15th Triennial World Congress, Barcelona, Spain

SUBSPACE IDENTIFICATION OF MULTIVARIABLE HAMMERSTEIN AND WIENER MODELS

Juan C. Gómez [*,1] **Enrique Baeyens** [**]

[*] *Laboratory for System Dynamics and Signal Processing*
FCEIA - Universidad Nacional de Rosario
Riobamba 245 Bis, 2000 Rosario, Argentina
e-mail: jcgomez@fceia.unr.edu.ar

[**] *Department of Systems Engineering and Automatic Control*
ETSII - Universidad de Valladolid
Paseo del Cauce s/n, 47011 Valladolid, Spain
e-mail: enrbae@eis.uva.es

Abstract: In this paper, subspace-based algorithms for the simultaneous identification of the linear and nonlinear parts of multivariable Hammerstein and Wiener models are presented. The proposed algorithms consist basically of two steps. The first one is a standard (linear) subspace algorithm applied to an equivalent linear system whose inputs (respectively outputs) are filtered (by the nonlinear functions describing the static nonlinearities) versions of the original inputs (respectively outputs). The second step consists in a 2-norm minimization problem which is solved via a Singular Value Decomposition. *Copyright ©2002 IFAC*

Keywords: Subspace Identification Methods, Hammerstein and Wiener models, Singular Value Decomposition.

1. INTRODUCTION

Modelling, identification, and control of nonlinear systems has been the subject of many research activities in the last decades. In contrast to linear models that approximate the system around a given operating point, nonlinear models are able to describe the global behaviour of the system over the entire operating range. One of the most frequently studied class of nonlinear models is the one that corresponds to the so-called *block-oriented* models, which consist of the interconnection of Linear Time-Invariant (LTI) systems and static (memoryless) nonlinearities. Among this class, two of the most frequently studied models are:

- the **Hammerstein model**, where the static nonlinearity is followed by a LTI system in a cascade connection (see (Narendra and Gallman, 1966), (Billings, 1980), (Billings and Fakhouri, 1982), (Eskinat *et al.*, 1991), (Boutayeb and Darouach, 1995), (Pearson and Pottmann, 2000) for different identification algorithms for Hammerstein models), and
- the **Wiener model**, in which the order of the linear and nonlinear blocks in the cascade connection is reversed (see for instance (Greblicki, 1994), (Wigren, 1993), for different identification methods for Wiener models).

These models have been successfully used to represent nonlinear systems in a number of practical applications in the areas of chemical processes (Eskinat *et al.*, 1991), (Kalafatis *et al.*, 1995), (Pearson and Pottmann, 2000), biological processes (Korenberg, 1973), signal processing (Stapleton and Bass, 1985), and control (Fruzzetti *et al.*, 1997).

In recent years, considerable amount of research has been devoted to the development of new identifica-

[1] Author to whom all correspondence should be addressed.

tion methods that are able to deliver reliable state-space models of multivariable LTI systems directly from input-output data, and that require a modest computational load without the need of iterative optimization procedures. These techniques have become collectively known as Subspace-based State Space IDentification (4SID) methods (see (Van Overschee and de Moor, 1994), (Viberg, 1995), and the references therein). The methods have their origin in state-space realization theory as developed in the sixties/seventies, and the main computational tools are QR and Singular Value Decomposition (SVD). Although there is a well developed theory for subspace methods for LTI systems, this is not the case for nonlinear systems. Among some recent contributions in this area, the works by Verhaegen and Westwick on subspace-based identification of MIMO Hammerstein and Wiener models (Verhaegen and Westwick, 1996), (Westwick and Verhaegen, 1996), and the works by Chen and coauthors (Chen and Maciejowski, 2000), and by Favoreel and coauthors (Favoreel *et al.*, 1999) on subspace identification of bilinear systems, can be mentioned.

In this paper, new subspace-based algorithms for the identification of Hammerstein and Wiener models are presented. The proposed algorithms consist of two basic steps. For the Hammerstein model, the first step is a standard (linear) subspace algorithm applied on an equivalent linear system whose inputs are filtered (by the nonlinear functions describing the static nonlinearities) versions of the original inputs, while the second step consists in a 2-norm minimization problem which is solved via an SVD. On the other hand, for the Wiener model, the first step is a standard (linear) subspace algorithm applied on an equivalent linear system whose outputs are filtered (by the nonlinear functions describing the inverse of the static nonlinearities) versions of the original outputs, while the second step is again a 2-norm minimization problem solved via an SVD.

2. HAMMERSTEIN MODEL IDENTIFICATION

2.1 *Problem Formulation*

A (multivariable) Hammerstein model is schematically represented in figure 1. The model consists of a zero-memory nonlinear element $\mathbf{N}(\cdot)$ in cascade with a Linear Time Invariant (LTI) system with state-space representation

$$x_{k+1} = Ax_k + Bv_k + \omega_k, \tag{1}$$

$$y_k = Cx_k + Dv_k + \nu_k, \tag{2}$$

where $y_k \in \mathbb{R}^m$, $x_k \in \mathbb{R}^n$, $v_k \in \mathbb{R}^p$, $\omega_k \in \mathbb{R}^n$, and $\nu_k \in \mathbb{R}^m$, are the LTI system output, state, input, process noise and output measurement noise vectors at time k, respectively, and where A, B, C and D are the (unknown) system matrices of appropriate dimensions. It will be assumed that the nonlinear zero-memory block can be described by a linear combination of basis functions in the form

$$v_k = \mathbf{N}(u_k) = \sum_{i=1}^{r} \alpha_i g_i(u_k), \tag{3}$$

where $g_i(\cdot) : \mathbb{R}^p \to \mathbb{R}^p$, $(i = 1, \cdots, r)$, are the assumed known basis functions, $\alpha_i \in \mathbb{R}^{p \times p}$, $(i = 1, \cdots, r)$ are unknown matrix parameters, and where $u_k \in \mathbb{R}^p$ is the Hammerstein model input vector at time k. Typically, the basis functions are polynomials [2], but they can also be basis functions generated by translations and dilations of a *mother* function (*e.g.*, wavelets, or Radial Basis Functions).

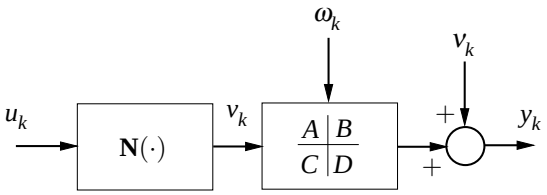

Fig. 1. Multivariable Hammerstein Model.

The identification problem is to estimate the unknown parameter matrices α_i, $(i = 1, \cdots, r)$, and A, B, C, D, characterizing the nonlinear and the linear parts of the system, respectively, and the model order n, from an N-point data record $\{u_k, y_k\}_{k=1}^{N}$ of observed input-output measurements.

2.2 *Subspace Identification Algorithm*

Substituting equation (3) into (1) and (2), these last two equations can be written as

$$x_{k+1} = Ax_k + \sum_{i=1}^{r} B\alpha_i g_i(u_k) + \omega_k, \tag{4}$$

$$y_k = Cx_k + \sum_{i=1}^{r} D\alpha_i g_i(u_k) + \nu_k. \tag{5}$$

It is clear from equations (4)-(5) that the parameterization (1)-(3) is not unique, since any parameter matrices $B\beta$, $D\beta$ and $\beta^{-1}\alpha_i$, for some nonsingular matrix $\beta \in \mathbb{R}^{p \times p}$, provide the same description (4)-(5). In other words, any identification experiment can not distinguish between the parameters B, D, α_i and $B\beta, D\beta, \beta^{-1}\alpha_i$, respectively. To obtain a one-to-one parameterization, i.e. for the system to be identifiable, additional contraints must be imposed on the parameters. A technique that can be used to obtain uniqueness is to normalize the parameter matrices α_i, that is to assume for instance that $\|\alpha_i\|_2 = 1$. A similar methodology was employed in (Bai, 1998) for a scalar Hammerstein-Wiener model. Under this assumption the parameterization (1)-(3) is unique.

[2] Any smooth function in an interval can be represented accurately by a polynomial of sufficiently high order.

Defining now $\widetilde{B} \triangleq \left[B\alpha_1, \cdots, B\alpha_r\right]$, $\widetilde{D} \triangleq \left[D\alpha_1, \cdots, D\alpha_r\right]$, and $U_k \triangleq \left[g_1(u_k)^T, \cdots, g_r(u_k)^T\right]^T$, equations (4) and (5) can be written as

$$x_{k+1} = Ax_k + \widetilde{B}U_k + \omega_k, \tag{6}$$

$$y_k = Cx_k + \widetilde{D}U_k + v_k. \tag{7}$$

Equations (6)-(7) can be interpreted as a state-space realization of a LTI system whose input U_k is a filtered (by the assumed known vector fields $g_i(\cdot)$) version of the original input u_k. It is clear then that any available subspace identification algorithm (such as the N4SID algorithm by Van Overschee and de Moor (Van Overschee and de Moor, 1994), the MOESP algorithm by Verhaegen (Verhaegen, 1994), or the CVA algorithm by Larimore (Larimore, 1990), can be employed to obtain estimates $\widehat{A}, \widehat{\widetilde{B}}, \widehat{C}$, and $\widehat{\widetilde{D}}$ of the system matrices $A, \widetilde{B}, C$, and $\widetilde{D}$, respectively, from input-output data.

Defining $\alpha \triangleq \left[\alpha_1, \alpha_2, \cdots, \alpha_r\right]^T$, matrices $\widetilde{B}$ and $\widetilde{D}$ can be written as $\widetilde{B} = B\alpha^T$, and $\widetilde{D} = D\alpha^T$, which can be expressed in a combined form as

$$\left[\widetilde{B}^T \widetilde{D}^T\right]^T \triangleq \Theta_{BD} = \begin{bmatrix} B \\ D \end{bmatrix} \alpha^T. \tag{8}$$

The problem now is how to compute estimates of the parameter matrices B, D and α from an estimate $\widehat{\Theta}_{BD}$ of the matrix Θ_{BD}. It is clear that the closest, in the 2-norm sense, estimates $\widehat{B}$, $\widehat{D}$ and $\widehat{\alpha}$ are such that

$$\left(\widehat{B}, \widehat{D}, \widehat{\alpha}\right) = \underset{B,D,\alpha}{\arg\min} \left\{ \left\| \widehat{\Theta}_{BD} - \begin{bmatrix} B \\ D \end{bmatrix} \alpha^T \right\|_2^2 \right\}. \tag{9}$$

The solution to this optimization problem is provided by the SVD (Golub and Van Loan, 1989) of the matrix $\widehat{\Theta}_{BD}$. The result is summarized in the following Theorem.

Theorem 2.1. Let $\widehat{\Theta}_{BD} \in \mathbb{R}^{(n+m)\times rp}$ have rank $s > p$, and let the 'economy-size' SVD of $\widehat{\Theta}_{BD}$ be given by

$$\widehat{\Theta}_{BD} = U_s \Sigma_s V_s^T = \sum_{i=1}^{s} \sigma_i u_i v_i^T \tag{10}$$

where Σ_s is a diagonal matrix containing the s nonzero singular values $(\sigma_i, i = 1, \cdots, s)$ of $\widehat{\Theta}_{BD}$ in nonincreasing order, and where the matrices $U_s = \left[u_1\ u_2\ \cdots\ u_s\right] \in \mathbb{R}^{(n+m)\times s}$ and $V_s = \left[v_1\ v_2\ \cdots\ v_s\right] \in \mathbb{R}^{rp\times s}$ contain only the first s columns of the unitary matrices $U \in \mathbb{R}^{(n+m)\times(n+m)}$ and $V \in \mathbb{R}^{rp\times rp}$ provided by the full SVD of $\widehat{\Theta}_{BD}$,

$$\widehat{\Theta}_{BD} = U\Sigma V^T, \tag{11}$$

respectively. Then, the matrices $\widehat{\alpha} \in \mathbb{R}^{rp\times p}$, $\widehat{B} \in \mathbb{R}^{n\times p}$, and $\widehat{D} \in \mathbb{R}^{m\times p}$ that minimize the norm

$$\left\| \widehat{\Theta}_{BD} - \begin{bmatrix} \widehat{B} \\ \widehat{D} \end{bmatrix} \widehat{\alpha}^T \right\|_2^2,$$

are given by

$$\left(\begin{bmatrix} \widehat{B} \\ \widehat{D} \end{bmatrix}, \widehat{\alpha} \right) = \left(U_1 \Sigma_1, V_1\right), \tag{12}$$

where $\Sigma_1 = \mathrm{diag}\left\{\sigma_1, \sigma_2, \cdots, \sigma_p\right\}$, $U_1 \in \mathbb{R}^{(n+m)\times p}$, and $V_1 \in \mathbb{R}^{rp\times p}$, are given by the following partition of the 'economy size' SVD in (10),

$$\widehat{\Theta}_{BD} = \begin{bmatrix} U_1\ U_2 \end{bmatrix} \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix} \begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix}, \tag{13}$$

and the approximation error is given by

$$\left\| \widehat{\Theta}_{BD} - \begin{bmatrix} \widehat{B} \\ \widehat{D} \end{bmatrix} \widehat{\alpha}^T \right\|_2^2 = \sigma_{p+1}^2. \tag{14}$$

Proof: The result is a direct application of Theorem 2.5.3 (pp. 72-73) in (Golub and Van Loan, 1989). ■

Based on this result, the nonlinear subspace identification algorithm can then be summarized as follows.

Algorithm 2.1.

Step 1: Compute estimates $\left(\widehat{A}, \widehat{\widetilde{B}}, \widehat{C}, \widehat{\widetilde{D}}\right)$ of the systems matrices $\left(A, \widetilde{B}, C, \widetilde{D}\right)$ in (6)-(7) using any available subspace algorithm for LTI systems.

Step 2: Based on the estimates $\widehat{\widetilde{B}}$ and $\widehat{\widetilde{D}}$ compute an estimate $\widehat{\Theta}_{BD}$ of the matrix Θ_{BD} defined in (8).

Step 3: Compute the 'economy size' SVD of $\widehat{\Theta}_{BD}$ as in Theorem 2.1 , and the partition of this decomposition as in equation (13).

Step 4: Compute the estimates of the parameter matrices B, D and α as $\begin{bmatrix} \widehat{B} \\ \widehat{D} \end{bmatrix} = U_1\Sigma_1$, and $\widehat{\alpha} = V_1$, respectively, with U_1, V_1 and Σ_1 defined as in Theorem 2.1. □

A similar procedure was suggested in (Rangan *et al.*, 1995), however the formulation here is more general, in the sense that any available subspace algorithm can be used as a first step, and proofs of consistency are provided that are not given in (Rangan *et al.*, 1995).

Under some assumptions on the persistency of excitation of the inputs (which depend on the particular subspace identification method used as the first step of the algorithm [3]) the estimates $\left(\widehat{A}, \widehat{\widetilde{B}}, \widehat{C}, \widehat{\widetilde{D}}\right)$ are consistent in the sense that they converge to the 'true' values as the number of data points $N \to \infty$. The convergence of the estimates $\widehat{\widetilde{B}}$ and $\widehat{\widetilde{D}}$ implies that of $\widehat{B}, \widehat{D}$ and $\widehat{\alpha}$. The result is summarized in the following Theorem.

Theorem 2.2. Let $\widehat{\widetilde{B}}$ and $\widehat{\widetilde{D}}$ be consistent estimates computed using the identification Algorithm 2.1. Then, under the uniqueness condition, the estimates $\widehat{B}, \widehat{D}$, and $\widehat{\alpha}$ provided by Algorithm 2.1 are also consistent, in the sense that $\widehat{B} \xrightarrow{\text{a.s.}} B, \widehat{D} \xrightarrow{\text{a.s.}} D$, and $\widehat{\alpha} \xrightarrow{\text{a.s.}} \alpha$, respectively, as $N \to \infty$.

[3] The reader is referred to (Van Overschee and de Moor, 1994), (Verhaegen, 1994), and (Larimore, 1990) for the consistency conditions for the N4SID, MOESP, and CVA algorithms, respectively.

Proof: See Appendix. ■

3. WIENER MODEL IDENTIFICATION

3.1 *Problem Formulation*

A (multivariable) Wiener model is schematically depicted in figure 2. The model consists of the cascade of a LTI system followed by a zero-memory nonlinear element with input-output characteristic given by $\mathbf{N}(\cdot)$. The LTI subsystem has a state-space representation of the form

$$x_{k+1} = Ax_k + Bu_k + \omega_k, \tag{15}$$
$$v_k = Cx_k + Du_k + \nu_k, \tag{16}$$

where A, B, C and D, are the system matrices of appropriate dimensions, and where $x_k \in \mathbb{R}^n$, $v_k \in \mathbb{R}^m$, $u_k \in \mathbb{R}^p$, and $\nu_k \in \mathbb{R}^m$, represent the LTI system state, output, input, and process noise vectors at time k, respectively. It will be assumed that the nonlinear func-

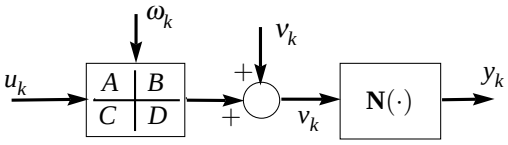

Fig. 2. Multivariable Wiener Model.

tion $\mathbf{N}(\cdot) : \mathbb{R}^m \to \mathbb{R}^m$ is invertible, and that its inverse $\mathbf{N}^{-1}(\cdot)$ can be described as

$$\mathbf{N}^{-1}(y_k) = \sum_{i=1}^{r} \alpha_i g_i(y_k) \tag{17}$$

where now $g_i(\cdot) : \mathbb{R}^m \to \mathbb{R}^m$, $(i = 1, \cdots, r)$, are known smooth vector fields, and $\alpha_i \in \mathbb{R}^{m \times m}$, $(i = 1, \cdots, r)$, are unknown matrix parameters. With this representation for the static nonlinearity, equation (16) can be written as

$$\alpha Y_k \triangleq \sum_{i=1}^{r} \alpha_i g_i(y_k) = Cx_k + Du_k + \nu_k, \tag{18}$$

where $\alpha \triangleq [\alpha_1, \cdots, \alpha_r]$, $Y_k \triangleq \left[g_1^T(y_k), \cdots, g_r^T(y_k)\right]^T$. The Wiener model can then be described as

$$x_{k+1} = Ax_k + Bu_k + \omega_k, \tag{19}$$
$$Y_k = \widetilde{C}x_k + \widetilde{D}u_k + \widetilde{\nu}_k, \tag{20}$$

with $\widetilde{C} \triangleq \alpha^\dagger C$, $\widetilde{D} \triangleq \alpha^\dagger D$, $\widetilde{\nu}_k \triangleq \alpha^\dagger \nu_k$, and where $\alpha^\dagger$ stands for the left pseudoinverse of α. Equations (19)-(20) can be interpreted as a state-space realization of a LTI system whose output Y_k is a filtered (by the assumed known vector fields $g_i(\cdot)$) version of the original output y_k. As in the case of the Hammerstein model, also here any available subspace identification algorithm can be employed to obtain estimates of the system matrices $A, B, \widetilde{C}$, and $\widetilde{D}$ from input-output data.

3.2 *Subspace Identification Algorithm*

Given estimates of the matrices $A, B, \widetilde{C}$, and $\widetilde{D}$, the problem is how to compute estimates of the matrices C, D and α. Proceeding in a similar way at what was done for the Hammerstein model, the best (in the mean squares sense) estimates of matrices C, D and α are such they minimize the norm

$$\left(\widehat{C}, \widehat{D}, \widehat{\alpha}^\dagger\right) = \underset{C,D,\alpha^\dagger}{\arg\min} \left\{ \left\| \left[\widehat{\widetilde{C}}\ \widehat{\widetilde{D}}\right] - \alpha^\dagger \left[C\ D\right] \right\|_2^2 \right\} \tag{21}$$

The solution to this minimization problem is provided by the SVD of the matrix $\left[\widehat{\widetilde{C}}\ \widehat{\widetilde{D}}\right]$. The result is summarized in the following Theorem, which is a restatement of Theorem 2.1.

Theorem 3.1. Let $\left[\widehat{\widetilde{C}}\ \widehat{\widetilde{D}}\right] \in \mathbb{R}^{mr \times (n+p)}$ have rank $s > m$, and let the 'economy-size' SVD of $\left[\widehat{\widetilde{C}}\ \widehat{\widetilde{D}}\right]$ be given by

$$\left[\widehat{\widetilde{C}}\ \widehat{\widetilde{D}}\right] = U_s \Sigma_s V_s^T = \sum_{i=1}^{s} \sigma_i u_i v_i^T \tag{22}$$

with similar definitions for the involved matrices as in Theorem 2.1. Then, the matrices $\widehat{\alpha}^\dagger \in \mathbb{R}^{mr \times m}$, $\widehat{C} \in \mathbb{R}^{m \times n}$, and $\widehat{D} \in \mathbb{R}^{m \times p}$ that minimize the norm

$$\left\| \left[\widehat{\widetilde{C}}\ \widehat{\widetilde{D}}\right] - \alpha^\dagger \left[C\ D\right] \right\|_2^2,$$

are given by

$$\left(\widehat{\alpha}^\dagger, \left[\widehat{C}\ \widehat{D}\right]\right) = \left(U_1, \Sigma_1 V_1^T\right), \tag{23}$$

where $\Sigma_1 = \mathrm{diag}\left\{\sigma_1, \sigma_2, \cdots, \sigma_m\right\}$, $U_1 \in \mathbb{R}^{mr \times m}$, and $V_1 \in \mathbb{R}^{(n+p) \times m}$, are given by the following partition of the 'economy size' SVD in (22),

$$\left[\widehat{\widetilde{C}}\ \widehat{\widetilde{D}}\right] = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix} \begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix}, \tag{24}$$

and the approximation error is given by

$$\left\| \left[\widehat{\widetilde{C}}\ \widehat{\widetilde{D}}\right] - \widehat{\alpha}^\dagger \left[\widehat{C}\ \widehat{D}\right] \right\|_2^2 = \sigma_{m+1}^2. \tag{25}$$

Proof: The proof is identical, *mutatis mutandi*, to the proof of Theorem 2.1, and therefore it is omitted. ■

Based on this result, the Subspace Identification Algorithm for the Wiener model can be summarized as follows

Algorithm 3.1.

Step 1: Compute estimates $\left(\widehat{A}, \widehat{B}, \widehat{\widetilde{C}}, \widehat{\widetilde{D}}\right)$ of the systems matrices $\left(A, B, \widetilde{C}, \widetilde{D}\right)$ in (19)-(20) using any available subspace algorithm for LTI systems.

Step 2: Compute the 'economy size' SVD of $\left[\widehat{\widetilde{C}}\ \widehat{\widetilde{D}}\right]$ as in Theorem 3.1 , and the partition of this decomposition as in equation (24).

Step 3: Compute the estimates of the parameter matrices C, D and α as $\left[\widehat{C}\ \widehat{D}\right] = \Sigma_1 V_1^T$, and

$\widehat{\alpha} = U_1^{\dagger}$, respectively, with U_1, V_1 and Σ_1 defined as in Theorem 3.1. □

The results on consistency of the estimates presented in Theorem 2.2 for the Hammerstein model can be straighforwardly extended to the Wiener model, and therefore they are omitted here.

4. SIMULATION EXAMPLES

To illustrate the proposed identification schemes, two simulation examples are presented in this section.

Example 4.1. **(Hammerstein Model)**
The nonlinear 'true' system consists of a third order linear discrete system with transfer function

$$G(z) = \frac{z^2 + 0.7z - 1.5}{z^3 + 0.9z^2 + 0.15z + 0.002}, \qquad (26)$$

preceded by a static nonlinearity described by a fourth order polynomial of the form

$$\mathbf{N}\left(u_k\right) = 0.8589u_k + 0.0149u_k^2 - 0.5113u_k^3 - 0.0263u_k^4. \qquad (27)$$

The nonlinear characteristic is shown in solid line in the left plot of figure 3.

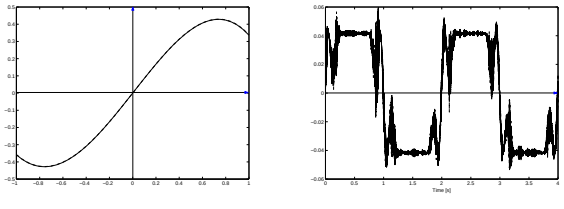

Fig. 3. Left Plot:True (solid line) and Estimated (dashed line) nonlinear characteristic (indistinguishable one from the other). Rigth Plot: Measured (solid line) and Estimated (dashed line) Outputs.

The system was excited with the following input signal

$$u_k = \sin(0.0005\pi k) + 0.5\sin(0.0015\pi k) + 0.3\sin(0.0025\pi k) + 0.1\sin(0.0035\pi k) + \gamma_k,$$

where γ_k is zero-mean Gaussian distributed white noise with standard deviation $\sigma = 0.001$. The output was corrupted with zero-mean coloured noise with spectrum $\Phi_v(\omega) = \frac{0.64\times 10^{-8}}{1.2-0.4\cos(\omega)}$.

Algorithm 2.1 was employed to identified the system from an $(N = 8001)$-point data record of observed input-output measurements. *Step 1* in Algorithm 2.1 was performed using the N4SID algorithm by Van Overschee and de Moor (Van Overschee and de Moor, 1994). A third order LTI subsystem was identified. The estimated transfer function was (compare with the 'true' transfer function (26))

$$\widehat{G}(z) = \frac{0.1797z^3 + 0.9632z^2 + 0.6599z - 1.5624}{z^3 + 1.2261z^2 + 0.2388z + 0.0019}.$$

On the other hand, a fourth order polynomial was used to represent the nonlinear part of the model. The estimated nonlinear model was (compare with the 'true' nonlinearity (27))

$$\widehat{\mathbf{N}}\left(u_k\right) = 0.8594u_k + 0.0108u_k^2 - 0.5107u_k^3 - 0.0222u_k^4.$$

The estimated nonlinear characteristic is represented in dashed line in the left plot of figure 3. It can be observed that it is indistinguishable from the true nonlinear characteristic.

Finally, the measured (solid line) and estimated (dashed line) outputs are represented in the right plot of figure 3, where a good agreement between them can be observed. Note the reader that in this case the system belongs to the model class. □

Example 4.2. **(Wiener Model)**
In this example, a Wiener model is identified based on the simulation data of a pH neutralization process in a constant volume stirring tank considered in (Henson and Seborg, 1994), which corresponds to a bench-scale plant at the University of California, Santa Barbara. The model was derived using the concept of reaction invariants, and it is highly nonlinear mainly due to the titration curve which models the output static characteristic (Henson and Seborg, 1994). The inputs to the system are the base flow rate (u_1) and the buffer flow rate (u_2), in liters per second, while the output (y) is the pH of the solution in the tank. The system was excited with band limited white noise around the nominal values of the base and buffer flow rates. The first six hundred data were used for the estimation of the model, while the following five hundred data were used for validation purposes. The estimation and validation input-output data are represented in the left plot of figure 4. A fifth order linear model was estimated, with eigenvalues at $\{0.9508 \pm 0.2885i, 0.9843 \pm 0.0157i, 0.9880\}$, while a third order polynomial was used to represent the nonlinear part of the model. The true and estimated output (validation data) are represented in the right plot of figure 4, where a good agreement between them can be observed. □

5. CONCLUDING REMARKS

In this paper, new subspace algorithms for the simultaneous identification of the linear and nonlinear parts of Hammerstein and Wiener models have been presented. The algorithms consist of two basic steps. The first one is a standard (linear) subspace algorithm, while the second one is a 2-norm minimization problem which is solved via an SVD.

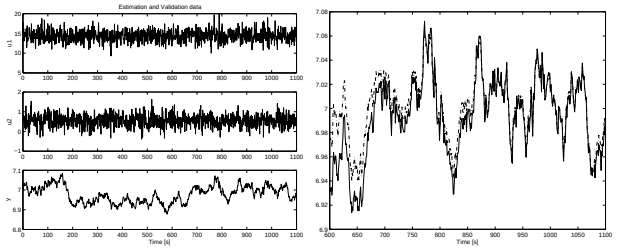

Fig. 4. Left Plot: Estimation (first 600 points) and Validation (remaining 500 points) Input-Output Data. Right Plot: True (solid line) and Estimated (dashdotted line) Output (Validation Data).

6. REFERENCES

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

Billings, S. (1980). Identification of nonlinear systems - A survey. *Proc. of IEEE, Part D* **127**, 272.

Billings, S. and S. Fakhouri (1982). Identification of systems containing linear dynamic and static nonlinear elements. *Automatica* **18**(1), 15–26.

Boutayeb, M. and M. Darouach (1995). Recursive identification method for MISO Wiener-Hammerstein model. *IEEE Trans. on Autom. Control* **40**(2), 287–291.

Chen, H. and J. Maciejowski (2000). Subspace identification method for combined deterministic-stochastic bilinear systems. In: *Proc. IFAC SYSID 2000*. Santa Barbara.

Eskinat, E., S. H. Johnson and W. L. Luyben (1991). Use of Hammerstein models in identification of nonlinear systems. *AIChE J.* **37**(2), 255–268.

Favoreel, W., B. de Moor and P. Van Overschee (1999). Subspace identification of bilinear systems subject to white inputs. *IEEE Trans. on Autom. Control* **44**(6), 1157–1165.

Fruzzetti, K., A. Palazoglu and K. McDonald (1997). Nonlinear model predictive control using Hammerstein models. *J. Process Control* **7**(1), 31–41.

Golub, G.H. and C.F. Van Loan (1989). *Matrix Computations*. 2nd ed.. The Johns Hopkins University Press. Baltimore.

Greblicki, W. (1994). Nonparametric identification of Wiener systems by orthogonal series. *IEEE Trans. on Autom. Control* **39**(10), 2077–2086.

Henson, M. and D. Seborg (1994). Adaptive nonlinear control of a pH neutralization process. *IEEE Trans. Contr. Sys. Tech.* **2**(3), 169–182.

Kalafatis, A., N. Arifin, L. Wang and W. R. Cluett (1995). A new approach to the identification of pH processes based on the Wiener model. *Chem. Eng. Sci.* **50**(23), 3693–3701.

Korenberg, M.J. (1973). Identification of biological cascades of linear and static nonlinear systems. In: *Proc. 16th. Midwest Symp. Circuit Theory*.

Larimore, W.E. (1990). Canonical variate analysis in identification, filtering, and adapt. control. In: *Proc. 29th IEEE CDC*. Hawaii, pp. 596–604.

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

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

Rangan, S., G. Wolodkin and K. Poolla (1995). New results for Hammerstein system identification. In: *Proc. 34th CDC*. New Orleans. pp. 697–702.

Stapleton, J. C. and S. C. Bass (1985). Adaptive noise cancellation for a class of nonlinear dynamic reference channels. *IEEE Trans. on Circuits and Systems* **CAS-32**, 143–150.

Van Overschee, Peter and Bart de Moor (1994). N4SID: Subspace algorithms for the identification of combined deterministic-stochastic systems. *Automatica* **30**(1), 75–93.

Verhaegen, M. (1994). Identification of the deterministic part of MIMO state space models given in innovations form from input-ouput data. *Automatica* **30**(1), 61–74.

Verhaegen, M. and D. Westwick (1996). Identifying MIMO Hammerstein systems in the context of subspace model identification methods. *Int. J. of Control* **63**(2), 331–349.

Viberg, Mats (1995). Subspace-based methods for the identification of linear time-invariant systems. *Automatica* **31**(12), 1835–1851.

Westwick, D. and M. Verhaegen (1996). Identifying MIMO Wiener systems using subspace model ident. methods. *Signal Processing* **52**, 235–258.

Wigren, T. (1993). Recursive prediction error identification using the nonlinear Wiener model. *Automatica* **29**(4), 1011–1025.

APPENDIX

Proof of Theorem 2.2 Let $\widehat{\Theta}_{BD}$ be a consistent estimate of Θ_{BD} defined in (8). Noting now that

$$\begin{aligned}
&\left\| \begin{bmatrix} \widehat{B} \\ \widehat{D} \end{bmatrix} \widehat{\alpha}^T - \begin{bmatrix} B \\ D \end{bmatrix} \alpha^T \right\|_2^2 = \\
&= \left\| \begin{bmatrix} \widehat{B} \\ \widehat{D} \end{bmatrix} \widehat{\alpha}^T - \widehat{\Theta}_{BD} + \widehat{\Theta}_{BD} - \begin{bmatrix} B \\ D \end{bmatrix} \alpha^T \right\|_2^2 \\
&\leq \left\| \begin{bmatrix} \widehat{B} \\ \widehat{D} \end{bmatrix} \widehat{\alpha}^T - \widehat{\Theta}_{BD} \right\|_2^2 + \left\| \widehat{\Theta}_{BD} - \begin{bmatrix} B \\ D \end{bmatrix} \alpha^T \right\|_2^2 \\
&= \sigma_{p+1}^2 + \left\| \widehat{\Theta}_{BD} - \begin{bmatrix} B \\ D \end{bmatrix} \alpha^T \right\|_2^2, \qquad (28)
\end{aligned}$$

and considering that Θ_{BD} is a rank-p matrix, then

$$\left\| \begin{bmatrix} \widehat{B} \\ \widehat{D} \end{bmatrix} \widehat{\alpha}^T - \begin{bmatrix} B \\ D \end{bmatrix} \alpha^T \right\|_2^2 \xrightarrow{\text{a.s.}} 0, \qquad (29)$$

as $N \to \infty$. Now, from the uniqueness of the decomposition $\begin{bmatrix} B \\ D \end{bmatrix} \alpha^T$, it can be concluded that $\widehat{B} \xrightarrow{\text{a.s.}} B$, $\widehat{D} \xrightarrow{\text{a.s.}} D$, and $\widehat{\alpha} \xrightarrow{\text{a.s.}} \alpha$, what ends the proof. ■