

Closed loop identification of MIMO Hammerstein models using LS-SVM^{*}

J.W. van Wingerden^{*} M. Verhaegen^{*}

^{*} Delft University of Technology, Delft, 2628 CD The Netherlands (Tel:
+3115 27 81720; e-mail: J.W.vanWingerden@tudelft.nl,
m.verhaegen@moesp.org).

Abstract: In this paper we present an algorithm to identify MIMO Hammerstein systems under open and closed-loop conditions. To do so, we formulate the optimized predictor based subspace identification algorithm in the dual space. In this dual space we utilize ideas from support vector machines to estimate the state sequence. With the state sequence known, we use the same machinery to estimate the system matrices and the static nonlinearity. The effectiveness of the approach is illustrated with a closed-loop simulation example.

1. INTRODUCTION

The identification of Linear Time-Invariant (LTI) and Hammerstein models is a well-studied subject and has attracted considerable attention in the past few years from both fundamental and application oriented research groups. From an identification perspective, early work in the input-output setting can be found in Narendra and Gallman [1966]. However, to deal with Multiple-Input Multiple-Output (MIMO) systems and to exploit the numerical properties of subspace techniques (which are solely depending on well-established techniques from linear algebra) the focus of this paper is on subspace-based Hammerstein system identification (see Verhaegen and Westwick [1996], Goethals et al. [2005a,b]).

For Hammerstein systems, researchers have only considered the open-loop situation, while from a practical point of view it is necessary to look at closed loop system identification. Since, system identification is often used to identify a part of a complex model, either because the complete model is too complex or parts of the model are well-understood and it is not necessary to identify these. In Figure 1 a sketch of a submodel is given where the user is interested in the dynamics between u and y . In such cases, typically the output of the unknown (sub)model is correlated with the input of the other (sub)model(s). For this situation it is well known that for the LTI case the projector type of subspace algorithms (e.g. MOESP [Verhaegen and Dewilde, 1992] and N4SID [Van Overschee and De Moor, 1996]) give biased estimates if the identification data is generated under closed-loop conditions. The main reason for the bias is the constraint that the noise and the input should be uncorrelated. This assumption is clearly violated if there is a feedback loop present or the models are bilateral coupled (as clearly explained by Ljung and McKelvey [1996]). Predictor-based subspace identification (e.g. PBSID [Chiuso and Picci, 2005] and SSARX [Jansson, 2005]) methods do not suffer from this drawback. These methods use high order ARX models

^{*} This research is supported by the Technology Foundation STW, applied science division of NWO and the technology program of the Ministry of Economic Affairs.

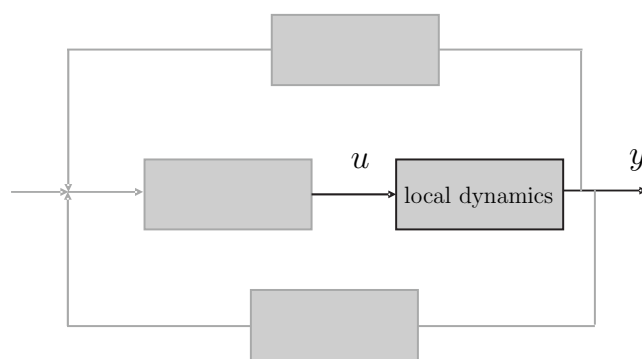


Fig. 1. Schematic representation of a coupled dynamic block.

to remove the correlation between the input and noise sequence.

In this paper we will present a framework to identify Hammerstein models using a predictor based subspace identification method entitled: Predictor Based Subspace Identification (PBSID_{opt}) (see Chiuso [2007a,b]). For this method we derive the dual problem which enables us to use the Least Squares Support Vector Machine (LS-SVM) framework (see Suykens et al. [2002]). In previous work this machinery was already used to identify Hammerstein models (Goethals et al. [2005a,b]). However, in this paper we give a different derivation of the algorithm and we develop the LS-SVM method for a closed-loop identification scheme. Furthermore, the derivation follows the derivation used for the identification of Linear Parameter-Varying (LPV) systems (although, different kernels are used [van Wingerden, 2008, van Wingerden and Verhaegen, 2009]).

The outline of this paper is as follows; we start in Section 2 with the problem formulation and assumptions. In Section 3 the basic idea behind the PBSID_{opt} identification scheme is presented. In Section 4 the problem is translated to the support vector machine framework. In Section 5, we show how we can obtain the system matrices and the static nonlinearity. In Section 6 a brief summary of the algorithm is given. In Section 7 a simulation example is presented.

We end this paper with our conclusions and state a number of directions for future research.

2. PROBLEM FORMULATION

In this section we present the problem formulation and the assumptions we make.

2.1 Problem formulation

For the derivation of the algorithm we consider the following Hammerstein system:

$$x_{k+1} = Ax_k + Bf(u_k) + Ke_k, \quad (1)$$

$$y_k = Cx_k + Df(u_k) + e_k, \quad (2)$$

where $x_k \in \mathbb{R}^n$, $u_k \in \mathbb{R}^r$, $y_k \in \mathbb{R}^\ell$, are the state, input and output vectors. $e_k \in \mathbb{R}^\ell$ denotes the zero mean white innovation process. The matrices $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times r}$, $C \in \mathbb{R}^{\ell \times n}$, $D \in \mathbb{R}^{\ell \times r}$, $K \in \mathbb{R}^{n \times \ell}$ are the local system, input, output, direct feed through, and the observer matrices and finally $f(\cdot)$ is a static smooth nonlinear function with as additional constraint that $f(0) = 0$. We can rewrite (1)-(2) in the predictor form as:

$$x_{k+1} = \tilde{A}x_k + \tilde{B}f(u_k) + Ky_k, \quad (3)$$

$$y_k = Cx_k + Df(u_k) + e_k, \quad (4)$$

with:

$$\tilde{A} = A - KC, \quad \tilde{B} = B - KD.$$

It is well-known that an invertible linear transformation of the state does not change the input-output behavior of a state-space system. Therefore, we can only determine the system matrices up to a similarity transformation $T \in \mathbb{R}^{n \times n}$: $T^{-1}AT$, $T^{-1}B$, $T^{-1}K$, CT , and D .

The identification problem can now be formulated as: given the input sequence u_k and the output sequence y_k over a time $k = \{0, \dots, N-1\}$; find all, if they exist, the system matrices A , B , K , C , and D up to a global similarity transformation and the nonlinear function $f(\cdot)$.

2.2 Assumptions and notation

Similar as in Jansson [2005], Chiuso [2007a] we define a past window denoted by p . This window is used to define the following stacked vector:

$$f(\bar{U}_P) = \begin{bmatrix} f(U_0) \\ f(U_1) \\ \vdots \\ f(U_{p-1}) \end{bmatrix}, \quad \bar{Y}_P = \begin{bmatrix} Y_0 \\ Y_1 \\ \vdots \\ Y_{p-1} \end{bmatrix},$$

and:

$$f(\bar{U}_F) = f(U_p), \quad \bar{Y}_F = Y_p,$$

with:

$$f(U_i) = [f(u_i), f(u_{i+1}), \dots, f(u_{i+N-p})], \quad (5)$$

$$Y_i = [y_i, y_{i+1}, \dots, y_{i+N-p}]. \quad (6)$$

We assume that the state sequence:

$$X_p = [x_p, \dots, x_{N-1}],$$

has full row rank and the extended observability matrix given by:

$$\Gamma = \begin{bmatrix} C \\ C\tilde{A} \\ \vdots \\ C\tilde{A}^{f-1} \end{bmatrix}, \quad (7)$$

has full column rank and f is referred to as the future window. We also define the extended controllability matrix:

$$\mathcal{K} = [\mathcal{K}_u, \mathcal{K}_y],$$

with:

$$\mathcal{K}_u = [\tilde{A}^{p-1}\tilde{B}, \tilde{A}^{p-2}\tilde{B}, \dots, \tilde{B}],$$

$$\mathcal{K}_y = [\tilde{A}^{p-1}K, \tilde{A}^{p-2}K, \dots, K].$$

In the next section we revise the idea of predictor based subspace identification and after that we introduce support vector machinery to estimate the system matrices for Hammerstein models.

3. PREDICTOR BASED SUBSPACE IDENTIFICATION

It is well known that the projector type of subspace algorithms (*e.g.* MOESP [Verhaegen and Dewilde, 1992] and N4SID [Van Overschee and De Moor, 1996]) give biased estimates if the identification data is generated under closed-loop conditions. The main reason for the bias is the constraint that for the projector type of algorithms the noise, e_k , and the input, u_k , should be uncorrelated. This assumption is clearly violated if there is a feedback loop present (as clearly explained by Ljung and McKelvey [1996]). Predictor-based subspace identification (*e.g.* PBSID [Chiuso and Picci, 2005] and SSARX [Jansson, 2005]) methods do not suffer from this drawback. These methods use high order ARX models to remove the correlation between the input and noise sequence. In this section we introduce the PBSID_{opt} algorithm and for now we assume that the nonlinearity is known. In the next section we formulate the dual problem and we explicitly estimate the nonlinearity.

3.1 Predictors

The first objective of the predictor based algorithms is to reconstruct the state sequence up to a similarity transformation. The state X_p is given by:

$$X_p = \tilde{A}^p X_0 + [\mathcal{K}_u, \mathcal{K}_y] \begin{bmatrix} f(\bar{U}_P) \\ \bar{Y}_P \end{bmatrix}.$$

The key approximation in this algorithm is that we assume that $\tilde{A}^j \approx 0$ for all $j \geq p$. It can be shown that if the system in (3)-(4) is uniformly exponentially stable, the approximation error can be made arbitrarily small by making p large [Knudsen, 2001, Chiuso and Picci, 2005, Chiuso, 2007a]. With this assumption the state X_p is approximately given by:

$$X_p \approx [\mathcal{K}_u, \mathcal{K}_y] \begin{bmatrix} f(\bar{U}_P) \\ \bar{Y}_P \end{bmatrix}. \quad (8)$$

In a number of other LTI subspace methods it is well known to make this step (*e.g.* N4SID, SSARX, PBSID,

PBSID_{opt}, and CVA). The input-output behavior is now approximately given by:

$$Y_F \approx C [\mathcal{K}_u, \mathcal{K}_y] \begin{bmatrix} f(\bar{U}_P) \\ \bar{Y}_P \end{bmatrix} + Df(U_F) + e_k.$$

If the matrix $\begin{bmatrix} f(\bar{U}_P)^T, \bar{Y}_P^T, f(\bar{U}_F)^T \end{bmatrix}^T$ has full row rank and we know the nonlinear function $f(\cdot)$, which we do not, the matrices $C\mathcal{K}$ and D can be estimated by solving the following linear regression problem:

$$\min_{C\mathcal{K}, D} \|Y_F - C\mathcal{K} \begin{bmatrix} f(\bar{U}_P) \\ \bar{Y}_P \end{bmatrix} - Df(\bar{U}_F)\|_F^2. \quad (9)$$

where $\|\cdot\|_F$ represents the Frobenius norm [Golub and Loan, 1996]. For finite p this linear problem will be biased due the approximation made in (8). In the LTI literature a number of papers appeared that studied the effect of the window size and although they proved the asymptotic properties of the algorithms (if $p \rightarrow \infty$ the bias disappears) it is hard to quantify the effect for finite p [Knudsen, 2001, Chiuso and Picci, 2005, Chiuso, 2007a].

3.2 Extended observability times extended controllability matrix

The product $\mathcal{K} \begin{bmatrix} f(U_P) \\ Y_P \end{bmatrix}$ that represents by definition the state sequence, X_p , can not be estimated directly. In the predictor based identification algorithms $C\mathcal{K}$ is used to construct the extended observability matrix times the extended controllability matrix. This matrix is given by (we assume $f = p$):

$$\Gamma\mathcal{K} = \begin{bmatrix} C\mathcal{K}_u & C\mathcal{K}_y \\ C\tilde{A}\mathcal{K}_u & C\tilde{A}\mathcal{K}_y \\ \vdots & \vdots \\ C\tilde{A}^{f-1}\mathcal{K}_u & C\tilde{A}^{f-1}\mathcal{K}_y \end{bmatrix},$$

$$= [\Gamma\mathcal{K}_u \quad \Gamma\mathcal{K}_y].$$

If we look into more detail at for instance the $\Gamma\mathcal{K}_u$ term we have (we assume $f = p$):

$$\Gamma\mathcal{K}_u = \begin{bmatrix} C\tilde{A}^{p-1}\tilde{B} & C\tilde{A}^{p-2}\tilde{B} & \dots & C\tilde{B} \\ C\tilde{A}^p\tilde{B} & C\tilde{A}^{p-1}\tilde{B} & \dots & C\tilde{A}\tilde{B} \\ \vdots & \vdots & \ddots & \vdots \\ C\tilde{A}^{p+f-2}\tilde{B} & C\tilde{A}^{p+f-3}\tilde{B} & \dots & C\tilde{A}^{p-1}\tilde{B} \end{bmatrix}.$$

This particular matrix is constructed in the PBSID algorithm. The following upper block diagonal matrix is used in the PBSID_{opt} algorithm (we assume $f = p$):

$$\Gamma\mathcal{K}_u \approx \begin{bmatrix} C\tilde{A}^{p-1}\tilde{B} & C\tilde{A}^{p-2}\tilde{B} & \dots & C\tilde{B} \\ 0 & C\tilde{A}^{p-1}\tilde{B} & \dots & C\tilde{A}\tilde{B} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & & & C\tilde{A}^{p-1}\tilde{B} \end{bmatrix}.$$

The zeros appear in this equation based on the approximation that $\tilde{A}^j \approx 0$ for all $j \geq p^1$. Observe that from the linear regression problem formulated in (9) we can construct this matrix. A similar thing can be done for the matrix $\Gamma\mathcal{K}_y$.

¹ Although, in Chiuso [2007b] they showed that these zeros appear in the solution based on the result of an optimization problem.

From the constructed matrix $\Gamma\mathcal{K}$ we can compute:

$$\Gamma\mathcal{K} \begin{bmatrix} f(\bar{U}_P) \\ \bar{Y}_P \end{bmatrix},$$

which equals by definition the extended observability times the state sequence, ΓX_p . By computing a Singular Value Decomposition (SVD) of this estimate we can estimate the state sequence and the order of the system. We will use the following SVD:

$$\widehat{\Gamma\mathcal{K}} \begin{bmatrix} f(\bar{U}_P) \\ \bar{Y}_P \end{bmatrix} = [U \quad U_\perp] \begin{bmatrix} \Sigma_n & 0 \\ 0 & \Sigma \end{bmatrix} \begin{bmatrix} \mathcal{V} \\ \mathcal{V}_\perp \end{bmatrix}, \quad (10)$$

where Σ_n is the diagonal matrix containing the n largest singular values and \mathcal{V} is the corresponding row space. Note that we can find the largest singular values by detecting a gap between the singular values [Verhaegen and Verdult, 2007]. The state is now estimated by:

$$\hat{X}_p = \Sigma_n \mathcal{V}. \quad (11)$$

It is well known that when the state, input, and output are known the system matrices can be estimated [Verhaegen and Verdult, 2007]. The approach presented so far assumed that we know the nonlinearity. In the next section we present the dual problem and show how we can solve the problem without knowing the nonlinear function.

4. SUPPORT VECTOR MACHINES (SVM)

In the previous section we showed that if we have a solution of the problem formulated in (9) we can reconstruct the state sequence. In this section we will use Support Vector Machines (SVM's) to solve the nonlinear problem in (9) and we show how we can construct the extended observability times extended controllability matrix.

4.1 Dual problem

The problem formulated in (9) is nonlinear. By exploiting ideas from the SVM framework we transform this nonlinear problem to the feature space where the problem becomes linear. The dual problem is given by:

$$\min_{\alpha} \|Y_F - \alpha W\|_F^2, \quad (12)$$

with $W = f(U_P)^T f(U_P) + Y_P^T Y_P + f(U_F)^T f(U_F)$ the kernel matrix, which is a square matrix of the size of the number of data points. This linear problem can be ill conditioned so regularization is necessary to obtain a solution:

$$\alpha = Y_F (W + I/\gamma)^{-1}, \quad (13)$$

where γ is the regularization parameter. It is well known that regularized problems result in biased estimates. The estimate of the dual problem, α , can be used to estimate \mathcal{K}_u and \mathcal{K}_y and elements therein. These matrices are given by:

$$C [\tilde{A}^{p-1}\tilde{B}, \tilde{A}^{p-2}\tilde{B}, \dots, \tilde{B}] = \alpha f(U_P)^T, \\ = [\alpha f(U_0)^T, \dots, \alpha f(U_{p-1})^T],$$

and a similar thing can be done for \mathcal{K}_y . It is again important to observe that we are not interested in this matrix but in the extended observability matrix times the state. This matrix is now given by:

$$\Gamma X_p = \begin{bmatrix} \alpha \left(\sum_{i=0}^{p-1} (W_{U_i, U_i} + W_{Y_i, Y_i}^l) \right) \\ \alpha \left(\sum_{i=1}^{p-1} (W_{U_i, U_{i-1}} + W_{Y_i, Y_{i-1}}^l) \right) \\ \vdots \\ \alpha \left(\sum_{i=f-1}^{p-1} (W_{U_i, U_{i-f-1}} + W_{Y_i, Y_{i-f-1}}^l) \right) \end{bmatrix}, \quad (14)$$

with $W_{a,b} = f(a)^T f(b)$ representing a nonlinear kernel while $W_{a,b}^l = a^T b$ represents a linear kernel. Observe that the kernel W is now given by:²

$$W = \sum_{i=0}^{p-1} (W_{U_i, U_i} + W_{Y_i, Y_i}^l) + W_{U_p, U_p}.$$

The kernels depending on y are linear and can be computed in a straightforward manner. The kernel dependent on u are nonlinear and we do not know the nonlinearity. For this situation we are going to replace these kernels by generic kernels like there are: linear, polynomial or gaussian radial basis functions. For example if we take radial basis functions in our kernels we have:

$$W_{U_i, U_j} = \dots \quad (15)$$

$$\begin{bmatrix} q(u_i, u_j) & q(u_i, u_{j+1}) & \dots & q(u_i, u_{j+N-p}) \\ q(u_{i+1}, u_j) & q(u_{i+1}, u_{j+1}) & \dots & q(u_{i+1}, u_{j+N-p}) \\ \vdots & \vdots & \ddots & \vdots \\ q(u_{i+N-p}, u_j) & q(u_{i+N-p}, u_{j+1}) & \dots & q(u_{i+N-p}, u_{j+N-p}) \end{bmatrix}$$

with $q(u_i, u_j) = \exp(-\|u_i - u_j\|_2^2 / \sigma)$. With the construction of the kernels and the solution of the dual problem we can construct the observability matrix times the state in (14). Similar as in the previous section we now can compute an SVD to recover the state sequence up to a similarity transformation.

4.2 Centering the kernels

The estimates so far can have a offset due to the fact that we have component-wise kernels. If we consider for example the element $D^{(1)}U_p^{(1)}$, where superscript (1) indicates the first column of D and the first row of U_p . The estimate of this element is given by $\alpha W_{U_p^{(1)}, U_p^{(1)}}$. If we would like to guarantee that $D^{(1)}f(0) = 0$ we have to subtract the offset term which is given by $\alpha W_{U_p^{(1)}, 0_{N-p, N-p}}$. So, the new estimate of $D^{(1)}U_p^{(1)}$ is now given by $\alpha (W_{U_p^{(1)}, U_p^{(1)}} - W_{U_p^{(1)}, 0_{N-p, N-p}})$. We have to do this for every element separately because we do not know how the offset enters the solution α .³

5. RECOVERY OF THE SYSTEM MATRICES

In linear subspace identification the problem is solved if the state is recovered because the equation in (1)-(2) are then linear. However, in the case of Hammerstein models

² Observe that this kernel is a component-wise kernel.

³ Observe that it is also possible to add this centering property to the definition of the kernels.

this is not true. However, we can use the same machinery as used before, SVM.

First we define the following matrices:

$$\begin{aligned} \hat{X}_{k+1} &= \hat{X}_p(:, 2 : \text{end}), \\ \hat{X}_k &= \hat{X}_p(:, 1 : \text{end} - 1), \\ f(U)_k &= f(\bar{U}_F)(:, 1 : \text{end} - 1), \\ Y_k &= \bar{Y}_F(:, 1 : \text{end} - 1), \\ \tilde{U}_k &= \bar{U}_F(:, 1 : \text{end} - 1). \end{aligned}$$

The primal problem for the system recovery is now given by

$$\min_{A, B, C, D} \left\| \begin{bmatrix} \hat{X}_{k+1} \\ Y_k \end{bmatrix} - \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} \hat{X}_k \\ f(U)_k \end{bmatrix} \right\|_F^2.$$

The dual problem is given by:

$$\begin{bmatrix} \hat{X}_{k+1} \\ Y_k \end{bmatrix} - \beta \left[\hat{X}_k^T \hat{X}_k + W_{\tilde{U}_k, \tilde{U}_k} \right] = 0, \quad (16)$$

and consequently:

$$\begin{aligned} A &= \beta(1 : n, :) \hat{X}_k^T, \\ B f(u_k) &= \beta(1 : n, :) W_{\tilde{U}_k, u_k}, \\ C &= \beta(n + 1 : \text{end}, :) \hat{X}_k^T, \\ D f(u_k) &= \beta(n + 1 : \text{end}, :) W_{\tilde{U}_k, u_k}. \end{aligned}$$

The dual problem should again be regularized and consequently we have to do a similar trick as in the previous paragraph to remove the offset.

Until now we estimate the product between the B matrix and the nonlinearity (and also the product between the D matrix and the nonlinear function). It should be clear that we can find the nonlinearity up to a square invertible matrix Q . Since:

$$BQQ^{-1}f(\cdot).$$

We can compute the nonlinearity up to a similarity transformation by using an SVD. Using this nonlinearity we can also compute the B and D matrix up to this square invertible matrix Q .

6. SUMMARY OF THE ALGORITHM

In this section we give a short summary of the proposed MIMO Hammerstein identification scheme.

Algorithm 1. (Hammerstein-PBSID_{opt} (kernel)). The algorithm can be summarized as follows:

- (1) Create the matrices W and W_{U_i, U_j} using (15),
- (2) Solve the linear problem given in (12). If desired regularized,
- (3) Construct ΓX_p using (14),
- (4) Compute the state sequence using (10) and (11),
- (5) With the estimated state, use the SVM machinery to obtain the system matrices and the static nonlinearity (see Section 5).

7. EXAMPLE

We have tested the proposed identification algorithm on a fourth order MIMO closed-loop model with $r = 2$,

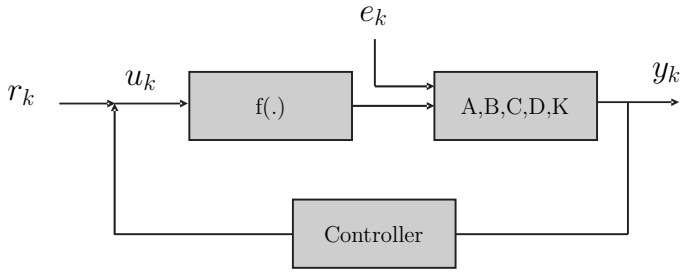


Fig. 2. Schematic representation of the simulation scheme. and $\ell = 2$. The collected data u_k, y_k , are used for the identification algorithm. We will use the following model (as illustrated in Figure 2):

$$\begin{aligned} x_{k+1} &= Ax_k + Bf(u_k) + Ke_k, \\ y_k &= Cx_k + e_k, \\ u_k &= F_b y_k + r_k, \end{aligned}$$

where F_b is the feedback controller and r_k the reference signal. The system matrices are given by:

$$\begin{aligned} A &= \begin{bmatrix} 0.67 & 0.67 & 0 & 0 \\ -0.67 & 0.67 & 0 & 0 \\ 0 & 0 & -0.67 & -0.67 \\ 0 & 0 & 0.67 & -0.67 \end{bmatrix}, \\ B &= \begin{bmatrix} 0.6598 & -0.5256 \\ 1.9698 & 0.4845 \\ 4.3171 & -0.4879 \\ -2.6436 & -0.3416 \end{bmatrix}, \\ C &= \begin{bmatrix} 0.3749 & 0.0751 & -0.5225 & 0.5830 \\ -0.8977 & 0.7543 & 0.1159 & 0.0982 \end{bmatrix}, \\ K &= \begin{bmatrix} -0.6968 & -0.1474 \\ 0.1722 & 0.5646 \\ 0.6484 & -0.4660 \\ -0.9400 & 0.1032 \end{bmatrix}, \\ F_b &= \begin{bmatrix} 0.5 & 0 \\ 0 & 0.5 \end{bmatrix}, \\ f(u_k) &= \begin{bmatrix} \text{sinc}(u_k^{(1)}) (u_k^{(1)})^2 \\ \sin(u_k^{(2)}) \end{bmatrix}, \end{aligned}$$

and $\text{cov}(e_k) = 0.04 \cdot I_\ell$. As reference signal we take a zero mean white noise signal with $\text{var}(r_k) = I_r$. For the identification experiment we used $N = 1000$ and $p = f = 10$. The collected data (u_k and y_k) is used to identify a Hammerstein model. The performance of the identified system is evaluated by looking at the value of the Variance-Accounted-For (VAF) on a data set different from the one used for identification. The VAF value is defined as:

$$\text{VAF}(y_k, \hat{y}_k) = \max \left\{ 1 - \frac{\text{var}(y_k - \hat{y}_k)}{\text{var}(y_k)}, 0 \right\} * 100\%$$

where \hat{y}_k denotes the output signal obtained by simulating the identified Hammerstein system, y_k is the output signal of the true Hammerstein system, and $\text{var}(\cdot)$ denotes the variance of a quasi-stationary signal. To investigate the sensitivity of the identification algorithm with respect to noise, a Monte-Carlo simulation with 100 runs was carried out. For each of the 100 simulations a different realization of the input u_k is used.

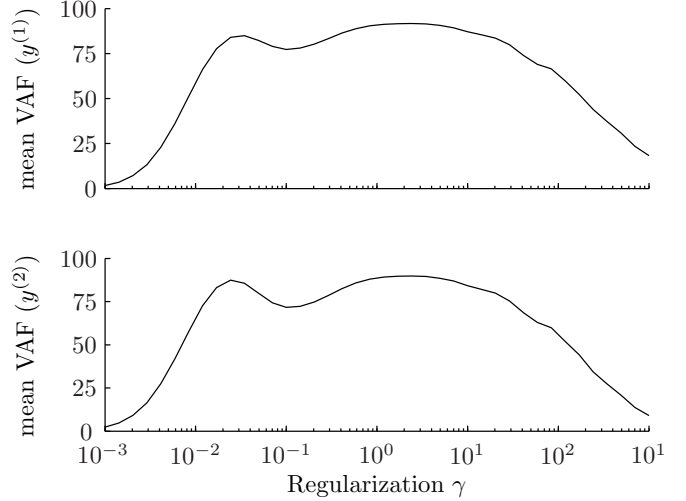


Fig. 3. The VAF for the validation data as function of the regularization parameter γ .

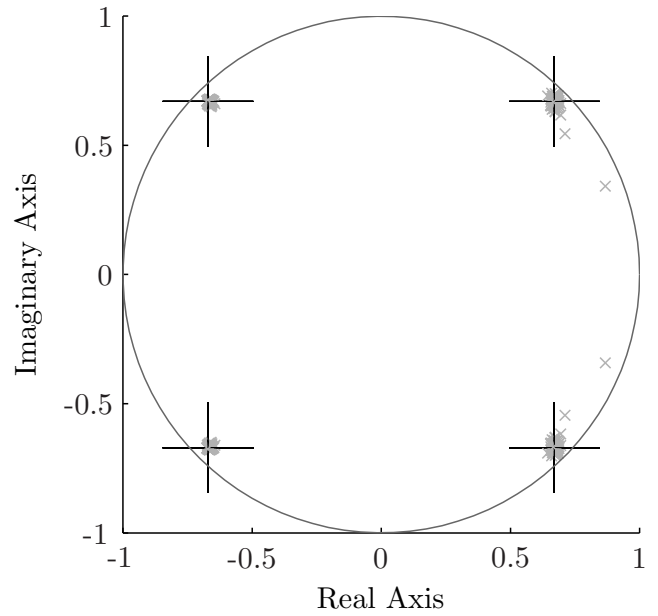


Fig. 4. True and estimated eigenvalues of A in one plot for 100 experiments and $\gamma = 2.4$.

The choice of γ was made by plotting the average validation VAF for 100 simulations against a logarithmic grid of different values of the parameter and choosing the value corresponding to a peak. From Fig. 3 we can conclude that $\gamma = 2.4$ is a motivated choice. With the regularization parameter we show in Figure 4 the poles of A for the 100 Monte Carlo simulation and we see that we have an almost unbiased estimate.

In Fig. 5 we show the nonlinearity by plotting the estimated product $CB(u_k)$ as function of u_k . We clearly see that we can estimate the nonlinearity quite accurately for different realizations of the noise.

8. CONCLUSIONS

In this paper we presented a novel algorithm to identify MIMO Hammerstein systems under open and closed-

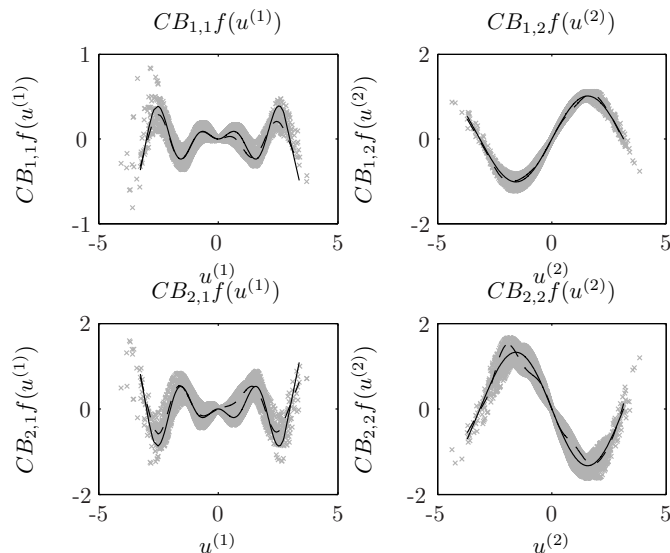


Fig. 5. Presentation of the product $CBf(u_k)$ based on 20 Monte Carlo simulations. The solid line represents the real system, the dashed line the estimate with the highest VAF and the dots are function evaluation of 20 Monte Carlo simulations ($\gamma = 2.4$).

loop conditions. To do so, we formulated the optimized Predictor Based Subspace Identification algorithm in the dual space. In this dual space we utilized the ideas from support vector machines to estimate the state sequence. With the state sequence known, we use the same support vector machine tools to estimate the system matrices and the static nonlinearity. The effectiveness of the approach was illustrated with a closed-loop example.

9. OUTLOOK

In the near future we are going to apply this technique to the identification of wind turbine dynamics, which have a strong Hammerstein nature in both the below and the above rated operating area (for more information see Manwell et al. [2002], Bianchi et al. [2007]). Furthermore, if we consider a number of LPV identification problems (see for instance Bianchi et al. [2005]) we can change the virtual system boundaries of the problem and consequently the problem can be reformulated as a closed-loop Hammerstein identification problem.

The algorithm proposed in this paper can be made recursive using similar techniques as presented in Houtzager et al. [2009], which enables us to do closed loop recursive Hammerstein identification, which is of interest for model based monitoring and fault detection.

REFERENCES

F.D. Bianchi, R.J. Mantz, and C.F. Christiansen. Gain scheduling control of variable-speed wind energy conversion systems using quasi-LPV models. *Control Engineering Practice*, 13(2):247–255, 2005.

F.D. Bianchi, H. De Battista, and R.J. Mantz. *Wind Turbine Control Systems; Principles, modelling and gain scheduling design*. Springer-Verlag London, 2007.

A. Chiuso. The role of vector autoregressive modeling in predictor based subspace identification. *Automatica*, 43(6):1034–1048, 2007a.

A. Chiuso. On the relation between CCA and predictor based subspace identification. *IEEE Transactions on Automatic Control*, 52(10):1795–1812, 2007b.

A. Chiuso and G. Picci. Consistency analysis of certain closed-loop subspace identification methods. *Automatica*, 41(3):377–391, 2005.

I. Goethals, K. Pelckmans, J. A. K. Suykens, and B. De Moor. Identification of MIMO Hammerstein models using least squares support vector machines. *Automatica*, 41(7):1263–1272, 2005a.

I. Goethals, K. Pelckmans, J. A. K. Suykens, and B. De Moor. Subspace identification of Hammerstein systems using least squares support vector machines. *IEEE Transactions on Automatic Control*, 50(10):1509–1519, 2005b.

G. H. Golub and C. F. van Loan. *Matrix computations*. The John Hopkins University Press, 1996.

I. Houtzager, J.W. van Wingerden, and M. Verhaegen. Fast-array recursive closed-loop system identification. In *Proceedings of the 15th IFAC Symposium on System Identification (SYSID)*, Saint-Malo, France, 2009.

M. Jansson. A new subspace identification method for open and closed loop data. In *Proceedings of the 16th IFAC World Congress*, Prague, Czech Republic, 2005.

T. Knudsen. Consistency analysis of subspace identification methods based on a linear regression approach. *Automatica*, 37(1):81–89, 2001.

L. Ljung and T. McKelvey. Subspace identification from closed loop data. *Signal Processing*, 52(2):209–215, 1996.

J.F. Manwell, J.G. McGowan, and A.L. Rogers. *Wind energy explained; Theory, design and application*. John Wiley and Sons, Ltd, Chichester, 2002.

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

J. A. K. Suykens, T. V. van Gestel, J. De Brabanter, B. De Moor, and J. Vandewalle. *Least squares support vector machines*. World Scientific, 2002.

P. Van Overschee and B. De Moor. *Subspace Identification for linear systems*. Kluwer Academic Publishers, 1996.

J. W. van Wingerden. *Control of Wind Turbines with ‘Smart’ Rotors: Proof of Concept & LPV Subspace Identification*. PhD thesis, Delft University of Technology, 2008.

J. W. van Wingerden and M. Verhaegen. Subspace identification of Bilinear and LPV systems for open and closed loop data. *Automatica*, 45(2):372–381, 2009.

M. Verhaegen and P. Dewilde. Subspace model identification part 1: The output-error state-space model identification class of algorithms. *International Journal of Control*, 56(5):1187–1210, 1992.

M. Verhaegen and V. Verdult. *Filtering and System Identification: An introduction*. Cambridge University Press, 2007.

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