Automatica 60 (2015) 201–209

Contents lists available at ScienceDirect

Automatica

journal homepage: www.elsevier.com/locate/automatica

Brief paper

Initial estimates for Wiener–Hammerstein models using phase-coupled multisines[☆]

Koen Tiels [1], Maarten Schoukens, Johan Schoukens

Department ELEC, Vrije Universiteit Brussel, 1050 Brussels, Belgium

ARTICLE INFO

Article history:
Received 19 May 2014
Received in revised form
13 April 2015
Accepted 7 July 2015

Keywords:
Block-oriented nonlinear system
Dynamic systems
Nonlinear systems
Phase-coupled multisines
System identification
Wiener–Hammerstein model

ABSTRACT

Block-oriented models are often used to model nonlinear systems. These models consist of linear dynamic (L) and nonlinear static (N) sub-blocks. This paper addresses the generation of initial estimates for a Wiener–Hammerstein model (LNL cascade). While it is easy to measure the product of the two linear blocks using a Gaussian excitation and linear identification methods, it is difficult to split the global dynamics over the individual blocks. This paper first proposes a well-designed multisine excitation with pairwise coupled random phases. Next, a modified best linear approximation is estimated on a shifted frequency grid. It is shown that this procedure creates a shift of the input dynamics with a known frequency offset, while the output dynamics do not shift. The resulting transfer function, which has complex coefficients due to the frequency shift, is estimated with a modified frequency domain estimation method. The identified poles and zeros can be assigned to either the input or output dynamics. Once this is done, it is shown in the literature that the remaining initialization problem can be solved much easier than the original one. The method is illustrated on experimental data obtained from the Wiener–Hammerstein benchmark system.

1. Introduction

Even if all physical dynamic systems behave nonlinearly to some extent, we often use linear models to describe them. If the nonlinear distortions get too large, a linear model is insufficient, and a nonlinear model is required.

One possibility is to use block-oriented models (Billings & Fakhouri, 1982; Giri & Bai, 2010), which combine linear dynamic (L) and nonlinear static (N), i.e. memoryless, blocks. Due to this highly structured nature, block-oriented models offer insight about the system to the user. This can be useful in e.g. fault detection, to detect in which part of the system a fault occurred, e.g. changing dynamics in only part of the model. Block-oriented models are preferred when there are localized nonlinearities in the system, thus leading to a sparse representation of the system in terms of interconnected blocks. Due to the separation between the dynamics and the nonlinearities, block-oriented models also allow for an easy discretization (i.e. the conversion from a continuous-time to a discrete-time representation). We refer the reader to Giri and Bai (2010) for an elaborated discussion. The simplest block-oriented models are the Wiener model (LN cascade) and the Hammerstein model (NL cascade). They can be generalized to a Wiener–Hammerstein model (LNL cascade, see Fig. 1). Applications of Wiener–Hammerstein models can mainly be found in biology (Bai, Cai, Dudley-Javorosk, & Shields, 2009; Dewhirst, Simpson, Angarita, Allen, & Newland, 2010; Korenberg & Hunter, 1986), but also in the modeling of RF power amplifiers (Isaksson, Wisell, & Rönnow, 2006).

Several identification methods have been proposed to identify 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). Wiener–Hammerstein systems are modeled as the cascade of well-selected Hammerstein models in Wills and Ninness (2012). The recursive identification of error-in-variables Wiener–Hammerstein systems is considered in Mu and Chen (2014). Both Chen and Fassois (1992) and Wills and Ninness (2012) indicate the importance of good initial estimates, but not how to obtain them. Sjöberg and Schoukens (2012) indicates the importance of good initial es-

☆ The material in this paper was partially presented at the 19th IFAC World Congress, August 24–29, 2014, Cape Town, South Africa. This paper was recommended for publication in revised form by Associate Editor Antonio Vicino under the direction of Editor Torsten Söderström.

E-mail addresses: Koen.Tiels@vub.ac.be (K. Tiels), Maarten.Schoukens@vub.ac.be (M. Schoukens), Johan.Schoukens@vub.ac.be (J. Schoukens).

[1] Tel.: +32 2 6293665; fax: +32 2 6292850.

http://dx.doi.org/10.1016/j.automatica.2015.07.020
0005-1098/

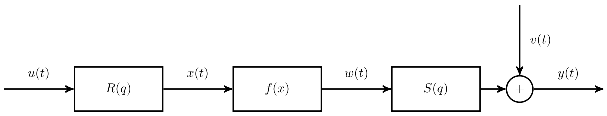

Fig. 1. A Wiener–Hammerstein system (R and S are linear dynamic systems and f is a nonlinear static system).

timates on an example. The optimization of the model parameters can either converge extremely slowly or get trapped in a local optimum, even if the correct number of poles and zeros is assigned to both the input and the output dynamics, leading to Wiener–Hammerstein models that only fit about as well as a linear model.

Some approaches obtain initial estimates by using specifically designed experiments. For example, Vandersteen, Rolain, and Schoukens (1997) proposes a series of experiments with large and small signal multisines. Weiss, Evans, and Rees (1998) uses only two experiments with paired multisines, but the approach requires the estimation of the Volterra kernels of the system. Crama and Schoukens (2005) proposes an iterative initialization scheme that only requires one experiment of a well-designed multisine excitation.

A major difficulty is the generation of good initial values for the two linear blocks $R(q)$ and $S(q)$ of the Wiener–Hammerstein system (see Fig. 1). An initial estimate for the static nonlinearity can be obtained using a simple linear regression if a basis function expansion, linear-in-the-parameters, for the nonlinearity is used, and if the dynamics are initialized. The poles and the zeros of both $R(q)$ and $S(q)$ can be obtained from the best linear approximation (BLA) (Pintelon & Schoukens, 2012) of the Wiener–Hammerstein system. To obtain initial estimates for $R(q)$ and $S(q)$, the poles and the zeros of the BLA should be split over the individual transfer functions $R(q)$ and $S(q)$. Several methods have been proposed to make this split. The brute-force method in Sjöberg, Lauwers, and Schoukens (2012) scans all possible splits, leading to an exponential scanning problem. The advanced method in Sjöberg et al. (2012) uses a basis function expansion for the input dynamics and one for the inverse of the output dynamics. A scanning procedure over the basis functions is proposed as well. Compared to the brute-force method, the number of scans is lower, but the computational time can still be large. The approach in Westwick and Schoukens (2012) not only uses the BLA, but also the so-called quadratic BLA (QBLA), a higher-order BLA from the squared input to the output residual of the first-order BLA. By doing so, the number of possible splits is reduced significantly. Due to the higher-order nature of the QBLA, however, long measurements are needed to obtain an accurate estimate. The nonparametric separation method proposed in Schoukens, Pintelon, and Rolain (2014a) avoids the pole/zero assignment problem completely, but also uses the QBLA.

The method proposed in Schoukens, Tiels, and Schoukens (2014b) and further developed in this paper uses again the first-order BLA. Using a well-designed excitation signal, the poles and the zeros of the input dynamics $R(q)$ shift with a frequency offset that can be chosen by the user, while the poles and the zeros of the output dynamics $S(q)$ remain invariant. Long measurement times can be avoided, because no use is made of higher-order BLAs. This paper generalizes the basic ideas in Schoukens et al. (2014b) from cubic nonlinearities to polynomial nonlinearities. Moreover, experimental results on the Wiener–Hammerstein benchmark system (Schoukens, Suykens, & Ljung, 2009) are reported.

The rest of this paper is organized as follows. The basic setup is described in Section 2. A brief overview of the BLA is presented in Section 3. The proposed method is presented in Section 4. The experimental results on the Wiener–Hammerstein benchmark system are reported in Section 5. Finally, the conclusions are drawn in Section 6.

2. Setup

This section introduces some notation. It also presents the considered Wiener–Hammerstein system and the assumptions.

2.1. Notation

Without loss of generality, discrete-time systems are considered. Hence, the integer t denotes the time as a number of samples. The results in this paper generalize to continuous-time systems with some minor modifications.

Notation 1 *($X(k)$ and $x(t)$)***.** The discrete Fourier transform (DFT) of a time domain signal $x(t)$ is denoted by $X(k) = X(e^{j\omega_k})$, and is given by

$$X(k) = \frac{1}{\sqrt{N}} \sum_{t=0}^{N-1} x(t) e^{-j2\pi \frac{k}{N} t}. \tag{1}$$

The inverse DFT is given by

$$x(t) = \frac{1}{\sqrt{N}} \sum_{k=-N/2+1}^{N/2} X(k) e^{j2\pi \frac{k}{N} t}. \tag{2}$$

Notation 2 (q^{-1})**.** The backward shift operator is denoted by q^{-1}, i.e. $q^{-1}x(t) = x(t-1)$.

Notation 3 ($O(\cdot)$)**.** The notation h is an $O(N^\alpha)$ indicates that for N big enough, $|h(N)| \leq cN^\alpha$, where c is a strictly positive real number.

Notation 4 ($(\cdot)^*$)**.** The complex conjugate of a complex number X is denoted by X^*.

2.2. The Wiener–Hammerstein system

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

$$\begin{aligned} x(t) &= R(q)u(t), \\ w(t) &= f(x(t)), \\ y(t) &= S(q)w(t) + v(t), \end{aligned} \tag{3}$$

where $R(q)$ and $S(q)$ are linear time-invariant (LTI) discrete-time transfer functions, i.e.

$$\begin{aligned} R(q) &= \frac{B_R(q)}{A_R(q)} = \frac{\sum_{l=0}^{n_R} b_{R,l} q^{-l}}{\sum_{l=0}^{m_R} a_{R,l} q^{-l}}, \\ S(q) &= \frac{B_S(q)}{A_S(q)} = \frac{\sum_{l=0}^{n_S} b_{S,l} q^{-l}}{\sum_{l=0}^{m_S} a_{S,l} q^{-l}}, \end{aligned} \tag{4}$$

and where $f(x)$ is a static nonlinear function. Only the input $u(t)$ and the noise-corrupted output $y(t)$ are available for measurement.

2.3. Assumptions

This paper addresses the generation of initial estimates for the linear dynamics $R(q)$ and $S(q)$. To do this, assumptions (A1)–(A4) are made.