www.ietdl.org

Published in IET Control Theory and Applications
Received on 14th January 2010
Revised on 26th July 2010
doi: 10.1049/iet-cta.2010.0023

ISSN 1751-8644

Semi-parametric identification of Wiener systems using a single harmonic input and retrospective cost optimisation

A.M. D'Amato[1] *B.O.S. Teixeira*[2] *D.S. Bernstein*[1]

[1]*Department of Aerospace Engineering, University of Michigan, Ann Arbor, MI, USA*
[2]*Department of Electronic Engineering, Universidade Federal de Minas Gerais, Belo Horizonte, MG, Brazil*
E-mail: brunoot@ufmg.br

Abstract: The authors present a two-step method for identifying single-input, single-output (SISO) Wiener systems. First, using a single harmonic input, they estimate a non-parametric model of the static non-linearity, which is assumed to be only piecewise continuous. Second, using the identified non-parametric map, the authors use retrospective cost optimisation to identify a parametric model of the linear dynamic system. This method is demonstrated on several examples of increasing complexity.

1 Introduction

Block-structured models are widely used for system identification [1–3]. These models provide useful information concerning the dynamic and static components of a system, and thus constitute grey-box models in which the block structure is ascribed physical meaning. The goal of system identification is to model the internal structure of each block from available data.

Among the most widely studied block-structured models are the Wiener [4–11] and Hammerstein [4, 7, 12–15] models. Each model structure involves a single linear dynamic block and a single non-linear static block. For these two-block structures, the difficulty of the identification problem typically depends on a priori assumptions made about the components, for example, finite impulse response (FIR)-against-infinite impulse response (IIR) dynamics, and invertible against non-invertible non-linearities [9]. Furthermore, identification of Wiener systems is generally considered to be more challenging than identification of Hammerstein systems because of the fact that the input to the non-linear block is available for Hammerstein systems but not for Wiener systems. In the present paper, we focus on Wiener systems.

The methods for identifying Wiener systems developed in [4, 6, 8, 11] assume that the non-linear block is invertible. To overcome this requirement, non-parametric probabilistic methods are used in [2]. Alternatively, frequency-domain methods that apply multiple harmonic inputs are employed in [5, 7]. In [7], the multiple harmonic inputs are assigned random phase shifts, and a non-parametric model of the non-linearity is obtained using the identified linear dynamic model, which is previously estimated in the frequency domain. In [5], the phase shift between the output of the linear dynamic block and the output is exploited in the frequency domain, for each harmonic input.

In the present paper, we develop a novel technique for identifying single-input, single-output (SISO) Wiener systems. The proposed approach is semi-parametric, which, as described in [2], refers to the fact that the non-linear block is estimated non-parametrically, whereas the linear dynamics are identified parametrically. To do this, we consider a two-step procedure. In the first step, we apply a single harmonic input signal, and measure the output once the trajectory of the system reaches harmonic steady state. We then examine the output of the system (which is not harmonic because of the non-linearity) relative to the input, and use the symmetry properties of these signals to estimate the non-harmonic phase shift. This estimate allows us to infer the phase shift of the unmeasured intermediate signal (i.e. the output of the linear block) and thus reconstruct this signal up to an arbitrary amplitude. By plotting the output against the reconstructed intermediate signal, we thus obtain a non-parametric approximation of the non-linear block of the system.

The second step of the algorithm uses a sufficiently rich signal to estimate the linear dynamics of the system. As we do not assume that the non-linear block is invertible, we do not have an estimate of the output of the linear block. To overcome this difficulty, we apply retrospective cost optimisation (RCO), which uses the available output signal (in this case, the output of the non-linear block) to recursively update the linear dynamics. This technique is inspired by retrospective-cost-based adaptive control [16–18], which is used for model updating in [14, 19, 20].

As alluded to above, the two-step identification algorithm described herein does not require invertibility of the non-linear block as assumed in [4, 6, 8, 11]. In fact, we do not require that the non-linear block be either one-to-one, onto or continuous, nor do we assume as in [6] that any specific value of the non-linearity be known.

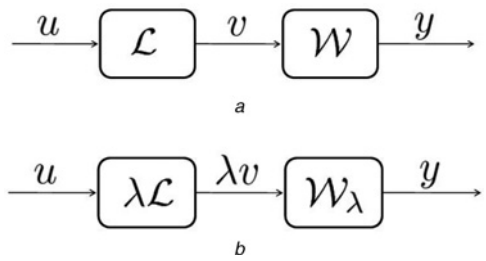

Fig. 1 *Block-structured Wiener models*

a Block-structured Wiener model, where u is the input, v is the intermediate signal, y is the output, $\mathcal{L}$ is a discrete-time linear time-invariant dynamic system and $\mathcal{W}$ is a static non-linearity
b An equivalent scaled model, where λ is a scaling factor and $\mathcal{W}_\lambda$ is a scaled-domain modification of $\mathcal{W}$ satisfying $\mathcal{W}_\lambda(\lambda v) = \mathcal{W}(v)$
The scaling factor λ is not identifiable

The contents of the paper are as follows. In Section 2 we define the Wiener identification problem. A method for non-parametric identification of the static non-linearity using a single harmonic input is presented in Section 3, whereas a method for parametric identification of the linear time-invariant dynamics using RCO is reviewed in Section 4. These methods are demonstrated on several examples of increasing complexity in Sections 5–7. Concluding remarks are presented in Section 8. A preliminary version of the results of this paper appears as in [21].

2 Problem formulation

Consider the block-structured Wiener model shown in Fig. 1*a*, where $\mathcal{L}$ is the SISO discrete-time linear time-invariant dynamic system

$$x(k+1) = Ax(k) + Bu(k) \tag{1}$$

$$v(k) = Cx(k) \tag{2}$$

with input $u(k) \in \mathbb{R}$ and intermediate signal $v(k) \in \mathbb{R}$, where k is the sample index, and $y(k) \in \mathbb{R}$ is the output given by

$$y(k) = \mathcal{W}(v(k)) \tag{3}$$

where $\mathcal{W}:\mathbb{R} \mapsto \mathbb{R}$ is the static non-linearity. We assume that $\mathcal{L}$ is asymptotically stable and $\mathcal{W}$ is piecewise continuous. Note that we do not assume that $\mathcal{W}$ is invertible, one-to-one, continuous or (as in [6]) $\mathcal{W}(0) = 0$. Also, we assume that $v(k)$ is not accessible, and that $x(0)$ is unknown and possibly non-zero.

Moreover, Fig. 1*b* shows the scaled-domain modification $\mathcal{W}_\lambda(v) \triangleq \mathcal{W}(v/\lambda)$ of $\mathcal{W}$, where λ is a non-zero real number. Therefore $\mathcal{W}_\lambda(\lambda v) = \mathcal{W}(v)$. Each value of λ scales both the gain of $\mathcal{L}$ and the domain of $\mathcal{W}$. However, λ is not identifiable.

3 Non-parametric identification of the static non-linearity

Consider the harmonic input signal

$$u(k) = A_0 \sin(\omega_0 k T_s) = A_0 \sin(\Omega_0 k) \tag{4}$$

where A_0 is the amplitude, ω_0 is the angular frequency in rad/s, T_s is the sample period in s/sample, and $\Omega_0 \triangleq \omega_0 T_s$ is the angular sample frequency in rad/sample. Since $\mathcal{L}$ is asymptotically stable, it follows that, for large values of k, the intermediate signal v is given approximately by the harmonic steady-state signal

$$v(k) = |G(e^{j\Omega_0})|A_0 \sin(\Omega_0 k + \angle G(e^{j\Omega_0})) \tag{5}$$

where $|G(e^{j\Omega_0})|$ and $\angle G(e^{j\Omega_0})$ are, respectively, the magnitude and phase shift of the frequency response of $G(\mathbf{z}) = C(\mathbf{z}I - A)^{-1}B$ at the angular sample frequency Ω_0. Therefore

$$y(k) = \mathcal{W}(|G(e^{j\Omega_0})|A_0 \sin(\Omega_0 k + \angle G(e^{j\Omega_0})) \tag{6}$$

Next, note that the continuous-time harmonic signal $\sin(\omega_0 t)$ is symmetric in the intervals $[0, \frac{1}{2}T_0]$ and $[\frac{1}{2}T_0, T_0]$ about the points $\frac{1}{4}T_0$ and $\frac{3}{4}T_0$, respectively, where $T_0 \triangleq 2\pi/\omega_0$ is the period of the harmonic input. To preserve symmetry for the sampled signal (4) about the points $\frac{1}{4}T_0$ and $\frac{3}{4}T_0$, we assume that

$$\Omega_0 = \frac{\pi}{2m} \tag{7}$$

where m is a positive integer. Thus $N_0 \triangleq 4m = T_0/T_s$ is the period of the discrete-time input (4). With this choice of Ω_0, the sampled signal $u(k)$ is symmetric in the intervals $[0, \frac{1}{2}N_0]$ and $[\frac{1}{2}N_0, N_0]$ about the points $\frac{1}{4}N_0$ and $\frac{3}{4}N_0$, respectively. Furthermore, assuming that $q \triangleq \angle G(e^{j\Omega_0})/\Omega_0$ is an integer, that is, $\angle G(e^{j\Omega_0})/\pi$ is an integer, the intermediate signal $v(k)$, which is shifted relative to $u(k)$ because of $\angle G(e^{j\Omega_0})$, is symmetric about $\frac{1}{4}N_0 + q$ in the interval $[q, \frac{1}{2}N_0 + q]$ and about $\frac{3}{4}N_0 + q$ in the interval $[\frac{1}{2}N_0 + q, N_0 + q]$. If q is not an integer, then $v(k)$ is only approximately symmetric.

Next, we note that the output signal y, which is not generally harmonic, possesses the same symmetry as v on the same intervals. By exploiting knowledge of this symmetry, we can identify the 'non-harmonic phase shift' of y relative to u, and thus the phase shift of v relative to u. Since y is not sinusoidal, the non-harmonic phase shift of y relative to u refers to the shifting of the symmetric portions of y relative to the symmetric portions of u. Knowledge of this non-harmonic phase shift allows us to determine v up to a constant multiple, specifically, v is a sinusoid that is shifted relative to u by a known number of samples.

To clarify the above discussion, we present two examples using $A_0 = 1$, $m = 18$ (so that $\Omega_0 = \pi/36$) and $G(\mathbf{z}) = 0.0685/(\mathbf{z} - 0.9164)$. First, consider the polynomial non-linearity $y = \mathcal{W}(v) = 0.6(v+1)^3 - 1$, which is neither even nor odd. Fig. 2*a* illustrates the resulting signals $u(k)$, $v(k)$ and $y(k)$ in harmonic steady state. Note that u is symmetric about the discrete-time index δ in the interval $[\delta - \frac{1}{4}N_0, \delta + \frac{1}{4}N_0]$ and about $\delta + \frac{1}{2}N_0$ in the interval $[\delta + \frac{1}{4}N_0, \delta_2 + \frac{3}{4}N_0]$. Likewise, v is symmetric about the discrete-time index ε in the interval $[\varepsilon - \frac{1}{4}N_0, \varepsilon + \frac{1}{4}N_0]$ and about $\varepsilon + \frac{1}{2}N_0$ in the interval $[\varepsilon + \frac{1}{4}N_0, \varepsilon + \frac{3}{4}N_0]$. It thus follows that y is symmetric about ε in the interval $[\varepsilon - \frac{1}{4}N_0, \varepsilon + \frac{1}{4}N_0]$ and about $\varepsilon + \frac{1}{2}N_0$ in the interval $[\varepsilon + \frac{1}{4}N_0, \varepsilon + \frac{3}{4}N_0]$.

Second, we consider the even polynomial non-linearity $y = \mathcal{W}(v) = v^2$. Fig. 2*b* illustrates the resulting signals $u(k)$, $v(k)$ and $y(k)$ in harmonic steady state. The signals u and v are equal to the signals shown in Fig. 2*a*. However, in addition to the two points of symmetry shown in Fig. 2*a*, note that y has two additional points of symmetry, specifically, y is symmetric

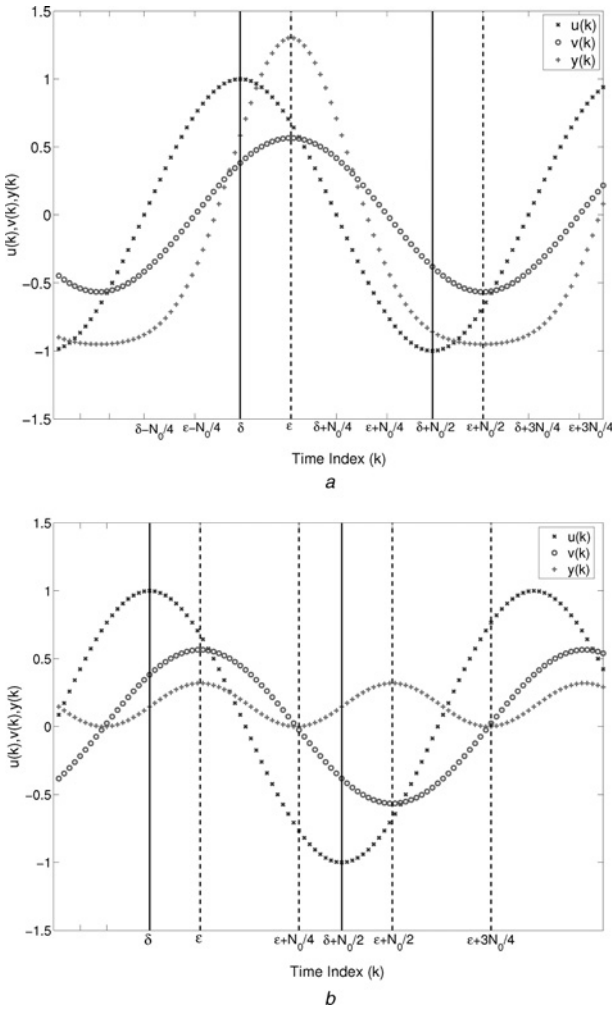

Fig. 2 *Illustration of the symmetry properties of the signals u, v and y given by (4)–(6), respectively*

a Non-even polynomial non-linearity $y = \mathcal{W}(v) = 0.6(v+1)^3 - 1$
b Even polynomial non-linearity $y = \mathcal{W}(v) = v^2$
The signals u and v are harmonic, whereas y is the output of the non-linear block $\mathcal{W}$ and thus is not harmonic. Note that, for both cases, u is symmetric about δ in the interval $[\delta - \frac{1}{4}N_0, \delta + \frac{1}{4}N_0]$ and about $\delta + \frac{1}{2}N_0$ in the interval $[\delta + \frac{1}{4}N_0, \delta + \frac{3}{4}N_0]$, whereas v and y are symmetric about ε in the interval $[\varepsilon - \frac{1}{4}N_0, \varepsilon + \frac{1}{4}N_0]$ and about $\varepsilon + \frac{1}{2}N_0$ in the interval $[\varepsilon + \frac{1}{4}N_0, \varepsilon + \frac{3}{4}N_0]$. In addition, for the case of an even polynomial non-linearity shown in *b*, y is also symmetric about $\varepsilon + \frac{1}{4}N_0$ in the interval $[\varepsilon, \varepsilon + \frac{1}{2}N_0]$ and about $\varepsilon + \frac{3}{4}N_0$ in the interval $[\varepsilon + \frac{1}{2}N_0, \varepsilon + N_0]$

about $\varepsilon + \frac{1}{4}N_0$ in the interval $[\varepsilon, \varepsilon + \frac{1}{2}N_0]$ and about $\varepsilon + \frac{3}{4}N_0$ in the interval $[\varepsilon + \frac{1}{2}N_0, \varepsilon + N_0]$.

3.1 Symmetry search algorithm

We now present an algorithm to determine ε from y. We then use ε to estimate the non-harmonic phase shift of y relative to u. For convenience, we assume that the harmonic steady state begins at $k = 0$.

Consider the signal y shown in Fig. 3, and let $n \geq 6m$ denote the width of the data window so that it includes at least one and a half periods. To encompass a complete signal period, we construct a sliding window with $N_0 + 1$ data points. The window is divided into quarters as shown in Fig. 3.

Next, for $k = 0, \ldots, n - N_0$, define

$$\beta_1(k) \triangleq \sum_{i=1}^{2m-1} |y(k+i-1) - y(k+2m-i+1)| \quad (8)$$

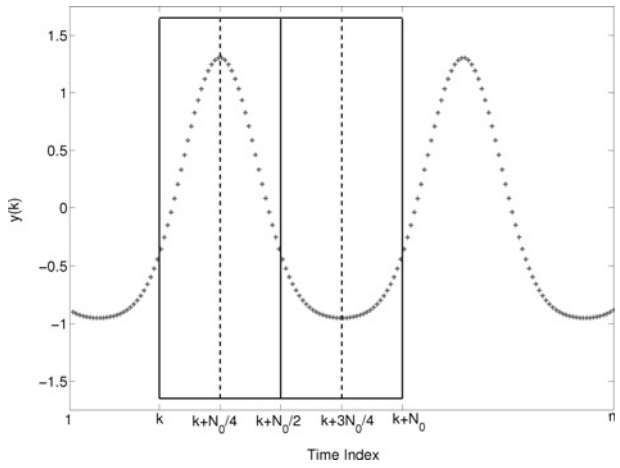

Fig. 3 *Illustration of the symmetry search algorithm*

The solid line box comprises the sliding window of length $N_0 + 1$ starting at time k, whereas the dashed lines indicate the windowed points of symmetry

which is the sum of the absolute difference in magnitude for each pair of candidate symmetric points in the first and second quarters about the point $k + \frac{1}{4}N_0$ for the sliding window starting at time step k. Likewise, for $k = 0, \ldots, n - N_0$, define

$$\beta_2(k) \triangleq \sum_{i=1}^{2m-1} |y(k+2m+i-1) - y(k+4m-i+1)| \quad (9)$$

for each pair of candidate symmetric points in the third and fourth quarters about the point $k + \frac{3}{4}N_0$. The values of β_1 and β_2 quantify the symmetry error about the points $k + \frac{1}{4}N_0$ and $k + \frac{3}{4}N_0$, respectively, for each allowable value of k. Thus, using (8) and (9), we define the 'symmetry error index'

$$\beta(k) \triangleq \beta_1(k) + \beta_2(k) \quad (10)$$

corresponding to the sliding window starting at point k, for $k = 0, \ldots, n - N_0$.

For $k = 0, \ldots, n - N_0$, let $k_0 < n - N_0$ be the minimiser of $\beta(k)$. We use knowledge of k_0 to determine the location of the points of symmetry ε and $\varepsilon + \frac{1}{2}N_0$ for the sliding window starting at point k_0. In particular, since k_0 is the starting point of the window that minimises β and since ε and $\varepsilon + \frac{1}{2}N_0$ are, respectively, the quarter point and three quarter point of the same window, it follows that

$$\varepsilon = k_0 + \frac{1}{4}N_0 \quad (11)$$

$$\varepsilon + \frac{1}{2}N_0 = k_0 + \frac{3}{4}N_0 \quad (12)$$

Note that, in general, $\beta(k_0) \neq 0$. However if $\angle G(e^{j\Omega_0})/\pi$ is an integer, then $\beta(k_0) = 0$, which indicates exact symmetry about $k_0 + \frac{1}{4}N_0$ in the interval $[k_0, k_0 + \frac{1}{2}N_0]$ and about $k_0 + \frac{3}{4}N_0$ in the interval $[k_0 + \frac{1}{2}N_0, k_0 + N_0]$.

To illustrate the symmetry search algorithm, we reconsider the example considered in Figs. 2*a* and 3, where $y = \mathcal{W}(v) = 0.6(v+1)^3 - 1$. Note that $\mathcal{W}$ is not even. Fig. 4*a* shows the values of β calculated for $y(k)$ on the interval $[k_0, k_0 + 2N_0]$. Since, in Fig. 4*a*, the data window of y is selected to start at $k_0 = \varepsilon - \frac{1}{4}N_0$, the minimum

doi: 10.1049/iet-cta.2010.0023

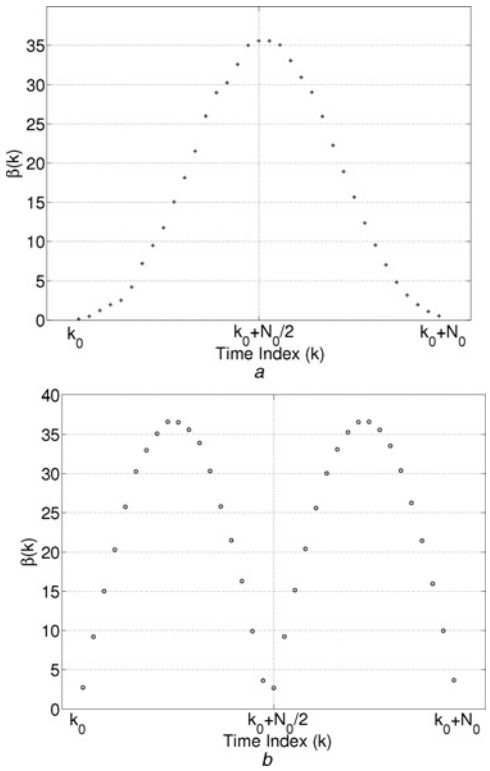

Fig. 4 *Illustration of the symmetry error index $\beta(k)$ given by (8)*

a Non-even polynomial
b Even polynomial
The values of $\beta(k)$ are shown for two static non-linearities

values of $\beta(k)$ occur at k_0 and $k_0 + N_0$, where $k_0 + N_0$ is the start of the next period and, thus, need not be considered. Thus, using the unique minimiser k_0 of $\beta(k)$, it follows that the locations of the points of symmetry are given by (11) and (12).

Next, for the even non-linearity $y = \mathcal{W}(v) = v^2$ considered in Fig. 2*b*, Fig. 4*b* shows the values of $\beta(k)$ calculated for $y(k)$ on the interval $[k_0, k_0 + 2N_0]$. In this case, the minimum values of $\beta(k)$ occur at k_0, $k_0 + \frac{1}{2}N_0$ and $k_0 + N_0$, where $k_0 + N_0$ is the start of the next period and, thus, need not be considered. Thus, using k_0, it follows that the locations of the points of symmetry are given by (11) and (12). Also, using $k_0 + \frac{1}{2}N_0$, we obtain two additional points of symmetry given by

$$\varepsilon + \frac{1}{4}N_0 = k_0 + \frac{1}{2}N_0 \tag{13}$$

$$\varepsilon + \frac{3}{4}N_0 = k_0 + N_0 \tag{14}$$

This ambiguity is due to the fact that ε and $\varepsilon + \frac{1}{2}N_0$ are the midpoints of two identical symmetric portions of y. Thus, the start of the data window within which the function has the symmetry properties illustrated in Fig. 3 can be taken as either k_0 or $k_0 + \frac{1}{2}N_0$. Note that the second minimiser $k_0 + \frac{1}{2}N_0$ appears only for even non-linearities.

3.2 Non-parametric approximation of the static non-linearity

Using δ, which is assumed to be known from the harmonic input u, and the estimate of ε obtained from y in Section 3.1, we now determine an estimate $\hat{\phi}$ of the non-harmonic phase shift of y relative to u by

$$\hat{\phi} \triangleq \Omega_0(\varepsilon - \delta) \tag{15}$$

which is an estimate of $\angle G(e^{j\Omega_0})$. Moreover, define the virtual signal

$$\tilde{v}(k) \triangleq A_0 \sin(\Omega_0 k + \hat{\phi}) \tag{16}$$

which is an approximation of the intermediate signal v given by (5) divided by the constant $|G(e^{j\Omega_0})|$. Note that, if $\hat{\phi} = \angle G(e^{j\Omega_0})$, then $|G(e^{j\Omega_0})|\tilde{v} = v$. Also, note that the amplitude of $\tilde{v}(k)$ is irrelevant because of the scaling factor λ shown in Fig. 1*b*.

Using $\tilde{v}$ and y, the non-parametric estimate of $\mathcal{W}$ is given by

$$\hat{\mathcal{W}} \triangleq \{(\tilde{v}(k_0), y(k_0)), (\tilde{v}(k_0 + 1), y(k_0 + 1)), \ \ldots, (\tilde{v}(n), y(n))\} \tag{17}$$

where each pair $(\tilde{v}(k), y(k))$, for $k = k_0, \ \ldots, n$, determines a value of the non-parametric estimate $\hat{\mathcal{W}}$ of $\mathcal{W}$.

Fig. 4 shows that, depending on the type of non-linearity, $\beta(k)$ has either one or two minima within each period. For a non-even polynomial non-linearity, $\beta(k)$ has one minimum within each period. Therefore, the estimate of the non-harmonic phase shift has two candidate values, namely, $\hat{\phi}$ and $\hat{\phi} + \pi$. For an even non-linearity, $\beta(k)$ has two minima within each period. Therefore the estimate of the non-harmonic phase shift has four candidate values, namely, $\hat{\phi}$, $\hat{\phi} + \frac{\pi}{2}$, $\hat{\phi} + \pi$ and $\hat{\phi} + \frac{3\pi}{2}$. However, for the even case, $\hat{\phi}$ and $\hat{\phi} + \pi$ yield the same non-parametric model $\hat{\mathcal{W}}$, whereas $\hat{\phi} + \frac{\pi}{2}$ and $\hat{\phi} + \frac{3\pi}{2}$ yield the same $\hat{\mathcal{W}}$.

Therefore for both non-even and even cases, there are two candidate non-parametric estimates of $\mathcal{W}$, both of which are constructed using (16) and (17). The correct non-parametric model will become apparent when identifying the dynamic block of the Wiener system.

4 Parametric identification of the linear time-invariant dynamics

Using the non-parametric model $\hat{\mathcal{W}}$ of $\mathcal{W}$, we now identify a model of $\mathcal{L}$ given by $\hat{\mathcal{L}}$ using RCO [20]. The RCO algorithm is presented in [17, 19, 20] together with guidelines for choosing its tuning parameters, namely, n_c, p and α.

Consider the adaptive feedback architecture for $\hat{\mathcal{L}}$ shown in Fig. 5, where $\hat{\mathcal{L}}_m$ denotes the initial model with input $w \in \mathbb{R}$ and output $\hat{v} \in \mathbb{R}$, and where $\hat{\mathcal{L}}_\Delta$ denotes the feedback delta model with inputs u, $\hat{v} \in \mathbb{R}$ and output w.

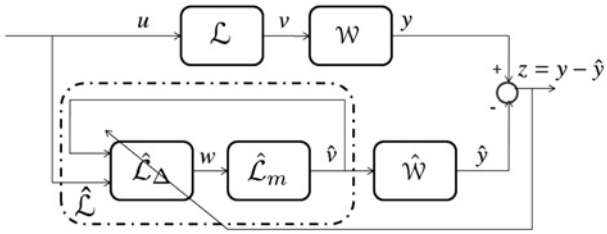

Fig. 5 *Identification architecture for Wiener models using RCO*

The goal is to adaptively tune $\hat{\mathcal{L}}_\Delta$ so that the performance variable

$$z(k) \triangleq y(k) - \hat{y}(k) \tag{18}$$

is minimised in the presence of the identification signal u. For simplicity, we choose $\hat{\mathcal{L}}_{\mathrm{m}}$ to be the one-step delay $1/\mathbf{z}$. Together, $\hat{\mathcal{L}}$ and $\hat{\mathcal{W}}$ comprise a 'semi-parametric model' of the Wiener system.

From Section 3.2, recall that there are two candidates for the non-parametric estimate of $\mathcal{W}$. Thus, we run RCO for each non-parametric estimate of $\mathcal{W}$ and obtain a corresponding parametric model of $\mathcal{L}$. Note that the performance variable z is calculated for both semi-parametric models. We choose the semi-parametric model whose performance variable has a smaller norm.

4.1 Retrospective cost optimisation

We now review the RCO adaptive control algorithm and show how it is used to identify linear time-invariant dynamic systems using $\hat{\mathcal{W}}$. A detailed discussion of RCO and as well as the theoretical foundations of the algorithm are found in [16, 17, 22].

RCO depends on several parameters that are selected a priori. Specifically, n_{c} is the estimated plant order, $p \geq 1$ is the data window size used to estimate $\hat{\mathcal{L}}_\Delta$ and μ is the number of Markov parameters of $\hat{\mathcal{L}}_\Delta$. The methodology for choosing these parameters is as follows. n_{c} is overestimated, that is, chosen to be greater than the expected order of $\hat{\mathcal{L}}$. From Section 4, recall that we assume that the controller $\hat{\mathcal{L}}_\Delta$ is placed in feedback with a unit delay. Therefore there is only one non-zero Markov parameter, so $\mu = 1$ in all example cases. The adaptive update law is based on a quadratic cost function, which involves a time-varying weighting parameter $\alpha(k) > 0$, referred to as the 'learning rate' since it affects the convergence speed of the adaptive control algorithm. In [17], RCO is presented for multiple-input multiple-output (MIMO) systems, where l_u, l_v, l_w and l_y denote the sizes of u, v, w and y, respectively. However, in this paper, we consider only the SISO system (1)–(3). For convenience, we keep the notation of [17] and set $l_u = l_v = l_w = l_y = 1$.

Let

$$\hat{\mathcal{L}}_{\mathrm{m}} \sim \begin{bmatrix} A & B \\ C & D \end{bmatrix}$$

as given by (1), (2), where $x(k) \in \mathbb{R}^{l_x}$, $A \in \mathbb{R}^{l_x \times l_x}$, $B \in \mathbb{R}^{l_x \times 1}$, $C \in \mathbb{R}^{1 \times l_x}$. Since $\hat{\mathcal{L}}_{\mathrm{m}}$ is set as unit delay, it follows that $A = 0_{l_x \times l_x}$, $B = 1_{l_x \times 1}$ and $C = 1_{1 \times l_x}$, yielding

$$\hat{v}(k) = w(k-1)$$

where $w(k-1)$ is the output of $\hat{\mathcal{L}}_\Delta$, which was obtained using RCO in the previous iteration. Note that, to compute (18), $\hat{y}(k)$ is assumed to be known. To accomplish that, we use the estimated intermediate signal $\hat{v}(k)$ with $\hat{\mathcal{V}}$ as follows. Note that, in general, $\hat{v}(k)$ is not in the set defined by (17). We thus suggest two methods by which this issue may be overcome. For simplicity, the first case is to use the closest value of $\tilde{v}(k)$ in the set (17) to $\hat{v}(k)$. Second, interpolation between the closest bounding values may be used. For convenience, henceforth, we use the first method.

Next, to compute $w(k)$ we use an exactly proper time-series controller of order n_{c} such that the control $w(k)$ is given by

$$w(k) = \sum_{i=1}^{n_{\mathrm{c}}} M_i(k) w(k-i) + \sum_{i=0}^{n_{\mathrm{c}}} N_i(k) \begin{bmatrix} \hat{v}(k-i) \\ u(k-i) \end{bmatrix} \tag{19}$$

where $M_i \in \mathbb{R}^{l_w \times l_w}$, $i = 1, \ldots, n_{\mathrm{c}}$ and $N_i \in \mathbb{R}^{l_w \times (l_v + l_u)}$, $i = 0, \ldots, n_{\mathrm{c}}$, are given by an adaptive update law. Note that the ARX model given in (19) is a model of $\hat{\mathcal{L}}_\Delta$. The control can be expressed as

$$w(k) = \theta(k)\psi(k) \tag{20}$$

where

$$\theta(k) \triangleq \begin{bmatrix} N_0(k) & \cdots & N_{n_{\mathrm{c}}}(k) & M_1(k) & \cdots & M_{n_{\mathrm{c}}}(k) \end{bmatrix}$$

is the 'controller parameter block matrix' and the 'regressor vector' $\psi(k)$ is given by

$$\psi(k) \triangleq \begin{bmatrix} \hat{v}(k) \\ \vdots \\ \hat{v}(k-n_{\mathrm{c}}) \\ u(k) \\ \vdots \\ u(k-n_{\mathrm{c}}) \\ w(k-1) \\ \vdots \\ w(k-n_{\mathrm{c}}) \end{bmatrix} \in \mathbb{R}^{n_{\mathrm{c}} l_w + (n_{\mathrm{c}}+1)(l_v + l_u)}$$

For positive integers p and μ, we define the 'extended performance vector' $Z(k)$ and the 'extended control vector' $W(k)$ by

$$Z(k) \triangleq \begin{bmatrix} z(k) \\ \vdots \\ z(k-p+1) \end{bmatrix}, \quad W(k) \triangleq \begin{bmatrix} w(k) \\ \vdots \\ w(k-p_{\mathrm{c}}+1) \end{bmatrix} \tag{21}$$

where $p_{\mathrm{c}} \triangleq \mu + p$.

From (20), it follows that the extended control vector $W(k)$ can be written as

$$W(k) \triangleq \sum_{i=1}^{p_{\mathrm{c}}} L_i \theta(k-i+1)\psi(k-i+1) \tag{22}$$

where

$$L_i \triangleq \begin{bmatrix} 0_{(i-1)l_w \times l_w} \\ I_{l_w} \\ 0_{(p_{\mathrm{c}}-i)l_w \times l_w} \end{bmatrix} \in \mathbb{R}^{p_{\mathrm{c}} l_w \times l_w} \tag{23}$$

We define the 'surrogate performance vector' $\hat{Z}(\hat{\theta}(k), k)$ by

$$\hat{Z}(\hat{\theta}(k), k) \triangleq Z(k) - \bar{B}_{zw}(W(k) - \hat{W}(k)) \tag{24}$$

doi: 10.1049/iet-cta.2010.0023

where

$$\hat{W}(k) \triangleq \sum_{i=1}^{p_c} L_i \hat{\theta}(k)\psi(k-i+1) \tag{25}$$

and $\hat{\theta}(k) \in \mathbb{R}^{l_w \times [n_c l_w + (n_c+1)(l_y+l_u)]}$ is the 'surrogate controller parameter block matrix'. The block-Toeplitz 'surrogate control matrix' $\bar{B}_{zw}$ is given by

$$\bar{B}_{zw} \triangleq \begin{bmatrix} 0_{l_z\times l_w} & \cdots & 0_{l_z\times l_w} & H_d & \cdots & H_\mu & 0_{l_z\times l_w} & \cdots & 0_{l_z\times l_w} \\ 0_{l_z\times l_w} & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & 0_{l_z\times l_w} \\ 0_{l_z\times l_w} & \cdots & 0_{l_z\times l_w} & 0_{l_z\times l_w} & \cdots & 0_{l_z\times l_w} & H_d & \cdots & H_\mu \end{bmatrix}$$

where the 'relative degree' d is the smallest positive integer i such that the ith Markov parameter $H_i = CA^{i-1}B$ of $\hat{\mathcal{L}}_m$ is non-zero. The leading zeros in $\bar{B}_{zw}$ account for the non-zero relative degree d. The algorithm places no constraints on either the value of d or the rank of H_d or $\bar{B}_{zw}$. For the SISO case when $\hat{\mathcal{L}}_m$ is a unit delay

$$\bar{B}_{zw} = \begin{bmatrix} 0 & 1 & 0 \end{bmatrix} \tag{26}$$

Furthermore, we define

$$D(k) \triangleq \sum_{i=1}^{n_c+\mu-1} \psi^T(k-i+1) \otimes L_i \tag{27}$$

$$f(k) \triangleq Z(k) - \bar{B}_{zw} W(k) \tag{28}$$

We now consider the cost function

$$J(\hat{\theta}, k) \triangleq \hat{Z}^T(\hat{\theta}, k) R_1(k) \hat{Z}(\hat{\theta}, k) + \mathrm{tr}[R_2(k)(\hat{\theta} - \theta(k))^T R_3(k)(\hat{\theta} - \theta(k))] \tag{29}$$

where $R_1(k) \triangleq I_{pl_z}$, $R_2(k) \triangleq \alpha(k) I_{n_c(l_w+(l_y+l_u))}$ and $R_3(k) \triangleq I_{l_w}$. Note that the cost function is quadratic in the retrospective term $\hat{Z}$, while the second term penalises the difference $\theta(k+1) - \theta(k)$; therefore R_2 and R_3 can be used to control how much the controller parameters will change in a given step.

Substituting (24) and (25) into (29), J is written as the quadratic form

$$J(\hat{\theta}, k) = c(k) + b^T \,\mathrm{vec}\,\hat{\theta} + (\mathrm{vec}\,\hat{\theta})^T A(k)\,\mathrm{vec}\hat{\theta} \tag{30}$$

where

$$A(k) = D^T(k)D(k) + \alpha(k)I \tag{31}$$

$$b(k) = 2D^T(k)f(k) - 2\alpha(k)\,\mathrm{vec}\,\theta(k) \tag{32}$$

$$c(k) = f(k)^T R_1(k) f(k) + \mathrm{tr}[R_2(k)\theta^T(k)R_3(k)\theta(k)] \tag{33}$$

Since $A(k)$ is positive definite, $J(\hat{\theta}, k)$ has the strict global minimiser

$$\hat{\theta}(k) = \frac{1}{2}\,\mathrm{vec}^{-1}(A(k)^{-1}b(k)) \tag{34}$$

The controller gain update law is

$$\theta(k+1) = \hat{\theta}(k) \tag{35}$$

such that $w(k)$ is computed using (20). The key feature of the adaptive control algorithm (20) is the surrogate performance variable $Z(k)$ based on the difference between the actual past control inputs $W(k)$ and the recomputed past control inputs based on the current control law $\hat{W}(k)$. The parameter α is chosen to be as small as possible while guaranteeing that $A(k)$ is positive definite.

5 Numerical examples: nominal case

To demonstrate semi-parametric model identification, we consider various static non-linearities. For each example, we choose G to have poles $0.34 \pm 0.87\jmath$, $-0.3141 \pm 0.9\jmath$, $0.05 \pm 0.3122\jmath$, -0.6875 and zeros $0.14 \pm 0.97\jmath$, $-0.12 \pm 0.62\jmath$, -0.89 with monic numerator and denominator. Also, $u(k)$ is chosen to be a realisation of zero-mean Gaussian white noise with standard deviation $\sigma_u = 3.5$.

Note that A_0 should, in practice, be chosen to be greater than the expected operating range of the Wiener system. This guarantees that the inputs to the model can be interpolated from the non-parametric map. For the following examples we choose m to be much larger than required. Although we show in Section 7.2 that very little performance gain is attained from choosing m large, it is visually easier to compare the identified non-parametric map to the true non-linearity when using more data points. Finally, the parameter $\alpha(k)$ discussed in the previous section is chosen as a constant value for all examples. We choose varying values for $\alpha(k)$ to demonstrate that the final estimate of the Wiener system is not sensitive to this parameter.

Example 1 (Non-even polynomial): Consider $\mathcal{W}$ defined by

$$y = \mathcal{W}(v) = v^3 + 4v + 7 \tag{36}$$

The parameters for non-parametric identification of $\mathcal{W}$ are $m = 500$ and $A_0 = 5$. Fig. 6 compares the true and identified non-linearities. The RCO parameters used to identify the linear dynamic system are set as $n_c = 9$, $p = 1$ and $\alpha = 1$. Fig. 7 shows the frequency response of the true

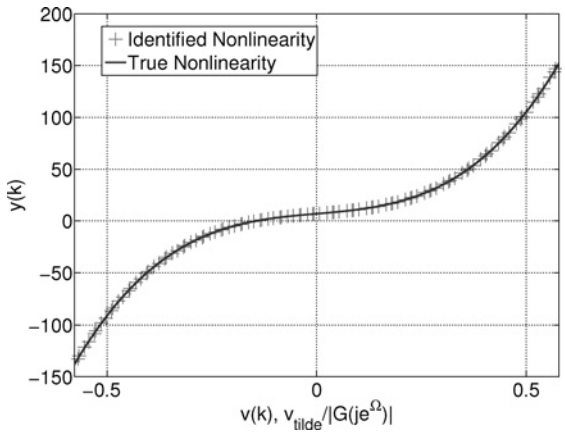

Fig. 6 *Identified non-linearity against true non-linearity (36), where $m = 500$ and $A_0 = 5$ (Example 1)*

The argument of the identified non-linearity is scaled by $1/|G(e^{\jmath\Omega_0})|$ to facilitate comparison with the true non-linearity

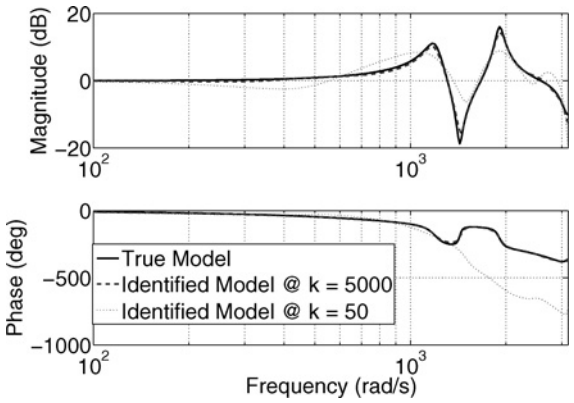

Fig. 7 *Frequency response comparison of the true G and the identified LTI system obtained using $\hat{\mathcal{W}}$ as an estimate of (36), where k is the number of data points used to determine the identified dynamic model*

The RCO controller order is $n_c = 9$ with $p = 1$ and $\alpha = 1$ (Example 1)

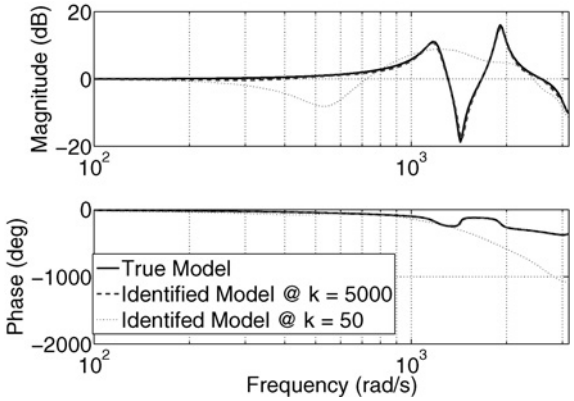

Fig. 9 *Frequency response comparison of the true G and the identified LTI system obtained using $\hat{\mathcal{W}}$ as an estimate of (37), where k is the number of data points used to determine the identified dynamic model*

The RCO controller order is $n_c = 9$ with $p = 1$ and $\alpha = 50$ (Example 2)

dynamic model G and the identified model using RCO with the identified non-linearity shown in Fig. 6.

Example 2 (Even polynomial): Consider $\mathcal{W}$ defined by

$$y(k) = \mathcal{W}(v) = 7v^4 + v^2 \tag{37}$$

The parameters for non-parametric identification of $\mathcal{W}$ are $m = 500$ and $A_0 = 5$. Fig. 8 compares the true and identified non-linearities. The RCO parameters used to identify the linear dynamic system are set as $n_c = 9$, $p = 1$ and $\alpha = 50$. Fig. 9 shows the frequency response of G and the identified model using RCO with the identified nonlinearity shown in Fig. 8.

Next, to illustrate the ambiguity discussed in Section 3.2, we select the incorrect non-harmonic phase shift, specifically, $\hat{\phi} + \pi/2$. Fig. 10 shows a comparison of the true and identified non-linearities. Note that the incorrect non-harmonic phase shift produces an erroneous non-parametric model of the non-linearity. Fig. 11 shows a frequency response comparison of G and the model identified using RCO with the identified non-linearity shown in Fig. 10.

To determine the appropriate phase shift $\hat{\phi}$ or $\hat{\phi} + \pi/2$, we examine the performance variable z given by (18), which

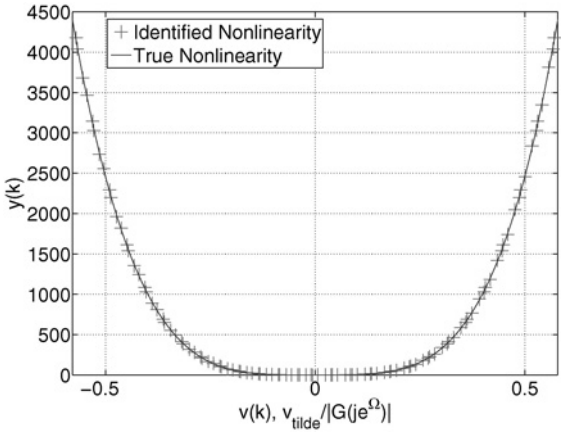

Fig. 8 *Identified non-linearity against true non-linearity (37), where $m = 500$ and $A_0 = 5$ (Example 2)*

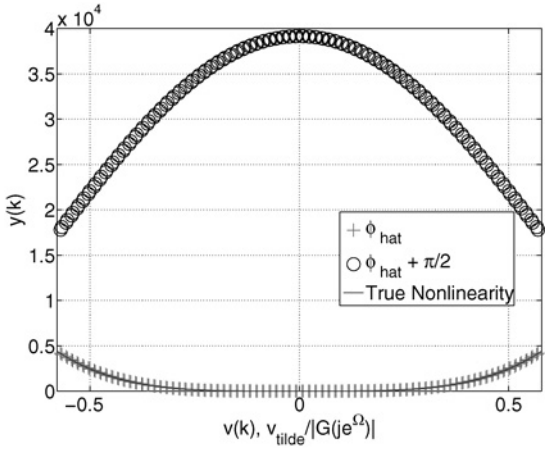

Fig. 10 *Identified non-linearity against true non-linearity (5.2), where $m = 500$ and $A_0 = 5$ (Example 2)*

Both candidate values for the non-harmonic phase shift, namely, $\hat{\phi}$ and $\hat{\phi} + (\pi/2)$, are used to build the two candidate identified non-linearities

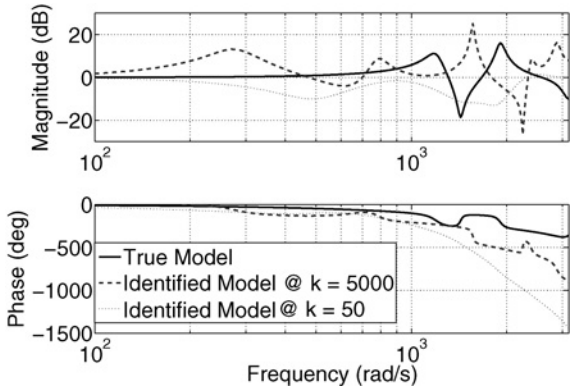

Fig. 11 *Frequency response comparison of the true G and the identified LTI system obtained using $\hat{\mathcal{W}}$ corresponding to the incorrect phase shift as an estimate of (37), where k is the number of data points used to determine the identified dynamic model*

The RCO controller order is $n_c = 9$ with $p = 1$, and $\alpha = 50$ (Example 2)

doi: 10.1049/iet-cta.2010.0023

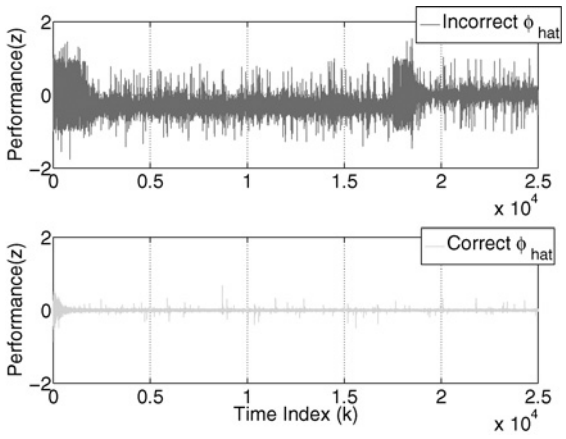

Fig. 12 *Retrospective optimisation performance comparison for Example 2*

The upper plot shows the performance variable z for the case in which the non-parametric model is generated using the incorrect candidate for the non-harmonic phase shift $\hat{\phi} + (\pi/2)$. The lower plot shows z for the case in which the correct candidate $\hat{\phi}$ is used

provides insight into which candidate value yields the correct semi-parametric model. The upper plot of Fig. 12 shows the RCO performance variable z for the incorrect non-parametric model of $\mathcal{W}$, whereas the lower plot shows the performance variable for the correct non-parametric model of $\mathcal{W}$. The correct semi-parametric model clearly outperforms the incorrect model.

6 Numerical examples: off-nominal cases

We now reconsider the Wiener system (1)–(3) with noise, as shown in Fig. 13. The input $u(k)$ is a realisation of zero-mean Gaussian white noise with standard deviation $\sigma_u = 3.5$, whereas $d_1(k) \in \mathbb{R}$ and $d_2(k) \in \mathbb{R}$ are unknown zero-mean Gaussian white disturbances with standard deviations σ_{d_1} and σ_{d_2}, respectively. The output

$$y(k) = \mathcal{W}(v(k)) + d_3(k) \tag{38}$$

has standard deviation σ_y about its mean, and $d_3(k) \in \mathbb{R}$ is an unknown zero-mean Gaussian white disturbance with standard deviation σ_{d_3}. The disturbance signals $d_1(k)$, $d_2(k)$ and $d_3(k)$ are process, input and output noise, respectively.

We now consider additional static non-linearities, where, for each example, we choose G as in Section 5.

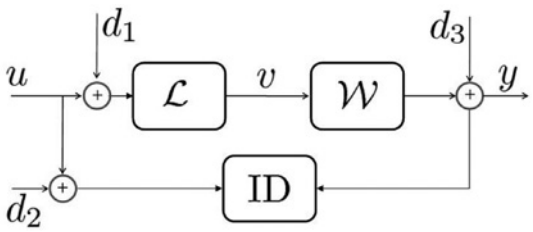

Fig. 13 *Block-structured Wiener model with process, input and output noise, where d_1, d_2 and d_3 are unknown zero-mean Gaussian disturbances*

Example 3 (Deadzone): Consider $\mathcal{W}$ defined by

$$y = \mathcal{W}(v) = \begin{cases} 0 & \text{if } |v| \leq 0.17 \\ v & \text{if } |v| > 0.17 \end{cases} \tag{39}$$

Furthermore, we consider process and output noise $\sigma_{d_1} = \frac{1}{15}\sigma_u$, $\sigma_{d_3} = \frac{1}{15}\sigma_y$ and $d_2 = 0$. For this problem, the parameters for non-parametric identification are $m = 250$ and $A_0 = 5$. In this example we also parameterise the estimated non-linearity for comparison with the non-parametric estimate. The parametric model is a 25th-order polynomial. Fig. 14 compares the true, non-parametric identified and parametric identified non-linearities. The RCO parameters used to identify the linear dynamic system are set as $n_c = 9$, $p = 1$ and $\alpha = 10$. Fig. 15*a* shows the frequency response of G and the identified model using RCO with the identified non-parametric model of the non-linearity, while Fig. 15*b* shows the frequency response of the identified model using RCO with the identified parametric model of the non-linearity. Fig. 16*a* compares the output of the Wiener system $y(k)$ and the output of the estimated semi-parametric Wiener model $\hat{y}(k)$ in response to a random input, whereas Fig. 16*b* compares the output of the Wiener system $y(k)$ and the output of the estimated parametric Wiener model $\hat{y}(k)$ in response to a random input. Fig. 16*c* shows the difference between the error in the semi-parametric Wiener model and the parametric Wiener model. Where the graph is negative, the semi-parametric model has superior performance, and where the graph is positive the parametric model is superior. In this case, it is not clear if the parametrisation of the non-linearity pays off.

Example 4 (Saturation): Consider $\mathcal{W}$ defined by

$$y = \mathcal{W}(v) = \begin{cases} 8.64(v + 0.23) - 3.98 & \text{if } 0.1 < v < 0.4 \\ 1.5 & \text{if } v \geq 0.4 \\ -1.2 & \text{if } v \leq 0.1 \end{cases} \tag{40}$$

Furthermore, we consider input noise $\sigma_{d_1} = \frac{1}{8}\sigma_u$ and

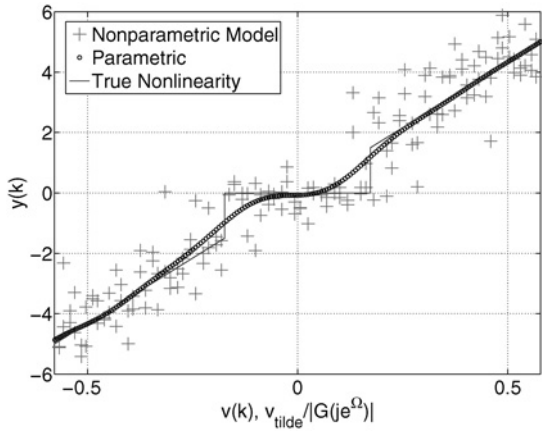

Fig. 14 *Identified non-linearity against true non-linearity (39), where $m = 250$ and $A_0 = 5$*

In this example, we also parameterise the estimated non-linearity using a 25th-order polynomial (Example 3)

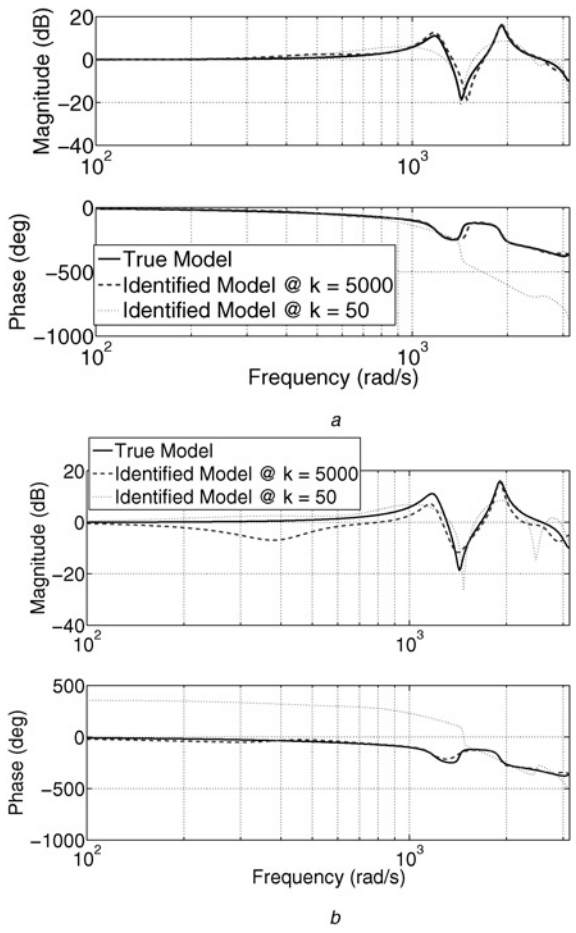

Fig. 15 *Frequency response comparison of the true G and the identified LTI system obtained*

a Non-parametric $\hat{\mathcal{W}}$ as an estimate of (39)
b Parametric 25th polynomial estimate of (39)
k is the number of data points used to determine the identified dynamic model. The RCO controller order is $n_c = 9$ with $p = 1$ and $\alpha = 10$ (Example 3)

$d_2 = d_3 = 0$. The parameters for non-parametric identification are $m = 150$ and $A_0 = 5$. Fig. 17 compares the true and identified non-linearities. The RCO parameters used to identify the linear dynamic system are set as $n_c = 9$, $p = 1$ and $\alpha = 1$. Fig. 18 shows the frequency response of G and the identified model using RCO with the identified non-linearity shown in Fig. 17.

Example 5 (Switch function): Consider $\mathcal{W}$ defined by

$$y = \mathcal{W}(v) = \begin{cases} 0 & \text{if } |v| = 0 \\ 8.64v + \text{sgn}(v)4.5 & \text{if } 0 < |v| \leq 1.5 \end{cases} \quad (41)$$

Furthermore, we consider process, input and output noise $\sigma_{d_1} = \frac{1}{15}\sigma_u$, $\sigma_{d_2} = \frac{1}{15}\sigma_w$ and $\sigma_{d_3} = \frac{1}{15}\sigma_y$.

The parameters for non-parametric identification are $m = 100$ and $A_0 = 5$. Fig. 19 compares the true and identified non-linearities. The RCO parameters used to identify the linear dynamic system are set as $n_c = 9$, $p = 1$ and $\alpha = 1$. Fig. 20 shows the frequency response of G and the identified model using RCO with the identified non-linearity shown in Fig. 19.

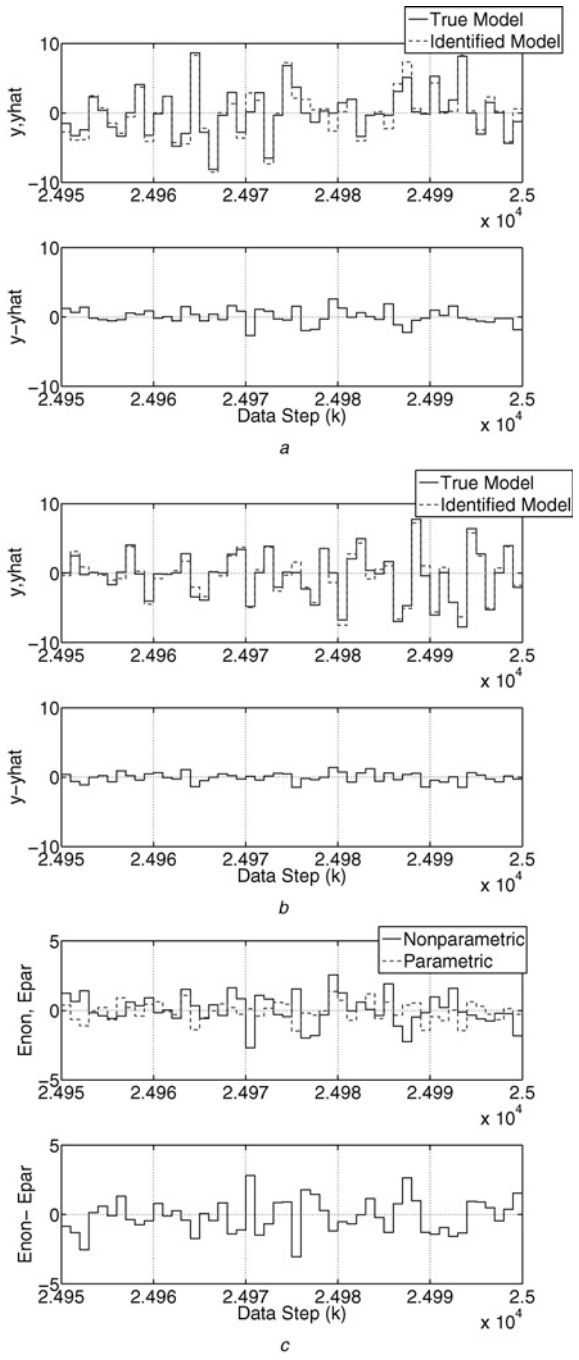

Fig. 16 *Performance comparison for Example 3*

a Non-parametric estimate for the non-linearity
b Parametric estimate of the non-linearity
In both *a* and *b*, the top plot shows the output of the Wiener system $y(k)$ and the output of the estimated system $\hat{y}(k)$, whereas the bottom plot shows the performance $z(k)$
c Difference between the error in the semi-parametric Wiener model and the parametric Wiener model is shown
Wherever the graph is negative, the semi-parametric model has superior performance, and wherever the graph is positive the parametric model is superior

Example 6 (Stairs function): Consider $\mathcal{W}$ defined by

$$y = \mathcal{W}(v) = \begin{cases} 0 & \text{if } |v| = 0 \\ \text{sgn}(v)1 & \text{if } 0 < |v| \leq 0.17 \\ \text{sgn}(v)3 & \text{if } 0.17 < |v| \leq 0.35 \\ \text{sgn}(v)4.5 & \text{if } 0.35 < |v| \leq 0.52 \\ \text{sgn}(v)6 & \text{if } 0.52 < |v| \end{cases} \quad (42)$$

doi: 10.1049/iet-cta.2010.0023

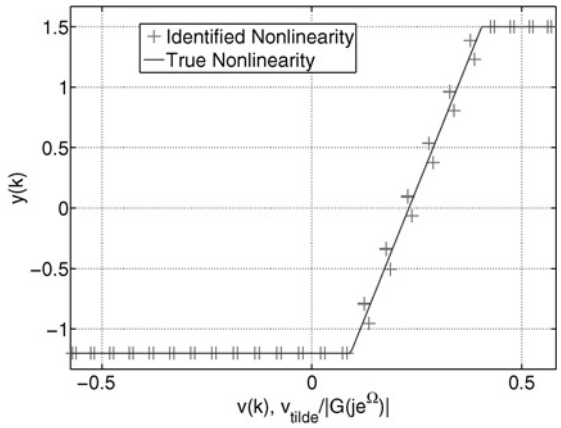

Fig. 17 *Identified non-linearity against true non-linearity (40) where $m = 150$ and $A_0 = 5$ (Example 4)*

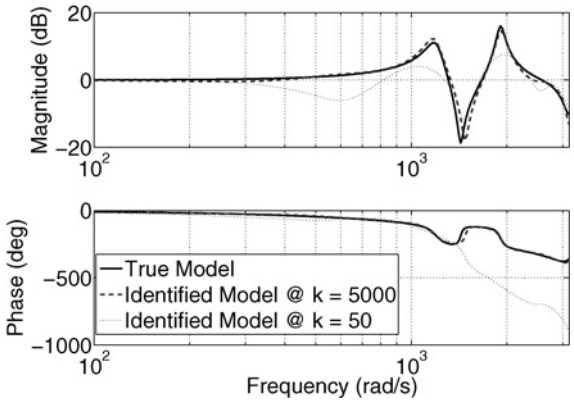

Fig. 18 *Frequency response comparison of the true G and the identified LTI system obtained using $\hat{\mathcal{W}}$ as an estimate of (40), where k is the number of data points used to determine the identified dynamic model*

The RCO controller order is $n_c = 9$ with $p = 1$ and $\alpha = 1$ (Example 4)

Furthermore, we consider process, input and output noise $\sigma_{d_1} = \frac{1}{8}\sigma_u$, $\sigma_{d_2} = \frac{1}{8}\sigma_w$ and $\sigma_{d_3} = \frac{1}{8}\sigma_y$. The parameters for non-parametric identification are $m = 75$ and $A_0 = 5$. Fig. 21 compares the true and identified non-linearities. The RCO parameters used to identify the linear dynamic system are set as $n_c = 9$, $p = 1$ and $\alpha = 1$. Fig. 22 is a frequency

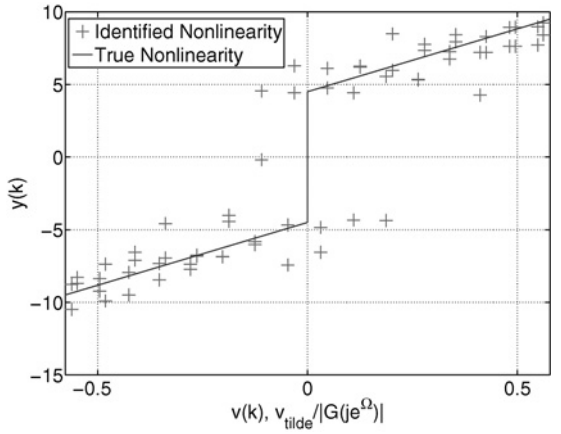

Fig. 19 *Identified non-linearity against true non-linearity (41), where $m = 100$ and $A_0 = 5$ (Example 5)*

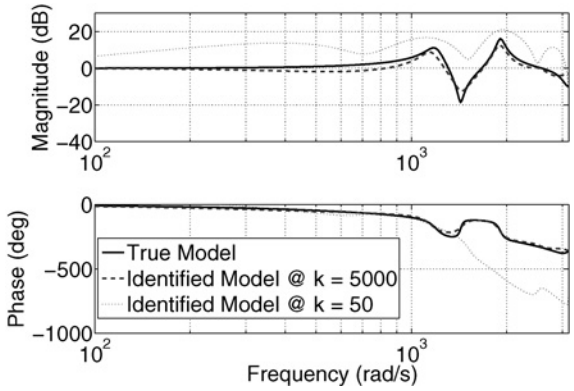

Fig. 20 *Frequency response comparison of the true G and the identified LTI system obtained using $\hat{\mathcal{W}}$ as an estimate of (41), where k is the number of data points used to determine the identified dynamic model*

The RCO controller order is $n_c = 9$ with $p = 1$ and $\alpha = 1$ (Example 5)

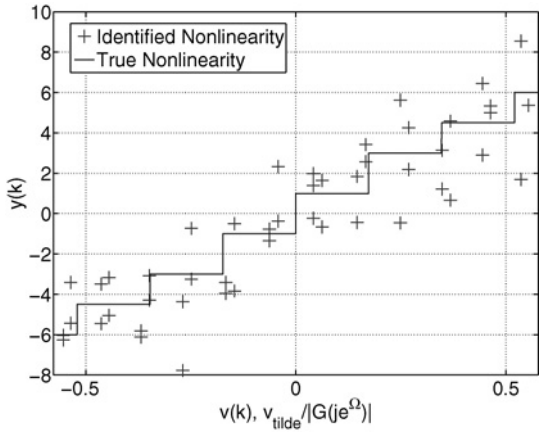

Fig. 21 *Identified non-linearity against true non-linearity (42), where $m = 75$ and $A_0 = 5$ (Example 6)*

response comparison of G and the system identified using RCO with the identified non-linearity shown in Fig. 21. Fig. 23 compares the output of the Wiener system $y(k)$ and the output of the estimated semi-parametric Wiener model $\hat{y}(k)$ in response to a random input.

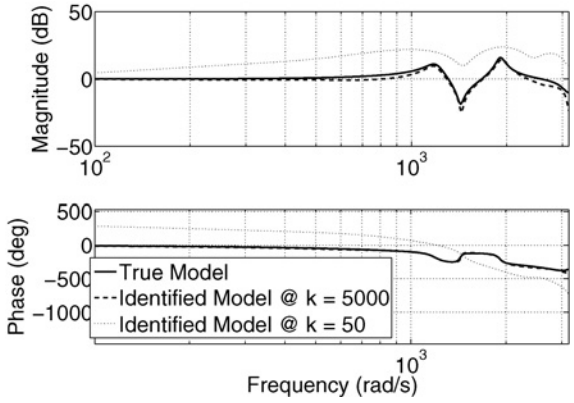

Fig. 22 *Frequency response comparison of the true G and the identified LTI system obtained using $\hat{\mathcal{W}}$ as an estimate of (42), where k is the number of data points used to determine the identified dynamic model*

The RCO controller order is $n_c = 9$ with $p = 1$ and $\alpha = 1$ (Example 6)

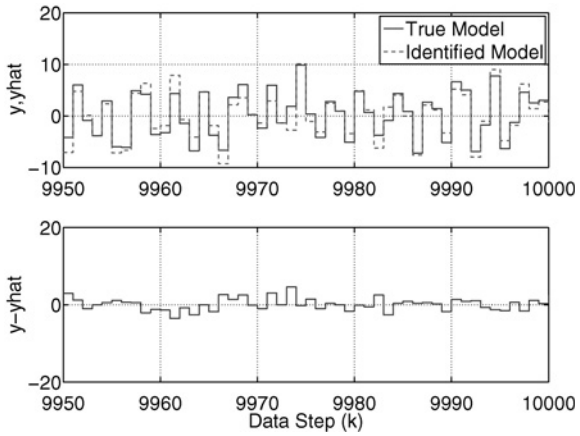

Fig. 23 *Performance comparison for Example 6 of the output of the Wiener system y(k), and the output of the estimated system ŷ(k)*

The bottom plot is the performance $z(k)$

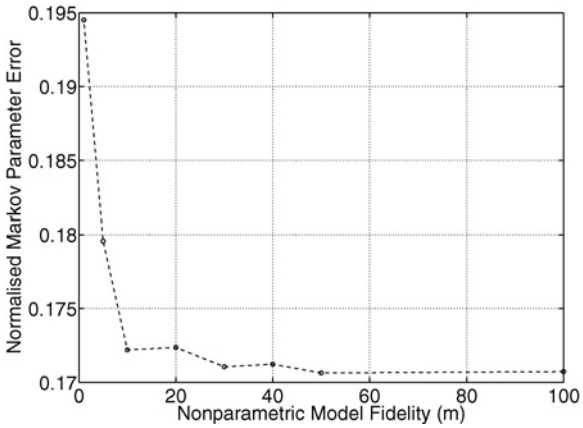

Fig. 25 *RMSE Markov parameter error for an increasing number of points in the non-parametric model*

For each value of m, a 100-run Monte Carlo simulation is performed

7 Numerical examples: error metrics

We now investigate the effect of systematically decreasing the amount of available output data that is used to identify the linear block of the Wiener system. Moreover, we investigate the effect of decreasing m, which determines the number of points in the non-parametric model, and therefore affects the fidelity of $\hat{\mathcal{W}}$.

To quantify the accuracy of the identified semi-parametric model, we compute the root-mean-square error (RMSE) for the first 15 Markov parameters of the true linear system and the identified linear system. The linear model is the same as in Sections 5 and 6, while $\mathcal{W}$ is given by (36).

7.1 Effect of disturbances

To evaluate the effect of σ_{d_1}, σ_{d_2} and σ_{d_3}, we decrease the number of available data points from 4000 to 10. For each case, we perform a 100-run Monte Carlo simulation with a signal-to-noise ratio of 10. We consider the effect of d_1, d_2 and d_3 individually, as well as the effect of all three noise signals, which may be uncorrelated or correlated. Furthermore we consider when d_1 and d_3 are correlated, and d_2 and d_3 are correlated.

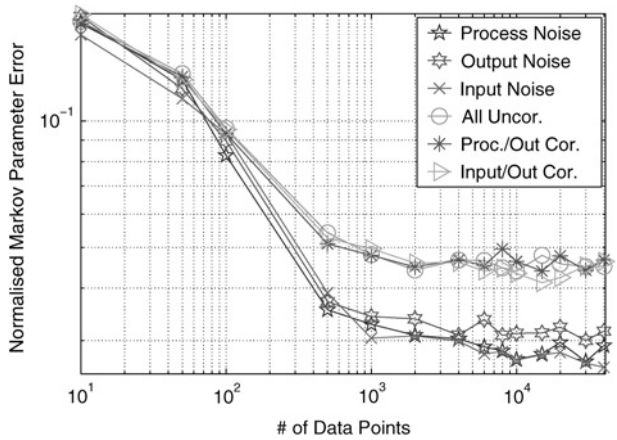

Fig. 24 *RMSE Markov parameter error against number of data points*

For each number of data points we perform a 100-run Monte Carlo simulation

Fig. 24 demonstrates the increase in error for decreasing amounts of available data. Furthermore, we see that the cases with correlated disturbances yield similar results compared to the case with uncorrelated disturbances.

7.2 Non-parametric model accuracy

We now perform a Monte Carlo simulation to evaluate how m affects the accuracy of the identified linear system. Specifically, we vary m from 1 to 100. For each value of m we average the result over 100 simulations. We consider the nominal case, that is, without noise.

Fig. 25 shows that RMSE generally decreases as m increases. Note that, for this example, only a slight decrease in RMSE is observed for $m \geq 20$.

8 Conclusions

In this paper we develop a two-step method to identify semi-parametric models for SISO discrete-time Wiener systems. We make only two assumptions about the system, namely, the linear dynamic block is assumed to be asymptotically stable, and the static non-linearity is assumed to be piecewise continuous. Furthermore, this method requires identification signals with specific properties for each of the two steps, as discussed as follows.

First, we choose a single harmonic input and measure the system output when the state trajectory is in harmonic steady state. By exploiting symmetry properties of these signals, we approximate the non-harmonic phase shift and, therefore estimate the intermediate signal. Using the estimate of the intermediate signal, a non-parametric model of the static non-linearity is obtained.

Second, using the identified non-parametric model, we use RCO to identify a parametric model of the dynamic system. As commonly assumed in the system identification literature, the identification signal for this step is assumed to be sufficiently persistent such that the dynamic linear system can be identified.

It is important to point out that the method investigated in this work does not require invertibility of the non-linearity, which is a common assumption in Wiener identification. However, the cost of removing this assumption is the need

doi: 10.1049/iet-cta.2010.0023

for two steps, and the requirement that the signal for the first step be a single harmonic. Furthermore, the user must wait until the system has reached a steady state before useful data can be obtained. On the other hand, from Section 1, recall that there are methods based on multiple harmonic inputs in the literature. Finally, it should be noted that, although a non-parametric model of the non-linearity was used in this discussion, the data that represent the non-parametric map could be parameterised.

The two-step method presented in this paper is effectively demonstrated on several examples of increasing complexity, including non-linearities in the form of both even and non-even polynomials, deadzone, saturation and discontinuity, and disturbances on the form of process, input and output noise.

9 Acknowledgments

This work was supported in part by NASA through grants NNX08AB92A and NNX08BA57A, USA, and by FAPEMIG and CNPq, Brazil.

10 References

1 Billings, S.A.: 'Identification of non-linear systems – a survey', *IEE Proc. D*, 1980, **127**, (6), pp. 272–285
2 Greblicki, W., Pawlak, M.: 'Nonparametric system identification' (Cambridge University Press, 2008)
3 Haber, R., Keviczky, L.: 'Nonlinear system identification – input–output modeling approach' (Kluwer Academic Publishers, 1999, vol. 1)
4 Aguirre, L.A., Coelho, M.C.S., Corrêa, M.V.: 'On the interpretation and practice of dynamical differences between Hammerstein and Wiener models', *IEE Proc. Control Theory Appl.*, 2005, **152**, (4), pp. 349–356
5 Bai, E.W.: 'Frequency domain identification of Wiener models', *Automatica*, 2003, **39**, (9), pp. 1521–1530
6 Bai, E.W., Reyland, J.: 'Towards identification of Wiener systems with the least amount of *a Priori* information: IIR cases', *Automatica*, 2009, **45**, pp. 956–964
7 Crama, P., Schoukens, J.: 'Initial estimates of Wiener and Hammerstein systems using multisine excitation', *IEEE Trans. Instrum. Meas.*, 2001, **50**, (6), pp. 1791–1795
8 Hagenblad, A., Ljung, L., Wills, A.: 'Maximum likelihood identification of Wiener models', *Automatica*, 2008, **44**, (11), pp. 2697–2705
9 Lacy, S., Bernstein, D.S.: 'Identification of FIR Wiener systems with unknown, non-invertible, polynomial nonlinearities', *Int. J. Control*, 2003, **76**, pp. 1500–1507
10 Lacy, S., Erwin, R.S., Bernstein, D.S.: 'Identification of Wiener systems with known noninvertible nonlinearities', *J. Dyn. Syst. Meas. Control*, 2001, **123**, pp. 566–571
11 Tötterman, S., Toivonen, H.T.: 'Support vector method for identification of Wiener models', *J. Process Control*, 2009, **19**, pp. 1174–1181
12 Bai, E.W.: 'Decoupling the linear and nonlinear parts in Hammerstein model identification', *Automatica*, 2004, **40**, (4), pp. 671–676
13 Ding, F., Shi, Y., Chen, T.W.: 'Auxiliary model-based least-squares identification methods for Hammerstein output-error systems', *Syst. Control Lett.*, 2007, **56**, (5), pp. 373–380
14 Palanthandalam-Madapusi, H., Renk, E.L., Bernstein, D.S.: 'Data-based model refinement for linear and Hammerstein systems using subspace identification and adaptive disturbance rejection'. Forty-fourth IEEE Proc. Conf. on Control Application, Toronto, Canada, August 2005, pp. 1630–1635
15 Wang, J., Zhang, Q., Ljung, L.: 'Revisiting Hammerstein system identification through the two-stage algorithm for bilinear parameter estimation', *Automatica*, 2009, **45**, (11), pp. 2627–2633
16 Hoagg, J.B., Santillo, M.A., Bernstein, D.S.: 'Discrete-time adaptive command following and disturbance rejection for minimum phase systems with unknown exogenous dynamics', *IEEE Trans. Autom. Control*, 2008, **53**, (4), pp. 912–928
17 Santillo, M.A., Bernstein, D.S.: 'A retrospective correction filter for discrete-time adaptive control of nonminimum phase systems'. Forty-seventh IEEE Proc. Conf. on Decision Control, Cancun, Mexico, December 2008, pp. 690–695
18 Venugopal, R., Bernstein, D.S.: 'Adaptive disturbance rejection using ARMARKOV system representations', *IEEE Trans. Control Syst. Tech.*, 2000, **8**, pp. 257–269
19 D'Amato, A.M., Bernstein, D.S.: 'LFT identification using retrospective cost optimization'. Proc. SYSID, Saint-Malo, France, July 2009, pp. 450–455
20 Santillo, M.A., D'Amato, A.M., Bernstein, D.S.: 'System identification using a retrospective correction filter for adaptive feedback model updating'. Proc. American Control Conf., St. Louis, MO, June 2009, pp. 4392–4397
21 D'Amato, A.M., Teixeira, B.O.S., Bernstein, D.S.: 'Semiparametric identification of Wiener systems using a single harmonic input and retrospective cost optimization'. Proc. American Control Conf., Baltimore, MD, June–July 2010, pp. 4498–4503
22 Hoagg, J.B., Bernstein, D.S.: 'Cumulative retrospective cost adaptive control with RLS-based optimization'. Proc. American Control Conf., Baltimore, MD, June–July 2010, pp. 4016–4021