldsymbol{\mathcal{T}} \approx \sum_{r=1}^{n_r} \begin{bmatrix}\boldsymbol{p}_r \\ 1\end{bmatrix} \circ \begin{bmatrix}\boldsymbol{p}_r \\ 1\end{bmatrix} \circ \begin{bmatrix}\boldsymbol{p}_r \\ 1\end{bmatrix} \circ \begin{bmatrix}\beta_{2r} \\ \beta_{3r}\end{bmatrix} \circ \boldsymbol{m}_r, \tag{20}$$

where,

$$m_{jr} = \phi_r^{[j]}, \quad j = 1, \ldots, n_{br}. \tag{21}$$

The one stacked with the $\boldsymbol{p}_r$ vector imposes a partial symmetry in the matrix, which is used during the decomposition step. The entries of $\boldsymbol{\mathcal{T}}$ are given by [1] [2]:

$$t_{j_1 j_2 j_3 2 j_4} = w_{j_1 j_2 j_3}^{[j_4]}, \tag{22}$$

$$t_{j_1 j_2 (n_{br}+1) 1 j_4} = \gamma_{j_1 j_2}^{[j_4]}, \tag{23}$$

$$t_{j_1 j_2 j_3 1 j_4} \approx \sum_{r=1}^{n_r} p_{j_1 r} p_{j_2 r} p_{j_3 r} \beta_{2r} \phi_r^{[j_4]}, \tag{24}$$

$$t_{j_1 j_2 (n_{br}+1) 2 j_4} \approx \sum_{r=1}^{n_r} p_{j_1 r} p_{j_2 r} \beta_{3r} \phi_r^{[j_4]}, \tag{25}$$

for $j_1, j_2, j_3, j_4 = 1, \ldots, n_{br}$.

The tensor $\boldsymbol{\mathcal{T}}$ is decoupled using the Tensorlab toolbox by Sorber et al. [2013b] which can handle missing entries and partial symmetry. This results in a decoupled parallel Wiener-Hammerstein model, as described by (13). The input dynamics, SISO polynomials, and output dynamics of the decoupled model are given by:

$$\tilde{H}^{[r]}(q) = \sum_{j=1}^{n_{br}} p_{jr} \hat{H}^{[j]}(q), \tag{26}$$

$$\tilde{f}^{[r]}(\tilde{x}_r(k)) = \beta_{2r} \tilde{x}_r^2(k) + \beta_{3r} \tilde{x}_r^3(k), \tag{27}$$

$$\tilde{S}^{[r]}(q) = \sum_{j=1}^{n_{br}} \phi_r^{[j]} \hat{S}^{[j]}(q), \tag{28}$$

for $r = 1, \ldots, n_r$. The resulting model has n_r branches, where n_r is set by the user.

Note that the optimization does not take into account the actual input/output data. The optimization is only done starting from the estimated polynomial coefficients. This allows us to generate relatively quickly decent starting values for the parameters of the decoupled structure. In a next step, these parameters can be further optimized starting from the input/output data.

4. MEASUREMENT EXAMPLE

The proposed method is illustrated on an experimental setup.

4.1 System and measurement setup

The device under test (DUT) is a 2-branch parallel Wiener-Hammerstein system. The front and back LTI blocks of each branch are third order IIR filters. The static nonlinearity of each branch is realized with a diode-resistor network.

The signals are generated by an arbitrary waveform generator (AWG), the Agilent/HP E1445A, at a sampling frequency of 625 kHz. An internal low-pass filter with a cut-off frequency of 250 kHz is used as a reconstruction filter for the input signal. The in- and output signals of the system are measured by the alias protected acquisition channels (Agilent/HP E1430A) at a sampling frequency of 78 kHz. The AWG and acquisition cards are synchronized to avoid leakage errors.

Finally, buffers with a high input impedance, added between the acquisition cards and the in- and output of the DUT, avoid that the circuit is loaded by the 50 Ohm input impedance of the acquisition card. The buffers are very linear ($\approx$ 85 dBc at full scale and 1 MHz) up to 10 V peak to peak, and have an input impedance of 1 MΩ and a 50 Ω output impedance.

[1] Due to symmetry $t_{j_1 j_2 (n_{br}+1) j j_4}$ is also equal to $t_{j_1 (n_{br}+1) j_2 j j_4}$ and $t_{(n_{br}+1) j_1 j_2 j j_4}$ for $j = 1, 2$.

[2] Note that the entries given by (24) and (25) are unknown, however, this issue can be handled by treating them as missing elements in the tensor to be decomposed.

4.2 Signal generation

The input signal $u(k)$ is a random phase multisine (see Pintelon and Schoukens [2012]) containing $N = 131072$ samples with a flat amplitude spectrum. The excited frequency band ranges from $\frac{f_s}{N}$ to $\frac{f_s}{2}$, viz.:

$$u(k) = A \sum_{n=1}^{N/2} \cos(2\pi n \frac{f_s}{N} k + \phi_n), \tag{29}$$

The phases ϕ_n are independent uniformly distributed random variables ranging from 0 to 2π. Twenty realizations of the multisines are used. The input signal is applied at 5 different rms values that are linearly distributed between 100 mV and 1 V.

The signals are measured at a sampling frequency of 78 kHz, which is 8 times slower than the sampling frequency at the generator side, resulting in $N = 16384$ samples per period.

4.3 Model estimation

First, the BLAs of the DUT are estimated at the different operating points (or input amplitudes in this case). The BLAs are parameterized using a discrete time LTI model of order 12 in both the numerator and denominator.

Starting from these BLAs, an initial 2-branch coupled parallel Wiener-Hammerstein model is estimated. The model has 2 parallel branches, and the MIMO static nonlinearity is modeled with a 3rd order multivariate polynomial. A static nonlinearity of low degree is used such that the number of parameters in the model description is limited. However, this simple static nonlinearity is able to provide sufficiently good initial values for the following steps.

Next, the low order multivariate polynomial is decoupled using the approach that is presented in Section 3.2, and a decoupled parallel 2-branch Wiener-Hammerstein model with 3rd order polynomial SISO static nonlinearities is constructed based on this result.

Finally, the static nonlinearities in the low order decoupled parallel Wiener-Hammerstein model are replaced by polynomials of order 15. A Levenberg-Marquardt optimization algorithm is applied on the high order decoupled 2-branch parallel Wiener-Hammerstein model to further optimize the parameters. In this step, other static nonlinear models such as artificial neural networks can also be used to model the static nonlinearities of the model.

4.4 Model validation

The model is validated using a different realization of the random phase multisines that are described in Section 4.2. The results of the validation are shown in Table 1 and Figure 3. Table 1 shows three figures of merit: the rms value of the simulation error $\text{rms}(e)$, the standard deviation of the simulation error σ_e, and the mean value of the simulation error μ_e, as defined below:

$$\text{rms}(e) = \sqrt{\frac{1}{N}\sum_{k=1}^{N} e^2(k)}, \tag{30}$$

$$\sigma_e = \sqrt{\frac{1}{N}\sum_{k=1}^{N} (e(k) - \mu_e)^2}, \tag{31}$$

$$\mu_e = \frac{1}{N}\sum_{k=1}^{N} e(k), \tag{32}$$

where $e(k)$ is the difference between the measured output $y(k)$ and the modeled output $\hat{y}(k)$.

The low order decoupled model has a clear performance loss compared to the coupled model. However, it provides a good starting value to further increase the static nonlinearity model complexity for further optimization. Increasing the model order in the coupled model is very costly in the number of parameters. Describing the coupled static nonlinearity with a 15th order multivariate polynomial would require 136 parameters, where only 32 parameters are needed to describe the 15th order polynomials of the decoupled static nonlinear blocks.

The high order decoupled model performs very well. Figure 3 shows that the model error is about 50 dB lower than the output spectrum for an excitation signal with an rms value of 550 mV. Furthermore, the model error is about 30 dB lower than the total distortion level of the output spectrum. The total distortion level includes the noise distortions and distortions due to the nonlinear behavior of the system (see Pintelon and Schoukens [2012]). This shows that the model describes the nonlinear behavior of the system very well. Finally, the model error is 10 to 20 dB higher than the noise floor. This suggests that the model quality can still be improved, for instance by increasing the order of the static nonlinearity even more.

The static nonlinearities that are present in the high order decoupled model are shown in Figure 4. They show a saturating behavior. This is what is to be expected since the static nonlinearities in the system are generated by electrical diode-resistor networks.

5. CONCLUSION

This paper presents an improved method to decouple a MIMO static nonlinearity described by multivariate polynomials into different SISO polynomial static nonlinearities using tensor decomposition methods. This decoupling approach is integrated with a parallel Wiener-Hammerstein identification method to obtain parallel Wiener-Hammerstein models with different SISO static nonlinearities rather than one MIMO static nonlinearity. The approach is applied on a measurement example to show the good performance of the proposed method.

REFERENCES

S.L. Baumgartner and W.J. Rugh. Complete identification of a class of nonlinear systems from steady state frequency response. *IEEE Trans. Circuits Syst.*, 22(9): 753–759, 1975.

S.A. Billings and S.Y. Fakhouri. Identification of nonlinear Sm systems. *International Journal of Systems Science*, 10(10):1401–1408, 1979.

Table 1. Validation error on a multisine signal.

	Validation error (mV)								
	decoupled model, 15th order			decoupled model, 3rd order			coupled model, 3rd order		
rms(u)	rms(e)	σ_e	μ_e	rms(e)	σ_e	μ_e	rms(e)	σ_e	μ_e
100	0.83	0.82	-0.15	12.08	8.69	8.39	4.92	3.24	3.71
325	0.75	0.70	-0.27	22.82	22.47	3.95	7.60	7.59	-0.35
550	0.66	0.66	0.01	31.05	30.95	2.45	11.92	11.27	-3.86
775	0.75	0.75	0.02	35.95	35.20	-7.33	13.38	12.94	-3.41
1000	1.30	1.28	0.17	49.02	48.00	-10.06	22.58	22.58	-0.09

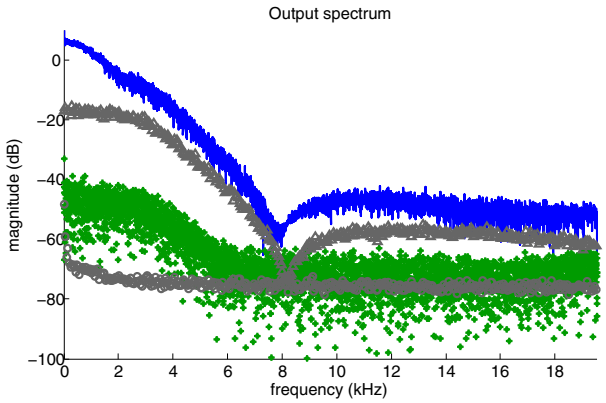

Fig. 3. Measured and modeled output spectrum of a validation dataset (rms(u) = 550 mV). The measured output and the model error are shown with the blue line and the green crosses respectively. The noise level and the total distortion level at the system output are shown with the bottom gray circles and the top gray triangles respectively.

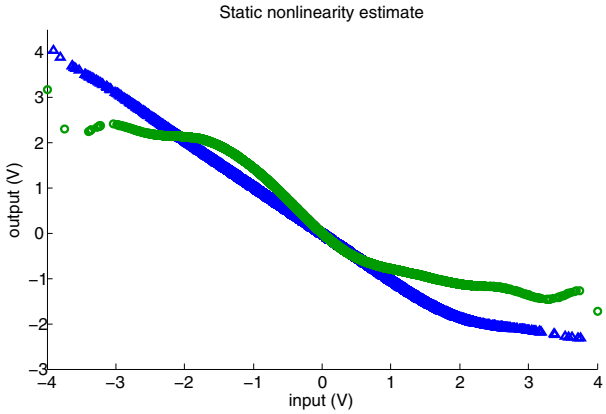

Fig. 4. The estimated static nonlinearity of the first and second branch of the parallel Wiener-Hammerstein model are shown in blue triangles and green circles respectively.

J. D. Carroll and J. J. Chang. Analysis of individual differences in multidimensional scaling via an N-way generalization of "Eckart-Young" decomposition. *Psychometrika*, 35(3):283–319, 1970.

G. Favier and T. Bouilloc. Parametric complexity reduction of Volterra models using tensor decompositions. In *17th European Signal Processing Conference (EUSIPCO)*, pages 2288–2292, Glasgow, Scotland, Aug. 2009.

F. Giri and E.W. Bai, editors. *Block-oriented Nonlinear System Identification*, volume 404 of *Lecture Notes in Control and Information Sciences*. Springer, Berlin Heidelberg, 2010.

R.A Harshman. Foundations of the PARAFAC procedure: Models and conditions for an "explanatory" multimodal factor analysis. *UCLA Working Papers in Phonetics*, 16: 1–84, 1970.

T. G. Kolda and B. W. Bader. Tensor decompositions and applications. *SIAM Rev.*, 51(3):455–500, 2009.

G. Palm. On representation and approximation of nonlinear systems Part II: Discrete Time. *Biological Cybernetics*, 34:49–52, 1979.

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

M. Schoukens and Y. Rolain. Crossterm elimination in parallel Wiener systems using a linear input transformation. *IEEE Trans. Instrum. Meas.*, 61(3):845–847, 2012a.

M. Schoukens and Y. Rolain. Parametric Identification of Parallel Wiener Systems. *IEEE Trans. Instrum. Meas.*, 61(10):2825–2832, 2012b.

M. Schoukens, R. Pintelon, and Y. Rolain. Parametric Identification of Parallel Hammerstein Systems. *IEEE Trans. Instrum. Meas.*, 60(12):3931–3938, 2011.

M. Schoukens, G. Vandersteen, and Y. Rolain. An identification algorithm for parallel Wiener-Hammerstein systems. In *52nd IEEE Conference on Decision and Control (CDC)*, Florence, Italy, Dec. 2013.

L. Sorber, M. Van Barel, and L. De Lathauwer. Optimization-based algorithms for tensor decompositions: canonical polyadic decomposition, decomposition in rank-$(L_r, L_r, 1)$ terms and a new generalization. *SIAM J. Optim.*, 23(2):695–720, 2013a.

L. Sorber, M. Van Barel, and L. De Lathauwer. Tensorlab v1.0. Available online, February 2013b. URL `http://esat.kuleuven.be/sista/tensorlab/`.

K. Tiels and J. Schoukens. From coupled to decoupled polynomial representations in parallel Wiener-Hammerstein models. In *52nd IEEE Conference on Decision and Control (CDC)*, Florence, Italy, Dec. 2013.

E.M. Wysocki and W.J. Rugh. Further results on the identification problem for the class of nonlinear systems Sm. *IEEE Trans. Circuits Syst.*, 23(11):664–670, 1976.