43rd IEEE Conference on Decision and Control
December 14-17, 2004
Atlantis, Paradise Island, Bahamas

Subspace Identification of Piecewise Linear Systems

Vincent Verdult and Michel Verhaegen

Abstract— **Subspace identification can be used to obtain models of piecewise linear state-space systems for which the switching is known. The models should not switch faster than the block size of the Hankel matrices used. The nonconsecutive parts of the input and output data that correspond to one of the local linear systems can be used to obtain the system matrices of that system up to a linear state transformation. The linear systems obtained in this way cannot be combined directly, because the state transformation is different for each of the local linear systems. The transitions between the local linear systems can be used to transform the models to the same state space basis. We show that the necessary transformations can be obtained from the data, if the data contains a sufficiently large number of transitions for which the states at the transition are linearly independent. An algorithm to determine the transformations is presented, and the sensitivity with respect to noise is investigated using a Monte-Carlo simulation.**

I. Introduction

Hybrid systems form an important class of dynamical systems for which analysis, verification and control techniques have been widely studied [1], [2], [3]. However, the identification of hybrid models from measured data is still in its infancy. The identification methods proposed in the literature focus on piecewise affine (PWA) and piecewise linear (PWL) systems. PWA and PWL systems form important subclasses of hybrid systems [4], [5], [6], [7], [8]. PWL systems were introduced by Sontag [9] as unifying models for describing interconnections between automata and linear systems. The few identification studies that are available focus on models in input-output form [10], [11], [12], [13], [14], [15], [16], expect for [17] which considers autonomous state-space systems. Although input-output models can be used to describe dynamical systems, they are not suitable for most dynamic analysis and control methods. Indeed, many existing hybrid and piecewise linear analysis and control methods are based on state-space models. Furthermore, state-space models are more suitable for dealing with multivariable inputs and outputs.

In this paper we present a subspace identification method for PWL systems with known switching. It has been argued that the identification of PWA and PWL systems for which the switching is known can be tackled using standard techniques [15]. However, we would like to point out that such a statement holds true for systems in input/output form, but *not* for systems in state-space form. The identification of state-space models provides an additional challenge of finding the unknown states of the system. Traditional prediction error models for state-space identification lead to nonlinear nonconvex optimization problems. Subspace identification techniques avoid such problems by utilizing techniques from linear algebra. At this time, subspace identification methods exist for linear time-invariant [18], [19], linear parameter-varying [20] and certain nonlinear systems (an overview is given in [21]).

A PWL system can be described by a set of state-space equations

$$x(k+1) = A_i x(k) + B_i u(k), \tag{1}$$
$$y(k) = C_i x(k) + D_i u(k), \tag{2}$$

for

$$\begin{bmatrix} x(k) \\ u(k) \end{bmatrix} \in \mathcal{P}_i,$$

where $u(k) \in \mathbb{R}^m$ is the input, $y(k) \in \mathbb{R}^\ell$ the output and $x(k) \in \mathbb{R}^n$ the state; and where $\mathcal{P}_i$, $i = 1, 2, \dots, M$ is a partition such that

$$\bigcup_{i=1}^{M} \mathcal{P}_i = \mathbb{R}^n \times \mathbb{R}^m, \qquad \text{and} \quad \mathcal{P}_i \cap \mathcal{P}_j = \emptyset \quad \text{for} \quad i \neq j.$$

The *local* linear state-space models described by (A_i, B_i, C_i, D_i), $i = 1, 2, \dots, M$ are switched according to the regions $\mathcal{P}_i$.

In this paper we assume that for each time-instant it is known in which partition the system operates: the switching is known. With this assumption we can rewrite the system description in the following form

$$x(k+1) = \sum_{i=1}^{M} p_i(k)\Big(A_i x(k) + B_i u(k)\Big), \tag{3}$$
$$y(k) = \sum_{i=1}^{M} p_i(k)\Big(C_i x(k) + D_i u(k)\Big), \tag{4}$$

where $p_i(k)$ is the switching signal which equals one if model i is active at time instant k. These switching signals satisfy for all k:

$$p_i(k) \in \{0, 1\} \quad \text{and} \quad \sum_{i=1}^{M} p_i(k) = 1.$$

The goal is to determine the system matrices (A_i, B_i, C_i, D_i) of the PWL system (3)–(4) from a finite number of measurements of the input $u(k)$, output $y(k)$ and switching signals $p_i(k)$. Note that there does not exist a unique set of matrices (A_i, B_i, C_i, D_i), $i = 1, 2, \dots, M$ that are compatible with the measurements of $u(k)$, $y(k)$ and $p_i(k)$: for each invertible matrix T the set of matrices $(T^{-1}A_iT, T^{-1}B_i, C_iT, D_i)$, $i = 1, 2, \dots, M$ is compatible

The authors are with the Delft Center for Systems and Control, Delft University of Technology, Mekelweg 2, 2628 CD Delft, The Netherlands v.verdult@dcsc.tudelft.nl m.verhaegen@dcsc.tudelft.nl

with the measurements. In fact this corresponds to a linear state transformation applied to the system (3)–(4). It is important to realize that this state transformation must be the *same* for all local state-space models. Without any other information of the system than the measurements mentioned above, we can only determine the system up to a linear state transformation. This is usually not considered to be a problem.

In the identification problem it is assumed that the noise is additive to the output:

$$y(k) = \sum_{i=1}^{M} p_i(k)\Big(C_i x(k) + D_i u(k) + v_i(k)\Big),$$

where $v_i(k)$ is a zero-mean sequence. The variance of the noise can be different for each local model. It is also assumed that all the local models are observable.

This paper is organized as follows. Section II describes the use of subspace identification to obtain the local models up to a linear state transformation. This state transformation is different for each local model. Section III discusses how the transitions between the models can be used to transform all local models to the same state basis. A simulation example is provided in Section IV.

II. Identification of local models

In this section we show that if the switching signals $p_i(k)$ are known the local models of the system (3)–(4) can be obtained up to a linear state transformation using subspace identification techniques.

For ease of notation we define the delay vectors:

$$\begin{aligned}
\overline{y}_s(k) &= \left[y(k)^T \quad y(k+1)^T \quad \ldots y(k+s-1)^T\right]^T, \\
\overline{u}_s(k) &= \left[u(k)^T \quad u(k+1)^T \quad \ldots u(k+s-1)^T\right]^T, \\
\overline{v}_s^{(i)}(k) &= \left[v_i(k)^T \quad v_i(k+1)^T \quad \ldots v_i(k+s-1)^T\right]^T.
\end{aligned}$$

If the system does not switch during the time interval $[k, k+s-1]$ and local model i is active during this interval, then the delay vectors can be related as follows

$$\overline{y}_s(k) = \Gamma_s^{(i)} x(k) + H_s^{(i)} \overline{u}_s(k) + \overline{v}_s^{(i)}(k), \qquad (5)$$

where

$$\Gamma_s^{(i)} = \begin{bmatrix} C_i \\ C_i A_i \\ C_i A_i^2 \\ \vdots \\ C_i A_i^{s-1} \end{bmatrix},$$

and

$$H_s^{(i)} = \begin{bmatrix} D_i & 0 & 0 & \cdots & 0 \\ C_i B_i & D_i & 0 & \cdots & 0 \\ C_i A_i B_i & C_i B_i & D_i & & 0 \\ \vdots & & \ddots & \ddots & \\ C_i A_i^{s-2} B_i & C_i A_i^{s-3} B_i & \cdots & C_i B_i & D_i \end{bmatrix}. \qquad (6)$$

The delay vectors can be used to build Hankel matrices, commonly used in subspace identification:

$$Y_{1,s,N} = \begin{bmatrix} \overline{y}_s(1) & \overline{y}_s(2) & \ldots & \overline{y}_s(N) \end{bmatrix}.$$

The matrix $U_{1,s,N}$ can be constructed in a similar way. Since the system (3)–(4) regularly switches from one local model to the other, only some columns of the Hankel matrices $Y_{1,s,N}$ and $U_{1,s,N}$ satisfy equation (5). To be precise: only the columns of which the rows span a time interval in which the system does not switch satisfy equation (5). The other columns contain data that correspond to more than one model; we will call these columns *mixed columns*. The structure of a Hankel matrix around a transition (switch) is illustrated in Figure 1.

We construct the local Hankel matrix $Y_{1,s,N}^{(i)}$ for the ith model by taking all the columns from $Y_{1,s,N}$ that correspond to the ith model, such that $Y_{1,s,N}^{(i)}$ does not contain any mixed columns. The local Hankel matrices $U_{1,s,N}^{(i)}$ and $V_{1,s,N}^{(i)}$ can be constructed in a similar way. For the ith model we can formulate a local data equation:

$$Y_{1,s,N}^{(i)} = \Gamma_s^{(i)} X_{1,N}^{(i)} + H_s^{(i)} U_{1,s,N}^{(i)} + V_{1,s,N}^{(i)}, \qquad (7)$$

where $X_{1,N}^{(i)}$ contains the states that correspond to the columns of $Y_{1,s,N}^{(i)}$ as in equation (5). To be able to use subspace identification, the PWL system should not switch faster than the block size $s > n$ of the Hankel matrices: if it does switch that fast all columns of the Hankel matrices will be mixed columns and the local Hankel matrices $Y_{1,s,N}^{(i)}$ and $U_{1,s,N}^{(i)}$ will be empty.

The local data equation (7) allows us to use the MOESP type of subspace algorithms [19], [22] to identify each local model up to a linear state transformation. Below, we will discuss the use of the PI-MOESP method [22] which can deal with nonwhite noises $v_i(k)$. First, the matrices A_i and C_i will be computed from an estimate of the matrix $\Gamma_s^{(i)}$, and then used in a second step to derive the B_i and D_i matrices. The PI-MOESP method starts with the computation of the following RQ factorization:

$$\begin{bmatrix} U^{(i)} \\ Z^{(i)} \\ Y^{(i)} \end{bmatrix} = \begin{bmatrix} R_{11}^{(i)} & 0 & 0 \\ R_{21}^{(i)} & R_{22}^{(i)} & 0 \\ R_{31}^{(i)} & R_{32}^{(i)} & R_{33}^{(i)} \end{bmatrix} \begin{bmatrix} Q_1^{(i)} \\ Q_2^{(i)} \\ Q_3^{(i)} \end{bmatrix}, \qquad (8)$$

where for ease of notation we dropped the subscripts of the Hankel matrices and where $Z^{(i)}$ is an instrumental variable matrix. This matrix is constructed as follows for each column of $U^{(i)}$ given by $\overline{u}_s(k)$ the corresponding column in $Z^{(i)}$ is taken as $\overline{u}_s(k-s)$. Note that with this construction the Hankel matrix $Z^{(i)}$ contains some mixed columns (compare Figure 1). This is not a problem under the standard assumption that the noise and input are independent, that is

$$\lim_{N^{(i)} \to \infty} \frac{1}{N^{(i)}} V^{(i)} \underbrace{\left[(U^{(i)})^T \quad (Z^{(i)})^T\right]} = 0, \qquad (9)$$

Model 1				Mixed columns			Model 2			
$\dots$	$y(k-6)$	$y(k-5)$	$y(k-4)$	$y(k-3)$	$y(k-2)$	$y(k-1)$	$\mathbf{y(k)}$	$\mathbf{y(k+1)}$	$\mathbf{y(k+2)}$	$\dots$
$\dots$	$y(k-5)$	$y(k-4)$	$y(k-3)$	$y(k-2)$	$y(k-1)$	$\mathbf{y(k)}$	$\mathbf{y(k+1)}$	$\mathbf{y(k+2)}$	$\mathbf{y(k+3)}$	$\dots$
$\dots$	$y(k-4)$	$y(k-3)$	$y(k-2)$	$y(k-1)$	$\mathbf{y(k)}$	$\mathbf{y(k+1)}$	$\mathbf{y(k+2)}$	$\mathbf{y(k+3)}$	$\mathbf{y(k+4)}$	$\dots$
$\dots$	$y(k-3)$	$y(k-2)$	$y(k-1)$	$\mathbf{y(k)}$	$\mathbf{y(k+1)}$	$\mathbf{y(k+2)}$	$\mathbf{y(k+3)}$	$\mathbf{y(k+4)}$	$\mathbf{y(k+5)}$	$\dots$

Fig. 1. Mixed columns in the Hankel matrix $Y_{1,4,N}$ in the vicinity of a model transition at time instant k.

where $N^{(i)}$ is the number of data points for the ith model. Note that from a practical point of view it is not realistic to let the number of data points for each model go to infinity. In practice a sufficiently large number will do. This is supported by the simulation examples provided in Section IV.

If in addition the input is persistently exciting such that the underbraced matrix in equation (9) has rank $2sm$, then it can be shown that [22]

$$\lim_{N^{(i)}\to\infty} \frac{1}{\sqrt{N^{(i)}}} R_{31}^{(i)} = \lim_{N^{(i)}\to\infty} \frac{1}{\sqrt{N^{(i)}}} \Big(\Gamma_s^{(i)} X^{(i)} (Q_1^{(i)})^T + H^{(i)} R_{11}^{(i)} \Big), \qquad (10)$$

$$\lim_{N^{(i)}\to\infty} \frac{1}{\sqrt{N^{(i)}}} R_{32}^{(i)} = \lim_{N^{(i)}\to\infty} \frac{1}{\sqrt{N^{(i)}}} \Gamma_s^{(i)} X^{(i)} (Q_2^{(i)})^T. \qquad (11)$$

Note that in the noise free case, these equations hold without the limits. The second equation shows that the column space of the matrix $R_{32}^{(i)}$ equals the column space $\Gamma_s^{(i)}$, asymptotically. Therefore, the matrices A_i and C_i (which determine $\Gamma_s^{(i)}$) can be retrieved up to a linear state transformation from the column space of $R_{32}^{(i)}$. This column space is determined by a singular value decomposition

$$R_{32}^{(i)} = \begin{bmatrix} \mathcal{U}^{(i)} & \mathcal{U}_{\perp}^{(i)} \end{bmatrix} \begin{bmatrix} \mathcal{S}^{(i)} & 0 \\ 0 & \mathcal{S}_{\perp}^{(i)} \end{bmatrix} \begin{bmatrix} \mathcal{V}^{(i)} \\ \mathcal{V}_{\perp}^{(i)} \end{bmatrix}, \qquad (12)$$

where $\mathcal{S}^{(i)}$ is a diagonal matrix containing the n dominant singular values. An estimate of the matrix C_i up to a linear transformation can be obtained as the first ℓ rows of $\mathcal{U}^{(i)}$:

$$\widehat{C}_i = \mathcal{U}^{(i)}(1\colon \ell, :\,),$$

where the Matlab-like notation $(1\colon \ell, :\,)$ refers to the first ℓ rows. The corresponding matrix $\widehat{A}_i$ follows by solving the set of linear equations:

$$\mathcal{U}^{(i)}(1\colon (s-1)\ell, :\,) A_i = \mathcal{U}^{(i)}(\ell+1\colon s\ell, :\,).$$

Given estimates of the matrices A and C, the corresponding matrices B and D are usually obtained by solving a linear least squares problem using the data and the estimates of A and C. However, such a procedure relies on using a consecutive batch of data points of the input and output, which is not available when a local model for a PWL systems needs to be determined: due to the switching nature of the PWL system we need to determine the system matrices of one local model from a number of different nonconsecutive batches of data points. Therefore, we determine matrices $\widehat{B}_i$ and $\widehat{D}_i$ using equation (10). This alternative procedure was explained in one of the first papers on MOESP methods. By construction $(\mathcal{U}_{\perp}^{(i)})^T \Gamma_s^{(i)} = 0$, hence an estimate $\widehat{H}_s^{(i)}$ of the matrix $H_s^{(i)}$ satisfies

$$(\mathcal{U}_{\perp}^{(i)})^T R_{31}^{(i)} (R_{11}^{(i)})^{-1} = (\mathcal{U}_{\perp}^{(i)})^T \widehat{H}_s^{(i)},$$

where $R_{11}^{(i)}$ is invertible because of the persistency of excitation assumption on the input. Because of the Toeplitz structure of the matrix $\widehat{H}_s^{(i)}$ —see equation (6)— this can be rewritten in a set of equations that are linear in $\widehat{B}_i$ and $\widehat{D}_i$. For further details we refer to the paper [22].

III. Determination of state transformations

The local models obtained from subspace identification cannot be combined directly because they have been determined up to a state transformation which is *different* for each local model. To be able to use the local models they must be transferred to the same state basis. The necessary state transformations can be obtained by comparing at each transition the state of the model before the transition with the state of the model after the transition. This idea was first formulated in [17] for autonomous systems. However, they used only one transition to determine each transformation. This leads to an underdetermined set of equations for which there are several solutions. These solutions are all compatible with the data used for identification, but *not* necessarily with another data set. Therefore, we propose to use several transitions to determine one state transformation. Below we will state conditions on the transitions such that a unique state transformation can be determined.

Before presenting the general algorithm we take a closer look at estimating the transformation from one particular type of transition. Consider all transitions from model p to model q. Let τ_j^{t} denote the time instant of the jth transition from model p to q, that is, the first time instant that model q becomes active. Let τ_j^{s} denote the start of the data segment in which model p is active, that precedes the transition j, and let τ_j^{e} denote the end of the data segment in which model q is active after the transition j. These time instances are illustrated in Figure 2. The state of model p before the transition can be estimated at time instant $\tau_j^{\mathrm{t}} - s$ as

$$\widehat{x}^p(\tau_j^{\mathrm{t}} - s) = (\Gamma_s^{(p)})^{\dagger} \Big(Y(:\,, \tau_j^{\mathrm{t}} - s) - H_s^{(p)} U(:\,, \tau_j^{\mathrm{t}} - s) \Big),$$

where for ease of notation we dropped again the subscripts of the Hankel matrices and where $(\Gamma_s^{(p)})^{\dagger}$ denotes the pseudo-inverse of $\Gamma_s^{(p)}$. The time instant $\tau_j^{\mathrm{t}} - s$ is used to

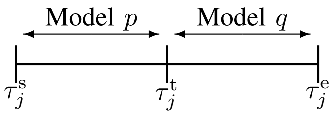

Fig. 2. Definition of the time instances τ_j^{s}, τ_j^{t}, and τ_j^{e}.

avoid using mixed columns of the Hankel matrices U and Y (this follows from Figure 1). Simulating model p we can obtain the state $\widehat{x}^p$ at the transition time τ_j^{t} as follows:

$$\widehat{x}^p(\tau_j^{\mathrm{t}}) = A_p^s \widehat{x}^p(\tau_j^{\mathrm{t}} - s) + \sum_{i=0}^{s-1} A_p^{s-i-1} B_p u(i + \tau_j^{\mathrm{t}} - s).$$

The state of model q after the transition follows as

$$\widehat{x}^q(\tau_j^{\mathrm{t}}) = (\Gamma_s^{(q)})^\dagger \Big(Y(\,:\,, \tau_j^{\mathrm{t}}) - H_s^{(q)} U(\,:\,, \tau_j^{\mathrm{t}}) \Big).$$

The estimates $\widehat{x}^p$ and $\widehat{x}^q$ are unbiased as long as the output noises $v_i(k)$ are zero mean. Now it follows that the final state of model p and the initial state of model q at the transition are related by

$$\widehat{x}^p(\tau_j^{\mathrm{t}}) = T_q \widehat{x}^q(\tau_j^{\mathrm{t}}), \qquad (13)$$

where $T_q \in \mathbb{R}^{n \times n}$ is the state transformation that needs to be applied to model q to bring it into the same state basis as model p. Not only the transitions from model p to model q can be used to obtain equations like (13) involving the state transformation T_q, in addition also the transitions from model q to model p can be used in a similar fashion. In a transition from model q to model p, first model q is active, thus

$$\widehat{x}^q(\tau_j^{\mathrm{t}} - s) = (\Gamma_s^{(q)})^\dagger \Big(Y(\,:\,, \tau_j^{\mathrm{t}} - s) - H_s^{(q)} U(\,:\,, \tau_j^{\mathrm{t}} - s) \Big).$$

After the transition model p is active:

$$\widehat{x}^p(\tau_j^{\mathrm{t}}) = (\Gamma_s^{(p)})^\dagger \Big(Y(\,:\,, \tau_j^{\mathrm{t}}) - H_s^{(p)} U(\,:\,, \tau_j^{\mathrm{t}}) \Big).$$

The state transformation T_q can be determined using several transitions from model p to q and vice versa to set up a set of linear equations:

$$\begin{bmatrix} \widehat{x}^p(\tau_1^{\mathrm{t}}) & \widehat{x}^p(\tau_2^{\mathrm{t}}) & \cdots \end{bmatrix} = T_q \underbrace{\begin{bmatrix} \widehat{x}^q(\tau_1^{\mathrm{t}}) & \widehat{x}^q(\tau_2^{\mathrm{t}}) & \cdots \end{bmatrix}}. \qquad (14)$$

We arrive at the important conclusion that from this set of equations T_q can be uniquely determined if the underbraced matrix is of rank n. This rank condition can be considered as a kind of persistency of excitation condition: we need a sufficient number of linearly independent states at the transition times to be able to determine the state transformation. Equation (14) can be solved in a least squares sense to obtain T_q. To take the estimation errors on the state into account the use of total least squares seems to be appropriate [23].

Given multiple models, say $p_1, p_2, \ldots, p_d$, that share the same state basis we can determine a transformation T_q that brings model q into the same basis using all available transitions from the models $p_1, p_2, \ldots, p_d$ to the model q and transitions from the model q to the models $p_1, p_2, \ldots, p_d$.

Step 1:

C	$\mathbf{1 \leftrightarrow 2}$	$1 \leftrightarrow 3$	$1 \leftrightarrow 4$	$2 \leftrightarrow 3$	$2 \leftrightarrow 4$	$3 \leftrightarrow 4$
T	**6**	4	3	4	2	1

Step 2:

C	$\mathbf{1 \leftrightarrow 3}$	$1 \leftrightarrow 4$
	$\mathbf{2 \leftrightarrow 3}$	$2 \leftrightarrow 4$
T	$\mathbf{4+4}$	$3+2$

Step 3:

C	$\mathbf{1 \leftrightarrow 4}$
	$\mathbf{2 \leftrightarrow 4}$
	$\mathbf{3 \leftrightarrow 4}$
T	$3+2+1$

Fig. 3. Example of the order in which the state transformations are determined. There are four local models. The bold entries indicate the selected combination of models. C: combination of models. T: total number of transitions for a particular combination of models. In step 1, models 1 and 2 are selected, because their combination has the largest number of transitions. In step 2, model 3 is selected, because it has the largest number of transitions with models 1 and 2. In step 3 only model 4 is left.

These observations form the basis for an algorithm that transforms all the local state-space models obtained from subspace identification into the same state-space basis. Assume that we have M models: $1, 2, \ldots, M$. The algorithm starts by determining for each combination (p, q) of models, the total number of transitions from p to q and from q to p. The combination with the largest number of transitions is chosen; let us assume that this is the combination $(1, 2)$. A transformation is computed to bring model 1 and 2 in the same basis. The next step is to determine the total number of transitions for each model i, $i = 3, 4, \ldots, M$ with *both* model 1 and 2. Again the model with the largest number of transitions is selected. The algorithm proceeds in this way until all models have been transformed. An example with four models is shown in Figure 3.

The general algorithm described in pseudo code is as follows:

```
1) Initialization: $\mathcal{M} = \{1, 2, \ldots, M\}$, $\mathcal{F} = \emptyset$.
2) Select model $q$ that is involved in the largest number
   of transitions. $\mathcal{M} = \mathcal{M} \backslash \{q\}$, $\mathcal{F} = \mathcal{F} \cup \{q\}$.
3) while $\mathcal{M} \neq \emptyset$
     a) Determine total number of switches:
        for each model $j \in \mathcal{M}$
          $T(j)$ equals the total number of switches
          from the models $p \in \mathcal{F}$ to the model $j$
        end
     b) Select the model $q \in \mathcal{M}$ with the largest number
        of transitions $T(q)$.
     c) Determine states at the transitions:
        for each model $p \in \mathcal{F}$
          for each transition $j$ with model $q$
```

```
            determine $\widehat{x}^p(\tau_j^{\mathrm{t}})$
            determine $\widehat{x}^q(\tau_j^{\mathrm{t}})$
        end
    end
    d) Determine state transformation $T_q$ as in equation (14).
    e) Apply state transformation to model $q$:
       $A_q \leftarrow T_q A_q T_q^{-1}$, $B_q \leftarrow T_q B_q$, $C_q \leftarrow C_q T_q^{-1}$.
    f) Adapt model sets: $\mathcal{M} = \mathcal{M}\backslash\{q\}$, $\mathcal{F} = \mathcal{F} \cup \{q\}$.
end
```

In this algorithm the set $\mathcal{M}$ contains the models for which the state transformation still needs to be determined. The set $\mathcal{F}$ contains the models that have been processed already. The order in which the models are processed is determined by the number of transitions with the already finished models.

IV. Simulation example

We have tested the proposed identification approach using simulated data from the following second-order PWL model:

$$
\begin{aligned}
A &= \left[\begin{array}{cc|cc|cc|cc} 0 & 0.8 & 0 & 0.5 & 0.8 & 0 & 0 & 0.4 \\ -0.8 & 0.5 & -0.5 & 0 & 0 & -0.3 & -0.4 & 0 \end{array}\right], \\
B &= \left[\begin{array}{c|c|c|c} 0.4 & 1 & 1 & 1 \\ 0 & 0.5 & 0 & 0 \end{array}\right], \\
C &= \left[\begin{array}{cc|cc|cc|cc} 1 & 0 & 1 & 0 & 1 & 1 & 1 & 0.5 \end{array}\right], \\
D &= 0.
\end{aligned}
$$

The system is excited using a zero-mean white noise input signal that has been filtered using a fourth-order Butterworth filter with a cut-off frequency of 0.8 times the Nyquist frequency. The switching sequence $p(k)$ is generated using pseudo random binary signals such that only one model is active at a certain time instant and such that the models do not switch faster than 20 time steps. Data sequences of $u(k)$, $y(k)$, and $p(k)$ of 2000 data points are used to identify a PWL system. The block size used in the subspace method equals 6. The performance of the identified system is evaluated by looking at the eigenvalues of the A_i matrices and 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:

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

where $\widehat{y}_k$ denotes the output signal obtained by simulating the identified PWL system, y_k is the output signal of the true PWL system, and $\operatorname{var}(\cdot)$ 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 $u(k)$ and switching signal $p(k)$ were used. The corresponding output signal $y(k)$ was simulated and zero-mean white noise was added to create three different noisy signals with signal-to-noise ratios of 15, 20, and 25 dB. The signal to noise ratio is defined as:

$$\text{SNR} = 20 \log_{10} \frac{\operatorname{var}(y(k))}{\operatorname{var}(v(k))},$$

where $v(k)$ is the noise sequence. For each of the three noisy signals a PWL model was identified. The results are summarized in Figures 4–9. The eigenvalues of the estimated models are compared to the true values in the first three figures; one figure for each signal-to-noise ratio. Figures 7–9 show the corresponding histograms of the VAF values on a fresh validation data set. As expected the quality of the models decreases with decreasing signal-to-noise ratio. The quality is still reasonable in all cases.

V. Conclusions

Multivariable PWL state-space systems with known switching can be identified using subspace identification methods. We have shown that the PI-MOESP method can be used on nonconsecutive parts of the input and output data that correspond to one of the switching linear systems as long as the PWL system does not switch faster than the block size used to build the Hankel matrices. The system matrices of the local models are obtained up to a linear state transformation. This state transformation is different for each of the local linear systems. To combine the local models they need to be transformed to the same state basis. We have discussed an algorithm that transforms the local systems into the same basis. It utilizes the transitions between the local linear systems. It requires a sufficiently large number of transitions for which the states at the transition are linearly independent.

References

[1] P. J. Antsaklis and A. Nerode, eds., "Special issue on hybrid systems," *IEEE Transactions on Automatic Control*, vol. 43, no. 4, Apr. 1998.

[2] A. S. Morse, C. C. Pantelides, S. Sastry, and J. M. Schumacher, eds., "Special issue on hybrid systems," *Automatica*, vol. 35, no. 3, Mar. 1999.

[3] A. J. van der Schaft and J. M. Schumacher, *An Introduction to Hybrid Dynamical Systems*, ser. Lecture Notes in Control and Information Sciences. London: Springer-Verlag, 2000, vol. 251.

[4] W. P. M. H. Heemels, B. De Schutter, and A. Bemporad, "Equivalence of hybrid dynamical models," *Automatica*, vol. 37, no. 7, pp. 1085–1091, July 2001.

[5] L. O. Chua and A.-C. Deng, "Canonical piecewise-linear representation," *IEEE Transactions on Circuits and Systems*, vol. 35, no. 1, pp. 101–111, Jan. 1988.

[6] T. A. M. Kevenaar and D. M. W. Leenaerts, "Comparison of piecewise-linear model descriptions," *IEEE Transactions on Circuits and Systems*, vol. 39, no. 12, pp. 996–1004, 1992.

[7] D. M. W. Leenaerts and W. M. G. van Bokhoven, *Piecewise Linear Modeling and Analysis*. Boston: Kluwer Academic Publishers, 1998.

[8] L. Vandenberghe, B. De Moor, and J. Vandewalle, "The generalized linear complementarity problem applied to the complete analysis of resistive piecewise-linear circuits," *IEEE Transactions on Circuits and Systems*, vol. 36, no. 11, pp. 1382–1391, Nov. 1989.

[9] E. D. Sontag, "Nonlinear regulation: The piecewise linear approach," *IEEE Transactions on Automatic Control*, vol. 26, no. 2, pp. 346–358, Apr. 1981.

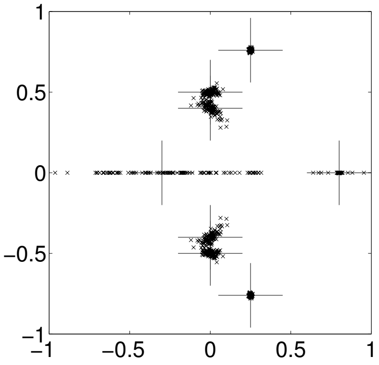

Fig. 4. Eigenvalues of the estimated A_i matrices in the complex plane, for 100 experiments with $\mathrm{SNR} = 25$ dB. The big crosses correspond to the real values of the eigenvalues of the matrices A_i.

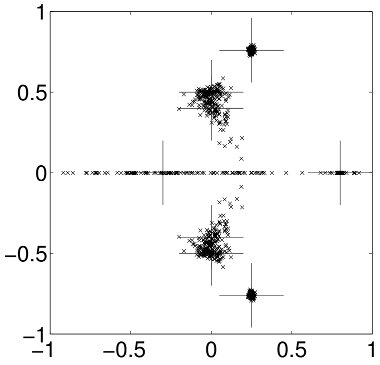

Fig. 5. Eigenvalues of the estimated A_i matrices in the complex plane, for 100 experiments with $\mathrm{SNR} = 20$ dB. The big crosses correspond to the real values of the eigenvalues of the matrices A_i.

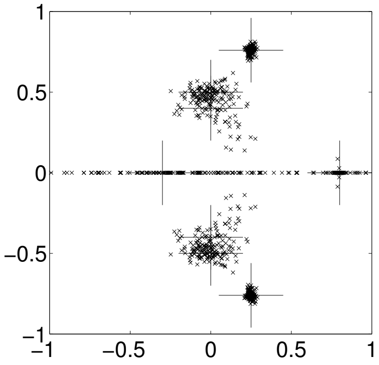

Fig. 6. Eigenvalues of the estimated A_i matrices in the complex plane, for 100 experiments with $\mathrm{SNR} = 15$ dB. The big crosses correspond to the real values of the eigenvalues of the matrices A_i.

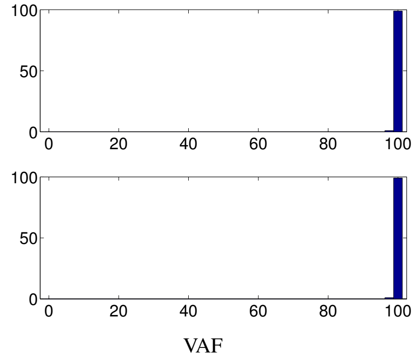

Fig. 7. Histogram of VAF values (%) obtained for validation data with the models estimated using a data set with $\mathrm{SNR} = 25$ dB. The range of VAF values from 0 to 100% is divided into bins of 5%. For each bin, it is shown how many data sets out of the total 100 resulted in VAF values that fall into that bin.

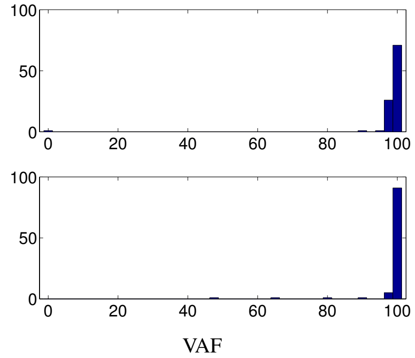

Fig. 8. Histogram of VAF values (%) obtained for validation data with the models estimated using a data set with $\mathrm{SNR} = 20$ dB. The range of VAF values from 0 to 100% is divided into bins of 5%. For each bin, it is shown how many data sets out of the total 100 resulted in VAF values that fall into that bin.

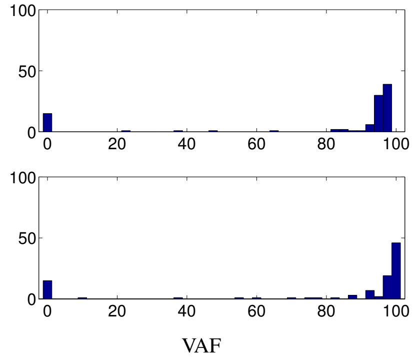

Fig. 9. Histogram of VAF values (%) obtained for validation data with the models estimated using a data set with $\mathrm{SNR} = 15$ dB. The range of VAF values from 0 to 100% is divided into bins of 5%. For each bin, it is shown how many data sets out of the total 100 resulted in VAF values that fall into that bin.

[10] A. Bemporad, A. Garulli, S. Paoletti, and A. Vicino, "A greedy approach to identification of piecewise affine models," in *Hybrid Systems: Computation and Control*, ser. Lecture Notes in Computer Science, O. Maler and A. Pnueli, Eds. Springer-Verlag, 2003, no. 2623, pp. 97–112.

[11] ——, "Set membership identification of piecewise affine models," in *Preprints of the IFAC Symposium on System Identification (SYSID)*, Rotterdam, The Netherlands, Aug. 2003, pp. 737–742.

[12] S. A. Billings and W. S. F. Voon, "Piecewise linear identification of non-linear systems," *International Journal of Control*, vol. 46, no. 1, pp. 215–235, 1987.

[13] G. Ferrari-Trecate, M. Muselli, D. Liberati, and M. Morari, "A clustering technique for the identification of piecewise affine systems," *Automatica*, vol. 39, no. 2, pp. 205–217, Feb. 2003.

[14] J. Ragot, G. Mourot, and D. Maquin, "Parameter estimation of switching piecewise linear system," in *Proceedings of the 42nd IEEE Conference on Decision and Control*, Maui, Hawaii, Dec. 2003, pp. 5783–5788.

[15] J. Roll, A. Bemporad, and L. Ljung, "Identification of piecewise affine systems via mixed-integer programming," *Automatica*, vol. 40, no. 1, pp. 37–50, Jan. 2004.

[16] R. Vidal, S. Soatta, Y. Ma, and S. Sastry, "An algebraic geometric approach to the identification of a class of linear hybrid systems," in *Proceedings of the 42nd IEEE Conference on Decision and Control*, Maui, Hawaii, Dec. 2003, pp. 167–172.

[17] A. Chiuso, R. Vidal, and S. Soatto, "Observability and identifiability of jump linear systems," in *Proceedings of the 41st IEEE Conference on Decision and Control*, Las Vegas, Nevada, Dec. 2002, pp. 3614–3619.

[18] P. Van Overschee and B. De Moor, *Subspace Identification for Linear Systems; Theory, Implementation, Applications*. Dordrecht, The Netherlands: Kluwer Academic Publishers, 1996.

[19] M. Verhaegen, "Identification of the deterministic part of MIMO state space models given in innovations form from input-output data," *Automatica*, vol. 30, no. 1, pp. 61–74, 1994.

[20] V. Verdult and M. Verhaegen, "Subspace identification of multivariable linear parameter-varying systems," *Automatica*, vol. 38, no. 5, pp. 805–814, 2002.

[21] V. Verdult, "Nonlinear system identification: A state-space approach," Ph.D. dissertation, University of Twente, Faculty of Applied Physics, Enschede, The Netherlands, 2002.

[22] M. Verhaegen, "Subspace model identification part 3: Analysis of the ordinary output-error state-space model identification algorithm," *International Journal of Control*, vol. 56, no. 3, pp. 555–586, 1993.

[23] S. Van Huffel and J. Vandewalle, *The Total Least Squares Problem: Computational Aspects and Analysis*, ser. Frontiers in Applied Mathematics. Philadelphia, Pennsylvania: SIAM, 1991, vol. 9.