An alternative approach to switching hyperplane estimation in PWA systems

Stefan Kersting and Martin Buss

Abstract— **Usually the identification of piecewise affine systems consists of two steps. First, subsystem parameters are identified before the state space partition is reconstructed. For the later step, it is common to label the states of a recorded trajectory and separate differently labeled states with the help of linear support vector machines. In this paper, we present an alternative approach to estimate switching hyperplanes in continuous time piecewise affine systems. The presented approach builds upon three components. First, a hybrid observer reconstructs the switching signal up to a bounded delay. Second, we remove the delay in the reconstructed switching signal by solving an optimization problem. We obtain more precise switching states that lie on a specific switching hyperplane. Third, the switching hyperplanes are fitted to the switching states with Total Least Squares. The paper illustrates the presented approach on a numerical example.**

I. INTRODUCTION

Many engineered and biological systems exhibit hybrid and nonlinear phenomena. The hybrid phenomena are frequently due to different modes of operation that the system undergoes. Other causes for hybrid behavior are physical limits, dead-zones or switching at thresholds. A popular way to deal with hybrid and nonlinear systems are piecewise affine (PWA) models, because they are known to be equivalent to certain classes of hybrid systems [1]. Furthermore, PWA models have an universal approximation property and can thus be used to approximate nonlinear systems. Therefore, research on the identification of PWA systems has gained a considerable amount of interest over the past decade. An extensive summary of this research effort is given in [2].

A PWA model is obtained by partitioning the state-input space into regions and assigning each region local, linear dynamics. Both subsystem dynamics and regions must be identified based on measurements that do not contain information about which subsystem generated the data. Most of the proposed approaches carry out the identification in two steps. First, the parameters of the linear subsystems are estimated either simultaneously or iteratively. Once the subsystem dynamics are known, the partitioning of the state-input space is established with standard, linear support vector machines (SVMs) [3].

The reconstruction of polyhedral regions with SVMs is successful in the case of piecewise autoregressive exogenous

(PWARX) systems with fixed regions. In that case, it is easy to correctly label regression vectors by finding the subsystem that predicts best the next measurement. Also, fixed regions allow us to collect a long time series of data-label pairs, which can be fed into the SVM in batch mode. For continuous-time PWA systems in state space form with time varying switching boundaries, however, the use of the standard SVM approach is not straightforward. First, the models only provide state derivatives which can generally not be measured. Therefore, observers must be used to determine the best fitting model. The observers introduce a certain delay and thus cause erroneous labels after each switch. Second, moving switching boundaries can cause inconsistent labels when the SVM categorization is carried out in batch mode.

In this paper, we propose an algorithm to recursively estimate and track time varying regions of continuous-time PWA systems in state space form. With the proposed algorithm it is possible to identify the switching mechanism of a class of hybrid systems. This information can be used to improve the tracking performance of observers for hybrid systems by predicting switches rather than reacting to switches. Also, tracking switching boundaries can be an enabling factor for adaptive control of PWA systems, which is so far restricted to fixed and known switching boundaries (e.g. [4]).

The proposed algorithm consists of three major steps. First, an online switching signal estimator [5] estimates the switching time and state up to a certain level of precision. At the same time, the applied observer records parts of the trajectory and determines the active models before and after the actual switch. The recorded trajectory thus contains only a single, not very precisely estimated switch. In the second step, an optimization problem is set up to find the actual switching state with greater precision. Repeating the first two steps for each switch leads to point clouds of switching states. Each switching state is thereby associated with two neighboring regions. Finally, the switching hyperplanes are fitted to these point clouds by Total Least Squares.

The remainder of this paper is structured as follows. Section II gives an introduction to PWA systems in state space form and formulates the problem. In section III, we present the new approach to hyperplane estimation consisting of three steps. We validate the proposed algorithm in section IV. Section V concludes the paper.

II. INTRODUCTION TO PWA SYSTEMS

We consider PWA systems in state space form with state $x \in \mathbb{R}^n$, control input $u \in \mathbb{R}^p$ and output $y \in \mathbb{R}^q$. In PWA systems, the state-input space $[x^T \ u^T]^T \in \mathbb{R}^{n+p}$ is

S. Kersting and M. Buss are with the Chair of Automatic Control Engineering and with the TUM Institute for Advanced Study, Technische Universität München, Lichtenbergstrasse 2a, 85748 Garching, Germany {stefan.kersting, mb}@tum.de

partitioned into $s \in \mathbb{N}^+$ polyhedral regions Ω_i. In order to obtain a well defined system, the partition is assumed to be complete, i.e. $\cup_{i=1}^s \Omega_i = \mathbb{R}^{n+p}$ and $\Omega_i \cap \Omega_j = \emptyset, \forall i \neq j$. A local, linear model with state matrix $A_i \in \mathbb{R}^{n \times n}$, input matrix $B_i \in \mathbb{R}^{n \times p}$, affine input vector $f_i \in \mathbb{R}^n$ and output matrix $C_i \in \mathbb{R}^{q \times n}$ is assigned to each region Ω_i. The piecewise constant, right continuous switching signal $\sigma(t) \in \mathcal{N} := \{1, \ldots, s\}$ indicates the currently active subsystem. The state-input vector and the regions Ω_i determine the active subsystem by

$$\sigma(t) = i \quad \text{if} \quad \begin{bmatrix} x(t) \\ u(t) \end{bmatrix} \in \Omega_i, \quad i \in \mathcal{N}.$$

With measurement noise $\nu \in \mathbb{R}^q$, the considered PWA systems in state space form are given by the switched system

$$\begin{aligned} \dot{x}(t) &= A_{\sigma(t)} x(t) + B_{\sigma(t)} u(t) + f_{\sigma(t)} \\ y(t) &= C_{\sigma(t)} x(t) + \nu(t). \end{aligned} \quad (1)$$

The borders of the polyhedral regions Ω_i are usually given in terms of hyperplanes in the state-input space. The points on a single hyperplane fulfill the condition

$$\sum_{k=1}^n \alpha_k x_k + \sum_{k=1}^p \beta_k u_k + \gamma = h \begin{bmatrix} x \\ u \\ 1 \end{bmatrix} = 0, \quad (2)$$

where $h = \begin{bmatrix} \alpha_1, \ldots, \alpha_n, \ \beta_1, \ldots, \beta_p, \ \gamma \end{bmatrix} \in \mathbb{R}^{1 \times (n+p+1)}$. If we replace the equality in (2) with inequalities ≤ 0 or < 0, equation (2) defines a half-space of $\mathbb{R}^{n+p}$ including or not including the hyperplane, respectively. In order to characterize the polyhedral region Ω_i, we define a matrix $\mathcal{H}_i \in \mathbb{R}^{\mu_i \times (n+p+1)}$ that combines $\mu_i \in \mathbb{N}^+$ such hyperplanes in its rows: $\mathcal{H}_i = \begin{bmatrix} h_{1,i}^T & \ldots & h_{\mu_i,i}^T \end{bmatrix}^T$. We also define a list $\preceq_{[i]}$ of μ_i operators $\leq$ and $<$ to decide which hyperplanes of $\mathcal{H}_i$ belong to Ω_i. For the polyhedral regions we thus have:

$$\Omega_i = \left\{ \begin{bmatrix} x \\ u \end{bmatrix} \in \mathbb{R}^{n+p} \middle| \mathcal{H}_i \begin{bmatrix} x \\ u \\ 1 \end{bmatrix} \preceq_{[i]} 0 \right\}. \quad (3)$$

A common problem in the identification of PWA systems is to estimate or reconstruct the polyhedral regions Ω_i once the subsystem dynamics are known [2]. With algorithms similar to the ones presented in [6] and [7], the subsystem dynamics of a switched system can be recursively determined; independently of the switching signal. The general problem discussed in this paper thus reads as follows.

Problem 1: Given the subsystem dynamics A_i, B_i, f_i and C_i, find the regions characterized by $\mathcal{H}_i$, $i \in \mathcal{N}$.

The most frequently used approach to Problem 1 is to generate labels for each measured state-input vector. The labeled state-input vectors are subsequently separated with standard SVMs, which yield the switching hyperplanes. In the discrete time case, labels are easily generated by predicting the next measurement for each subsystem and finding the subsystem that predicts best the next measurement. To apply the same intuitive approach in continuous time, we would design an observer for each subsystem to estimate the measured output:

$$\dot{\hat{x}}_i = (A_i - L_i C_i) \hat{x}_i + L_i y + B_i u + f_i \quad (4)$$

where $L_i \in \mathbb{R}^{n \times q}$ are the observer gains to ensure that the estimate $\hat{x}_i$ remains bounded; especially if x is not generated by the i-th subsystem. The labels would then be generated by

$$\hat{\sigma}(x(t), u(t)) = \arg \min_i |y(t) - C_i \hat{x}_i(t)|^2. \quad (5)$$

Due to the limited convergence rate of the observers, the obtained labels do not coincide well with the true switching signal. We thus propose a novel approach to identify the polyhedral regions more precisely. Furthermore, our algorithm works recursively and can thus also track time-varying polyhedral regions. The following problem formulation describes this more challenging task.

Problem 2: Given the subsystem dynamics A_i, B_i, f_i and C_i, find and track the time varying regions characterized by $\mathcal{H}_i(t)$.

III. Hyperplane Estimation

The solution to Problem 2 consists of three steps. In the first step, a hybrid observer estimates the switching signal online. Due to the inherently reactive nature of observers to estimation errors, the estimated switching signal obtained in the first step is delayed compared with the true switching signal. However, the applied observer allows to record parts of the output trajectories before and after the switch, which can subsequently help to improve the estimated switching time. In the second step, we define and solve an optimization problem based on the recorded trajectory, which gives us a more precise estimate of the switching time and state. After collecting several switching states, the third step fits a hyperplane to these states. The following three subsections describe the three steps in greater detail.

A. Hybrid switching state observers

The first step towards estimating the polyhedral regions of PWA systems is the reconstruction of the unknown switching signal $\sigma(t)$ as well as the state $x(t)$ based on measurements of the noisy output $y(t)$. This alone constitutes a challenging task and has attracted a considerable amount of research on observability of hybrid and switched systems (e.g. [8]–[10]) and put forth various switching signal estimators and observers for hybrid and switched systems.

In [10], the switching signal is assumed to be known and the challenge is to estimate the states of a switched system with state jumps even if some of the subsystems are not observable. In [11] and [12], the inputs and outputs of the switched system are numerically differentiated to generate residuals. Only the residual which belongs to the active subsystem is zero. Another approach to estimate the switching times of an autonomous, switched system online is presented in [13] and [14] and based on algebraic tools and distribution theory, which is needed to handle derivatives of non smooth dynamics. The observer for hybrid systems presented in [15] consists, as most hybrid observers, of two parts. The first part estimates the current mode of the system and the second part estimates the continuous states of the system.

We found the observer for switched linear systems presented in [5] to be the most suitable for our endeavor, because it allows for a straightforward recording of input and output data before and after a switch. The estimation algorithm in [5] is based on the idea of a minimum distance criterion [16]. With the proposed algorithm, we estimate switches of the PWA system in two stages. Note that the overall hyperplane estimation consists of three steps, whereas the online switching signal estimator has two stages. The first stage of the algorithm applies at the beginning or whenever a switch occurs. This stage is called **estimation** and has a minimum duration of δ, in which the current mode of the PWA system is correctly determined and the state is estimated up to a certain noise level. Once the current mode is correctly estimated, the algorithm switches to its second stage: The **detection** stage aims to detect a switch within a maximum delay Δ. After detecting a switch, the algorithm carries on with the estimation stage and so forth.

We now recap the two stages estimation and detection of the online switching signal estimator in sufficient detail for implementation. For the rigorous proofs of the algorithm, we refer the reader to [5]. Three assumptions arise for the algorithm to operate in the anticipated way. For a better readability, we only discuss here the practical implications of the three assumptions and list the detailed assumptions in the appendix. The distinguishability of any two subsystems is guaranteed by Assumption 1. Assumption 2 gives constraints for the variables δ, Δ, $\mathcal{X}_\Delta$ and $\mathcal{S}_\Delta$ and highlights their relationship to the input u and measurement noise ν. Furthermore, they depend on a quantity that expresses the maximum distance between two subsystem parameterizations. For a more detailed treatment of distance measures between dynamical systems we refer to [17]. A minimum dwell time of length $\delta + \Delta$ is demanded by Assumption 3.

Estimation of the active mode: The first stage of the algorithm is active when the algorithm starts and whenever a switch has been detected. Assume that we have just detected the j-th switch at time t_j, then we set the estimated switching time $\hat{t}_j = t$. At the initialization of the algorithm we have $\hat{t}_0 = 0$. While the PWA system (1) evolves over the time-interval $[\hat{t}_j, \hat{t}_j + \delta]$, we integrate in real-time the following differential equations with initial conditions $\breve{x}_i(\hat{t}_j) = 0$ and $z_i(\hat{t}_j) = 0$ for all subsystems $i \in \mathcal{N}$:

$$\dot{\breve{x}}_i = A_i \breve{x}_i + B_i u + f_i \tag{6}$$

$$\dot{z}_i = -A_i^T z_i + C_i^T (y - C_i \breve{x}_i). \tag{7}$$

At the same time, we record $u(\tau)$ and $y(\tau)$, $\tau \in [\hat{t}_j, \hat{t}_j + \delta]$. Note here that we use the ˇ symbol ($\breve{x}$) to characterize an open loop observer in contrast to the ˆ symbol ($\hat{x}$) for the standard Luenberg observer in (4).

With the states $z_i(\hat{t}_j + \delta)$, obtained in (7), we calculate better initial conditions X_i^* for $\breve{x}_i$ at time instance $\hat{t}_j$:

$$X_i^* = U_i^{-1}(\delta) e^{A_i^T \delta} z_i(\hat{t}_j + \delta), \tag{8}$$

where $U_i(t) := \int_0^t \phi_i^T(\tau)\phi_i(\tau)d\tau$ is the observability Gramian for mode i and $\phi_i(t) := C_i e^{A_i t}$. With the improved initial conditions $\breve{x}_i(\hat{t}_j) = X_i^*$ and $s_i(\hat{t}_j) = 0$ and the stored

values of u and y, we integrate the following differential equations once more over the interval $[\hat{t}_j, \hat{t}_j + \delta]$ (for convenience, this is assumed to happen instantaneously at $\hat{t}_j$):

$$\dot{\breve{x}}_i = A_i \breve{x}_i + B_i u + f_i \tag{9}$$

$$\dot{s}_i = |y - C_i \breve{x}_i|^2. \tag{10}$$

The integral of (10) basically sums up the difference between the measured outputs y and the estimated outputs of the open loop observer (9). We evaluate at time $\hat{t}_j + \delta$ the costs $J_i(X_i^*) = s_i(\hat{t}_j + \delta)$. The currently active mode is the one with minimum cost:

$$\hat{\sigma}_j := \hat{\sigma}(\hat{t}_j + \delta) = \arg \min_{i \in \mathcal{N}} J_i(X_i^*). \tag{11}$$

We also reset the state of another Luenberg observer to the corresponding best estimate $\hat{x}(\hat{t}_j) = X_{\hat{\sigma}_j}^*$. If Assumptions 1 and 2 hold, one can show that the estimated mode $\hat{\sigma}_j$ is correct and that the state estimate $\hat{x}(\hat{t}_j)$ is equal to the true state $x(\hat{t}_j)$ up to certain noise levels [5, Theorem 1]:

$$\hat{x}(\hat{t}_j) = x(\hat{t}_j) + U_{\hat{\sigma}_j}^{-1}(\delta) \int_{\hat{t}_j}^{\hat{t}_j + \delta} \phi_{\hat{\sigma}_j}^T (\tau - \hat{t}_j) \nu(\tau) d\tau. \tag{12}$$

Detection of mode switches: During this stage, we estimate the state of the system with two observers, both parametrized for the $\hat{\sigma}_j$-th subsystem. One of the observers is an open loop observer ($\breve{x}$), while the other one is a standard Luenberg observer ($\hat{x}$). With the recorded data of u and y, we catch up from $\hat{t}_j$ to the current time $\hat{t}_j + \delta$ by integrating

$$\dot{\breve{x}} = A_{\hat{\sigma}_j} \breve{x} + B_{\hat{\sigma}_j} u + f_{\hat{\sigma}_j} \tag{13}$$

$$\dot{\hat{x}} = (A_{\hat{\sigma}_j} - L_{\hat{\sigma}_j} C_{\hat{\sigma}_j})\hat{x} + B_{\hat{\sigma}_j} u + f_{\hat{\sigma}_j} + L_{\hat{\sigma}_j} y \tag{14}$$

$$\dot{s} = |y - C_{\hat{\sigma}_j} \breve{x}|^2 \tag{15}$$

with initial conditions $\breve{x}(\hat{t}_j) = \hat{x}(\hat{t}_j) = X_{\hat{\sigma}_j}^*$ (as in (8)) and $s(\hat{t}_j) = 0$. After catching up to $t = \hat{t}_j + \delta$, we integrate (13) - (15) over time frames of length Δ in real-time. For that, we reset the state of the open loop observer $\breve{x}(t) = \hat{x}(t)$ and the starting time of the Δ-frame $t_{\text{up}} = t$. After each Δ-frame ($t - t_{\text{up}} > \Delta$), we again reset the open-loop observer state $\breve{x}(t) = \hat{x}(t)$ and the starting time $t_{\text{up}} = t$. We detect a switch when one of the following two conditions is fulfilled:

$$|\breve{x}(t) - \hat{x}(t)| > \mathcal{X}_\Delta, \qquad s(t) - s(t - \Delta) > \mathcal{S}_\Delta, \tag{16}$$

where the thresholds $\mathcal{X}_\Delta$ and $\mathcal{S}_\Delta$ are given by Assumption 2 in the appendix. If Assumptions 1 to 3 hold, the detection algorithm is guaranteed to find an estimate $\hat{t}_{j+1}$ of the next switch at t_{j+1} within the interval $[t_{j+1}, t_{j+1} + \Delta]$ [5, Theorem 2]. After detecting a switch, the algorithm returns to the estimation stage.

The estimated switching time $\hat{t}_{j+1}$ has a maximum delay of Δ. To eliminate this delay, we propose to record u and y over each time frame Δ as well as the value of $\hat{x}$ at the beginning of each interval. As the memory of control systems is usually limited and the PWA system might not switch its mode for a long time, it suffices to store only the last two recorded intervals. Together with the recorded values during the estimation stage, we collect three connected parts of the system trajectory: the first element of length Δ before

the switch, the second element also of length Δ containing the estimated switching time $\hat{t}_{j+1}$ and the third element of length δ after the estimated switching time. Figure 1 sketches the recorded output trajectory and highlights the elements of length Δ and δ. We use these recordings in the next step of the proposed hyperplane estimation algorithm to estimate the switching state more precisely.

B. Optimized switching states

We now formulate an optimization problem to find a more precise estimate of the true switching time and state. There exist various approaches for optimal control of switching systems (see [18]–[20] and references therein). They usually aim to determine the optimal sequence and time instances in which a given set of system dynamics needs to be activated in order to minimize a certain cost function. Our optimization problem is easier as we know the switching sequence and only need to optimize for a single switching time.

We have records of the inputs u and outputs y of the PWA system over a time interval $[t_0, t_e]$ (see Fig. 1). Furthermore, we know that the system switches its mode during this period from $\hat{\sigma}_j$ to $\hat{\sigma}_{j+1}$. Finally, we have the estimated state $\hat{x}(t_0)$. The output trajectory for a specific switching time τ is $\check{y}(t) = C_{\hat{\sigma}_j}\check{x}(t)$ if $t \leq \tau$ or $\check{y}(t) = C_{\hat{\sigma}_{j+1}}\check{x}(t)$ if $t > \tau$ where

$$\dot{\check{x}}(t) = \begin{cases} A_{\hat{\sigma}_j}\check{x}(t) + B_{\hat{\sigma}_j}u(t) + f_{\hat{\sigma}_j} & \text{,if } t \leq \tau \\ A_{\hat{\sigma}_{j+1}}\check{x}(t) + B_{\hat{\sigma}_{j+1}}u(t) + f_{\hat{\sigma}_{j+1}} & \text{,if } t > \tau \end{cases}$$

and $\check{x}(t_0) = \hat{x}(t_0)$. The squared error between the estimated output trajectory $\check{y}$ and the recorded output trajectory y serves as a cost function. This cost function solely depends on the parameter τ and must be minimized to find the optimal switching time $\bar{t}_{j+1}$:

$$\bar{t}_{j+1} = \arg \min_{\substack{\tau \\ \text{s.t. } t_0 \leq \tau \leq t_e}} \int_{t_0}^{t_e} |y(t) - \check{y}(t)|^2 dt. \quad (17)$$

No real-time constraints arise for this constrained optimization problem as the estimation and detection algorithm in III-A does not rely on the solution. Standard optimization techniques such as golden section search and parabolic interpolation are suitable for the optimization problem.

The optimal switching time $\bar{t}_{j+1}$ is used to calculate a more precise estimate of the switching state $\hat{x}(\bar{t}_{j+1})$ and input $u(\bar{t}_{j+1})$, which are used in the subsequent step to estimate the switching hyperplanes.

C. Hyperplane fitting with Total Least Squares

By performing the previous two steps over multiple switches, we generate sets of labeled switching states. The
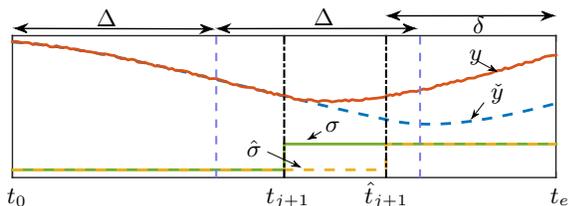

Fig. 1. Sketch of the recorded time frame containing the actual switching time t_{j+1} as well as the time instance $\hat{t}_{j+1}$ when the switch was detected.

label uniquely specifies the switching hyperplane that the state belongs to. Assume that we detected a switch from region k to region l in the first step of the algorithm. Hence, we know that there must exist a switching hyperplane that separates Ω_k from Ω_l and the estimated switching state lies on this hyperplane. After we collected multiple switching states for switches from region k to l and vice versa, all that is left do is to fit a hyperplane to this set of points.

With the hyperplane equation (2) and m collected switching states, fitting a hyperplane to the data points can be interpreted as finding the solution to an overdetermined system of equations $\mathcal{A}X \approx \mathcal{B}$. Here $\mathcal{A} \in \mathbb{R}^{m \times (n+p)}$ and $\mathcal{B} \in \mathbb{R}^m$ consist of given data and $X \in \mathbb{R}^{n+p}$ contains the unknown hyperplane parameters. We typically have more data points than unknowns ($m > n + p + 1$) and must assume that $\mathcal{A}$ and $\mathcal{B}$ contain uncertainties. Therefore, an exact solution to the overdetermined system does not exist. With Total Least Squares however it is possible to find an approximate solution by correcting the system of equations with slack variables $\mathcal{E}_\mathcal{A}$ and $\mathcal{E}_\mathcal{B}$. This leads to the optimization problem

$$\min_{\substack{X, \mathcal{E}_\mathcal{A}, \mathcal{E}_\mathcal{B} \\ \text{s.t. } (\mathcal{A} + \mathcal{E}_\mathcal{A})X = \mathcal{B} + \mathcal{E}_\mathcal{B}}} \left\| \begin{bmatrix} \mathcal{E}_\mathcal{A} & \mathcal{E}_\mathcal{B} \end{bmatrix} \right\|_F \quad (18)$$

where $\| \cdot \|_F$ is the Frobenius norm. An overview of the history of Total Least Squares as well as some recent advances in the field are given in [21].

For our case, we set $\gamma = 1$ and let X contain the remaining parameters of the hyperplane that are to be found, i.e. $X = [\alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_p]^T$. Each row of $\mathcal{A}$ corresponds to one of the obtained switching states (e.g. $[\hat{x}^T(\bar{t}_j), u^T(\bar{t}_j)]$) and $\mathcal{B}$ is a column vector with all elements equal to -1.

In order to find the solution to the Total Least Squares problem (18), we compute the singular value decomposition $\begin{bmatrix} \mathcal{A} & \mathcal{B} \end{bmatrix} = U\Sigma V^T$ and obtain the block matrix V with $V_{11} \in \mathbb{R}^{(n+p) \times (n+p)}$, $V_{12} \in \mathbb{R}^{(n+p) \times 1}$, $V_{21} \in \mathbb{R}^{1 \times (n+p)}$ and $V_{22} \in \mathbb{R}$. A solution to (18) exists if V_{22} is nonzero and the optimal parameter vector $\hat{X}$ is given by $\hat{X} = -V_{12}V_{22}^{-1}$ [21, Algorithm 1]. In turn, the hyperplane is $\hat{h} = [\hat{X}^T, 1]$.

Since any two switches are assumed to be separated by a dwell time, we also assume that the time between two switches suffices to recompute the singular value decomposition. If the dwell time is too short or if the computational cost is too large, one could resort to a recursive implementation of Total Least Squares (see [22] and references therein).

IV. NUMERICAL VALIDATION

We now validate our algorithm with an exemplary PWA system. While the algorithm was also successfully applied to more complex systems, note that greater system complexity might make it impossible to fulfill the required assumptions. We consider here the PWA system discussed in [23] and [24] with the following three subsystem dynamics:

$$A_1 = \begin{bmatrix} 0 & 1 \\ -2 & -1 \end{bmatrix} \quad A_2 = \begin{bmatrix} 0 & 1 \\ -2.5 & -1 \end{bmatrix} \quad A_3 = \begin{bmatrix} 0 & 1 \\ -1.5 & -1 \end{bmatrix}$$

$$f_1 = \begin{bmatrix} 0 & 0.2 \end{bmatrix}^T \quad f_2 = \begin{bmatrix} 0 & 0.4 \end{bmatrix}^T \quad f_3 = \begin{bmatrix} 0 & -0.3 \end{bmatrix}^T$$

and $B = \begin{bmatrix} 0 & 1.5 \end{bmatrix}^T$ and $C = \begin{bmatrix} 1 & 0 \end{bmatrix}$. The hyperplane h_{12} separates subsystem 1 and 2, while h_{13} separates subsystem 1 and 3. For the first $250\,\mathrm{s}$, the hyperplanes are constant with $h_{12,\mathrm{I}} = \begin{bmatrix} -1 & 0 & 0 & 2 \end{bmatrix}$ and $h_{13,\mathrm{I}} = \begin{bmatrix} 1 & 0 & 0 & 2 \end{bmatrix}$. The PWA system is excited by a sinusoidal input signal with varying amplitude: $u(t) = (9 + 3\cos(0.015t))\sin(0.1t)$. The output $y(t)$ is affected by additive white noise with zero mean and $\nu_{\max} = 0.1$. For the online estimation of the switching signal $\hat{\sigma}$ according to section III-A, the observer gains L_i are chosen such that the poles of the Luenberg observers $A_i - L_i C_i$ lie at -3 and -3.5. This gives sufficiently fast convergence without amplifying the measurement noise. The remaining design parameters are $\delta = 0.8$, $\Delta = 0.8$, $\mathcal{S}_\Delta = 0.003$ and $\mathcal{X}_\Delta = 0.04$.

Figure 2 shows the signals estimated with the proposed algorithm during the first $250\,\mathrm{s}$. The dash-dotted trajectory represents the true state of the PWA system. The solid, colored trajectory shows the estimated state obtained with the hybrid state observer. The hybrid state observer also delivers an estimate of the switching signal $\hat{\sigma}(t)$, which is included in the figure through color coding. The red (middle), green (right) and blue (left) parts of the trajectory belong to the subsystems 1, 2 and 3, respectively. Note that the state estimation is precise as long as the estimated switching signal is correct. Whenever the PWA system switches at $x_1 = \pm 2$, the hybrid observer requires an additional time interval of maximum length $\delta + \Delta$ to detect the switch and estimate the current mode. During this transient phase, the estimated trajectory deviates from the true trajectory. Such deviations cannot be avoided in hybrid observers with switching signal reconstruction.

A linear SVM is trained with the colored/labeled trajectory elements. The hyperplanes obtained with this traditional approach are visualized by dashed lines in Fig. 2. Note that the SVM hyperplanes deviate from the true switching hyperplanes by a few degrees. This is due to the delayed detection of switches. As proposed in section III-B, we find a more precise estimate of the switching time by solving the simple optimization problem (17). We use the Matlab function `fminbnd` which on average solved the problem within $443.6\,\mathrm{ms}$ on a quad core (i5-2500K,3.30Ghz). Figure 3 shows how the optimization step improves the estimated switching time from $\hat{t}_j$ (obtained online) to $\bar{t}_j$. Circles and triangles highlight the optimized switching states in Fig. 2.
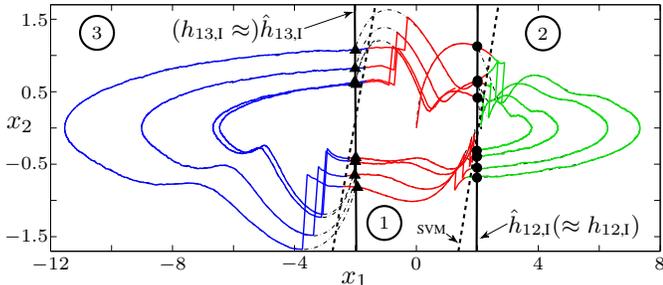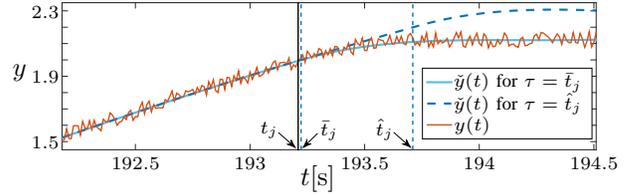

Fig. 3. Improving the estimated switching time from $\hat{t}_j$ to $\bar{t}_j$ by minimizing the squared error between measured output $y(t)$ and estimated output $\tilde{y}(t)$.

In the final step, presented in section III-C, the estimated hyperplanes $\hat{h}_{12}$ and $\hat{h}_{13}$ are fitted to the optimized switching states by Total Least Squares. We obtain

$$\hat{h}_{12,\mathrm{I}} = \begin{bmatrix} -1.0024 & 0.0107 & 0.0000 & 2.0000 \end{bmatrix},$$
$$\hat{h}_{13,\mathrm{I}} = \begin{bmatrix} 1.0023 & 0.0149 & 0.0000 & 2.0000 \end{bmatrix},$$

which is very close to $h_{12,\mathrm{I}}$ and $h_{13,\mathrm{I}}$. Compared with the SVM-approach (dotted lines), the proposed approach (solid lines) delivers a much better approximation of the true hyperplanes.

After $t = 250\,\mathrm{s}$, we change the hyperplanes to $h_{12,\mathrm{II}} = \begin{bmatrix} -0.2 & 0.2 & 0 & 0.1 \end{bmatrix}$ and $h_{13,\mathrm{II}} = \begin{bmatrix} 1 & 0 & 0 & 3 \end{bmatrix}$ in order to analyze the tracking properties of the presented algorithm. The parameter changes resemble a rotation of h_{12} and a translation of h_{13} and could be caused by aging or wear in a practical system. The tracking ability was gained by calculating the Total Least Squares solution only for the last $m = 6$ switching states per hyperplane.

Figure 4 shows how the algorithm tracks the changing hyperplanes. The estimated state space trajectory from $t = 250\,\mathrm{s}$ to $t = 500\,\mathrm{s}$ is given by a colored, solid trajectory. The intensity of the trajectory increases with time. The same applies for the estimated switching states and hyperplanes. The switching states for $t < 250\,\mathrm{s}$ are given by gray circles and triangles, while the estimated hyperplanes at $t = 250\,\mathrm{s}$ are shown as dashed, gray lines. The intensity of the estimated switching states then shifts from gray to black as time goes on. The figure also shows the time varying estimates $\hat{h}_{12}(t)$ and $\hat{h}_{13}(t)$ from $t = 250\,\mathrm{s}$ to $t = 500\,\mathrm{s}$. Arrows indicate the movement of the estimates. Intermediate estimates are shown by lines of varying width and intensity. The thinner and lighter a line, the older is the corresponding estimate of the switching hyperplane. The final estimates are again very close to the true parameters $h_{12,\mathrm{II}}$ and $h_{13,\mathrm{II}}$:

$$\hat{h}_{12,\mathrm{II}}(t=500) = \begin{bmatrix} -0.2020 & 0.2008 & 0.0000 & 0.1000 \end{bmatrix},$$
$$\hat{h}_{13,\mathrm{II}}(t=500) = \begin{bmatrix} 0.9974 & 0.0045 & 0.0000 & 3.0000 \end{bmatrix}.$$

Fig. 2. State space trajectory of the real system (dash-dotted) and the online-switching signal observer (colored) for $250\,\mathrm{s}$. Circles and triangles mark the estimated switches $1 \leftrightarrow 2$ and $1 \leftrightarrow 3$ obtained by optimization.
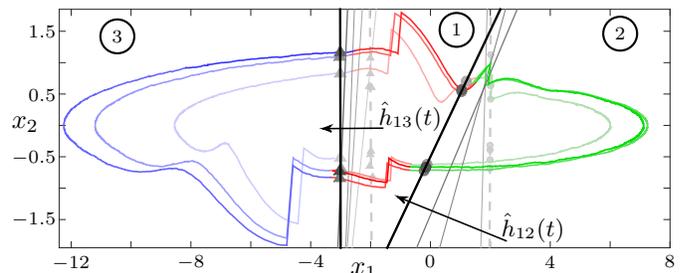

Fig. 4. Tracking time-varying hyperplanes with the proposed algorithm.

V. Conclusion

We present an alternative approach to estimate switching hyperplanes of continuous-time PWA systems in state space form. We show that the existing SVM approach delivers shifted estimates in the continuous-time case. The proposed algorithm overcomes this limitation in three steps. First, a hybrid state observer obtains rough estimates of the switching signal and the full state of the PWA system. After every detected switch, the accuracy of the estimated switching time is improved by solving a simple optimization problem. A switching hyperplane is then fitted to multiple switching states by Total Least Squares. In comparison to the SVM approach, our algorithm delivers much more accurate estimates. Furthermore, our algorithm easily tracks parametric changes of time-varying switching hyperplanes, which is not possible in the SVM case. Guidlines for finding feasible design parameters are subject to future work.

Appendix

We repeat and discuss here the assumptions for the hybrid observer in section II-A, which were originally presented in [5]. A general premise for mode detection in switched systems is related to joint observability.

Assumption 1: The joint observability Gramian

$$W_{i,j}(t) := \int_0^t \begin{bmatrix} \phi_i^T(\tau) \\ \phi_j^T(\tau) \end{bmatrix} \begin{bmatrix} \phi_i(\tau) & \phi_j(\tau) \end{bmatrix} d\tau,$$

of two subsystems i and j with $\phi_i(t) := C_i e^{A_i t}$ is nonsingular for any $t \neq 0$ and $i \neq j$. $\diamond$

Let $\omega_{\min}(t) = \min_{i \neq j} \lambda_{\min}(W_{i,j}(t))$ be the minimum eigenvalue of the observability Gramian for all subsystem combinations and a time period t. Furthermore, we introduce the approximations $\check{\mu}, \hat{\mu} \geq 1$ and $\check{\lambda}, \hat{\lambda} \geq 0$ such that $\|e^{A_i t}\| \leq \check{\mu} e^{\check{\lambda} t}$ and $\|e^{(A_i - L_i C_i)t}\| \leq \hat{\mu} e^{-\hat{\lambda} t}$, $\forall t \geq 0, \forall i \in \mathcal{N}$. Finally, let $L_{\max} = \max_{i \in \mathcal{N}} \|L_i\|$ and $C_{\max} = \max_{i \in \mathcal{N}} \|C_i\|$. Then the intervals δ and Δ can be related to the input and noise of the PWA system by the following assumption.

Assumption 2: The signals u and ν are uniformly bounded, i.e. $|u(t)| \leq u_{\max}$, $|\nu(t)| \leq \nu_{\max}, \forall t \geq 0$. Moreover, there exist positive constants δ and Δ, such that

$$\omega_{\min}(\delta)|x(t)|^2 > \left(u_{\max} \sqrt{N_u(\delta)} + 2\nu_{\max} \sqrt{\delta} \right)^2,$$

$$\omega_{\min}(\Delta)|x(t)|^2 > \frac{1}{2} \left(u_{\max} \sqrt{N_u(\Delta)} + 2\nu_{\max} \sqrt{\Delta} + 2\nu_{\max} C_{\max}(\check{\mu} e^{\check{\lambda}\Delta} + 1)^2 \hat{E}(\delta) \sqrt{\Delta} \right)^2,$$

where $h_i(t) := \phi_i(t) B_i$ and $U_i(t) := \int_0^t \phi_i^T(\tau)\phi_i(\tau)d\tau$ and

$$N_u(\delta) := \max_{i,j \in \mathcal{N}, i \neq j} \int_0^\delta \left(\int_0^\tau \|h_i(s) - h_j(s)\| ds \right)^2 d\tau,$$

$$M_{\max}(\delta) := \max_{i \in \mathcal{N}} M_i(\delta) := \max_{i \in \mathcal{N}} \int_0^\delta \|U_i^{-1}(\delta)\phi_i^T(\tau)\| d\tau,$$

$$\hat{E}(\delta) := \max \left\{ M_{\max}(\delta), \frac{L_{\max}}{\hat{\lambda}} \right\}.$$

The threshold $\mathcal{N}_u(\delta)$ resembles a maximum distance between any two subsystem parameterizations. Finally define the following thresholds $\mathcal{X}_\Delta := \nu_{\max}(\check{\mu} e^{\check{\lambda}\Delta} + 1)\hat{E}(\delta)$ and $\mathcal{S}_\Delta := \nu_{\max}^2(1 + C_{\max}(\check{\mu} e^{\check{\lambda}\Delta} + 2)\hat{E}(\delta))^2 \Delta$. $\diamond$

Assumption 3: The switching signal σ has a dwell time of $\Delta + \delta$ and there is no switch in the initial time interval $[0, \delta]$. $\diamond$

References

[1] W. P. M. H. Heemels, B. d. Schutter, and A. Bemporad, "Equivalence of hybrid dynamical models," *Automatica*, vol. 37, pp. 1085–1091, 2001.

[2] A. Garulli, S. Paoletti, and A. Vicino, "A Survey on Switched and Piecewise Affine System Identification," in *System Identification*. IFAC, Elsevier, 2012, pp. 344–355.

[3] S. Paoletti, A. L. Juloski, G. Ferrari-Trecate, and R. Vidal, "Identification of hybrid systems: a tutorial," *European Journal of Control*, vol. 13, no. 2-3, pp. 242–260, 2007.

[4] M. Di Bernardo, U. Montanaro, and S. Santini, "Hybrid model reference adaptive control of piecewise affine systems," *IEEE Trans. Autom. Control*, vol. 58, pp. 304–316, 2013.

[5] C. Lee, Z. Ping, and H. Shim, "On-line Switching Signal Estimation of Switched Linear Systems with Measurement Noise," in *European Control Conference (ECC)*, 2013, pp. 2180–2185.

[6] M. J. Feiler and K. Narendra, "Simultaneous Identification and Control of Time-Varying Systems," in *Proc. of the 45th IEEE CDC*, 2006, pp. 1093–1098.

[7] ——, "Identification and Control Using Multiple Models," in *Proc. of the 14th Yale Workshop on Adaptive and Learning Systems*, 2008.

[8] S. Chaib, D. Boutat, A. Benali, and J. P. Barbot, "Observability of the discrete state for dynamical piecewise hybrid systems," *Nonlinear Anal.: Theory, Methods & Applicat.*, vol. 63, no. 3, pp. 423–438, 2005.

[9] E. d. Santis, M. Di Benedetto, and G. Pola, "On observability and detectability of continuous-time linear switching systems," in *Proc. of the 42nd IEEE CDC*, 2003, pp. 5777–5782.

[10] A. Tanwani, H. Shim, and D. Liberzon, "Observability for Switched Linear Systems: Characterization and Observer Design," *IEEE Transactions on Automatic Control*, vol. 58, no. 4, pp. 891–904, 2013.

[11] M. Fliess, C. Join, and W. Perruquetti, "Real-time estimation for switched linear systems," in *Proc. of the 47th IEEE CDC*, 2008, pp. 941–946.

[12] ——, "Real-time estimation of the switching signal for perturbed switched linear systems," *3rd IFAC Conference on Analysis and Design of Hybrid Systems*, 2009.

[13] Y. Tian, T. Floquet, L. Belkoura, and W. Perruquetti, "Switching time estimation for linear switched systems: an algebraic approach," in *Proc. of the 48th IEEE CDC*, 2009, pp. 3909–3913.

[14] ——, "Algebraic switching time identification for a class of linear hybrid systems," *Nonl. Anal.: Hybrid Syst.*, vol. 5, pp. 233–241, 2011.

[15] A. Balluchi, L. Benvenuti, M. D. Di Benedetto, and A. Sangiovanni-Vincentelli, "The design of dynamical observers for hybrid systems: Theory and application to an automotive control problem," *Automatica*, vol. 49, pp. 915–925, 2013.

[16] G. Battistelli, "On adaptive stabilization of mode-observable switching linear systems," in *Proc. 18th IFAC World Congr.*, 2011, pp. 356–361.

[17] S. Kersting and M. Buss, "Metrics to Compare and Control Dynamical Systems," in *European Control Conference (ECC)*, 2014, pp. 738–743.

[18] H. Axelsson, Y. Wardi, M. Egerstedt, and E. Verriest, "A Provably Convergent Algorithm for Transition-Time Optimization in Switched Systems," in *Proc. of the 44th IEEE CDC*, 2005, pp. 1397–1402.

[19] M. Egerstedt, Y. Wardi, and H. Axelsson, "Transition-time optimization for switched-mode dynamical systems," *IEEE Transactions on Automatic Control*, vol. 51, no. 1, pp. 110–115, 2006.

[20] X. C. Ding, Y. Wardi, and M. Egerstedt, "On-line Optimization of Switched-Mode Dynamical Systems," *IEEE Transactions on Automatic Control*, vol. 54, no. 9, pp. 2266–2271, 2009.

[21] I. Markovsky and S. van Huffel, "Overview of total least-squares methods," *Signal Processing*, vol. 87, no. 10, pp. 2283–2302, 2007.

[22] S. Rhode, F. Bleimund, and F. Gauterin, "Recursive Generalized Total Least Squares with Noise Covariance Estimation," in *Proc. of the 19th IFAC World Congr.*, 2014, pp. 4637–4643.

[23] S. Kersting and M. Buss, "Concurrent learning adaptive identification of piecewise affine systems," in *Proc. of the 53rd IEEE CDC*, 2014, pp. 3930–3935.

[24] ——, "Online identification of piecewise affine systems," in *UKACC 10th International Conference on CONTROL*, 2014, pp. 90–95.