

An Algebraic Geometric Approach to the Identification of a Class of Linear Hybrid Systems*

René Vidal¹

Stefano Soatto²

Yi Ma³

Shankar Sastry¹

Abstract

We propose an algebraic geometric solution to the identification of a class of linear hybrid systems. We show that the identification of the model parameters can be decoupled from the inference of the hybrid state and the switching mechanism generating the transitions, hence we do not constraint the switches to be separated by a minimum dwell time. The decoupling is obtained from the so-called *hybrid decoupling constraint*, which establishes a connection between linear hybrid system identification, polynomial factorization and hyperplane clustering. In essence, we represent the number of discrete states n as the degree of a homogeneous polynomial p and the model parameters as factors of p . We then show that one can estimate n from a rank constraint on the data, the coefficients of p from a linear system, and the model parameters from the derivatives of p . The solution is closed form if and only if $n \leq 4$. Once the model parameters have been identified, the estimation of the hybrid state becomes a simpler problem. Although our algorithm is designed for noiseless data, we also present simulation results with noisy data.

1 Introduction

Hybrid systems are dynamical models that can be used to describe continuous phenomena that exhibit discontinuous behavior. For instance, the continuous trajectory of a bouncing ball results from the alternation between free fall and elastic contact. However, hybrid systems 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 models.

In this paper we look at the problem of modeling input/output data by piecewise linear (hybrid) models: Given input/output data, we want to simultaneously estimate the number of underlying linear models, the parameters of each model, and the hybrid state (continuous and discrete).

¹Department of EECS, University of California, Berkeley CA 94720, USA, e-mail: {rvidal,sastry}@eecs.berkeley.edu.

²Department of Computer Science, University of California, Los Angeles CA 90095, USA, e-mail: soatto@ucla.edu.

³Department of ECE, University of Illinois, Urbana-Champaign IL 61801, USA, e-mail: yima@uiuc.edu.

*Work funded by grants ONR N00014-00-1-0621, NSF-ECS 0200511, NSF-STC CENS and UIUC ECE startup.

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 [17]. After a decade-long hiatus, the problem has recently been enjoying considerable interest (see [3, 9, 15, 16, 19, 20] and references therein). Related work has also appeared in the machine learning community (see [6, 7, 8, 12, 13, 14] 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; [10] proposes a moving horizon estimator that, under some 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: [11] assumes that the number of models is known, and proposes an identification algorithm that combines clustering, regression and classification techniques; [5] uses mixed-integer quadratic programming; [4] uses a greedy approach for initialization and then iterates between assigning data points to models and computing the model parameters.

In this paper, we propose a novel algebraic geometric solution to the identification of a class of linear hybrid systems. We derive a polynomial constraint p , the so-called *hybrid decoupling constraint*, which decouples the identification of the model parameters from the inference of the hybrid state and the switching mechanism generating the transitions, hence we do not constraint the switches to be separated by a minimum dwell time. By representing the number of discrete states n as the degree of the polynomial p and the model parameters as factors of p , we cast the identification problem as a polynomial factorization problem, which we solve using simple linear-algebraic techniques: we estimate n from a rank constraint on the data, the coefficients of p from a linear system, and the model parameters from the derivatives of p . The solution is closed form if and only if $n \leq 4$. Once the model parameters have been identified, the estimation of the

hybrid state becomes a simpler problem. Although our algorithm is designed for noiseless data, we also present simulation results with noisy data.

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* $x_t \in \mathbb{R}^{n_a}$ 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

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

where $A(i) \in \mathbb{R}^{n_a \times n_a}$, $B(i) \in \mathbb{R}^{n_a \times q}$ and $C(i) \in \mathbb{R}^{p \times n_a}$, for $i \in \{1, 2, \dots, n\}$. Furthermore, we assume that each linear system can be written in ARX form¹

$$y_t = \sum_{j=1}^{n_a} a_j(\lambda_{t-1})y_{t-j} + \sum_{j=1}^{n_c} c_j(\lambda_{t-1})u_{t-j} \quad (2)$$

where $n_c \leq n_a$ is the degree of the input.

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

We call a JLS whose linear dynamics can be written in ARX form a Piece-Wise ARX (PWARX) model and consider the following identification/filtering problem.

Problem 1 *Let $\{u_t, y_t\}_{t=0}^T$ be input/output data generated by a PWARX model with known dimension of the state space n_a and degree of the input n_c . Estimate the number of discrete states n , the model parameters $\{a_\ell(i)\}_{\ell=1, \dots, n_a}^{i=1, \dots, n}$ and $\{c_\ell(i)\}_{\ell=1, \dots, n_c}^{i=1, \dots, n}$, and the hybrid state $\{x_t, \lambda_{t-1}\}_{t=n_a}^T$.*

3 Identification of linear hybrid systems in PWARX form

In this section, we present an algebraic geometric solution to the identification of linear hybrid systems in PWARX form. Most of our development will concentrate on the case of single-input single-output (SISO) systems, *i.e.* $u_t \in \mathbb{R}$ and $y_t \in \mathbb{R}$. However, we show in Remark 5 that our approach can be easily extended to the multiple-input multiple-output (MIMO) case.

¹For a single SISO linear time-invariant system, both the state-space and the ARX representations are equivalent. For SISO JLS, the ARX representation is more restrictive.

We notice from (2) that if we let $K \doteq n_a + n_c + 1$,

$$\begin{aligned} \mathbf{x}_t &= [u_{t-n_c}, \dots, u_{t-1}, y_{t-n_a}, \dots, y_{t-1}, -y_t]^T \in \mathbb{R}^K \text{ and} \\ \mathbf{b}_i &= [c_{n_c}(i), \dots, c_1(i), a_{n_a}(i), \dots, a_1(i), 1]^T \in \mathbb{R}^K, \end{aligned}$$

for $i = 1, \dots, n$, then we have that for all $t \geq n_a$ there exists a discrete state $\lambda_{t-1} = i \in \{1, \dots, n\}$ such that the point \mathbf{x}_t lies on the hyperplane

$$\mathbf{b}_i^T \mathbf{x}_t = 0. \quad (3)$$

Therefore, the input/output data $\{u_t, y_t\}_{t=0}^T$ generated by a SISO PWARX model can be mapped into a set of points $\{\mathbf{x}_t\}_{t=n_a}^T$ lying on a collection of hyperplanes $\mathcal{H}_i = \{\mathbf{z} : \mathbf{b}_i^T \mathbf{z} = 0\} \subset \mathbb{R}^K$, for $i = 1, \dots, n$. Furthermore, each hyperplane corresponds to one of the linear models, because the normal to the i^{th} hyperplane, \mathbf{b}_i , encodes the model parameters of the i^{th} ARX model. Hence, switching from one linear model to another is equivalent to jumping from one hyperplane to another. Similarly, staying in the same discrete state is equivalent to drawing samples $\{\mathbf{x}_t\}$ from the same hyperplane.

With this interpretation, the identification problem is reduced to estimating the number of hyperplanes (number of discrete states) and their normals (the model parameters), from sample points (input/output data) lying on those hyperplanes (generated by the PWARX model). In order to do so, we first need to decouple the identification of the model parameters from the filtering of the hybrid state and the identification of the 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 derivatives of the hybrid decoupling constraint. 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 $t \geq n_a$ there exists an i such that $\mathbf{b}_i^T \mathbf{x}_t = 0$. Therefore, the following constraint 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

$$\prod_{i=1}^n (\mathbf{b}_i^T \mathbf{x}_t) = 0. \quad (4)$$

We call equation (4) the *hybrid decoupling constraint (HDC)*, since it will allow us to identify the model parameters $\{\mathbf{b}_i\}_{i=1}^n$ independently from the filtering of the discrete state $\{x_t, \lambda_{t-1}\}$ and regardless of the mechanism generating the transitions (JLS, JMLS, or PWAS). Notice that in the absence of knowledge about the switching mechanism, the HDC in (4) encodes all the information we can obtain from input/output data.

3.2 Identification of the model parameters

Thanks to the HDC we can concentrate on the identification of the number of models n and the model parameters $\{\mathbf{b}_i\}_{i=1}^n$ from input/output data without having to worry about knowing the value of the hybrid state or the type of switching mechanism. The identification problem is then equivalent to solving for the number of discrete states n and the model parameters $\{\mathbf{b}_i\}_{i=1}^n$ from the HDC. To this end, notice that the HDC

$$p_n(\mathbf{z}) \doteq \prod_{i=1}^n (\mathbf{b}_i^T \mathbf{z}) = 0 \quad (5)$$

is a homogeneous polynomial of degree n in $K = n_a + n_c + 1$ variables. Therefore it can be written as

$$p_n(\mathbf{z}) \doteq \sum h_{n_1, \dots, n_K} z_1^{n_1} \cdots z_K^{n_K} = \mathbf{h}^T \boldsymbol{\nu}_n(\mathbf{z}) = 0, \quad (6)$$

where $h_I \in \mathbb{R}$ represents the coefficient of the monomial $\mathbf{z}^I = z_1^{n_1} z_2^{n_2} \cdots z_K^{n_K}$ with $0 \leq n_j \leq n$ for $j = 1, \dots, K$, and $n_1 + n_2 + \cdots + n_K = n$; $\boldsymbol{\nu}_n : \mathbb{R}^K \rightarrow \mathbb{R}^{M_n}$ is the Veronese map of degree n which is defined as:

$$\boldsymbol{\nu}_n : [z_1, \dots, z_K]^T \mapsto [\dots, \mathbf{z}^I, \dots]^T, \quad (7)$$

with I chosen in the degree-lexicographic order; and

$$M_n = \binom{n+K-1}{K-1} = \binom{n+K-1}{n} \quad (8)$$

is the total number of independent monomials. One can show [18] that the vector $\mathbf{h} \in \mathbb{R}^{M_n}$ is simply a vector representation of the symmetric tensor product of the individual model parameters $\{\mathbf{b}_i\}_{i=1}^n$, i.e.

$$\sum_{\sigma \in \mathfrak{S}_n} \mathbf{b}_{\sigma(1)} \otimes \mathbf{b}_{\sigma(2)} \otimes \cdots \otimes \mathbf{b}_{\sigma(n)}, \quad (9)$$

where \mathfrak{S}_n is the permutation group of n elements. We will refer to \mathbf{h} as the *hybrid model parameters*, since it is a combination of the individual model parameters.

If we now apply the HDC (6) to the input/output data $\{\mathbf{x}_t\}_{t=n_a}^T$, then we obtain the following system of linear equations on the hybrid model parameters \mathbf{h}

$$L_n \mathbf{h} \doteq \begin{bmatrix} \nu_n(\mathbf{x}_{n_a})^T \\ \nu_n(\mathbf{x}_{n_a+1})^T \\ \vdots \\ \nu_n(\mathbf{x}_T)^T \end{bmatrix} \mathbf{h} = \mathbf{0} \in \mathbb{R}^{T-n_a+1}. \quad (10)$$

We are now interested in determining under what conditions we can solve for the number of discrete states n and the hybrid model parameters \mathbf{h} from (10). To this end, notice that if n was known, we could immediately recover \mathbf{h} from the linear system $L_n \mathbf{h} = \mathbf{0}$. However, since the linear system in (10) depends explicitly on n , we cannot estimate \mathbf{h} directly without knowing n in advance. The following theorem (see [18]) shows that the estimation of the number of discrete states n is very much related to the conditions under which the solution for \mathbf{h} from (10) is unique (recall that $h_{M_n} = 1$).

Theorem 1 (Number of discrete states) *Given input/output data $\{u_t, y_t\}_{t=0}^T$ generated by a PWARX model, let $L_i \in \mathbb{R}^{(T-n_a+1) \times M_i}$ be the matrix defined in (10), but computed with the Veronese map ν_i of degree i . If T is large enough ($T \geq M_n + n_a - 2$, if n is known) and if the points $\{\mathbf{x}_t\}_{t=n_a}^T$ are in general position on the hyperplanes $\{\mathbf{b}_i^T \mathbf{z} = 0\}_{i=1}^n$ with at least $K-1$ points per hyperplane, then:*

$$\text{rank}(L_i) \begin{cases} > M_i - 1, & i < n, \\ = M_i - 1, & i = n, \\ < M_i - 1, & i > n. \end{cases} \quad (11)$$

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

$$n = \min\{i : \text{rank}(L_i) = M_i - 1\}. \quad (12)$$

Remark 1 (Identifying n and \mathbf{h} from noisy data)

In the presence of noise, we can still solve for the hybrid model parameters \mathbf{h} from (10) in a least-squares sense. We let \mathbf{h} be the eigenvector of $L_n^T L_n$ associated with its smallest eigenvalue, and then normalize \mathbf{h} so that $h_{M_n} = 1$. However, we cannot directly estimate n from (12), because the matrix L_i may be full rank. In the noisy case we declare the rank of L_i to be r if $\sigma_{r+1}/\sigma_r < \epsilon$, where σ_ℓ is the ℓ^{th} singular value of L_i and $\epsilon > 0$ is a pre-specified threshold. We have found this simple criterion to work well in our experiments. One can also use this estimate of n to initialize iterative techniques, such as the one in [4].

Theorem 1 and the linear system in (10) allow us to determine the number of discrete states n and the hybrid model parameters \mathbf{h} , respectively, from input/output data $\{u_t, y_t\}_{t=0}^T$. The rest of the problem is to recover the model parameters $\{\mathbf{b}_i\}_{i=1}^n$ from \mathbf{h} . To this end, let us consider the partial derivative of $p_n(\mathbf{z})$ in (5)

$$Dp_n(\mathbf{z}) = \frac{\partial p_n(\mathbf{z})}{\partial \mathbf{z}} = \sum_{i=1}^n \prod_{\ell \neq i} (\mathbf{b}_\ell^T \mathbf{z}) \mathbf{b}_i. \quad (13)$$

If \mathbf{z} belongs to hyperplane \mathcal{H}_i , then we have $\mathbf{b}_i^T \mathbf{z} = 0$. Therefore, since the K^{th} entry of \mathbf{b}_i is equal to one, after replacing $\mathbf{b}_i^T \mathbf{z} = 0$ into (13) we obtain

$$\mathbf{b}_i = \frac{Dp_n(\mathbf{z})}{e_K^T Dp_n(\mathbf{z})} \Big|_{\mathbf{z} \in \mathcal{H}_i}, \quad (14)$$

where $e_K = [0, \dots, 0, 1]^T \in \mathbb{R}^K$. Therefore, we can estimate the model parameters directly from the derivatives of $p_n(\mathbf{z})$ at a collection of n points $\{\mathbf{z}_i \in \mathcal{H}_i\}_{i=1}^n$ lying on the n hyperplanes. However, since the value of the discrete state λ_i is unknown, we do not know which data points $\{\mathbf{x}_t\}_{t=n_a}^T$ belong to which hyperplane. In order to find the set of points $\{\mathbf{z}_i \in \mathcal{H}_i\}_{i=1}^n$, let us consider a line with base point \mathbf{z}_0 and direction \mathbf{v} , $\mathcal{L} = \{\mathbf{z}_0 + \alpha \mathbf{v}, \alpha \in \mathbb{R}\}$. If $\mathbf{z}_0 \neq \mathbf{0}$, \mathbf{z}_0 is not parallel to \mathbf{v} , and $\mathbf{b}_i^T \mathbf{v} \neq 0$, then the line \mathcal{L} must intersect

the union of all hyperplanes $\cup_{i=1}^n \mathcal{H}_i = \{\mathbf{z} : p_n(\mathbf{z}) = 0\}$ at n distinct points

$$\mathbf{z}_i = \mathbf{z}_0 + \alpha_i \mathbf{v} \in \mathcal{H}_i \cap \mathcal{L} \quad i = 1, \dots, n, \quad (15)$$

where $\{\alpha_i\}$ are the roots of the univariate polynomial

$$q_n(\alpha) = p_n(\mathbf{z}_0 + \alpha \mathbf{v}). \quad (16)$$

We are left with choosing the parameters \mathbf{x}_0 and \mathbf{v} for the line \mathcal{L} . The base point \mathbf{x}_0 can be chosen as any nonzero vector in \mathbb{R}^K . Given \mathbf{z}_0 , the direction \mathbf{v} must be chosen not parallel to \mathbf{z}_0 and such that $\mathbf{b}_i^T \mathbf{v} \neq 0$, for all $i = 1, \dots, n$. Since the latter constraint is equivalent to $p_n(\mathbf{v}) \neq 0$, and p_n is known, we can immediately choose \mathbf{v} even though we do not know the model parameters $\{\mathbf{b}_i\}_{i=1}^n$. We have shown the following:

Theorem 2 (Identifying the model parameters)

Given input/output data $\{u_t, y_t\}_{t=0}^T$, $T \geq M_n + n_a - 2$, generated by a PWARX model with n discrete states, the model parameters $\{\mathbf{b}_i\}_{i=1}^n$ can be computed from the the hybrid model parameters $\mathbf{h} \in \mathbb{R}^{M_n}$ as follows:

1. Choose $\mathbf{z}_0 \neq 0$ and \mathbf{v} such that $\mathbf{v} \neq \gamma \mathbf{z}_0$ and $p_n(\mathbf{v}) \neq 0$.
2. Solve for the n roots $\{\alpha_i\}_{i=1}^n$ of $q_n(\alpha) = p_n(\mathbf{z}_0 + \alpha \mathbf{v})$ and obtain the model parameters $\{\mathbf{b}_i\}_{i=1}^n$ as

$$\mathbf{b}_i = \frac{Dp_n(\mathbf{z}_i)}{e_K^T Dp_n(\mathbf{z}_i)}, \quad (17)$$

where $\mathbf{z}_i = \mathbf{z}_0 + \alpha_i \mathbf{v}$ for $i = 1, \dots, n$.

Hence, the identification of the model parameters can be obtained in closed form if and only if $n \leq 4$, because the only nonlinear step is to solve for the roots of q_n .

Remark 2 (Identifying $\{\mathbf{b}_i\}_{i=1}^n$ from noisy data)

In the presence of noise, we can still estimate the normal vectors $\{\mathbf{b}_i\}_{i=1}^n$ as in Theorem 2. However, the quality of the estimates will depend on the choice of the parameters \mathbf{z}_0 and \mathbf{v} . In this case, one can choose multiple $(\mathbf{z}_0, \mathbf{v})$ satisfying the above conditions, obtain the model parameters for each choice, and let $\{\mathbf{b}_i\}_{i=1}^n$ be the parameters that better reconstruct \mathbf{h} . Alternatively, one can obtain $\{\mathbf{z}_i\}_{i=1}^n$ as points in the data set that minimize a certain distance to the hyperplanes. We refer the interested reader to [18] for further details.

3.3 Filtering of the hybrid state

Given the number of discrete states n and the model parameters $\{\mathbf{b}_i\}_{i=1}^n$, we now show how to reconstruct the hybrid state trajectory $\{x_t, \lambda_{t-1}\}$ from input/output data $\{\mathbf{x}_t\}_{t=n_a}^T$. We first notice that for each time $t \geq n_a$ there exists a generally unique² i such that $\mathbf{b}_i^T \mathbf{x}_t = 0$. Therefore, the discrete state can be trivially identified as:

$$\lambda_{t-1} = \arg \min_{i=1, \dots, n} (\mathbf{b}_i^T \mathbf{x}_t)^2. \quad (18)$$

²In principle, it is possible that a data point \mathbf{x}_t belongs to more than one hyperplane $\mathbf{b}_i^T \mathbf{z} = 0$. However, the set of all such points is a zero measure set on the variety $\{\mathbf{z} : p_n(\mathbf{z}) = 0\}$.

Furthermore, since the parameters of the PWARX model $\{a_\ell(\lambda_{t-1})\}_{\ell=1}^{n_a}$ and $\{c_\ell(\lambda_{t-1})\}_{\ell=1}^{n_c}$ are now known for all $t \geq n_a$, we can find a state-space realization (A_t, B_t, C_t) of the model. Therefore, we can express the output y_t directly as a function of the continuous state at time $t = n_a$, x_{n_a} , and the (known) input u_t as

$$\begin{bmatrix} y_{n_a} \\ y_{n_a+1} \\ y_{n_a+2} \\ \vdots \\ y_t \end{bmatrix} = \begin{bmatrix} C_{n_a} \\ C_{n_a+1} A_{n_a} \\ C_{n_a+2} A_{n_a+1} A_{n_a} \\ \vdots \\ C_t A_{t-1} \cdots A_{n_a} \end{bmatrix} x_{n_a} + \begin{bmatrix} 0 & 0 & \cdots & 0 \\ C_{n_a+1} B_{n_a} & 0 & \cdots & 0 \\ C_{n_a+2} A_{n_a+1} B_{n_a} & C_{n_a+2} B_{n_a+1} & & \\ \vdots & & \ddots & \\ C_t A_{t-1} \cdots A_{n_a+1} B_{n_a} & & & C_t B_{t-1} \end{bmatrix} \begin{bmatrix} u_{n_a} \\ u_{n_a+1} \\ u_{n_a+2} \\ \vdots \\ u_{t-1} \end{bmatrix}. \quad (19)$$

From (19) we can solve for x_{n_a} uniquely, provided that the matrix multiplying x_{n_a} has full rank. Given x_{n_a} , u_t and λ_t , the continuous state trajectory $\{x_t\}$ can be trivially recovered from equation (1).

Remark 3 (No minimum dwell time) Notice that our algorithm does not require that the switching times be separated by a minimum dwell time.

Remark 4 (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 [4, 11] for specific algorithms.

Remark 5 (MIMO systems)

Notice that our algorithm for SISO systems can also be applied to MIMO systems. To see this, let us first consider the case of multiple-input single-output (MISO) systems. In this case one can write the PWARX model as

$$y_t = \sum_{j=1}^{n_a} a_j(\lambda_{t-1}) y_{t-j} + \sum_{j=1}^{n_c} c_j^1(\lambda_{t-1}) u_{t-j}^1 + \cdots + \sum_{j=1}^{n_c} c_j^q(\lambda_{t-1}) u_{t-j}^q,$$

where n_c^j is the degree of input $j = 1, \dots, q$. In this case, one can define the data points and the normals to the hyperplanes as

$$\mathbf{x}_t = [u^{1T}, \dots, u^{qT}, y_{t-n_a}, \dots, y_{t-1}, -y_t]^T \in \mathbb{R}^K \text{ and } \mathbf{b}_i = [\mathbf{c}^{1T}(i), \dots, \mathbf{c}^{qT}(i), a_{n_a}(i), \dots, a_1(i), 1]^T \in \mathbb{R}^K,$$

where $K = n_a + \sum_{j=1}^q n_c^j + 1$, $u^j = [u_{t-n_c^j}^j, \dots, u_{t-1}^j]^T \in \mathbb{R}^{n_c^j}$ and $\mathbf{c}^j(i) = [c_{n_c^j}^j(i), \dots, c_1^j(i)]^T \in \mathbb{R}^{n_c^j}$ for $j = 1, \dots, q$ and $i = 1, \dots, n$. Therefore, if the dimensions of the PWARX model n_a and n_c^j are known for all $j = 1, \dots, q$, then the algorithm for SISO systems can also be applied to MISO systems. Furthermore, in the case of MIMO systems one can apply the MISO algorithm to each one of the q outputs independently.

4 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 following ARX model with zero-mean additive Gaussian noise w_t

$$y_t = a_1(\lambda_{t-1})y_{t-1} + a_2(\lambda_{t-1})y_{t-2} + c_1(\lambda_{t-1})u_{t-1} + w_t.$$

The dimension of the state space is $n_a = 2$, the degree of the input is $n_c = 1$ and the discrete state is $\lambda_t \in \{1, 2, 3\}$. In each one of the 1000 trials, 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

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

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. The standard deviation σ of w_t was chosen in the range $[0, 0.01]$ to simulate a measurement error of about 1%.

We compare the polynomial factorization algorithm (PFA) of [21], which identifies the model parameters based on algebraically factoring the polynomial $p_n(z)$, with our polynomial differentiation algorithm (PDA), which identifies the parameters from the derivatives of $p_n(z)$. Figure 1 shows the mean error on the estimation of the model parameters,³ the continuous state,⁴ and the discrete state,⁵ respectively, as a function of σ .⁶ Both the model parameters and the continuous state are perfectly estimated when $\sigma = 0$. For $\sigma > 0$, the estimation error increases approximately linearly with the amount of noise. Notice that the discrete state is incorrectly estimated approximately 7 – 8% of the times for $\sigma = 0.01$. Notice also that PDA performs better than PFA, especially in the estimation of the continuous state where it achieves about half of the error.

³The 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 the number of trials.

⁴The error between the estimated continuous state \hat{x}_t and the true continuous state x_t was computed as $\|\hat{x}_t - x_t\|$, averaged over the number of data points and the number of trials.

⁵The error between the estimated discrete state $\hat{\lambda}_t$ and the true discrete state λ_t was computed as the percentage of times in which $\hat{\lambda}_t \neq \lambda_t$, averaged over the number of trials.

⁶In all the trials, the number of discrete states was correctly estimated as $n = 3$ by using a threshold of $\epsilon = 3 \times 10^{-3}$ on the singular values of L_n to compute its rank (see Remark 1).

Figures 2 and 3 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. However, the continuous state is estimated with a small error throughout the whole time interval, in spite of the errors in the identification of the discrete state.

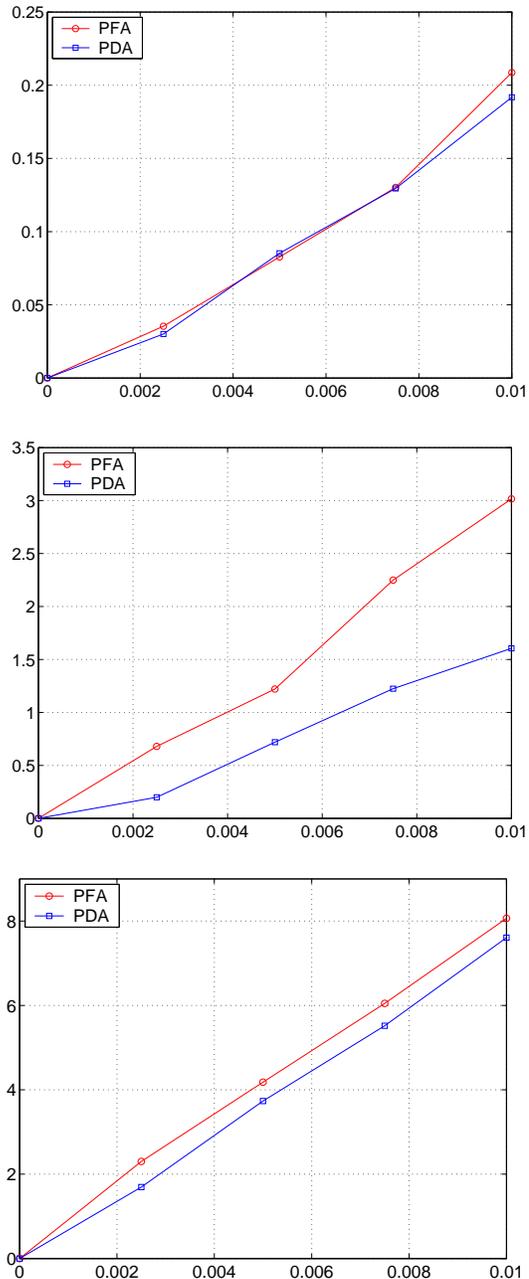


Figure 1: Mean error over for the identification of the model parameters (top), the continuous state (middle) and the discrete state (bottom) as a function of the standard deviation of the measurement error σ . Since neither PFA nor PDA impose stability constraints on the estimation of the parameters of the JLS, about 150 trials yield unstable systems. We considered those 150 trials as outliers and computed the mean error over the remaining 850 trials.

5 Conclusions and open issues

We have proposed an algebraic geometric solution to the identification of a class of linear hybrid systems. We showed that the identification of the model parameters can be decoupled from the inference of the hybrid state and from the switching mechanism generating the transitions. By representing the number of discrete states n as the degree of a polynomial p and the model parameters as factors of p , we casted the identification problem as a polynomial factorization problem, which we solved using simple linear-algebraic techniques. We showed that one can estimate n from a rank constraint on the data, the coefficients of p from a linear system, and the model parameters from the derivatives of p . We presented simulation results evaluating the performance of the algorithm with noisy data. Open issues include a detailed analysis of the robustness of the algorithm with noisy data, as well as extending it to situations in which the dimension of the state space is unknown, or the models have different dimensions. Other extensions include imposing stability or causality constraints on the estimation of the model parameters.

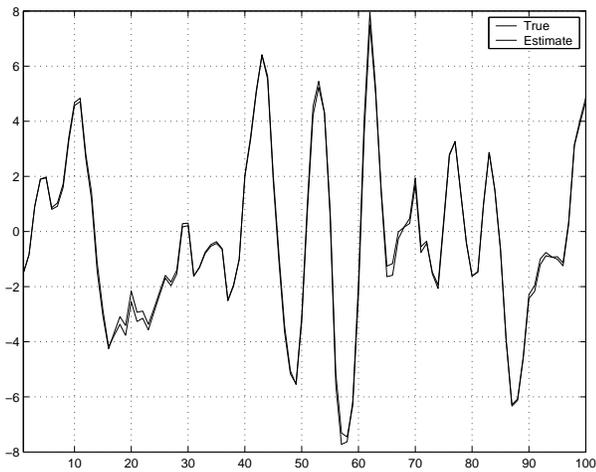


Figure 2: Evolution of the first entry of the continuous state x_t and its estimate \hat{x}_t for $0 \leq t \leq 100$.

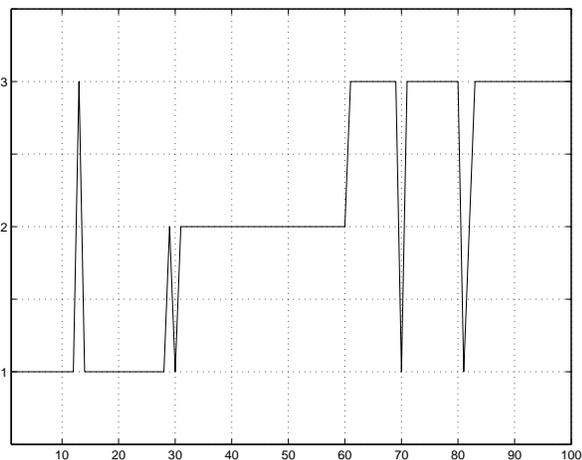


Figure 3: Evolution of the estimated discrete state $\hat{\lambda}_t$.

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] 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.
- [4] A. Bemporad, A. Garulli, S. Paoletti, and A. Vicino. A greedy approach to identification of piecewise affine models. In *Hybrid Systems: Computation and Control*, LNCS, pages 97–112. Springer Verlag, 2003.
- [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, D. Mignone, and M. Morari. Moving horizon estimation for hybrid systems. In *Proceedings of the American Control Conference*, pages 1684–1688, 2000.
- [11] G. Ferrari-Trecate, M. Muselli, D. Liberati, and M. Morari. A clustering technique for the identification of piecewise affine systems. *Automatica*, 39(2):205–217, 2003.
- [12] Z. Ghahramani and G. E. Hinton. Variational learning for switching state-space models. *Neural Computation*, 12(4):963–996, 1998.
- [13] K. Murphy. Switching Kalman filters. Technical report, U. C. Berkeley, 1998.
- [14] 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.
- [15] A. Sun, S. S. Ge, and T. H. Lee. Controllability and reachability criteria for switched linear systems. *Automatica*, 38:775–786, 2002.
- [16] 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.
- [17] J. K. Tugnait. Detection and estimation for abruptly changing systems. *Automatica*, 18(5):607–615, 1982.
- [18] R. Vidal. *Generalized Principal Component Analysis (GPCA): an Algebraic Geometric Approach to Subspace Clustering and Motion Segmentation*. PhD thesis, University of California, Berkeley, August 2003.
- [19] 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.
- [20] R. Vidal, A. Chiuso, S. Soatto, and S. Sastry. Observability of linear hybrid systems. In *Hybrid Systems: Computation and Control*, LNCS, pages 526–539. Springer Verlag, 2003.
- [21] R. Vidal, Y. Ma, and S. Sastry. Generalized principal component analysis. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2003.