ADCHEM 2006

International Symposium on Advanced Control of Chemical Processes
Gramado, Brazil – April 2-5, 2006

IDENTIFICATION OF UNCERTAIN WIENER SYSTEMS

Jose Figueroa [*,1] **Silvina Biagiola** [*]
Osvaldo Agamennoni [*]

[*] *Departamento de Ingeniería Eléctrica y de Computadoras, Universidad Nacional del Sur, Av. Alem 1253; (8000) Bahía Blanca, Argentina*

Abstract: A significant research work has been carried out on modeling, identification and control of processes represented by Wiener models. These models include a cascade connection of a linear time invariant system and a static nonlinearity. Several approaches can be found in the literature to perform the identification process. In this article, we describe a parametric description for the system, that allows to describe the uncertainty as a set of parameters. The proposed algorithm is illustrated through a pH neutralization process.

Keywords: Wiener Models, Process Control, Uncertainty

1. INTRODUCTION

Nonlinear model-based control has been widely diused among the chemical engineering community. The use of models based entirely on fundamental process understanding has the advantage of possessing a clear physical interpretation. However, these models tend to be highly complex making impossible their application in popular model-based control strategies (Pottmann and Pearson, 1998).

On the other hand, purely empirical models (black-box), based entirely on input/output data, lack of physical interpretation. However, they are known to be "successful" and to have good flexibility.

A third approach is used when some physical insight is available, but several parameters remain to be determined from observed data. In this category, Pearson and Pottmann (2000), include three model structures: the Wiener model, the Hammerstein model and the feedback block-oriented model. These models are built from the combination of two components: a static (memoryless) nonlinearity $N(.)$ and a linear time invariant (LTI) system $H(z)$.

In this paper we are interested in Wiener models: a cascade connection of $H(z)$ followed by the static nonlinearity $N(.)$. The use of these models has been treated in literature in dierent contexts (Pearson and Pottmann, 2000; Lussón *et al.*, 2003; Biagiola *et al.*, 2004). Some representation and identification algorithms for uncertain Wiener Models will be presented. The goal is to obtain a nominal model of the process plus a parametric description of the uncertainty, which is the main contribution of this work. For this purpose, Laguerre polynomials are used to model the linear dynamic block, and a piecewise linear (PWL) representation of the nonlinear static block is provided. This modeling approach shows to be advantageous due to its simplicity, easy use and good application results. Moreover, the model

[1] Corresponding author. Email: figueroa@uns.edu.ar. Phone: +54 291 4595101 ext. 3325. FAX: +54 291 4595154.
This work was financially supported by the CONICET, CIC and the Universidad Nacional del Sur.

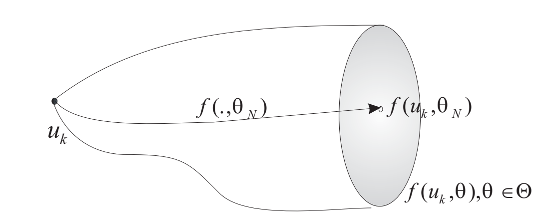

Fig. 1. Model under uncertainties

uncertainty can be easily mapped on to the model parameters.

The paper is organized as follows. In Section 2, general concepts about models and uncertainties are introduced. In Section 3 some usual descriptions and identification techniques of Wiener systems are reviewed. The proposed uncertainty model is presented in Section 4 and an algorithm for parameter uncertainty characterization is introduced. In Section 5, the results are evaluated on the basis of a simulation of a pH neutralization process. Final remarks are addressed in Section 6.

2. PROCESS INFORMATION, MODELS AND UNCERTAINTIES

Let us consider that process data are available in the form of two sets of process inputs ($\mathbf{u} = \{u_0, u_1, \cdots, u_N\}$) and outputs ($\mathbf{y} = \{y_0, y_1, \cdots, y_N\}$). Then, we aim at finding a mathematical model which approximates these data. This is performed in a two steps procedure.

In the first step, a "type model" is selected. We use the previous knowledge about the process:

$$\hat{y}_{k+1} = F(\hat{y}_k, \cdots, \hat{y}_{k-N_y}, u_k, \cdots, u_{k-N_u}, \theta) \quad (1)$$

where the predicted output at time $k+1$ depends of the previous inputs and predicted outputs and of the set of parameters (θ) to be determined.

In the second step, the parameters (θ) are computed to minimize the di erence between the process and model outputs ($y_k - \hat{y}_k$) to any time. This is usually performed by minimizing the least squared error. In what follows we denote this set of parameters as *nominal parameters* θ_N.

When the interest aims at obtaining an uncertainty related with this nominal model, a typical approach is to define a set of possible models to represent all the process behaviours. This is performed by considering a set of model parameters such that when these parameters θ are used, the whole set of exciting inputs $\mathbf{u}$ is "mapped" onto an output set which contains the set of the output data (see Fig. 1). In this way, we assume the same format for all the possible models in the uncertain set. This models family is defined in terms of a set of parameters.

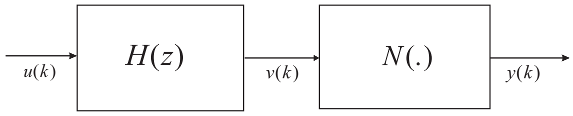

Fig. 2. The Wiener model structure.

3. WIENER MODEL IDENTIFICATION

3.1 Model Description

Figure 2 depicts a Wiener model. It consists of a LTI system $H(z)$ followed by a static nonlinearity $N(.)$. That is, the linear model $H(z)$ maps the input sequence $\{u(k)\}$ into the intermediate sequence $\{v(k)\}$, and the overall model output is $y(k) = N(v(k))$. In the following, there is no loss of generality in assuming $H(1) = 1$, since that any other value of this gain can be included in the nonlinear block (Pearson and Pottmann, 2000).

One of the most common choices for the representation of the linear block are the *Rational Transfer Functions* (Pearson and Pottmann, 2000; Figueroa *et al.*, 2004). Another usual option are the *Linear State Space Models* (Lussón *et al.*, 2003). A drawback of these models is that we need a large number of parameters to describe a system with a slow impulse response or a damped system. Alternative representations, where prior knowledge about the dominant poles can be used, are the *Laguerre and Kautz Models*. For example, the Laguerre model describes the transfer function $H(z)$ with the following basis function expansion,

$$H(z) = \sum_{i=0}^{N_L} h_i L_i(z, a) \quad (2)$$

$$L_i(z, a) = \frac{\sqrt{1 - a^2}}{z - a} \left(\frac{1 - az}{z - a} \right)^{i-1} \quad (3)$$

where the parameters of the model are the coe cients h_i and a is a filter coe cient chosen a priori. The nonlinear block $N(.)$ is, in general, a real-value function of one variable, $y = N(v)$. We describe the nonlinear function as

$$y = \sum_{i=0}^{N_n} \tilde{f}_i \tilde{B}_i(v) \quad (4)$$

where the basis functions $\tilde{B}_i(v)$ have been predetermined, the values $\tilde{f}_i$ are the parameters that should be computed and N_n will be referred to as "order" of the nonlinearity. Once the basis functions $\tilde{B}_i$ are fixed, the output is a linear function of the parameters. This allows us to use a linear regression to estimate the parameters. The two basic advantages of this approach are the low complexity and the uniqueness of the solution. Some possible choices for the basis functions are *Power*

Series, *Chebyshev Polynomials*, *Sigmoid Neural Networks* or *Piecewise Linear Function* (PWL). In particular, the PWL functions have proved to be a very powerful tool in the modeling and analysis of nonlinear systems. The general formulation of PWL functions allows us to represent a non-linear system through a set of linear expressions, each of them valid in a certain operation region. To make this approximation, the domain of variables is partitioned into a set of non-empty regions ^{i}, such that $= \bigcup_{i=1}^{\sigma} {}^{i}$. In each of these regions the non-linear function is approximated using a linear (a ne) representation. These functions allow a systematic and accurate treatment of the approximating functions. It can be proved (Julián *et al.*, 1999) that any nonlinear continuous function $N(v)$: ^{m} ^{1} can be uniquely represented using PWL functions in the form of Eq. (4) as:

$$\tilde{B}_i(v) = (v, \beta_i) \quad (5)$$

where β_i are given parameters that define the partition of the domain of v, and are functions that involve nested absolute values. In this paper we use an orthonormal description of the basis due to its local properties.

3.2 Nominal Model Identification

Di erent methods for Wiener models identification have been reported, and they can be grouped in three main approaches. The first one is an iterative algorithm for Hammerstein models identification (Narendra and Gallman, 1966). If the system is adequately parameterized, then the prediction error can be linearly separated into each set of parameters (the those of the linear and the nonlinear blocks). The estimation is then performed by minimizing alternatively, with respect to each set of parameters.

A second approach, based on correlation techniques (Billings and Fakhouri, 1978), relies on a separation principle, but with the rather restrictive requirement on the input to be white noise.

A recent approach for the identification of block-oriented models is based on least squares estimation and singular value decomposition (Bai, 1998). Due to the particular parameterization used, this method applies only for single input/single output systems. Gómez and Baeyens (2004) performed a more general parameterization to deal with multiple input/ multiple output (MIMO) systems. This approach will be herein followed for nominal model identification.

Let us assume that an input-output data set is available, noted as u_k and y_k, respectively. To obtain these data sets, several aspects should be taken into account. For example, the process should be persistently excited in the whole domain of the nonlinear block, such that all the relevant dynamics is captured.

From Fig. 2, the signal v_k can be written as

$$v_k = H(z) \bullet u_k, \text{ as well as } v_k = N^{-1}(y_k) \quad (6)$$

Equating both sides of these equations (with the inclusion of an error function $\epsilon(k)$ to allow for modeling error) the following equation is obtained

$$\sum_{i=0}^{N_n} f_i B_i(y_k) = h_0 l_0(u_k) + \sum_{i=1}^{N_l} h_i l_i(u_k) + \epsilon(k) \quad (7)$$

or, equivalently,

$$\epsilon(k) = \sum_{i=0}^{N_n} f_i B_i(y_k) - h_0 l_0(u_k) - \sum_{i=1}^{N_l} h_i l_i(u_k) \quad (8)$$

which is a linear regression. Defining

$$\theta = [f_0, f_1, \cdots, f_{N_n}, h_1, h_2, \cdots, h_{N_l}]^T \quad (9)$$

$$\phi = [B_0(y_k), B_1(y_k), \cdots, B_{N_n}(y_k), -l_1(u_k), -l_2(u_k), \cdots, -l_{N_l}(u_k)]^T, \quad (10)$$

Then, Eq. (8) can be written as

$$\epsilon(k) = \theta^T \phi - l_0(u_k) \quad (11)$$

Now, an estimate $\hat{\theta}$ of θ can be computed by minimizing a quadratic criterion on the prediction errors $\epsilon(k)$ (i.e. the least squares estimate). It is well known that this estimate is given by:

$$\hat{\theta} = \left({}_{N} {}^{T}_{N} \right)^{-1} {}_{N} \quad (12)$$

where $= [-l_0(u_1), \cdots, -l_0(u_N)]^T$ and $= [\phi(1), \cdots, \phi(N)]$ are formed using the set of the N data available from the process.

Now, estimates of the parameters $\hat{f}_i$ $(i = 0, \cdots, N_n)$, $\hat{h}_0 = 1$ and $\hat{h}_i$ $(i = 1, \cdots, N_l)$ can be computed by partitioning the estimate $\hat{\theta}$, according to the definition of θ in (9). It is important to remark that we are identifying the inverse of the nonlinearity, which is frequently used in many control applications.

4. UNCERTAINTY CHARACTERIZATION

In this section we develop an algorithm, based on the ideas of Section 2, to characterize the uncertainties of the model obtained in Section 3. We introduce a set of parameters H for the linear dynamic block and a set F for the parameters of the inverse of the nonlinear block:

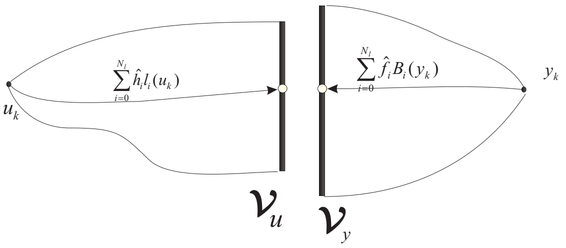

Fig. 3. Uncertainty sets in Wiener Model

$$\mathcal{H} = \left\{ h : h = \hat{h} + \delta^h, h_i^l \leq \delta_i^h \leq h_i^u \, 1 \leq i \leq N_l \right\} \quad (13)$$

$$\mathcal{F} = \left\{ f : f = \hat{f} + \delta^f, f_i^l \leq \delta_i^f \leq f_i^u \, 1 \leq i \leq N_n \right\} \quad (14)$$

To define these bounds, let us define some sets. Given the input data u_k, the linear uncertain system defined by H maps at some specific time k over a set

$$\mathrm{V}_u = \left\{ v : v = \sum_{i=0}^{N_l} h_i l_i(u_k), h \quad \mathrm{H} \right\} \qquad (15)$$

Given an input u_k, the Laguerre term of order i, $l_i(u_k)$ is a real number and the set V_u takes the form of $\mathrm{V}_u = \{v : v_l \quad v \quad v_u\}$.

On the other hand, if we consider the uncertain description of the parameters in F, a given output y_k maps at some specific time k over a set

$$\mathrm{V}_y = \left\{ v : v = \sum_{i=0}^{N_n} f_i B_i(y_k), f \quad \mathrm{F} \right\} \qquad (16)$$

This situation is showed in Fig. 3. From this picture it is clear that the parameters set will describe the uncertainties description of Section 2 if $\mathrm{V}_y \quad \mathrm{V}_u = \quad$. In this way, the point u_k is mapped onto V_u through H. Then, since $\mathrm{V}_y \quad \mathrm{V}_u = \quad$, this point will be mapped in y_k through the inverse of F. Then, it is only necessary to compute the parameters bounds to satisfy this condition. The nominal linear model parameters $\hat{h}_i$ can be written as a vector, by considering that the Laguerre basis $l_i(u_k)$ are a set of real numbers for each input u_k. Let $l(u_k)$ be the vector which i^{th} entry is the Laguerre basis $l_i(u_k)$. Then, the expression of the linear model is

$$\hat{v}(k) = \hat{h}^T l(u_k). \qquad (17)$$

In a similar way, the PWL basis $B_i(y_k)$ are a set of positive real numbers for each output y_k. $B(y_k)$ is the vector whose i^{th} entry is the PWL basis $B_i(y_k)$. Then, the linear model expression is:

$$v(k) = \hat{f}^T B(y_k). \qquad (18)$$

In the following, let us analyze the bounds on the parameters.

4.1 Uncertainty concentrated in the linear block

In this case, let us assume that the uncertainty is concentrate in the linear block. Then, we are looking for the uncertain model that maps the set of data $\mathbf{u}$ to the set $\mathbf{v} = \hat{f}^T B(\mathbf{y})$. To define an uncertain model that allows to describe the complete set of data, we should compute the set $\left\{ h : h = \hat{h} + \delta^h, h_i^l \quad \delta_i^h \quad h_i^u \right\}$. Now, since that the entries of $l(u_k)$ could be positive or negative, it is possible to split the vector $l(u_k)$ by defining $l^+(u_k) = max(l(u_k), 0)$ and $l^-(u_k) = min(l(u_k), 0)$. Then, forming the vector $\gamma = \left[-(l^-(u_k))^T, (l^+(u_k))^T \right]^T$, we can compute the uncertainties bounds as

$$\min_{h^l, h^u} \sum_{i=1}^{N_l} \left(h_i^l + h_i^u \right) \qquad (19)$$

s.t.

$$\left[(h^l)^T, (h^u)^T \right] \gamma \quad e(k), \text{ if } e(k) \quad 0; k = 1, \cdots, N$$

$$-\left[(h^l)^T, (h^u)^T \right] \gamma \quad e(k), \text{ if } e(k) \quad 0; k = 1, \cdots, N$$

$$h_i^l, h_i^u \quad 0$$

$$\text{where} \qquad e(k) = \hat{c}^T B(y_k) - \hat{h}^T l(u_k) \qquad (20)$$

4.2 Uncertainty concentrated in the nonlinear block

In this case, let us assume that the uncertainty is concentrated in the nonlinear stationary block. Then, we are looking for the uncertain model that maps the set of data $\mathbf{y}$ to the set $\mathbf{v} = \hat{h}^T l(\mathbf{u})$. Then, to define an uncertain model that allows to describe the complete set of data, we should compute the set $\left\{ f : f = \hat{f} + \delta^f, f_i^l \quad \delta_i^f \quad f_i^u \right\}$. Now, since that the entries of $B(y_k)$ are positive, we can compute the upper bound uncertainties as

$$\min_{f^u} \sum_{i=1}^{N_n} f_i^u \qquad (21)$$

$$s.t. \quad (f^u)^T B(y_k) \quad e(k), k = 1, \cdots, N$$

$$f_i^u \quad 0$$

and the lower bound as

$$\min_{f^l} \sum_i f_i^l \qquad (22)$$

$$s.t. \quad -(f^l)^T B(y_k) \quad e(k), k = 1, \cdots, N$$

$$f_i^l \quad 0$$

4.3 Uncertainty in both the linear and nonlinear blocks

In this case, we consider the most general case, where uncertainty is present in both models. Note that the intersection of the uncertainties in the linear and nonlinear models should be non empty. This can be solved as:

$$\min_{h^l,h^u,f^l,f^u} \sum_i \left(h_i^l + h_i^u + f_i^l + f_i^u\right)$$
$$\text{s.t.} \left[-(h^l)^T, -(h^u)^T, (f^u)^T\right] \begin{bmatrix} \gamma \\ B(y_k) \end{bmatrix} \quad e(k),$$
$$\text{if } e(k) \quad 0; k = 1, \cdots, N$$
$$\left[-(h^l)^T, -(h^u)^T, (f^l)^T\right] \begin{bmatrix} \gamma \\ B(y_k) \end{bmatrix} \quad e(k),$$
$$\text{if } e(k) \quad 0; k = 1, \cdots, N$$

5. PROCESS DESCRIPTION

To illustrate the identification procedure, simulation results were obtained. The example consists of the neutralization reaction between a strong acid (HA) and a strong base (BOH) in the presence of a bu er agent (BX) (Galán, 2000). The neutralization takes place in a CSTR with a constant volume V. An acidic solution with a time-varying flow $q_A(t)$ of composition $x_{1i}(t)$ is neutralized using an alkaline solution with flow $q_B(t)$ of known composition made up of base x_{2i} and bu er agent x_{3i}. For this specific case, under some assumptions, the dynamic behavior of the process can be described considering the state variables: $x_1 = [A^-]$, $x_2 = [B^+]$ and $x_3 = [X^-]$. Then, the mathematical model of the process is:

$$\dot{x}_1 = q_A/V\, x_{1i} - (q_A + q_B)/V\, x_1 \quad (23)$$
$$\dot{x}_2 = q_B/V\, x_{2i} - (q_A + q_B)/V\, x_2 \quad (24)$$
$$\dot{x}_3 = q_B/V\, x_{3i} - (q_A + q_B)/V\, x_3 \quad (25)$$
$$F(x,\xi) \quad \xi + x_2 + x_3 - x_1 - K_w/\xi$$
$$- x_3/[1 + (K_x\xi/K_w)] = 0 \quad (26)$$

where $\xi = 10^{-pH}$. The parameters of the system are addressed in Table 1. Using this model a set

Table 1. Neutralization Parameters

Parameter	Value
x_{1i}	0.0012 $mol\ HCL/l$
x_{2i}	0.0020 $mol\ NaOH/l$
x_{3i}	0.0025 $mol\ NaHCO_3/l$
K_x	10^{-7} mol/l
K_w	10^{-14} mol^2/l^2
q_A	1 l/m
V	2.5 l

of data is generated by simulating 2000 samples with a sample time $T_s = 0.5$. A random signal uniformly distributed in $[0,1]$ is applied to the manipulated variable q_B, this input changes each five samples. A random gaussian noise with zero media and variance 0.5 is added to the measured pH. Before proceeding with the identification, the steady values are removed from input $(q_B = 0.5)$ and output $(pH = 7.7182)$, respectively.

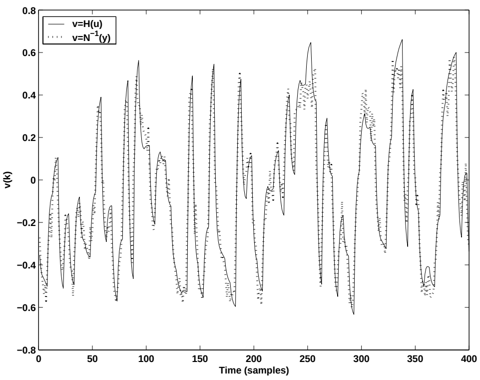

Fig. 4. Simulation for the nominal Wiener model

In a first step, we compute a nominal Wiener Model as described in Section 3. We consider three Laguerre polynomials (i.e. $N_l = 3$) with $a = 0.7$ to represent the linear model and a PWL with 8 sections partition to describe the nonlinear static gain. The identification is performed using a set of 1000 data, and the remaining data are used for validation. Figure 4 shows a set of these results, restricted to 400 samples (half for identification and half for validation). Two curves are shown: the signal $v(k)$ as the output of the linear block and as the output of the inverse of the nonlinear block $N^{-1}(y(k))$. The parameters are:

$$h^T = \left[1 - 0.2022\ 0.1386\right]$$
$$f^T = [-0.660\ -0.445\ -0.416\ -0.389\ -0.374$$
$$-0.303\ -0.042\ 0.132\ 0.204\ 0.219\ 0.557]$$

for the linear and the nonlinear blocks, respectively.

In a second step, we assume the uncertainty is concentrated in the linear block. By solving the problem described in Section 4.1, the uncertainty (see Fig. 5) in the parameters is described by:

$$h^u = \left[0.5320\ 0.120\ 0.315\right]$$
$$h^l = \left[0.427\ 0.174\ 0.319\right]$$

The case with uncertain nonlinear parameters is now considered. Solving the problem of Section 4.2, the parameter bounds (see Fig. 6) are:

$$f^u = [0.000\ 0.083\ 0.060\ 0.074\ 0.056\ 0.135$$
$$0.293\ 0.355\ 0.216\ 0.478\ 0.053\ ]^T$$
$$f^l = [0.000\ 0.137\ 0.260\ 0.000\ 0.273\ 0.304$$
$$0.404\ 0.054\ 0.295\ 0.206\ 0.079]^T$$

Finally, let us consider the case with uncertainty in both blocks. Solving the problem of Section 4.3, the parameter bounds (see Fig. 7) are:

$$f^u = [0.029\ 0.156\ 0.082\ 0.131\ 0.124\ 0.147$$

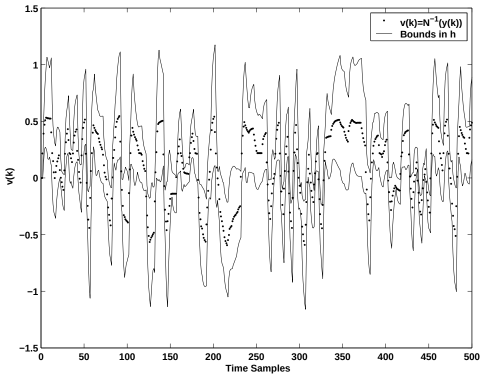

Fig. 5. Uncertainty in linear parameters

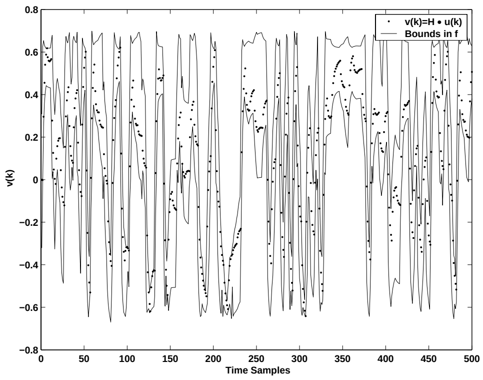

Fig. 6. Uncertainty in nonlinear parameters

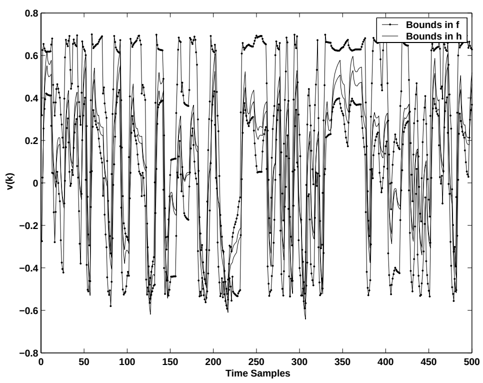

Fig. 7. Uncertainty in linear and nonlinear parameters

$$
\begin{aligned}
& \quad 0.312\ \ 0.341\ \ 0.215\ \ 0.479\ \ 0.053]^T \\
f^l &= [0.000\ \ 0.000\ \ 0.174\ \ 0.000\ \ 0.133\ \ 0.231 \\
& \quad 0.342\ \ 0.055\ \ 0.267\ \ 0.163\ \ 0.106]^T \\
h^u &= \left[\,0.000\ 0.000\ 0.046\,\right] \\
h^l &= \left[\,0.0833\ 0.000\ 0.000\,\right]
\end{aligned}
$$

6. CONCLUSIONS

In this article, identification and robustness analysis of Wiener systems are considered. Different representations had been compared in terms of robust modeling capabilities. PWL functions were used to represent the nonlinear gain, with benefits due to its good approximation level. The simultaneous identification approach herein used showed a slight advantage in terms of approximation errors. These errors exhibit a linear dependence on the model parameters, which reduces the complexity of the identification formulation.

REFERENCES

Bai, E. (1998). An optimal two-stage identification algorithm for Hammerstein-Wiener nonlinear systems. *Automatica* **34, 3**, 333–338.

Biagiola, S., O. Agamennoni and J.L. Figueroa (2004). H_∞ control of a Wiener type system. *International Journal of Control* **77, 6**, 572–583.

Billings, S. and S. Fakhouri (1978). Identification of nonlinear systems using the correlation techniques. *Proceedings IEE* **125**, 691–697.

Figueroa, J.L., J. Cousseau and R. de Figueiredo (2004). A simplicial canonical piecewise linear adaptive filter. *Circuits, System and Signal Processing* **23**, 365–386.

Galán, O. (2000). *Robust multi-linear model-based control for nonlinear plants*. PhD thesis. University of Sydney, Australia.

Gómez, J.C. and E. Baeyens (2004). Identification of block-oriented nonlinear systems using orthonormal bases. *Journal of Process Control* **14**, 685–697.

Julián, P., A.C. Desages and O.E. Agamennoni (1999). High level canonical piecewise linear representation using a simplicial partition. *IEEE Transactions on Circuits and Systems* **CAS-46**, 463–480.

Lussón, A., O. Agamennoni and J.L. Figueroa (2003). A nonlinear model predictive control scheme based on Wiener piecewise linear models. *Journal of Process Control* **13**, 655–666.

Narendra, K.S. and P.G. Gallman (1966). An iterative method for the identification of nonlinear systems using a Hammerstein model. *IEEE Transactions on Automatic Control* **AC-11**, 546–550.

Pearson, R.K. and M. Pottmann (2000). Gray-box identification of block-oriented nonlinear models. *Journal of Process Control* **10**, 301–315.

Pottmann, M. and R.K. Pearson (1998). Block-oriented NARMAX models with output multiplicities. *AIChE Journal* **44, 1**, 131–140.