Proceedings 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

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

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.

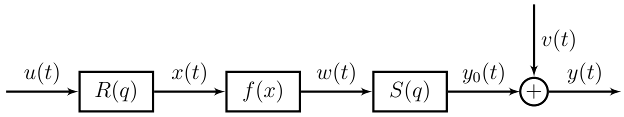

Fig. 1. A Wiener-Hammerstein system (R and S are linear dynamic systems and f is a nonlinear static system).

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