AN ALGEBRAIC GEOMETRIC APPROACH TO THE IDENTIFICATION OF LINEAR HYBRID SYSTEMS

by

René Vidal, Stefano Soatto and Shankar Sastry

Memorandum No. UCB/ERL M03/10

8 March 2003

AN ALGEBRAIC GEOMETRIC APPROACH TO THE IDENTIFICATION OF LINEAR HYBRID SYSTEMS

by

René Vidal, Stefan Soatto and Shankar Sastry

Memorandum No. UCB/ERL M03/10

8 March 2003

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

An Algebraic Geometric Approach to the Identification of Linear Hybrid Systems

René Vidal[1] Stefano Soatto[2] Shankar Sastry[3]

Abstract

We propose an algebraic geometric solution to the identification of linear hybrid systems. Our solution establishes a connection between hybrid systems identification and polynomial factorization: we represent the number of discrete states as the degree of a homogeneous polynomial and the model parameters as roots (factors) of such a polynomial. This interpretation allows us to decouple the identification of the model parameters from the inference of the hybrid state and from the switching mechanism generating the transitions. We first derive a rank constraint on the input/output data from which one can estimate the number of models N. We then estimate the model parameters using a linear algebraic polynomial factorization technique that is closed-form for $N \leq 4$. Once the model parameters have been identified, the estimation of the hybrid state and of the switching parameters becomes a simpler problem. We present simulation experiments that validate the performance of the algorithm.

1 Introduction

Hybrid dynamical models can be used to describe continuous phenomena that exhibit discontinuous behavior due to sudden changes of dynamics. For instance, the continuous trajectory of a bouncing ball results from the alternation between free fall and elastic contact. However, hybrid dynamical models can also be used to approximate a phenomenon that does not itself exhibit discontinuous behavior, by concatenating different models from a simple class. For instance, a non-linear dynamical system can be approximated by switching among various linear dynamical models.

In this paper we look at the problem of modeling input/output data by piecewise linear (hybrid) dynamical models: Given input/output data, we want to simultaneously estimate the number of underlying linear models, the parameters of each model, the hybrid state (continuous and discrete), and possibly the switching mechanism that governs the transitions from one linear model to another.

Work on filtering and identification of hybrid systems first appeared in the seventies; a review of the state of the art as of 1982 can be found in [20]. After a decade-long hiatus, the problem has recently been enjoying considerable interest (see [19, 18, 9, 4, 5, 21] and references therein). Related work has also appeared in the machine learning community (see [8, 17, 16, 6, 7, 13] and references therein). When the model parameters and the switching mechanism are *known*, the identification problem reduces to the design of observers for the hybrid state; [1] considers the case in which the discrete state is further known and proposes a Luenberger observer for the continuous state; [2] combines location observers with Luenberger observers to design a hybrid observer that identifies the discrete location in a finite number of steps and converges exponentially to the continuous state; [12] proposes a moving horizon estimator that, under certain conditions, is asymptotically convergent and can be implemented via mixed-integer quadratic programming. When the model parameters and the switching mechanism are *unknown*, the identification problem becomes even more challenging: [10, 11] assume that the number of models is known, and propose an identification algorithm that combines clustering, regression and classification techniques; [3] also rely on a clustering algorithm that is optimal with respect to deviations from the model.

We propose an algebraic geometric solution to the identification of linear hybrid systems: we represent the number of discrete states as the degree of a homogeneous polynomial and the model parameters as roots (factors) of such a polynomial. This interpretation allows us to decouple the identification of the model parameters from the inference of the hybrid state. In Section 3 we concentrate on SISO systems. We derive a rank constraint on the input/output data from which one can estimate the number of models N. We then estimate the model parameters using a linear algebraic polynomial factorization technique that is closed-form for $N \leq 4$. Once the model parameters have been identified, the estimation of the hybrid state and of the switching parameters becomes a simpler problem. In Section 4, we extend our results to MIMO systems and in Section 5 we present simulation experiments.

[1]Department of EECS, University of California, Berkeley CA 94720, USA, e-mail: rvidal@eecs.berkeley.edu.
[2]Department of Computer Science, University of California, Los Angeles CA 90095, USA, e-mail: soatto@ucla.edu.
[3]Department of EECS, University of California, Berkeley CA 94720, USA, e-mail: sastry@eecs.berkeley.edu.

2 Problem statement

We consider a class of discrete-time hybrid systems, known as *linear hybrid systems*, whose evolution is determined by a collection of linear models with *continuous state* $x_t \in \mathbb{R}^n$ connected by switches, indexed by a number of *discrete states* $\lambda_t \in \{1, 2, \dots, N\}$.

The evolution of the continuous state x_t is described by the linear system

$$\begin{aligned} x_{t+1} &= A(\lambda_t)x_t + B(\lambda_t)u_t & (1)\\ y_t &= C(\lambda_t)x_t, & (2)\end{aligned}$$

where $A(k) \in \mathbb{R}^{n\times n}$, $B(k) \in \mathbb{R}^{n\times q}$ and $C(k) \in \mathbb{R}^{p\times n}$, for $k \in \{1, 2, \dots, N\}$.

The evolution of the discrete state λ_t can be described in a variety of ways. In *Jump-linear systems (JLS)* λ is a deterministic but unknown input that is piecewise constant, right-continuous and finite-valued. In *Jump-Markov linear systems (JMLS)* λ is an irreducible Markov chain governed by the transition probabilities $\pi(k, k') \doteq P(\lambda_{t+1} = k'|\lambda_t = k)$. One can also have the switching mechanism governed by the value of the continuous state, as in *Piecewise affine systems* (PWAS), where λ is a piecewise constant function defined by a polyhedral partition of the state space. We take the least restrictive model (JLS), so that our results will apply also to other switching mechanisms.

Given input/output data $\{u_t, y_t\}$ generated by a linear hybrid system $\Sigma = \{A(k), B(k), C(k); k = 1, \dots, N\}$, we focus our attention on how to infer the *model parameters* $A(k), B(k), C(k)$ and the *hybrid state* $\{x_t, \lambda_t\}$.

3 Identification of SISO linear hybrid systems in PWARX form

In this section, we assume that each linear system is SISO and written in controllable canonical form

$$x_{t+1} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0\\ 0 & 0 & 1 & \cdots & 0\\ \vdots & & & \ddots & \vdots\\ 0 & \cdots & \cdots & 0 & 1\\ a_n & a_{n-1} & \cdots & \cdots & a_1 \end{bmatrix} x_t + \begin{bmatrix} 0\\0\\ \vdots\\ 0\\ 1\end{bmatrix} u_t \quad (3)$$

$$y_t = \begin{bmatrix} 0 & \dots & 0 & c_{n_u} & \cdots & c_1 \end{bmatrix} x_t. \quad (4)$$

Furthermore, we assume that they can be written in ARX form[1]

$$\begin{aligned} y_t = a_1 y_{t-1} + a_2 y_{t-2} + \cdots + a_n y_{t-n} +\\ c_1 u_{t-1} + c_2 u_{t-2} + \cdots + c_{n_u} u_{t-n_u} \end{aligned} \quad (5)$$

[1]For a single SISO linear time-invariant system, both the state-space and the ARX representations are equivalent. For JLS, the ARX representation is slightly more restrictive. We will address the relationship between these two representations from an identification point of view in future research.

where $n_u \leq n$ is the degree of the input. By letting

$$\begin{aligned} x_t &= (u_{t-n_u}, \dots, u_{t-1}, y_{t-n}, \dots y_{t-1}, -y_t)^T \in \mathbb{R}^{n+n_u+1}\\ \ell &= (c_{n_u}, \cdots, c_1, a_n, \dots a_1, 1)^T \in \mathbb{R}^{n+n_u+1}\end{aligned}$$

we write the ARX model of a single linear system as the following hyperplane in $\mathbb{R}^{n+n_u+1}$

$$\ell^T x_t = 0 \qquad t \geq n. \quad (6)$$

with ℓ representing the normal to the hyperplane.

If we denote the model parameters for each one of the N SISO linear systems as $\{\ell_i\}_{i=1}^N$, the identification and filtering problem can be stated as follows.

Problem 1 *Let $\{u_t, y_t\}_{t=0}^T$ be input/output data generated by a SISO linear hybrid system with known dimension of the state space n and degree of the input n_u. Estimate the number of discrete states N, the model parameters $\{\ell_i\}_{i=1}^N$ and the hybrid state $\{x_t, \lambda_t\}_{t=0}^T$.*

In the following sections, we give an algebraic geometric solution to this problem. The key idea is to consider the input/output data generated by a SISO linear hybrid system as points in $\mathbb{R}^K$, with $K = n + n_u + 1$, lying on one of the hyperplanes $\{\ell_i^T z = 0\}_{i=1}^N$. The identification problem is then reduced to estimating the number of hyperplanes (number of discrete states) and their normals (the model parameters), from sample data points (input/output data) on those hyperplanes. In order to so, we first need to decouple the identification of the model parameters from the filtering of the hybrid state and the identification of switching parameters. We show how to do so in Section 3.1 where we derive the so-called *hybrid decoupling constraint.* In Section 3.2 we show how to recover the number of discrete states from a rank constraint on the input/output data and the model parameters from the factorization of a homogeneous polynomial. Section 3.3 shows how to recover the hybrid state.

3.1 Decoupling identification from filtering

At a given time, t, the discrete state λ_t takes one out of N possible values $\{1, 2, \dots, N\}$. In other words, for all $t \geq n$ there exists an i such that $\ell_i^T x_t = 0$. Therefore, the following constraints must be satisfied by the model parameters and the input/output data regardless of the value of the discrete state and regardless of the switching mechanism generating the evolution of the discrete state

$$\boxed{\prod_{i=1}^N (\ell_i^T x_t) = 0.} \quad (7)$$

We call equation (7) the *hybrid decoupling constraint* (HDC), since it will allow us to identify the model parameters $\{\ell_i\}_{i=1}^N$ independently from the filtering of the discrete state $\{x_t, \lambda_t\}$ and regardless of the mechanism generating the transitions (JLS, JMLS, or PWAS). Notice that the HDC constraint is the only thing we can say from input/output data for a JLS in the absence of knowledge of the switching mechanism.

3.2 Identification of the model parameters

The HDC allows us to concentrate on the identification of the number of models N and the model parameters $\{\boldsymbol{\ell}_i\}_{i=1}^N$ from input/output data without the need to know the hybrid state. This is done by solving for N and the vectors $\boldsymbol{\ell}_i$'s from the HDC. To this end, notice that the HDC

$$p_N(\boldsymbol{z}) \doteq \prod_{i=1}^{N}(\boldsymbol{\ell}_i^T \boldsymbol{z}) = 0 \tag{8}$$

is a homogeneous polynomial of degree N in $K \doteq n + n_u + 1$ variables. Therefore it can be written as

$$p_N(\boldsymbol{z}) \doteq \sum h_{n_1,\dots,n_K} z_1^{N_1} \cdots z_K^{N_K} = \boldsymbol{h}^T \nu_N(\boldsymbol{z}) = 0, \tag{9}$$

where $h_I \in \mathbb{R}$ represents the coefficient of the monomial $\boldsymbol{z}^I = z_1^{N_1} z_2^{N_2} \cdots z_K^{N_K}$ with $0 \le N_j \le N$, $j = 1, \dots, K$, and $N_1 + N_2 + \cdots + N_K = N$; $\nu_N : \mathbb{R}^K \to \mathbb{R}^{M_N}$ is the Veronese map of degree N defined as:

$$\nu_N: \ [z_1, \dots, z_K]^T \mapsto [\dots, z^I, \dots]^T, \tag{10}$$

with I chosen in the degree-lexicographic order, and

$$M_N = \binom{N+K-1}{K-1} = \binom{N+K-1}{N} \tag{11}$$

is the total number of different monomials. One can show that the vector $\boldsymbol{h} \in \mathbb{R}^{M_N}$ is the symmetric tensor product of the individual model parameters $\{\boldsymbol{\ell}_i\}_{i=1}^N$, *i.e.*

$$\boldsymbol{h} = \sum_{\sigma \in \mathfrak{S}_N} \boldsymbol{\ell}_{\sigma(1)} \otimes \boldsymbol{\ell}_{\sigma(2)} \otimes \cdots \otimes \boldsymbol{\ell}_{\sigma(N)} \tag{12}$$

where $\mathfrak{S}_N$ is the permutation group of N elements. Hence we will refer to $\boldsymbol{h}$ as the *hybrid model parameters.*

We now show how to estimate the hybrid model parameters from input/output data $\{\boldsymbol{x}_t\}_{t=n}^{n+T-1}$. After applying equation (9) to the data, we obtain the following system of linear equations on the hybrid model parameters $\boldsymbol{h}$

$$L_N\, \boldsymbol{h} \doteq \begin{bmatrix} \nu_N(\boldsymbol{x}_n)^T \\ \nu_N(\boldsymbol{x}_{n+1})^T \\ \vdots \\ \nu_N(\boldsymbol{x}_{n+T-1})^T \end{bmatrix} \boldsymbol{h} = 0 \in \mathbb{R}^T. \tag{13}$$

We are now interested in determining whether there exists a unique solution for $\boldsymbol{h}$ from (13), *i.e.* under what conditions we have $\mathrm{rank}(L_N) = M_N - 1$. It turns out that the uniqueness of $\boldsymbol{h}$ is very much related to the estimation of the number of models N as shown by the following theorem.

Theorem 1 (Number of discrete states) *Given input/output data $\{\boldsymbol{x}_t\}_{t=n}^{T+n-1}$, $T \ge M_N - 1$, generated by a linear hybrid system, let $L_i \in \mathbb{R}^{T \times M_i}$ be the matrix defined in (13), but computed with the Veronese map ν_i of degree i. If the data points are in general position on the subspaces $\{\boldsymbol{\ell}_i^T \boldsymbol{z} = 0\}_{i=1}^N$ and at least $K-1$ points lie on each subspace, then:*

$$\mathrm{rank}(L_i) \begin{cases} > M_i - 1, & i < N, \\ = M_i - 1, & i = N, \\ < M_i - 1, & i > N. \end{cases} \tag{14}$$

Therefore, the number of discrete states N is given by:

$$\boxed{N = \min\{i : \mathrm{rank}(L_i) = M_i - 1\}.} \tag{15}$$

Remark 1 *In the presence of noise, one cannot directly estimate N from (15), because the matrix L_N is always full rank. In practice we declare the rank of L_N to be r if $\sigma_{r+1}/\sigma_r < \epsilon$, where σ_i is the i-th singular value of L_N and $\epsilon > 0$ is a pre-specified threshold.*

Theorem 1 and the linear system in equation (13) allow us to determine the number of discrete states N and the hybrid model parameters $\boldsymbol{h}$, respectively, from input/output data $\{\boldsymbol{x}_t\}_{t=n}^{T+n-1}$. The rest of the problem becomes now how to recover the model parameters $\{\boldsymbol{\ell}_i\}_{i=1}^n$ from $\boldsymbol{h}$. From equation (9) we have that recovering $\{\boldsymbol{\ell}_i\}_{i=1}^n$ from $\boldsymbol{h}$ is equivalent to factoring a given a homogeneous polynomial $p_N(\boldsymbol{z})$ of degree N into N *distinct* polynomials $(\boldsymbol{\ell}_i^T \boldsymbol{z})$ of degree 1. We show now that this polynomial factorization problem has a unique solution that can be obtained by solving for the roots of a polynomial of degree N in *one* variable, plus solving a linear system in N variables [22].

Solving for the last 2 entries of each $\boldsymbol{\ell}_i$. Consider the last $N+1$ coefficients of $p_N(\boldsymbol{z})$:

$$[h_{0,\dots,0,N,0}\ ,\ h_{0,\dots,0,N-1,1}\ ,\ \cdots\ ,\ h_{0,\dots,0,0,N}]^T \in \mathbb{R}^{N+1},$$

which define the following homogeneous polynomial of degree N in the two variables z_{K-1} and z_K:

$$\sum h_{0,\dots,0,N_{K-1},N_K} z_{K-1}^{N_{K-1}} z_K^{N_K} = \prod_{i=1}^{N}(\ell_{iK-1} z_{K-1} + \ell_{iK} z_K).$$

Since by construction $\ell_{iK} = 1$, then $h_{0,\dots,0,0,N} = 1$. Thus if we let $w = z_K / z_{K-1}$ we have that:

$$\prod_{i=1}^{N}(\ell_{iK-1} z_{K-1} + \ell_{iK} z_K) = 0 \Leftrightarrow \prod_{i=1}^{N}(\ell_{iK-1} + w) = 0.$$

Hence the N roots of the polynomial

$$q_N(w) = h_{0,\dots,0,N,0} + h_{0,\dots,0,N-1,1} w + \cdots + w^N \tag{16}$$

are exactly $w_i = -\ell_{iK-1}$, for all $i = 1, \ldots, n$. Therefore we obtain the last two entries of each ℓ_i as:

$$(\ell_{iK-1}\ ,\ \ell_{iK}) = (-w_i\ ,\ 1). \tag{17}$$

Solving for the first $K-2$ entries of each ℓ_i. We now show how to compute the first $K-2$ entries of each ℓ_i. We assume that we have computed $\{\ell_{ij}\}_{i=1}^N$, $j = J+1, \ldots K$ for some J, starting with $J = K-2$, and show how to linearly solve for $\{\ell_{iJ}\}_{i=1}^N$. To this end, we consider the coefficients of $p_N(z)$ which are *linear* in z_J. These coefficients are of the form $h_{0,\ldots,0,1,N_{J+1},\ldots,N_K}$ and are linear in ℓ_{iJ}. To see this, notice that the polynomial $\sum h_{0,\ldots,0,1,N_{J+1},\ldots,N_K} z_{J+1}^{N_{J+1}} \cdots z_K^{N_K}$ is equal to the partial derivative of $p_N(z)$ with respect to z_J evaluated at $z_1 = z_2 = \cdots = z_J = 0$. Since

$$\frac{\partial}{\partial z_J}\left(\prod_{i=1}^N (\ell_i^T z)\right) = \sum_{i=1}^N \ell_{iJ}\left(\prod_{k=1}^{i-1}(\ell_k^T z)\prod_{k=i+1}^N (\ell_k^T z)\right),$$

after evaluating at $z_1 = z_2 = \cdots = z_J = 0$ we obtain

$$\sum h_{0,\ldots,0,1,N_{J+1},\ldots,N_K} z_{J+1}^{N_{J+1}} \cdots z_K^{N_K} = \sum_{i=1}^N \ell_{iJ}\, g_i^J(z), \tag{18}$$

where

$$g_i^J(z) = \prod_{k=1}^{i-1}\left(\sum_{j=J+1}^K \ell_{kj} z_j\right)\prod_{k=i+1}^N\left(\sum_{j=J+1}^K \ell_{kj} z_j\right) \tag{19}$$

is a homogeneous polynomial of degree $N-1$ in the last $K-J$ variables in z. Let $\mathcal{V}_i^J$ be the vector of coefficients of the polynomial $g_i^J(z)$. From equation (18) we get

$$\begin{bmatrix}\mathcal{V}_1^J & \mathcal{V}_2^J & \cdots & \mathcal{V}_N^J\end{bmatrix}\begin{bmatrix}\ell_{1J}\\ \ell_{2J}\\ \vdots\\ \ell_{NJ}\end{bmatrix} = \begin{bmatrix}h_{0,\ldots,0,1,N-1,0,\ldots,0}\\ h_{0,\ldots,0,1,N-2,1,\ldots,0}\\ \vdots\\ h_{0,\ldots,0,1,0,0,\ldots,N-1}\end{bmatrix} \tag{20}$$

from which we can linearly solve for the unknowns $\{\ell_{iJ}\}_{i=1}^N$. Notice that the vectors $\{\mathcal{V}_i^J\}_{i=1}^N$ are known, because they are functions of ℓ_{ij}, $j \geq J+1$, which are known.

We have shown the following.

Theorem 2 (Estimating the model parameters) *Given input/output data $\{x_t\}_{t=0}^{T+n-1}$, $T \geq M_N - 1$, generated by a linear hybrid system with N discrete states, the model parameters $\{\ell_i\}_{i=1}^N$ can be computed as follows:*

1. *Solve for the hybrid model parameters $h \in \mathbb{R}^{M_N}$ from the null space of $L_N \in \mathbb{R}^{T\times M_N}$ (13).*
2. *Solve for the model parameters $\{\ell_i\}_{i=1}^N$ by factoring $p_N(z) = h^T \nu_N(z)$ as $(\ell_1^T z)(\ell_2^T z)\cdots(\ell_N^T z)$. The factorization problem is equivalent to solving for the roots of a polynomial of degree N in one variable plus $K-2$ linear systems in N variables.*

Example 1 *If $N = 2$ and $K = 3$, then*

$$\begin{aligned} p_2(z) &= (\ell_{11}z_1 + \ell_{12}z_2 + z_3)(\ell_{21}z_1 + \ell_{22}z_2 + z_3)\\ &= (\ell_1^T z)(\ell_2^T z) = [z_1^2, z_1z_2, z_1z_3, z_2^2, z_2z_3, z_3^2]h\\ &= \underbrace{(\ell_{11}\ell_{21})}_{h_{2,0,0}}z_1^2 + \underbrace{(\ell_{11}\ell_{22}+\ell_{12}\ell_{21})}_{h_{1,1,0}}z_1z_2 + \underbrace{(\ell_{11}+\ell_{21})}_{h_{1,0,1}}z_1z_3 +\\ &\quad \underbrace{(\ell_{12}\ell_{22})}_{h_{0,2,0}}z_2^2 + \underbrace{(\ell_{12}+\ell_{22})}_{h_{0,1,1}}z_2z_3 + \underbrace{(\ 1\ )}_{h_{0,0,2}}z_3^2. \end{aligned}$$

Since $h \in \mathbb{R}^6$ is known, so is the second order polynomial

$$q_2(w) = h_{0,2,0} + h_{0,1,1}w + w^2 = (\ell_{12} + w)(\ell_{22} + w).$$

Thus we can obtain $-\ell_{12}$ and $-\ell_{22}$ from the roots w_1 and w_2 of $q_2(w)$. Finally, we notice that $h_{1,1,0}$ and $h_{1,0,1}$ are linear functions of the remaining unknowns ℓ_{11} and ℓ_{21}. Thus we can compute them from the linear system

$$\begin{bmatrix}\ell_{22} & \ell_{12}\\ 1 & 1\end{bmatrix}\begin{bmatrix}\ell_{11}\\ \ell_{21}\end{bmatrix} = \begin{bmatrix}h_{1,1,0}\\ h_{1,0,1}\end{bmatrix} \tag{21}$$

Notice that this linear system has a unique solution if and only if $\ell_{22} - \ell_{12} \neq 0$, i.e. if $w_1 \neq w_2$.

3.3 Filtering of the hybrid state

Given the number of discrete states N and the model parameters $\{\ell_i\}_{i=1}^n$, we now show how to reconstruct the hybrid state trajectory $\{x_t, \lambda_t\}$ from input/output data $\{x_t\}_{t=n}^{T+n-1}$. To this end, we first notice that at each time t there exists a unique[2] i such that $\ell_i^T x_t = 0$. Therefore, the discrete state can be trivially identified as:

$$\boxed{\lambda_t = \arg\min_i (\ell_i^T x_t)^2.} \tag{22}$$

Furthermore, since the model parameters ℓ_i are known, the entire evolution of the state space parameters $(A_t, C_t) = (A(\lambda_t), C(\lambda_t))$ is known. Therefore, we can express the output y_t directly as a function of the initial continuous state x_0 and the (known) input u_t as

[2] In principle, it is possible that a data point x_t belongs to more than one subspace $\ell_i^T z = 0$. However, the set of all such points is a set of measure zero on the variety $\{z : p_N(z) = 0\}$. Unique association of each data point to one subspace is guaranteed if the observability conditions described in [21] are satisfied.

$$
\begin{bmatrix} y_0 \\ y_1 \\ y_2 \\ \vdots \\ y_t \end{bmatrix} = \begin{bmatrix} C_0 \\ C_1A_0 \\ C_2A_1A_0 \\ \vdots \\ C_tA_{t-1}\cdots A_0 \end{bmatrix} x_0 + \tag{23}
$$
$$
\begin{bmatrix} 0 & 0 & \cdots & 0 \\ C_1B_0 & 0 & \cdots & 0 \\ C_2A_1B_0 & C_2B_1 & & \\ \vdots & & \ddots & \\ C_tA_{t-1}\cdots A_1B_0 & & & C_tB_{t-1} \end{bmatrix} \begin{bmatrix} u_0 \\ u_1 \\ u_2 \\ \vdots \\ u_{t-1} \end{bmatrix}
$$

from which we can solve for x_0 uniquely, provided that the matrix multiplying x_0 has full rank. This is guaranteed if the observability conditions of [21] are satisfied. Given x_0, u_t and λ_t, the continuous state trajectory $\{x_t\}$ can be trivially recovered from equation (1).

Remark 2 (Inferring the switching parameters) *Once the model parameters and the hybrid state have been identified, the problem of estimating the switching parameters, e.g. the partition of the state space for PWAS, becomes a simpler problem. We refer interested readers to [3, 10] for specific algorithms.*

4 Identification of MIMO systems

Let us first consider the case of a linear hybrid system with a single output but multiple inputs $u_t \in \mathbb{R}^q$, $q > 1$. In principle, the transfer function for each linear system is given by

$$
G(z) = \begin{bmatrix} \frac{n_1(z)}{d_1(z)} & \frac{n_2(z)}{d_2(z)} & \cdots & \frac{n_q(z)}{d_q(z)} \end{bmatrix}. \tag{24}
$$

However, we can always compute a common denominator, $d(z) = 1 - a_1z^{-1} - \cdots a_nz^{-n}$, so that, without loss of generality, each ARX model reads.

$$
\begin{aligned}
y_t = {} & a_1y_{t-1} + a_2y_{t-2} + \cdots + a_ny_{t-n} + \\
& c_1^1u_{t-1}^2 + c_2^2u_{t-2}^2 + \cdots + c_{n_u^1}^1 u_{t-n_u^1}^1 + \cdots \\
& c_1^q u_{t-1}^q + c_2^q u_{t-2}^q + \cdots + c_{n_u^q}^q u_{t-n_u^q}^q
\end{aligned}
$$

where $u_t = (u_t^1, u_t^2, \cdots, u_t^q)^T \in \mathbb{R}^q$ and n_u^i is the degree of input $i = 1, \ldots, q$. The above ARX model can be written as

$$
\boldsymbol{\ell}^T \boldsymbol{x}_t = (\boldsymbol{c}^{1T}, \ldots, \boldsymbol{c}^{qT}, \boldsymbol{a}^T)(\boldsymbol{u}_t^{1T}, \ldots, \boldsymbol{u}_t^{qT}, \boldsymbol{y}_t^T)^T = 0 \tag{25}
$$

where $\boldsymbol{x}_t \in \mathbb{R}^K$ and $\boldsymbol{\ell} \in \mathbb{R}^K$, with $K = n + \sum_{j=1}^{1} n_u^j + 1$,

$$
\begin{aligned}
\boldsymbol{c}^j &= (c_{n_u^j}^j, \cdots, c_1^j)^T \in \mathbb{R}^{n_u^j} && j = 1, \cdots, q \\
\boldsymbol{a} &= (a_n, \cdots, a_1, 1)^T \in \mathbb{R}^{n+1} \\
\boldsymbol{u}_t^j &= (u_{t-n_u^j}^j, \cdots, u_{t-1}^j)^T \in \mathbb{R}^{n_u^j} && j = 1, \cdots, q \\
\boldsymbol{y}_t &= (y_{t-n}, \cdots, y_{t-1}, -y_t)^T \in \mathbb{R}^{n+1}.
\end{aligned}
$$

Therefore, the *same* identification procedure for SISO systems can be directly applied to MISO systems, as long as the degree of the numerators and denominator, n_u^j and n respectively, are known.

Consider now the more general case in which each linear system is MIMO with transfer function of the form

$$
G(z) = \begin{bmatrix} \frac{n_{11}(z)}{d_{11}(z)} & \cdots & \frac{n_{1q}(z)}{d_{1q}(z)} \\ \vdots & & \vdots \\ \frac{n_{p1}(z)}{d_{p1}(z)} & \cdots & \frac{n_{pq}(z)}{d_{pq}(z)} \end{bmatrix} \in \mathbb{R}^{p \times q}(z). \tag{26}
$$

Again, without loss of generality, we can assume that there is a common denominator $d^i(z)$ of degree n^i, for each output $i = 1, \ldots, p$. Similarly to (25), we obtain the following system of equations on the data points

$$
\mathcal{L}^T \boldsymbol{x}_t = \begin{bmatrix} \boldsymbol{c}^{11} & \cdots & \boldsymbol{c}^{p1} \\ \vdots & & \vdots \\ \boldsymbol{c}^{1q} & \cdots & \boldsymbol{c}^{pq} \\ \boldsymbol{a}^1 & & 0 \\ \vdots & \ddots & \vdots \\ 0 & & \boldsymbol{a}^p \end{bmatrix}^T \begin{bmatrix} \boldsymbol{u}_t^1 \\ \vdots \\ \boldsymbol{u}_t^q \\ \boldsymbol{y}_t^1 \\ \vdots \\ \boldsymbol{y}_t^p \end{bmatrix} = 0 \tag{27}
$$

where $\mathcal{L} \in \mathbb{R}^{p \times K}$ with $K = \sum_{i=1}^p (n^i + 1) + \sum_{j=1}^q n_u^j$, $\boldsymbol{a}^i \in \mathbb{R}^{n^i+1}$ is the parameters of the i-th output y_t^i, $i = 1, \ldots, p$, and $\boldsymbol{y}_t^i = (y_{t-n^i}^i, \cdots, y_{t-1}^i, -y_t^i)^T \in \mathbb{R}^{n^i+1}$.

Therefore, the input/output data generated by a linear system lies in a $(K-p)$-dimensional subspace of $\mathbb{R}^K$. This implies that the identification of MIMO linear hybrid systems is equivalent to the identification of a collection of N $(K-p)$-dimensional subspaces of $\mathbb{R}^K$ from sample points on those subspaces, without knowing which sample points belong to which subspaces.

The case $p = 1$ (MISO systems) corresponds to the identification of a collection of N hyperplanes. Since the algebraic variety associated with N hyperplanes can be uniquely described with a single homogeneous polynomial, as described in Section 3 one can reduce the identification problem to estimating such a polynomial from data and then factorizing it to obtain the model parameters.

The case $p > 1$ is more complex, because there are multiple polynomials describing the algebraic variety associated with a collection of N subspaces. Multiple polynomials arise in the case of subspaces because one can change the basis for the orthogonal complement to each subspace[3]. In the case of hyperplanes the basis for the orthogonal complement is uniquely defined by the normal to the subspace.

[3]Since the coefficients of the polynomials representing all the subspaces must lie in the null space of the data matrix L_N in (13), we conclude that $\text{rank}(L_N) \le M_N - 1$.

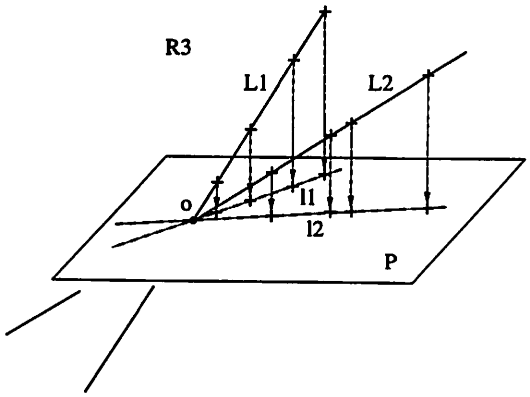

Figure 1: Samples on two 1-dimensional subspaces L_1, L_2 in $\mathbb{R}^3$ projected onto a 2-dimensional plane P. The membership of each sample (labeled as "+") is preserved through the projection.

It turns out that one can avoid estimating multiple polynomials by first projecting the data onto a lower-dimensional linear space. Consider, for example, the problem of estimating two lines L_1 and L_2 in $\mathbb{R}^3$ from sample data points (see Figure 1). In this case, one can project the lines onto a plane P not orthogonal to the plane containing the lines, and the segmentation of the data is preserved. In the plane, the two lines are represented by a *single* polynomial. We showed in [22] that, for subspaces of arbitrary dimension, one can always project the data onto a lower-dimensional subspace in which the data is represented by a unique homogeneous polynomial. Furthermore, the set of projections that fail to preserve the segmentation of the data is only a zero-measure set [14].

The question is now how to choose a projection onto a lower-dimensional subspace that preserves the segmentation. While this may be hard in general, because we do not know the basis for each one of the subspaces, in the case of hybrid systems identification we are looking for a collection of subspaces whose orthogonal complement has a particular structure given by the matrix $\mathcal{L}$ in equation (27). More specifically, we can choose a collection of p projections $\{\pi_i\}_{i=1}^p$ of the form

$$\pi_i : \mathbb{R}^K \to \mathbb{R}^{K_i}$$
$$x_t \mapsto (u_t^1, \cdots, u_t^q, y_t^i)$$

where $K_i = n^i + 1 + \sum_{j=1}^q n_u^j$. Each one of those projections allows us to identify the model parameters $(c^{i1}, \cdots, c^{iq}, a^i)$ by applying our MISO identification algorithm to the input/output data $(u_t^1, \cdots, u_t^q, y_t^i)$. Once *all* the model parameters have been identified, the filtering of the discrete state can be trivially obtained, similarly to (22). The filtering of the continuous state may also be obtained as before, but requires to first find a state-space realization of each MIMO linear system. See [15] for further details on how to obtain a realization of a linear system given a transfer function representation.

5 Experiments

We present simulation results on the identification of 1000 randomly chosen JLS with $N = 3$ discrete states. Each linear system is described by the ARX model

$$y_t = a_1(\lambda_t)y_{t-1} + a_2(\lambda_t)y_{t-2} + c_1(\lambda_t)u_{t-1} + w_t \quad (28)$$

where the dimension of the state space is $n = 2$, the degree of the input is $n_u = 1$ and the discrete state is $\lambda_t \in \{1,2,3\}$. For each trial, the model parameters (a_1, a_2) for each discrete state were randomly chosen so that the poles of each linear system are uniformly distributed on the annulus $0.8 \leq \|z\| \leq 1 \subset \mathbb{C}$. The model parameter c_1 for each discrete state was chosen according to a zero-mean unit variance Gaussian distribution. The value of the discrete state was chosen as

$$\lambda_t = \begin{cases} 1 & 1 \leq t \leq 30 \\ 2 & 31 \leq t \leq 60 \\ 3 & 61 \leq t \leq 100 \end{cases} \quad (29)$$

The initial value of the continuous state was randomly drawn from a zero-mean Gaussian distribution with variance $\Sigma = I_2$. The input sequence was drawn from a zero-mean unit variance Gaussian distribution. We added zero-mean Gaussian noise with standard deviation $\sigma \in [0, 0.01]$, w_t, to simulate a measurement error of about 1%. Notice that the algorithm is designed for the noise-free case.

Figure 2 shows the mean error on the estimation of the model parameters[4], continuous state[5], and discrete state[6], respectively, as a function of σ. Both the model parameters and the continuous state are correctly estimated with an error that increases approximately linearly with the amount of noise. Notice that the discrete state is incorrectly estimated approximately 8% of the times for $\sigma = 0.01$. Notice also that there is no error for $\sigma = 0$.

Figures 3 and 4 show the reconstruction of the state trajectory for a particular trial with $\sigma = 0.01$. Notice that there are 5 time instances in which the estimates of the discrete state are incorrect. Notice however, that the continuous state is estimated with a small error throughout the whole interval, in spite of erroneous identification of the discrete state.

[4]The error between the estimated model parameters $(\hat{a_1}, \hat{a_2}, \hat{c_1})$ and the true model parameters (a_1, a_2, c_1) was computed as $\|(\hat{a_1}, \hat{a_2}, \hat{c_1}) - (a_1, a_2, c_1)\|$, averaged over the number of models and trials.

[5]The error between the estimated continuous state $\hat{x}_t$ and the true continuous state (a_1, a_2, c_1) was computed as $\|\hat{x}_t - x_t\|$, averaged over the number of data points and trials.

[6]The error between the estimated discrete state $\hat{\lambda}_t$ and the true discrete state λ_t was computed as the number of times in which $\hat{\lambda}_t \neq \lambda_t$, averaged over the number of trials.

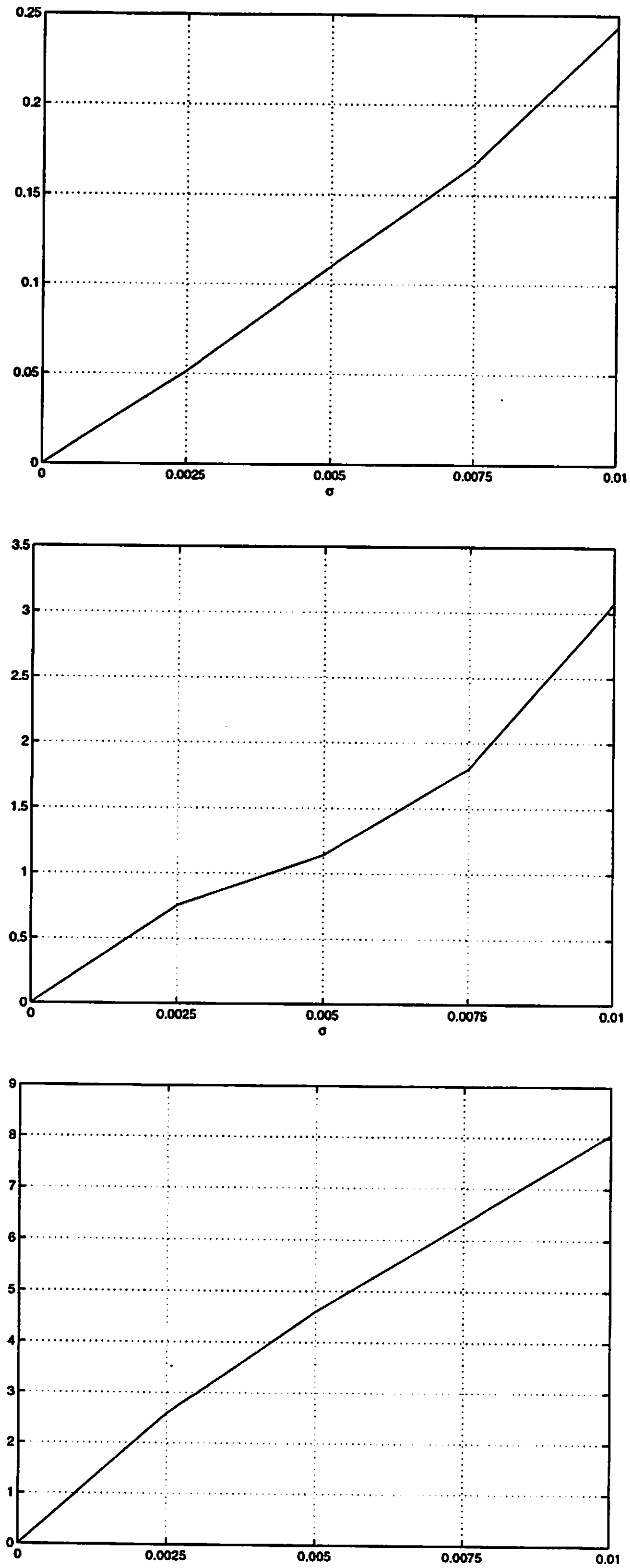

Figure 2: Mean error over 1000 trials for the identification of the model parameters (top), the continuous state (middle) and the discrete state (bottom) as a function of the standard deviation of the measurement error σ.

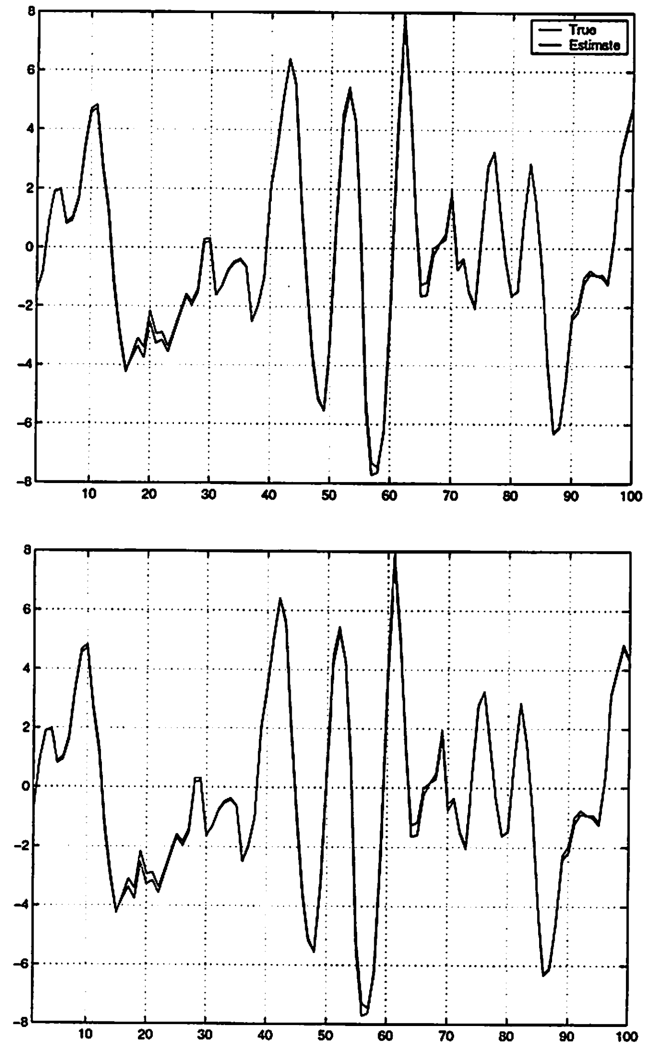

Figure 3: Evolution of the first (top) and second (bottom) entry of the continuous state x_t and its estimate $\hat{x_t}$.

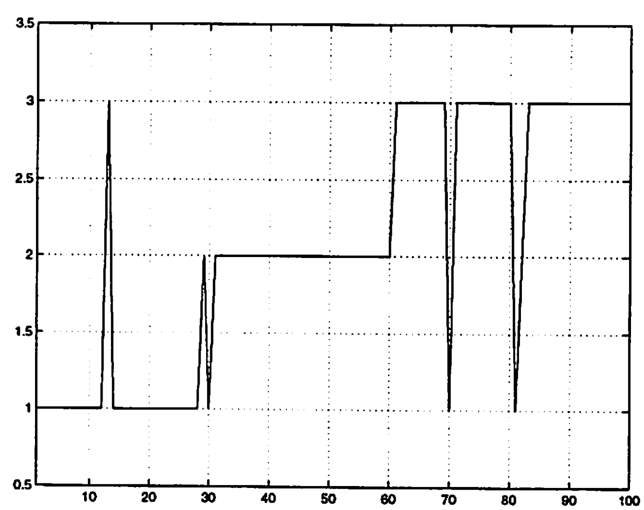

Figure 4: Evolution of the estimated discrete state $\hat{\lambda}_t$.

6 Conclusions

We have proposed an algebraic geometric solution to the identification of linear hybrid systems. We showed that the identification problem of the model parameters can be decoupled from the inference of the hybrid state and made independent from the switching mechanism generating the transitions. The decoupling was made possible by the hybrid decoupling constraint, which establishes an equivalence between linear hybrid systems identification and the estimation of a collection of subspaces from data points on those subspaces. Then, we derived a rank constraint on the input/output data from which one can estimate the number of models N. Given N, the estimation of the model parameters is equivalent to a factorization problem in the space of homogeneous polynomials, which can be solved using linear algebraic techniques. Once the model parameters have been identified, the filtering of the hybrid state and the switching parameters can be obtained using standard techniques. We presented simulation results evaluating the performance of the algorithm.

References

[1] A. Alessandri and P. Coletta. Design of Luenberger observers for a class of hybrid linear systems. In *Hybrid Systems: Computation and Control*, volume 2034 of *LNCS*, pages 7–18. Springer Verlag, 2001.

[2] A. Balluchi, L. Benvenuti, M. Di Benedetto, and A. Sangiovanni-Vincentelli. Design of observers for hybrid systems. In *Hybrid Systems: Computation and Control*, volume 2289 of *LNCS*, pages 76–89. Springer Verlag, 2002.

[3] Bemporad, A. Garulli, S. Paoletti, and A. Vicino. A greedy approach to identification of piecewise affine models. In *Hybrid Systems: Computation and Control*, LNCS. Springer Verlag, 2003.

[4] A. Bemporad, G. Ferrari, and M. Morari. Observability and controllability of piecewise affine and hybrid systems. *IEEE Transactions on Automatic Control*, 45(10):1864–1876, October 2000.

[5] A. Bemporad, J. Roll, and L. Ljung. Identification of hybrid systems via mixed-integer programming. In *Proc. of IEEE Conf. on Decision and Control*, pages 786–792, 2001.

[6] M. Billio, A. Monfort, and C.P. Robert. Bayesian estimation of switching ARMA models. *Journal of Econometrics*, (93):229–255, 1999.

[7] A. Blake, B. North, and M. Isard. Learning multi-class dynamics. *Advances in Neural Information Processing Systems*, 11:389–395, 1999. MIT Press.

[8] A. Doucet, A. Logothetis, and V. Krishnamurthy. Stochastic sampling algorithms for state estimation of jump Markov linear systems. *IEEE Transactions on Automatic Control*, 45(1):188–202, 2000.

[9] J. Ezzine and A. H. Haddad. Controllability and observability of hybrid systems. *International Journal of Control*, 49(6):2045–2055, 1989.

[10] G. Ferrari-Trecate, M. Muselli, D. Liberati, and M. Morari. A clustering technique for the identification of piecewise affine systems. In *Hybrid Systems: Computation and Control*, volume 2034 of *LNCS*, pages 218–231. Springer Verlag, 2001.

[11] G. Ferrari-Trecate, M. Musselli, D. Liberati, and M. Morari. Identification of piecewise affine and hybrid systems. In *Proceedings of the American Control Conference*, pages 3521–3526, 2001.

[12] M. Morari G. Ferrari-Trecate, D. Mignone. Moving horizon estimation for hybrid systems. In *Proceedings of the American Control Conference*, pages 1684–1688, 2000.

[13] Z. Ghahramani and G. E. Hinton. Variational learning for switching state-space models. *Neural Computation*, 12(4):963–996, 1998.

[14] M. Hirsch. *Differential Topology.* Springer, 1976.

[15] T. Kailath. *Linear Systems.* Prentice Hall, 1980.

[16] K. Murphy. Switching Kalman filters. Technical report, U. C. Berkeley, 1998.

[17] V. Pavlovic, J. M. Rehg, T. J. Cham, and K. P. Murphy. A dynamic Bayesian network approach to figure tracking using learned dynamic models. In *Proc. of the Intl. Conf. on Comp. Vision*, pages 94–101, 1999.

[18] A. Sun, S. S. Ge, and T. H. Lee. Controllability and reachability criteria for switched linear systems. *Automatica*, 38:775–786, 2002.

[19] F. Szigeti. A differential algebraic condition for controllability and observability of time varying linear systems. In *Proc. of IEEE Conference on Decision and Control*, pages 3088–3090, 1992.

[20] J. K. Tugnait. Detection and estimation for abruptly changing systems. *Automatica*, 18(5):607–615, 1982.

[21] R. Vidal, A. Chiuso, and S. Soatto. Observability and identifiability of jump linear systems. In *Proc. of IEEE Conference on Decision and Control*, pages 3614–3619, 2002.

[22] R. Vidal, Y. Ma, and S. Sastry. Generalized principal component analysis. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2003.