IDENTIFICATION OF APPROXIMATIVE NONLINEAR STATE-SPACE MODELS BY SUBSPACE METHODS

Andreas Schrempf [**] **Vincent Verdult** [***]

[*] *Department of Design and Control of Mechatronical Systems, Johannes Kepler University Linz, 4040 Linz, Austria*
[**] *University of Applied Sciences - Upper Austria, Medical Technology, Krankenhausstrasse 26, 4020 Linz, Austria*
[***] *Delft Center for Systems and Control, Delft University of Technology, Mekelweg 2, 2628 CD Delft, The Netherlands*

Abstract: A subspace identification algorithm for state-affine state-space systems which allows to approximate nonlinear systems arbitrarily well is derived. The proposed algorithm depends on an approximation step where a detailed approximation error analysis is provided. A special case is presented in which this approximation error vanishes. To tackle higher-order systems and ill-posed problems a regularized kernel method is proposed. The algorithm is evaluated by a simulation study.

Keywords: Subspace identification, state-affine systems

1. INTRODUCTION

Consider the generic nonlinear, discrete-time system

$$\xi(j+1) = F\left(\xi(j), u(j)\right), \tag{1}$$

$$y(j) = H\left(\xi(j)\right), \tag{2}$$

which describes the dynamic behavior of the system of interest. The states are denoted by $\xi \in \mathbb{R}^n$, the inputs by $u \in \mathbb{R}^m$ and the outputs by $y \in \mathbb{R}^p$. Many practical applications require an accurate description of the dynamical behavior of the system which can be obtained either by modeling using first principles, by identification or a combination of both approaches.

If the functions F and H of the system (1)–(2) are linear, we can use several well-established linear system identification techniques to obtain a mathematical description of the system [Ljung, 1999]. Among these methods, linear subspace identification methods [Van Overschee and De Moor, 1996, Verhaegen, 1994] are quite popular, because they are suitable for multivariable systems and can be used to generate an initial starting point for the classical iterative prediction-error methods. This combination of subspace and prediction-error methods is a powerful tool for determining an linear system from input and output measurements. Unfortunately, in many applications linear systems do not provide an accurate description of the underlying real system. Therefore, identification methods for other descriptions, like the nonlinear system (1)–(2), are needed.

In this paper we describe a subspace identification method for state-affine nonlinear systems. State-affine systems can be used to approximate general nonlinear systems up to an arbitrary accuracy. The subspace identification technique that we present builds upon previous results on subspace identification for bilinear [Favoreel et al., 1999, Chen and Maciejowski, 2000, Verdult and Verhaegen, 2003] and LPV systems [Verdult and Verhaegen, 2002], but also incorporates structural information about the system [Schrempf and Del Re, 2001] which allows to derive an improved identification algorithm [Schrempf, 2004].

The paper is organized as follows: In Section 2 the general approximation capabilities of state-affine systems are discussed which motivates the choice of the model structure. A subspace identification algorithm is proposed in Section 3 where a detailed approximation error analysis is provided. Finally the

proposed identification algorithm is evaluated by a simulation study in Section 4.

2. STATE-AFFINE SYSTEMS

A state-affine state-space system can be written as follows

$$x(j+1) = Ax(j) + Nf(u(j)) \otimes x(j) + Bu(j) + Mg(u(j)), \quad (3)$$

$$y(j) = Cx(j), \quad (4)$$

with the states $x \in \mathbb{R}^n$, the inputs $u \in \mathbb{R}^m$ and the outputs $y \in \mathbb{R}^p$ where $f : \mathbb{R}^m \mapsto \mathbb{R}^q$ and $g : \mathbb{R}^m \mapsto \mathbb{R}^r$ are vectors of monomials of degree up to ρ. By explicitly writing the linear terms we will stress out that the state-affine system has a nontrivial linearization around the origin, which is a technical requirement for the proposed identification algorithm.

Fliess [1982] shows that on a finite time interval and with bounded inputs, any input-output system which depends continuously on the inputs can be arbitrarily well approximated by a state-affine state-space system (3)–(4). The restriction to finite intervals of time originates from the compactness requirements stated by the Stone-Weierstrass approximation theorem.

The requirement for a finite interval of time can be replaced by a fading memory assumption. The concept of fading-memory basically states that the output of the system is asymptotically independent of the initial state. Boyd and Chua [1985] show that every nonlinear system represented by an input-output map with fading memory can be arbitrarily well approximated by a bounded polynomial input-output map. Sontag [1979] shows that a bounded polynomial input-output map is realizable by an observable state-affine system of the form (3)–(4). This yields the following

Corollary 1. Every input-output system, with bounded inputs and fading memory can be approximated arbitrarily well by an observable state-affine state-space system of form (3)–(4).

Fading memory of the input-output system and exponential stability of the corresponding state-space system are related. Zang and Iglesias [2003] state that if an autonomous affine state-space system is exponentially stable and fulfills some additional Lipschitz like conditions, then the system has fading memory. For more general state-space systems like (1)–(2) this connection is not understood completely.

Due to its general approximation capabilities state-affine state-space systems are proposed for approximation of the generic nonlinear system (1)–(2) and will be considered in the sequel.

3. IDENTIFICATION

The aim is to determine a state-affine approximation of the nonlinear system (1)–(2) using a finite number N_s measurements of its inputs and outputs. We need to determine the system matrices A, B, C, M, and N of the state-affine system (3)–(4). A subspace identification method is presented which is a direct approach that obtains the state-space system directly [Gatt and Kalouptsidis, 2002, Verdult and Verhaegen, 2002]. Alternatively, an indirect approach could be used where the state-affine state-space system is obtained via realization of an identified difference equation [Diaz and Desrochers, 1988, Constanza et al., 1983].

A purely deterministic setting will be considered which allows to analyze the proposed subspace method with respect to deterministic approximation errors. Due to these errors, the estimates of the system matrices will be biased. These biased estimates can be improved using a final optimization step: the system matrices obtained by subspace identification can be used as an initial starting point for a nonconvex prediction-error type of identification method as discussed by Lee and Poolla [1996], Verdult [2002] and Schrempf and Del Re [2001].

3.1 Assumptions

We require the state-affine system (3)–(4) to be
(A1) strongly locally observable, and
(A2) uniform exponentially stable.
Corollary 1 states that the state-affine realization is observable which implies local observability [Albertini and D'Alessandro, 1996] and thus (A1) is fulfilled. Together with (A2) a rigorous approximation error analysis is possible, where according to the close connection of fading memory to exponential stability (A2) is assumed not to be too restrictive.

3.2 Data equations

Subspace identification relies on relations between structured matrices constructed from the outputs, inputs and states. We define:

$$U_j := [u(j), u(j+1), \ldots, u(j+c-1)] \in \mathbb{R}^{m \times c}$$
$$Y_j := [y(j), y(j+1), \ldots, y(j+c-1)] \in \mathbb{R}^{p \times c}$$
$$X_j := [x(j), x(j+1), \ldots, x(j+c-1)] \in \mathbb{R}^{n \times c}$$
$$\mathcal{F}(U_j) := [f(u(j)), \ldots, f(u(j+c-1))] \in \mathbb{R}^{q \times c}$$
$$\mathcal{G}(U_j) := [g(u(j)), \ldots, g(u(j+c-1))] \in \mathbb{R}^{r \times c}$$

where c denotes the number of columns of the block matrices. We also recursively define:

$$Y_{j|j} := Y_j,$$
$$Y_{k+j|j} := \begin{bmatrix} Y_{k+j} \\ Y_{k+j-1|j} \end{bmatrix} \in \mathbb{R}^{(k+1)p \times c},$$
$$U_{j|j} := \begin{bmatrix} U_j \\ \mathcal{G}(U_j) \end{bmatrix} \in \mathbb{R}^{(m+r) \times c},$$
$$U_{k+j|j} := \begin{bmatrix} U_{k+j} \\ \mathcal{G}(U_{k+j}) \\ U_{k+j-1|j} \end{bmatrix} \in \mathbb{R}^{(k+1)(m+r) \times c},$$

$$\widetilde{U}_{j|j} := \begin{bmatrix} U_j \\ \mathcal{G}(U_j) \end{bmatrix} \in \mathbb{R}^{(m+r)\times c},$$

$$\widetilde{U}_{k+j|j} := \begin{bmatrix} U_{k+j} \\ \mathcal{G}(U_{k+j}) \\ \widetilde{U}_{k+j-1|j} \\ \mathcal{F}(U_{k+j}) \odot \widetilde{U}_{k+j-1|j} \end{bmatrix} \in \mathbb{R}^{q_k(m+r)/q\times c},$$

$$X_{j|j} := [\mathcal{F}(U_j) \odot X_j] \in \mathbb{R}^{qn\times c},$$
$$X_{k+j|j} := \begin{bmatrix} \mathcal{F}(U_{k+j}) \odot X_{k+j} \\ X_{k+j-1|j} \end{bmatrix} \in \mathbb{R}^{(k+1)qn\times c},$$
$$\widetilde{F}_{j|j} := \mathcal{F}(U_j) \in \mathbb{R}^{q\times c},$$
$$\widetilde{F}_{k+j|j} := \begin{bmatrix} \widetilde{F}_{k+j-1|j} \\ \mathcal{F}_1(U_{k+j}) \\ \mathcal{F}_1(U_{k+j}) \odot \widetilde{F}_{k+j-1|j} \\ \vdots \\ \mathcal{F}_q(U_{k+j}) \\ \mathcal{F}_q(U_{k+j}) \odot \widetilde{F}_{k+j-1|j} \end{bmatrix} \in \mathbb{R}^{q_k\times c}.$$

where $q_k = (q+1)^{k+1} - 1$. With the previous definitions the data equations given in the following lemmas can be obtained. These relations can be proven using a simple induction argument as given by Schrempf [2004] and Verdult [2002].

Lemma 1. The "future" state sequence of the system (3)–(4) is related for $k \geq 1$ to the block input, block output matrix and the state sequence as follows:

$$X_{k+j} = A^k X_j + \widetilde{A}_k \widetilde{F}_{k+j-1|j} \odot X_j + B_k \widetilde{U}_{k+j-1|j} \tag{5}$$

where $A_k := [A^k, \widetilde{A}_k]$ and

$$\begin{aligned} A_1 &:= [A, N_1, N_2, \ldots, N_q], \\ A_k &:= [AA_{k-1}, N_1 A_{k-1}, \ldots, N_q A_{k-1}], \\ B_1 &:= [B, M], \\ B_k &:= [B_1, AB_{k-1}, N_1 B_{k-1}, \ldots, N_q B_{k-1}]. \end{aligned}$$

Lemma 2. The block output matrix of the system (3)–(4) depends for $k \geq 1$ on the state sequence, the extended state sequence and the block input matrix as follows:

$$Y_{k+j|j} = \Gamma_k X_j + G_k X_{k+j-1|j} + H_k U_{k+j-1|j} \tag{6}$$

with

$$G_1 := \begin{bmatrix} CN \\ 0 \end{bmatrix},\ G_k := \begin{bmatrix} CN & CA\bar{A}_{k-1} \\ 0 & G_{k-1} \end{bmatrix},$$
$$\bar{A}_1 := N,\ \bar{A}_k := \begin{bmatrix} N & A\bar{A}_{k-1} \end{bmatrix},$$
$$H_1 := \begin{bmatrix} CB & CM \\ 0 & 0 \end{bmatrix},\ H_k := \begin{bmatrix} CB & CM & CAB_{k-1} \\ 0 & 0 & H_{k-1} \end{bmatrix}$$

and the extended observability matrix

$$\Gamma_0 := C, \qquad \Gamma_k := \begin{bmatrix} CA^k \\ \Gamma_{k-1} \end{bmatrix}.$$

These data equations form now the basis for the estimation of the state sequence as shown in the following subsection.

3.3 Identification method

Data equations (5) and (6) form the basis for the development of the subspace identification algorithm, where first the state sequence X_k is estimated from measurements and then the estimate of the state sequence $\widehat{X}_k$ is used to obtain the system matrices as the solution to the following set of least-squares problems

$$\min_C \quad \left\| Y_k - C\widehat{X}_k \right\|_2^2$$

$$\min_{A,N,B,M} \left\| \widehat{X}_{k+1} - [A, N, B, M] \begin{bmatrix} \widehat{X}_k \\ \mathcal{F}(U_k) \odot \widehat{X}_k \\ U_k \\ \mathcal{G}(U_k) \end{bmatrix} \right\|_2^2$$

3.3.1. Approximation of the data equations

In order to be able to treat the identification problem with subspace identification techniques, the state sequence must be expressed in terms of the input and output measurements in such a way that the data equations remain linear in the parameters. We propose to use the following approximating

$$G_k X_{k+j-1|j} = L_k Z_{j+k-1|j} + \Xi^t_{k+j-1|j} \tag{7}$$

where L_k denotes a unknown parameter matrix, $Z_{j+k-1|j}$ a matrix which can be constructed from input-output measurements and $\Xi^t_{k+j-1|j}$ accounts for this approximation error. In order to derive a suitable approximation, consider the matrix N which can be rewritten as follows

$$\begin{aligned} N = \widetilde{N}_1 (I_q \otimes C) + \widetilde{N}_2 (I_q \otimes CA) + \cdots \\ + \widetilde{N}_n \left(I_q \otimes CA^{n-1}\right). \end{aligned}$$

Note that this is always possible since the system is assumed to be strongly locally observable. In this case the data-matrix $Z_{k+j|j}$ can be constructed as follows

$$Z_{j|j} := \mathcal{F}(U_j) \odot Y_j \in \mathbb{R}^{pq},$$
$$Z_{k+j|j} := \begin{bmatrix} Z_{k+j|j+1} \\ \mathcal{F}(U_j) \odot \begin{bmatrix} Y_{k+j|j} \\ Z_{k+j-1|j} \\ U_{k+j-1|j} \end{bmatrix} \end{bmatrix} \in \mathbb{R}^{d_Z(k)\times c},$$

with $d_Z(k) = q_k p + kq(p+m+r)$. The latter equation relies on the fact that $\Xi^t_{k+j-1|j}$ vanishes if the state-affine system is of special (triangular) structure which is a crucial result for the proposed method and will be further discussed in Subsection 3.3.3.

3.3.2. Estimating the state sequence

Choosing $j = 0$ and $j = k$ in the data equations (5) and (6), which corresponds to split the block matrices into two parts each of block size k, the following set of equations is obtained

$$Y_{k|0} = \Gamma_k X_0 + L_k Z_{k-1|0} + H_k U_{k-1|0} + \Xi_{k-1|0}, \tag{8}$$
$$X_k = A^k X_0 + \widetilde{A}_k \widetilde{F}_{k-1|0} \odot X_0 + B_k \widetilde{U}_{k-1|0}, \tag{9}$$
$$Y_{2k|k} = \Gamma_k X_k + L_k Z_{2k-1|k} + H_k U_{2k-1|k} + \Xi_{2k-1|k}. \tag{10}$$

Hereby $\Xi_{k+j-1|j}$ denotes the approximation error, which consists of two parts, $\Xi^t_{k+j-1|j}$ due to approximation of the data equations and $\Xi^a_{k+j-1|j}$ due to approximating nonlinear system (1)–(2) by a state-affine system. Observe now that the "past" state sequence X_0 can be simply obtained from equation (8)

$$X_0 = (\Gamma_k)^\dagger \left(Y_{k|0} - L_k Z_{k-1|0} - H_k U_{k-1|0} - \Xi_{k-1|0}\right).$$

Inserting this equation into (9) and subsequently into (10), we see that the future output can be written as the linear combination

$$Y = \Theta_X \Phi_X + \Theta_Z \Phi_Z + \Theta_U \Phi_U + \Xi \tag{11}$$

with $Y := Y_{2k|k}$, $\Phi_Z := Z_{2k-1|k}$, $\Phi_U := U_{2k-1|k}$ and

$$\Phi_X := \begin{bmatrix} Y_{k|0} \\ Z_{k-1|0} \\ \widetilde{U}_{k-1|0} \\ \widetilde{F}_{k-1|0} \odot \begin{bmatrix} Y_{k|0} \\ Z_{k-1|0} \\ U_{k-1|0} \end{bmatrix} \end{bmatrix}.$$

where Ξ accounts for all approximation errors. Hereby the fact was used that $U_{k-1|0}$ is included in the row space of $\widetilde{U}_{k-1|0}$. By comparing equations (10) and (11) it can be concluded, that

$$\Gamma_k X_k \approx \Theta_X \Phi_X, \tag{12}$$

where by exploiting this relation the state sequence can be recovered up to an unknown similarity transformation from the row space of matrix $\Theta_X \Phi_X$. The singular value decomposition

$$\Theta_X \Phi_X = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix} \begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix} \tag{13}$$

can be used for this purpose. Let n be the number of dominant singular vales of Σ stored in $\Sigma_1 \in \mathbb{R}^{n \times n}$ which provides an estimate of the system order. Now the future state sequence can be estimated as $\widehat{X}_k = \Sigma_1^{1/2} V_1^T$.

3.3.3. Approximation error analysis

The estimate of the state sequence $\widehat{X}_k$ is influenced by Ξ where we have

$$\begin{aligned} \|\Xi\| \leq \|\Xi_{2k-1|k}\| + \|\Gamma_k A^k (\Gamma_k)^\dagger \Xi_{k-1|0} \\ + \Gamma_k \widetilde{A}_k \widetilde{F}_{k-1|0} \odot (\Gamma_k)^\dagger \Xi_{k-1|0}\| \end{aligned} \tag{14}$$

From Corollary 1, we know that $\Xi^a_{k+j|j}$ can be arbitrarily small and hence we assume that Ξ depends mainly on $\Xi^t_{k+j-1|j} = \Delta G_k X_{k+j-1|j}$ where ΔG_k can be obtained by substitution $N \leftarrow \Delta N$ with

$$\Delta N := \sum_{i=0}^{n-2} \sum_{j=i+2}^{n} C A^i \widetilde{N}_j \left(I_q \otimes C A^{j-1}\right) \tag{15}$$

in Lemma 1 and Lemma 2.

The approximation error $\Xi^t_{k+j-1|j}$ vanishes if the state-affine system is of *triangular structure*, that is, if the system is such that

$$C A^i N = \lambda_i \left(I_q \otimes \Gamma_i\right) \tag{16}$$

holds for $i = 0, 1, \ldots, n-1$ where λ_i denotes a parameter matrix of suitable dimension.

To illustrate this requirement, consider (15) for the single output case where the state-affine system is assumed to be in canonical observable form, that is, $CA^i = e_{i+1}^T$ for $i = 0, \ldots, n-1$, where e_j^T denotes the $j-th$ unity vector. Then, all terms of ΔN given by $e_{i+1}^T \widetilde{N}_j \left(I_q \otimes e_j^T\right)$ vanish if the first $j-1$ rows of $\widetilde{N}_j$ are zero. In that special case the matrix N is of lower triangular structure which motivates to denote such systems *triangular systems*.

A detailed proof for the multiple output case can be found in the thesis by Schrempf [2004]. This proof shows via induction that the choice of $Z_{k+j|j}$ given in Subsection 3.3.1 together with the structural requirement (16) results is an exact solution to the deterministic identification problem.

In the nontriangular case the approximation error will not vanish, but if the state-affine system is uniform exponentially stable, similarly to Verdult and Verhaegen [2002] it can be shown that the second term of (14) decreases with increasing block size k. Similar results can be obtained for the case where measurement noise is acting on the system: the bias due to the noise terms will decrease with increasing block size k.

3.4 Computational efficient algorithm

To be able to recover the state sequence X_k from equation (12) we need to estimate the parameter matrix Θ_X in (11) which can be obtained by solving the following least squares problem:

$$\widehat{\Theta} = \arg\min_{\Theta} \|Y - \Theta\Phi\|_F^2, \tag{17}$$

where $\|\cdot\|_F$ denotes the Frobenius norm and where $\Phi^T = [\Phi_X^T, \Phi_Z^T, \Phi_U^T]$ and $\widehat{\Theta} = [\widehat{\Theta}_X, \widehat{\Theta}_Z, \widehat{\Theta}_U]$.

Unfortunately, due to the multiplicative nonlinearity of state-affine systems, the number of rows of the data matrices grows exponentially with block size k. It easily exceeds the number of measurements N_s (number of columns) for higher-order systems. To avoid computations with data matrices having a huge number of rows, a kernel algorithm can be used to solve the least squares problem (17). This algorithm only uses the kernel matrix $\Phi^T\Phi \in \mathbb{R}^{N_s \times N_s}$ and not the matrix Φ directly. In this way, computations in the large dimensional space spanned by the columns of Φ are avoided. The kernel method can be derived by assuming that the solution is of the form $\Theta = \Psi\Phi^T$. The least squares problem (17) becomes

$$\min_{\Psi} \|Y - \Psi\Phi^T\Phi\|_F^2. \tag{18}$$

It is easy to show that the assumption $\Theta = \Psi\Phi^T$ does not change the solution of the least squares problem.

A detailed discussion of these Kernel methods for subspace identification is provided by Verdult and Verhaegen [2003].

For the subspace identification method discussed above, we can estimate the state sequence from the row space of the matrix $\widehat{\Psi}\Phi_X^T\Phi_X$ similar to equation (12), because

$$\begin{aligned}\Theta\Phi &= \Psi(\Phi_X^T\Phi_X + \Phi_Z^T\Phi_Z + \Phi_U^T\Phi_U) \\ &= \begin{bmatrix}\Psi\Phi_X^T & \Psi\Phi_Z^T & \Psi\Phi_U^T\end{bmatrix}\begin{bmatrix}\Phi_X \\ \Phi_Z \\ \Phi_U\end{bmatrix}.\end{aligned}$$

We would like to stress that to compute the estimate $\widehat{\Psi}$, only the kernel matrix $\Phi^T\Phi$ is needed.

In order to be able to approximate a nonlinear system, state-affine systems require in many cases high dimensional data matrices. One reason for this is the polynomial expansion of the nonlinearities. Such high dimensional data matrices can easily lead to an ill-posed problem since only a limited number of samples is available and since the input is often not persistently exciting for such a high dimension. Therefore, regularization of the least-squares problem (18) is highly recommended. Tikhonov regularization is a commonly used method. In combination with the kernel method it results in the following least squares problem

$$\min_{\Psi}\Big(\lambda^2\|\Psi\Phi^T\|_F^2 + \|Y - \Psi\Phi^T\Phi\|_F^2\Big),$$

where λ is the regularization parameter. A proper choice of this parameter is crucial. Standard techniques involve generalized cross-validation and the L-curve criterion [Hansen, 1999].

4. SIMULATION EXAMPLE

The proposed identification algorithm will be illustrated by means of a simulation example. Consider the following nonlinear, discrete-time system

$$\begin{aligned}x_1(j+1) &= x_1(j) + \frac{8}{50}x_2(j) - \frac{1}{10}\tanh(x_1(j))x_1(j)^2 \\ x_2(j+1) &= -2x_1(j) + \frac{3}{20}x_2(j) + u(j) + \frac{3}{5}u(j)^2 \\ &\quad + \frac{1}{10}\tanh(x_1(j))x_1(j)^2 \\ y(j) &= x_1(j) + v(j)\end{aligned}$$

where as input sequence a filtered zero-mean white noise sequence with amplitudes in the range $[-1.5, 1.5]$ was used to simulate the output of the system. Zero-mean white-noise measurement noise was added such that the signal-to-noise ratio amounts approximately 10dB. A state-affine system with $\rho = 2$ was used for approximation.

A Monte-Carlo simulation study with 50 independent runs was performed where each identified state-affine model was validated with a fresh independent data set where the variance-accounted-for value, VAF $= \max\{\text{var}(\widehat{y} - y)/\text{var}(y), 0\}$, defines the model quality.

Table 1. Bias of estimated matrices $\widehat{A}$ and the mean and variance (between brackets) of the VAF values for different block sizes k and $N_s = 500$.

block size k	$\|eig(A) - eig(\widehat{A})\|$	**VAF**
2	0.1842	0.88 $(3.5 \cdot 10^{-2})$
3	0.2361	0.82 $(7.7 \cdot 10^{-2})$
4	0.1236	0.93 $(8.6 \cdot 10^{-3})$
5	0.0765	0.96 $(3.0 \cdot 10^{-3})$

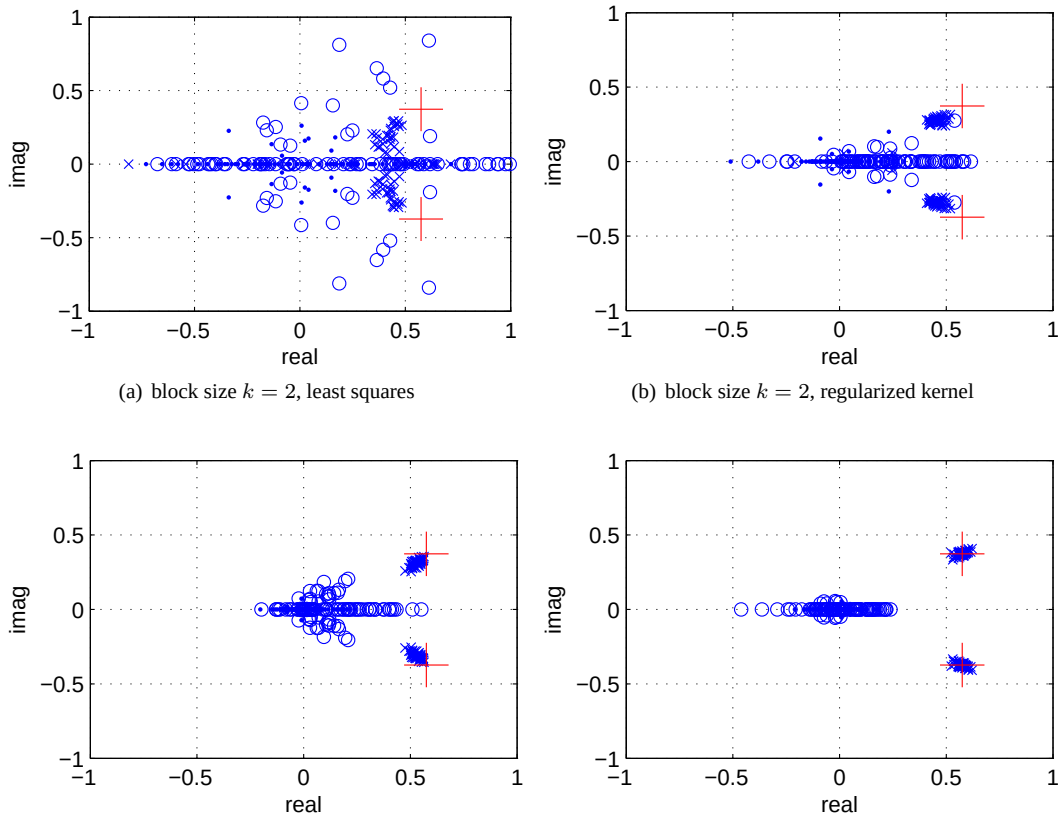
The results for fixed block size $k = 2$ are depicted in Figure 1a for the least squares approach, and in Figure 1b for the kernel method with Tikhonov regularization. The regularization parameter was obtained from the L-curve criterion. The estimates can be improved significantly by using the kernel method with regularization. Figure 1c shows the results when using a block size $k = 5$ which allows to decrease the estimation bias. The result after a final prediction error type optimization is depicted in Figure 1d. Since such an optimization problem is nonconvex, there is in general no guarantee that the optimization algorithm converges to the optimal parameter values. However, for the performed simulation experiments the good initial estimates provided by the proposed identification algorithm result in models with excellent performance.

Table 1 summarizes the results for the performed simulation experiments where the bias of the eigenvalues of estimated matrix $\widehat{A}$ and the VAF values for the proposed identification algorithm are compared for different block size k. A clear reduction of the estimation bias can be observed by increasing the block size.

5. CONCLUSIONS

A subspace identification algorithm for state-affine state-space systems was proposed which allows to approximate a generic state-space system. For the proposed identification algorithm a detailed approximation error analysis was performed. It was shown that if the state-affine system is of triangular structure the deterministic approximation error vanishes. In the general case, the estimation bias can be reduced by increasing the block size k. Kernel methods can be used to deal with the huge dimension of the involved data matrices.

When using approximative models like state-affine models the identification problem is in many cases ill posed since the corresponding data equations contain nearly redundant information or the input may be not rich enough. By incorporating regularization techniques this problem can be tackled.



(a) block size $k = 2$, least squares

(b) block size $k = 2$, regularized kernel

(c) block size $k = 5$, regularized kernel

(d) after optimization

Figure 1. The eigenvalues of the identified matrix $\widehat{A}$ are indicated by $\times$ whereas the eigenvalues of the identified quadratical submatrices of $\widehat{N}$ are indicated by $\cdot$ and $\circ$ respectively. The eigenvalues of matrix A of the original system are denoted by the big crosses.

REFERENCES

F. Albertini and D. D'Alessandro. *Remarks on the observability of nonlinear discrete time systems*, pages 155–162. System Modeling and Optimization (Prague 1995). Chapman and Hall, London, 1996.

S. Boyd and L. O. Chua. Fading Memory and the problem of Approximating nonlinear opperators with volterra series. *IEEE Transactions on Circuits and Systems*, 32(11):1150–1161, 1985.

Huixin Chen and Jan Maciejowski. An improved subspace identification method for bilinear systems. In *Proceedings of the 39th IEEE Conference on Decision and Control*, Sydney, Australia, December 2000.

V. Constanza, B.W. Dickinson, and Johnson. Universal Approximations of Discrete-time Control Systems over Finite Time. *IEEE Transactions on Automatic Control*, 28(4):439–451, 1983.

H. Diaz and A. Desrochers. Modeling of Nonlinear Discrete-time Systems from Input-Output Data. *Automatica*, 24(5):629–641, 1988.

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

M. Fliess. On the approximation of nonlinear systems by some simple state-space models. *IFAC Identification and System Parameter Estimation*, pages 511–514, 1982.

G. Gatt and N. Kalouptsidis. Identification of discrete-time state-affine state space models using cumulants. *Automatica*, 38: 1663–1681, 2002.

P. C. Hansen. Regularization tools version 3.0 for matlab 5.2. *Numerical Algorithms*, 20(2–3):195–196, 1999.

L. H. Lee and K. Poolla. Identification of linear parameter-varying systems via LFTs. *Proceedings of the 35th IEEE Conference on Decision and Control*, 2:1545–1550, 1996.

L. Ljung. *System Identification, Theory for the User*. Prentice Hall, second edition, 1999.

A. Schrempf. *Identification of extended state-affine systems*. PhD thesis, Johannes Kepler University, Linz, Austria, 2004.

A. Schrempf and L. Del Re. Linear in parameters identifiability of extended bilinear systems. *Proceedings of the European Control Conference ECC 2001*, pages 1–6, 2001.

E. D. Sontag. *Polynomial Response Maps*. Springer, 1979.

Peter Van Overschee and Bart De Moor. *Subspace Identification for Linear Systems; Theory, Implementation, Applications*. Kluwer Academic Publishers, Dordrecht, The Netherlands, 1996. ISBN 0-7923-9717-7.

Vincent Verdult. *Nonlinear System Identification: A State-Space Approach*. PhD thesis, University of Twente, Enschede, The Netherlands, 2002. URL `http://www.dcsc.tudelft.nl/~vverdult/`.

Vincent Verdult and Michel Verhaegen. Subspace identification of multivariable linear parameter-varying systems. *Automatica*, 38 (5):805–814, 2002.

Vincent Verdult and Michel Verhaegen. A kernel method for subspace identification of multivariable bilinear systems. In *Proceedings of the 42nd IEEE Conference on Decision and Control*, Maui, Hawaii, December 2003.

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

G. Zang and P. A. Iglesias. Fading memory and stability. *Journal of the Franklin Institute*, (340):489–502, 2003.