2017 IEEE 56th Annual Conference on Decision and Control (CDC)
December 12-15, 2017, Melbourne, Australia

On the identification of Wiener systems with polynomial nonlinearity

Giulio Bottegal, Ricardo Castro-Garcia, Johan A. K. Suykens

Abstract— **In this paper we introduce a new method for Wiener system identification that relies on the data collected on two separate experiments. In the first experiment, the system is excited with a sine signal at fixed frequency and phase shift. Using the steady state response of the system, we estimate the static nonlinearity, which is assumed to be a polynomial. In the second experiment, the system is fed with a persistently exciting input, which allows to identify the linear time-invariant block composing the Wiener structure. We show that the estimation of the static nonlinearity reduces to the solution of a least squares problem, and we provide an expression for the asymptotic variance of the estimated polynomial coefficients. The effectiveness of the method is demonstrated through numerical experiments.**

I. Introduction

The Wiener system is a cascaded system composed of a linear time-invariant (LTI) system followed by a static nonlinear function [1]. Wiener systems find application in different areas of science and engineering, e.g., chemical processes [2], [3], and biological systems [4]. Furthermore, Wiener models can approximate any nonlinear system with arbitrarily high accuracy [5].

Identification of Wiener systems has been object of research for many years, see [6]; a brief overview is reported in the following. Maximum likelihood identification is analyzed in [7], showing that estimation of the parameters requires the solution of a nonlinear optimization problem. The complexity of the nonlinear problem can be reduced using separable least-squares [8] or recursive identification schemes [9]. In [10] and [11], authors discuss semi-parametric approaches relying on a Bayesian model of the static nonlinearity and on a parametric model of the LTI block. In particular, [10] proposes to estimate the LTI block and the static nonlinearity via a joint maximum-a-posteriori/maximum-likelihood criterion, requiring the solution of a nonlinear optimization problem. The method proposed in [11] provides a minimum mean square estimate of the system using Markov Chain Monte Carlo techniques, which turn out computationally demanding. Several contributions discuss fully nonparametric kernel methods [12], [13], [14], [15], relying on the assumption that the input is an i.i.d. sequence. Other approaches rely on frequency-domain techniques [16], instrumental variables [17], Wiener G-functionals [18], and subspace methods [19], while some other techniques are tailored to special cases of Wiener systems, e.g. systems with quantized output data [20]. Experiment design techniques specifically tailored for Wiener system identification are discussed in [21] and [22].

The subject of this paper is to propose a new identification procedure for Wiener systems that reduces the computational burden of maximum likelihood/prediction error techniques by separating the identification of the static nonlinearity from the identification of the LTI block. We assume that the user has the freedom to design two different experiments where the system is fed with specific inputs. Moreover, we assume that the static nonlinearity can be well represented by a polynomial of known order.

The first experiment of the procedure consists in feeding the system with a sine signal, with user-choice frequency and phase delay. In this experiment, the focus is on the nonlinearity. We show that the estimation of the polynomial coefficients is obtained following a procedure that essentially reduces to a least-squares estimation. As part of the contributions of the paper, we provide an expression for the asymptotic variance of the estimated coefficients, showing that this does not depend on the frequency and phase of the input sine. The second experiment uses the estimated polynomial to identifiy the LTI block by means of a modified version of the standard prediction error method (PEM) for linear output-error (OE) models [23]. Here the system is fed with a persistently exciting input. In this way, the computational burden of this second step reduces to the one of PEM for OE models. The proposed method is tested via numerical simulations that show its effectiveness compared to PEM for Wiener systems. We note that the idea of feeding a Wiener system with a sine signal was also explored in previous work. In [16], the phase delay introduced by the LTI block is estimated by comparing the frequency content of the output and the input. In [24], the phase delay is estimated using a geometric approach. The method presented in this paper only requires estimating the sign of the phase delay.

The method presented in this paper is a special case – tailored to polynomial nonlinearities– of the two-experiment approach to Wiener system identification discussed in [25].

The paper is organized as follows. In Section II we define

The research leading to these results has received funding from the European Research Council under the European Unions Seventh Framework Programme (FP7/2007-2013) / ERC AdG A-DATADRIVE-B (290923), and under the Advanced Research Grant SYSDYNET, under the European Union's Horizon 2020 research and innovation programme (grant agreement No 694504). This paper reflects only the authors' views and the Union is not liable for any use that may be made of the contained information. Research Council KUL: CoE PFV/10/002 (OPTEC), BIL12/11T; PhD/Postdoc grants Flemish Government: FWO: projects: G.0377.12 (Structured systems), G.088114N (Tensor based data similarity); PhD/Postdoc grant iMinds Medical Information Technologies SBO 2015 IWT: POM II SBO 100031 Belgian Federal Science Policy Office: IUAP P7/19 (DYSCO, Dynamical systems, control and optimization, 2012-2017)

Giulio Bottegal (g.bottegal@tue.nl) is with the Department of Electrical Engineering, TU Eindhoven, The Netherlands, and with KU Leuven, ESAT-STADIUS, Leuven, Belgium. Ricardo Castro-Garcia (ricardo.castro@esat.kuleuven.be) and Johan A.K. Suykens (johan.suykens@esat.kuleuven.be) are with KU Leuven, ESAT-STADIUS, Belgium.)

the problem under study. Section III describes the proposed method including for polynomial nonlinearity identification. Section IV illustrates the results of some numerical experiments. Section V gives some conclusions.

II. Wiener system identification using two experiments

A. The Wiener system

We consider the following single-input single-output (SISO) system, also called a Wiener system (see Fig. 1 for a schematic representation of Wiener systems):

$$\begin{aligned} x_t &= G(q^{-1})u_t \\ y_t &= f(x_t) + e_t\,. \end{aligned} \tag{1}$$

In the former equation, $G(q^{-1})$ represents the transfer function of a causal LTI subsystem, driven by the input u_t, where q^{-1} denotes the time shift operator, namely $q^{-1}u_t = u_{t-1}$. In the latter equation, y_t is the result of a static nonlinear transformation, denoted by $f(\cdot)$, of the signal x_t, and e_t is white noise with unknown variance σ^2. The problem under study is to estimate the LTI subsystem and the nonlinear function from a set of input and output measurements. We shall consider the following standing assumptions throughout the paper.

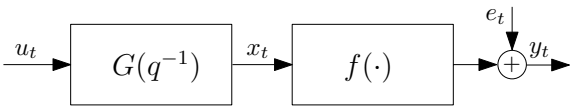


Fig. 1. Block scheme representation of a Wiener system.

Assumption 1: The transfer function of the LTI subsystem admits the parametric representation

$$G(q^{-1}) = \frac{b_0 + b_1 q^{-1} + \ldots + b_m q^{-m}}{1 + a_1 q^{-1} + \ldots + a_n q^{-n}}\,, \tag{2}$$

where the polynomial orders m and n are known.

Assumption 2: The static nonlinearity is a polynomial of known order p, i.e.

$$f(x) = c_0 + c_1 x + \ldots + c_p x^p\,. \tag{3}$$

Therefore, the identification problem we discuss in this paper is to estimate the parameter vectors $\theta := \begin{bmatrix} b_0 & b_1 & \ldots & b_m & a_1 & \ldots & a_n \end{bmatrix}$ and $c := \begin{bmatrix} c_0 & c_1 & \ldots & c_p \end{bmatrix}$.

B. Identification via a two-experiment procedure

In this paper we consider the case where the user has the freedom to design the input signal u_t. In particular, we are interested in the case where the user has the possibility to run two separate experiments, each having a particular signal u_t as input. The goal of this paper is to describe an identification technique for the system (1) that is linked to a particular choice of these experiments. The proposed procedure consists of the two following steps:

1) Feed the system with a sinusoid at a prescribed frequency. Use the steady-state data to estimate the nonlinear function $f(\cdot)$.
2) Feed a system with a persistently exciting input signal and identify the LTI subsystem using the information gathered on the first step regarding the static nonlinearity.

We briefly describe the second step of the proposed procedure, assuming that an estimate of the polynomial coefficients, $\hat{c}$, is available after the first step of the procedure, we can set up a PEM-based identification criterion as follows

$$\hat{\theta} = \arg\min_{\theta} \frac{1}{N_2} \sum_{t=1}^{N_2} \left(y_t - \sum_{i=0}^{p} \hat{c}_i (G(q^{-1})u_t)^i \right)^2\,, \tag{4}$$

where N_2 is the number of samples collected during the second experiment. Note that this is a mild generalization of the standard PEM, requiring only to account, in the optimization process, for the known polynomial part. This does not make the solution of (4) harder than a standard PEM applied to an output-error model, because in both cases we have to face a nonlinear optimization problem, and in both cases gradient-based methods can be easily applied [23].

As opposed to the aforementioned second step, the first step can be more involved and requires a more thorough analysis. We shall focus on this step in the remainder of the paper.

III. Estimating the nonlinearity

In this section we discuss the first step of the procedure, namely the estimation of the static nonlinearity.

Our approach is based on feeding the system with the following input signal

$$u_t = \sin(\omega t + \phi_0)\,,$$

where ω is an user-prescribed frequency and ϕ_0 is a known phase delay which we can assume equal to 0 without loss of generality. Then, after the transient effect of $G(q^{-1})$ has vanished, we have that

$$x_t = A_\omega \sin(\omega t + \phi_\omega)\,,$$

where $A_\omega > 0$ and $\phi_\omega \leq 0$ are the gain and the phase delay of the LTI subsystem $G(q^{-1})$ at the frequency ω [23, Ch. 2]. Due to the structural non-identifiability of Wiener systems, A_ω can not be determined (see Remark 1 below). We thus drop it and define a new signal

$$\bar{x}_t = \sin(\omega t + \phi_\omega)\,,$$

which is parameterized by the unknown quantity ϕ_ω. Accordingly, we write the output of the system as

$$\begin{aligned} y_t &= f(\sin(\omega t + \phi_\omega)) + e_t & (5)\\ &= \sum_{i=0}^{p} c_i \sin^i(\omega t + \phi_\omega) + e_t\,. & (6) \end{aligned}$$

Then, the problem under study, that is to estimate the vector c, is coupled with the problem of estimating ϕ_ω.

We note that there is an ambiguity in determining ϕ_ω, since

$$\sin(\omega t + \phi_\omega) = (-1)^k \sin(\omega t + \phi_\omega + k\pi),$$

for any $k \in \mathbb{Z}_-$. Therefore, we have to restrict our domain of search of ϕ_ω, as stated by the following assumption.

Assumption 3: Let $\mathcal{I} = (-\pi, 0]$. The phase delay induced by the LTI system at frequency ω is such that $\phi_\omega \in \mathcal{I}$.

In the following subsection, we describe our approach to tackle to the problem of estimating the polynomial nonlinearity assuming that the number of collected samples of y_t (at its steady state) is equal to N_1.

Remark 1: Since we are estimating the static nonlinearity using the signal $\bar{x}_t$ instead of x_t, we are obtaining a scaled (in the x-axis) version of $f(\cdot)$, that is, we are estimating $f(x/A_\omega)$ instead of $f(x)$. This scaling effect is compensated in the second phase of the method, because (4) will return the estimate $A_\omega G(q^{-1})$ instead of $G(q^{-1})$. Then, we need additional information (e.g., on the LTI system gain, see [26]) to uniquely recover $G(q^{-1})$ and $f(\cdot)$; this lack of identifiability is a well known issue in block oriented system identification. However, if the focus is on output prediction (as is in the experiments of Section IV), rescaling of the two blocks is not required.

Remark 2: In practice, the quantity A_ω crucially determines the quality of the identified static nonlinearity. Because we are assuming additive white noise, the average signal-to-noise ratio (SNR) is $A_\omega^2/2\sigma^2$; then optimal results are obtained where A_ω is maximum. However, A_ω is not known in advance, and to determine its largest value one would have to perform a preliminary experiment swiping the frequency spectrum to detect where $|G(\omega)|$ is maximum.

A. The proposed estimation procedure

For ease of exposition, we assume that p is even (the case p odd runs along the same lines of reasoning). We recall that (see e.g. [16]), for $i \in \mathbb{N}$,

$$\sin^{2i}(x) = \frac{1}{2^{2i}}\binom{2i}{i} + \frac{(-1)^i}{2^{2i-1}}\sum_{k=0}^{i-1}(-1)^k\binom{2i}{k}\cos(2(i-k)x),$$

$$\sin^{2i+1}(x) = \frac{(-1)^i}{4^i}\sum_{k=0}^{i}(-1)^k\binom{2i+1}{k}\sin((2i+1-2k)x).$$

Combining these expressions with standard trigonometric addition formulas, we can rewrite the nonlinear function as

$$\begin{aligned} f(\bar{x}_t) = \sum_{i=0}^{p/2} c_{2i}\Big[&\frac{1}{2^{2i}}\binom{2i}{i} + \frac{(-1)^i}{2^{2i-1}}\sum_{k=0}^{i-1}(-1)^k\binom{2i}{k} \\ &\times [\cos(2(i-k)\phi_\omega)\cos(2(i-k)\omega t) \\ &- \sin(2(i-k)\phi_\omega)\sin(2(i-k)\omega t)]\Big] \\ + \sum_{i=0}^{p/2-1} c_{2i+1}&\frac{(-1)^i}{4^i}\sum_{k=0}^{i}(-1)^k\binom{2i+1}{k} \\ &\times [\cos((2i+1-2k)\phi_\omega)\sin((2i+1-2k)\omega t) \\ &+ \sin((2i+1-2k)\phi_\omega)\cos((2i+1-2k)\omega t)]\,. \end{aligned} \tag{7}$$

This equation is particularly interesting because it permits to express $f(\bar{x}_t)$ as a linear combination of sines and cosines with frequency $q\omega_t$, where $q = 0, \ldots, p$. The coefficients of this linear combination are also a function of sines and cosines of the unknown phase delay ϕ_ω. Let

$$k_0 := \sum_{i=0}^{p/2} c_{2i}\frac{1}{2^{2i}}\binom{2i}{i}$$

$$k_{s,q} := \sum_{i=q/2}^{p/2} c_{2i}\frac{1}{2^{2i-1}}\binom{2i}{i-q/2}\sin(q\phi_\omega) \qquad (q \text{ even})$$

$$k_{c,q} := -\sum_{i=q/2}^{p/2} c_{2i}\frac{1}{2^{2i-1}}\binom{2i}{i-q/2}\cos(q\phi_\omega) \qquad (q \text{ even})$$

$$\begin{aligned} k_{s,q} := (-1)^{2i-(q-1)/2}&\sum_{i=(q-1)/2}^{p/2-1} c_{2i+1}\frac{1}{4^i}\binom{2i+1}{i-(q-1)/2} \\ &\times \cos(q\phi_\omega) \end{aligned} \qquad (q \text{ odd})$$

$$\begin{aligned} k_{c,q} := (-1)^{2i-(q-1)/2}&\sum_{i=(q-1)/2}^{p/2-1} c_{2i+1}\frac{1}{4^i}\binom{2i+1}{i-(q-1)/2} \\ &\times \sin(q\phi_\omega) \end{aligned} \qquad (q \text{ odd})$$

Let also, for $q = 1, \ldots, p$,

$$k_q = \sum_{i=q/2}^{p/2} c_{2i}\frac{1}{2^{2i-1}}\binom{2i}{i-q/2}, \qquad (q \text{ even})$$

$$k_q = \sum_{i=(q-1)/2}^{p/2-1} c_{2i+1}\frac{1}{4^i}\binom{2i+1}{i-(q-1)/2}, \qquad (q \text{ odd})$$

so that there exists a matrix $M \in \mathbb{R}^{p+1\times p+1}$ such that

$$\mathbf{k} = Mc, \quad \mathbf{k} := \begin{bmatrix} k_0 & \ldots & k_p \end{bmatrix}^T. \tag{8}$$

Then we can write

$$\begin{aligned} \bar{\mathbf{k}} &= \begin{bmatrix} k_0 & k_{s,1} & k_{c,1} & \ldots & k_{s,p} & k_{c,p} \end{bmatrix}^T \\ &= \begin{bmatrix} \gamma_0 k_0 & \gamma_{s,1} k_1 \sin(\phi_\omega) & \gamma_{c,1} k_1 \cos(\phi_\omega) & \ldots \\ \ldots & \gamma_{s,p} k_p \sin(p\phi_\omega) & \gamma_{c,p} k_p \cos(p\phi_\omega) \end{bmatrix}^T, \end{aligned} \tag{9}$$

where γ_0 and the $\gamma_{s,i}$, $\gamma_{c,i}$, $i = 1, \ldots, p$, are equal to -1 or 1 and are known in advance. Therefore, if we are able to estimate the vector $\bar{\mathbf{k}}$ and the phase delay ϕ_ω, we can also estimate $\mathbf{k}$ and consequently c. To this end, we define

$$\psi_t^T := [1 \;\; \cos(\omega t) \;\; \sin(\omega t) \;\ldots\; \sin(p\omega t) \;\; \cos(p\omega t)], \tag{10}$$

so that we can rewrite (7) as $f(\bar{x}_t) = \psi_t^T \bar{\mathbf{k}}$, and express the measurement equation via the linear regression model

$$y_t = \psi_t^T \bar{\mathbf{k}} + e_t\,,$$

which also represents the Fourier expansion of the output signal. We can then obtain the least-squares estimate

$$\widehat{\bar{\mathbf{k}}} = \left(\sum_{t=1}^{N_1}\psi_t\psi_t^T\right)^{-1}\sum_{t=1}^{N_1}\psi_t y_t\,. \tag{11}$$

Using this estimate, we now show how to recover the coefficients of the polynomial. From (9), the absolute value of each k_i, $i = 1, \ldots, p$, can be reconstructed via

$$|\hat{k}_i| = \sqrt{\hat{k}_{s,i}^2 + \hat{k}_{c,i}^2}\,. \tag{12}$$

Using the estimates $\hat{k}_{s,1}$ and $\hat{k}_{c,1}$, one can recover the phase delay ϕ_ω as

$$\hat{\phi}_\omega = \tan^{-1}\left(\frac{\hat{k}_{s,1}}{\hat{k}_{c,1}}\right), \tag{13}$$

where uniqueness of the solution is guaranteed in $\mathcal{I} = (-\pi, 0]$. Using $\hat{\phi}_\omega$, we can uniquely recover the sign of the coefficients k_i, exploiting the knowledge of the coefficients $\gamma_{s,i}$, $\gamma_{c,i}$ introduced in (9). Finally, using the matrix M defined in (8), the estimate $\hat{c}$ can be recovered via

$$\hat{c} = M^{-1}\hat{\mathbf{k}}\,, \tag{14}$$

where M is introduced in (8).

We summarize the procedure for the estimation of the polynomial nonlinearity in Algorithm 1.

Algorithm 1 Polynomial static nonlinearity estimation

Input: $\{y_t\}_{t=1}^{N_1}$, ω
Output: $\hat{c}_0, \ldots, \hat{c}_p$

1: Construct the regressors (10) and compute the least-squares estimate (11)
2: Recover the absolute value of the coefficients k_i, $i = 0, \ldots, p$ using (12)
3: Estimate the phase shift via (13) and the sign of the coefficients k_i, $i = 0, \ldots, p$
4: Recover the coefficients c_i, $i = 0, \ldots, p$, using (14)

We observe that this procedure requires to recover the phase delay ϕ_ω. However, this can be done with one simple operation, namely (13). Moreover, it is not required to have a highly accurate estimate of ϕ_ω. In fact, what we need is just that $\hat{\phi}_\omega$ lies in the same orthant of ϕ_ω, so that we can correctly determine the sign of the coefficients k_i, $i = 0, \ldots, p$. As for the asymptotic performance of this estimation procedure, we have the following result showing that the asymptotic variance of the procedure *does not depend on the choice of* ω.

Proposition 1: The procedure outlined above gives consistent estimates of the coefficients c_i, $i = 0, \ldots, p$. Furthermore, the asymptotic covariance of their estimates is equal to

$$\frac{\sigma^2}{N_1} M^{-1} D M^{-T}\,, \tag{15}$$

where $D = \operatorname{diag}\{1, 2, \ldots, 2\}$.

Proof: See the Appendix.

Remark 3: A similar approach to Wiener system identification with polynomial nonlinearity was proposed in [16, Sec. 4]. The main difference between that approach and the one proposed in this paper lies in the way the phase delay is estimated. In [16] the phase delay is estimated by comparing the DFTs of the output and the input, leading to a non-consistent phase estimate [27]. An additional advantage of the method proposed in this paper is that it is required to estimate only the sign of the phase delay, arguably leading to estimates of nonlinearity coefficients with higher accuracy.

Example 1: Consider the third-order polynomial nonlinearity

$$f(x) = c_0 + c_1 x + c_2 x^2 + c_3 x^3\,.$$

We have

$$\begin{aligned} f(\bar{x}_t) = c_0 &+ \frac{c_2}{2} + (c_1 + \frac{3}{4}c_3)\cos(\phi_\omega)\sin(\omega t) \\ &+ (c_1 + \frac{3}{4}c_3)\sin(\phi_\omega)\cos(\omega t) \\ &+ \frac{c_2}{2}\sin(2\phi_\omega)\sin(2\omega t) - \frac{c_2}{2}\cos(2\phi_\omega)\cos(2\omega t) \\ &- \frac{c_3}{4}\cos(3\phi_\omega)\sin(3\omega t) - \frac{c_3}{4}\sin(3\phi_\omega)\cos(3\omega t)\,, \end{aligned}$$

and

$$\begin{bmatrix} \gamma_0 \\ \gamma_{s,1} \\ \gamma_{c,1} \\ \gamma_{s,2} \\ \gamma_{c,2} \\ \gamma_{s,3} \\ \gamma_{c,3} \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ -1 \\ -1 \\ -1 \end{bmatrix}, \; M = \begin{bmatrix} 1 & 0 & 1/2 & 0 \\ 0 & 1 & 0 & 3/4 \\ 0 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 1/4 \end{bmatrix}.$$

The asymptotic covariance of the coefficient estimates, computed via (15), is then

$$\frac{\sigma^2}{N_1}\begin{bmatrix} 3 & 0 & -4 & 0 \\ 0 & 20 & 0 & -24 \\ -4 & 0 & 8 & 0 \\ 0 & -24 & 0 & 32 \end{bmatrix},$$

which shows that, for the case $p = 3$, the odd coefficients are estimated with lower accuracy. Not surprisingly, the estimates of the even coefficients and the estimates of the odd coefficients are uncorrelated.

IV. Numerical experiments

The proposed approach is tested on two simulated Wiener systems, which we refer to as S1 and S2. The LTI block of S1 is obtained using the Matlab command `cheby2(3,5,0.2)`, which returns a third-order system with stopband edge frequency 0.2π and 5 dB of stopband attenuation at the passband value. The static nonlinearity is the third-order polynomial $f(x) = x^3$. The parameters characterizing S1 are then

$$\begin{aligned} a &= \begin{bmatrix} 2.46 & 2.26 & -0.77 \end{bmatrix}, \\ b &= 10^{-2}\begin{bmatrix} 0.47 & 1.42 & 1.42 & 0.47 \end{bmatrix}, \\ c &= \begin{bmatrix} 0 & 0 & 0 & 1 \end{bmatrix}. \end{aligned}$$

As for S2, the LTI part is obtained using the Matlab command `cheby2(4,18,0.2)`, while the nonlinear part is the third-order polynomial $f(x) = x + 5x^2 - 0.5x^3$. Thus, S2 is described by the parameters

$$a = \begin{bmatrix} -2.49 & 2.53 & -1.18 & 0.22 \end{bmatrix},$$

$$b = \begin{bmatrix} 0.11 & -0.21 & 0.28 & -0.21 & 0.11 \end{bmatrix},$$
$$c = \begin{bmatrix} 0 & 1 & 5 & -0.5 \end{bmatrix}.$$

The two systems are depicted in Fig. 2. The variance of the output noise is obtained by matching different signal-to-noise ratios (SNR). In particular, we test the proposed methods for SNR $= 10$, 20, 40 dB. Thus, we have in total 6 experimental conditions. For each experimental condition we generate 100 Monte Carlo runs, each with a different noise realization. The performance of the methods is evaluated by assessing the accuracy in tracking the output of a noiseless test set $y_{t,\text{test}}$ of length $N = 500$, obtained by feeding the systems with an i.i.d. Gaussian sequence. We use the normalized mean absolute error (NMAE), defined as

$$\%\text{NMAE} = \frac{100}{N} \frac{\sum_{t=1}^{N} |y_{t,\text{test}} - \hat{y}_{t,\text{test}}|}{|\max(y_{t,\text{test}}) - \min(y_{t,\text{test}})|}, \qquad (16)$$

and the normalized mean square error (NMSE), defined as

$$\%\text{NMSE} = 100 \frac{\sqrt{\sum_{t=1}^{N} \left(y_{t,\text{test}} - \hat{y}_{t,\text{test}}\right)^2}}{\sqrt{\sum_{t=1}^{N} \left(y_{t,\text{test}} - \text{mean}(y_{t,\text{test}})\right)^2}}. \qquad (17)$$

To test our method, which we refer to as **2-E** (two-experiment), we first feed the system with the signal $u_t^1 = \sin(\omega t)$, where ω is equal to 0.02 or 0.05, keeping $N_1 = 500$ transient-free samples. In the second stage, we use random white Gaussian noise with unit variance as input u_t^2, collecting $N_2 = 500$ samples. The method is compared with a single-stage PEM-based estimator, implemented through the Matlab command `nlhw` (see [28] for details); we refer to this method as **NLHW**. To get a fair comparison, the data used by **NLHW** are obtained by feeding the system with a random white Gaussian sequence of length $N_1 + N_2$ with variance equal to the variance of the sequence $[u_1^1 \ \dots \ u_{N_1}^1 \ u_1^2 \ \dots \ u_{N_2}^2]$.

In Table I we report the medians of the results obtained in the 6 experiments. As can be seen, on S1 the performance of the three methods are comparable to each other, while **NLHW** fails in providing a good model of S2. The motivation is that the method often falls into a local minimum of the cost function related to the PEM optimization problem. Separating the identification of the LTI block from the static nonlinearity avoids this issue; thus, the proposed method **2-E** gives reliable results for both S1 and S2, and for both the choices of ω.

TABLE I
MEDIAN VALUES OF %NMAE AND %NMSE OVER 100 MONTE CARLO RUNS OBTAINED ON THE 6 TESTED EXPERIMENTAL CONDITIONS.

SNR (dB)	Method	S1		S2	
		NMAE	NMSE	NMAE	NMSE
10	**2-E** $\omega = 0.02$	0.21	5.49	0.54	8.27
	2-E $\omega = 0.05$	0.2	5.1	0.49	6.87
	NLHW	0.19	5.28	3.3	39.63
20	**2-E** $\omega = 0.02$	0.18	4.18	0.35	5.44
	2-E $\omega = 0.05$	0.16	4.07	0.33	4.67
	NLHW	0.16	3.71	3.72	51.12
40	**2-E** $\omega = 0.02$	0.14	3.65	0.28	4.23
	2-E $\omega = 0.05$	0.13	3.33	0.25	3.49
	NLHW	0.09	2.25	3.61	44.35

V. CONCLUSIONS

We have proposed a new method for Wiener system identification. The main idea underlying the method is that we can separate the estimation of the static nonlinear function from the identification of the LTI block composing the Wiener system. To do so, we have to excite the system using two different inputs. The first input is a sinusoid, which, after the transient effect of the LTI system has vanished, permits to estimate the coefficients of the polynomial constituting the static nonlinearity by a relatively simple least-squares-based procedure. Using the information on the static nonlinearity, we use a persistently exciting input to identify the LTI block. The proposed method is shown to compare favorably with a PEM-based method for Wiener system identification.

Future challenges are to extend the two-experiment approach to more involved model structures, such as Wiener-Hammerstein and Hammerstein-Wiener systems.

APPENDIX

A. Proof of Proposition 1

Let

$$\Gamma = \begin{bmatrix} \gamma_0 & 0 & & \dots & 0 \\ 0 & \gamma_{s,\,1} \sin(\phi_\omega) & 0 & \dots & 0 \\ 0 & \gamma_{c,\,1} \cos(\phi_\omega) & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & \dots & & 0 & \gamma_{s,\,p} \sin(p\phi_\omega) \\ 0 & \dots & & 0 & \gamma_{c,\,p} \cos(p\phi_\omega) \end{bmatrix}, \qquad (18)$$

so that (9) can be rewritten as $\bar{\mathbf{k}} = \Gamma \mathbf{k}$. We note that the Moore-Penrose pseudoinverse of Γ is $\Gamma^{\#} = (\Gamma^T \Gamma)^{-1} \Gamma^T = \Gamma^T$, so that $\mathbf{k} = \Gamma^T \bar{\mathbf{k}}$.

Since we do not need to estimate ϕ_ω but only its sign, we can conclude that the covariance matrix of the estimated coefficients c_i, $i = 0, \dots, p$ can be calculated as the same as the one obtained via least-squares, that is

$$\begin{aligned} \mathbb{E}[(\hat{c} - c)(\hat{c} - c)^T] &= M^{-1} \mathbb{E}[(\hat{\mathbf{k}} - \mathbf{k})(\hat{\mathbf{k}} - \mathbf{k})^T] M^{-T} \\ &= M^{-1} \Gamma^T \mathbb{E}[(\hat{\bar{\mathbf{k}}} - \bar{\mathbf{k}})(\hat{\bar{\mathbf{k}}} - \bar{\mathbf{k}})^T] \Gamma M^{-T} \\ &= \sigma^2 M^{-1} \Gamma^T \left(\sum_{t=1}^{N_1} \psi_t \psi_t^T \right)^{-1} \Gamma M^{-T}. \end{aligned}$$

Recalling the structure of ψ_t given in (10), it is straightforward to check that, as N_1 grows large,

$$\frac{1}{N_1} \sum_{t=1}^{N_1} \psi_t \psi_t^T \longrightarrow \text{diag} \left\{ 1, \frac{1}{2}, \dots, \frac{1}{2} \right\} := \bar{D}^{-1}, \qquad (19)$$

where $\bar{D}$ has size $2p + 1 \times 2p + 1$. Equation (15) follows from the fact that

$$\Gamma^T \bar{D} \Gamma = \Gamma^T \Gamma D = D.$$

The consistency of the estimates follows from the consistency of the least-squares estimates (11). This concludes the proof.

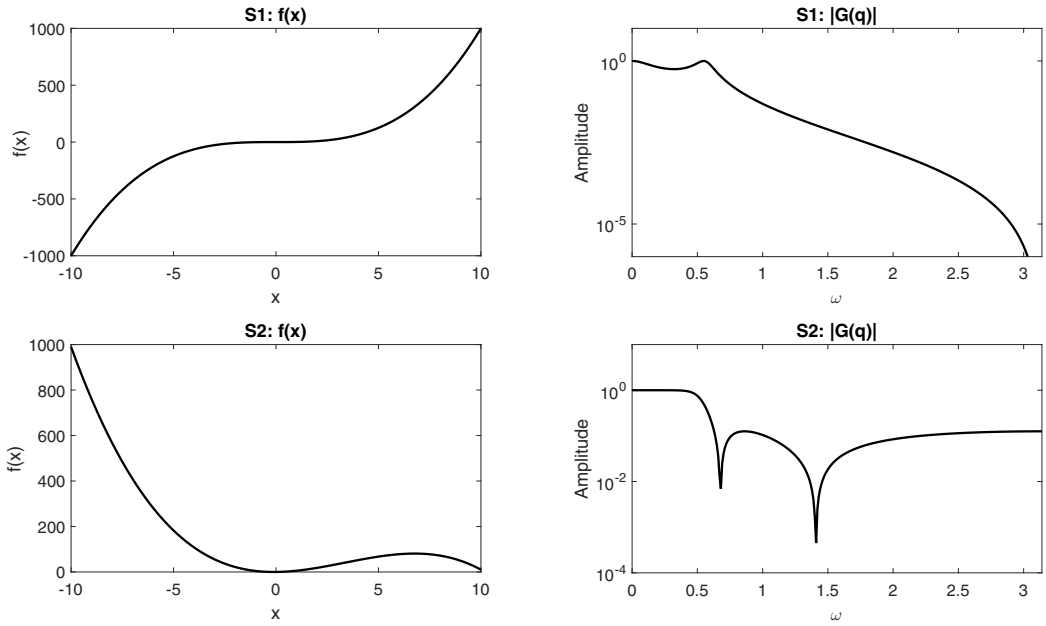


Fig. 2. Blocks composing S1 and S2. (Up) S1. (Down) S2. (Left) Nonlinear blocks. (Right) Magnitude of the transfer functions of the LTI blocks.

REFERENCES

[1] N. Wiener, "Nonlinear problems in random theory," *Nonlinear Problems in Random Theory. Cambridge, Massachusetts, USA: The MIT Press, August 1966.(Paper)*, p. 142, 1966.

[2] Y. Zhu, "Distillation column identification for control using Wiener model," in *American Control Conference, 1999. Proceedings of the 1999*, vol. 5. IEEE, 1999, pp. 3462–3466.

[3] A. Kalafatis, N. Arifin, L. Wang, and W. Cluett, "A new approach to the identification of pH processes based on the Wiener model," *Chemical Engineering Science*, vol. 50, no. 23, pp. 3693–3701, 1995.

[4] I. Hunter and M. Korenberg, "The identification of nonlinear biological systems: Wiener and Hammerstein cascade models," *Biological cybernetics*, vol. 55, no. 2-3, pp. 135–144, 1986.

[5] S. Boyd and L. Chua, "Fading memory and the problem of approximating nonlinear operators with Volterra series," *IEEE Transactions on circuits and systems*, vol. 32, no. 11, pp. 1150–1161, 1985.

[6] F. Giri and E.-W. Bai, Eds., *Block-oriented nonlinear system identification*. Springer, 2010, vol. 1.

[7] A. Hagenblad, L. Ljung, and A. Wills, "Maximum likelihood identification of Wiener models," *Automatica*, vol. 44, no. 11, pp. 2697–2705, 2008.

[8] J. Bruls, C. Chou, B. Haverkamp, and M. Verhaegen, "Linear and nonlinear system identification using separable least-squares," *European Journal of Control*, vol. 5, no. 1, pp. 116–128, 1999.

[9] T. Wigren, "Recursive prediction error identification using the nonlinear Wiener model," *Automatica*, vol. 29, no. 4, pp. 1011–1025, 1993.

[10] G. Pillonetto, "Consistent identification of Wiener systems: A machine learning viewpoint," *Automatica*, vol. 49, no. 9, pp. 2704–2712, 2013.

[11] F. Lindsten, T. B. Schön, and M. I. Jordan, "Bayesian semiparametric Wiener system identification," *Automatica*, vol. 49, no. 7, pp. 2053–2063, 2013.

[12] W. Greblicki, "Nonparametric approach to Wiener system identification," *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, vol. 44, no. 6, pp. 538–545, 1997.

[13] ——, "Nonparametric identification of Wiener systems," *IEEE Transactions on information theory*, vol. 38, no. 5, pp. 1487–1493, 1992.

[14] P. Wachel and G. Mzyk, "Direct identification of the linear block in Wiener system," *International Journal of Adaptive Control and Signal Processing*, vol. 30, no. 1, pp. 93–105, 2016.

[15] G. Mzyk, "A censored sample mean approach to nonparametric identification of nonlinearities in Wiener systems," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 54, no. 10, pp. 897–901, 2007.

[16] E.-W. Bai, "Frequency domain identification of Wiener models," *Automatica*, vol. 39, no. 9, pp. 1521–1530, 2003.

[17] A. Janczak, "Instrumental variables approach to identification of a class of MIMO Wiener systems," *Nonlinear Dynamics*, vol. 48, no. 3, pp. 275–284, 2007.

[18] K. Tiels and J. Schoukens, "Identifying a Wiener system using a variant of the Wiener G-functionals," in *Decision and Control and European Control Conference (CDC-ECC), 2011 50th IEEE Conference on*. IEEE, 2011, pp. 5780–5785.

[19] D. Westwick and M. Verhaegen, "Identifying MIMO Wiener systems using subspace model identification methods," *Signal Processing*, vol. 52, no. 2, pp. 235–258, 1996.

[20] G. Bottegal, H. Hjalmarsson, and G. Pillonetto, "A new kernel-based approach to system identification with quantized output data," *Automatica*, vol. 85, pp. 145–152, 2017.

[21] M. Gevers, M. Caenepeel, and J. Schoukens, "Experiment design for the identification of a simple Wiener system," in *Decision and Control (CDC), 2012 IEEE 51st Annual Conference on*. IEEE, 2012, pp. 7333–7338.

[22] K. Mahata, J. Schoukens, and A. De Cock, "Information matrix and D-optimal design with Gaussian inputs for Wiener model identification," *Automatica*, vol. 69, pp. 65–77, 2016.

[23] L. Ljung, *System identification : theory for the user*, ser. Prentice Hall information and system sciences series. Upper Saddle River (NJ): Prentice Hall PTR, 1999.

[24] F. Giri, Y. Rochdi, and F.-Z. Chaoui, "An analytic geometry approach to Wiener system frequency identification," *IEEE Transactions on Automatic Control*, vol. 54, no. 4, pp. 683–696, 2009.

[25] G. Bottegal, R. Castro-Garcia, and J. A. K. Suykens, "A two-experiment approach to Wiener system identification," *Automatica (submitted for possible publication)*, 2017.

[26] E. W. Bai, "An optimal two stage identification algorithm for Hammerstein-Wiener nonlinear systems," *Automatica*, vol. 34, pp. 333–338, 1998.

[27] F. Giri, Y. Rochdi, and F. Chaoui, "Comment on Frequency domain identification of Wiener models, by EW Bai, Automatica 39 (2003), 1521–1530," *Automatica*, vol. 44, no. 5, pp. 1451–1455, 2008.

[28] L. Ljung, Q. Zhang, P. Lindskog, A. Iouditski, and R. Singh, "An integrated system identification toolbox for linear and nonlinear models," in *Proceedings of the 4th IFAC Symposium on System Identification, Newcastle, Australia*, 2006.