16th IFAC Symposium on System Identification
The International Federation of Automatic Control
Brussels, Belgium. July 11-13, 2012

A survey on switched and piecewise affine system identification

Andrea Garulli * **Simone Paoletti** * **Antonio Vicino** *

* *Dipartimento di Ingegneria dell'Informazione, Università di Siena, Via Roma 56, 53100 Siena, Italy*

Abstract: Recent years have witnessed a growing interest on system identification techniques for switched and piecewise affine models. These model classes have become popular not only due to the universal approximation properties of piecewise affine functions, but also because the proposed identification procedures have proven to be effective in problems involving complex nonlinear systems with large data sets. This paper presents a review of recent advances in this research field, including theoretical results, algorithms and applications.

1. INTRODUCTION

Switched affine systems are defined as collections of dynamical affine systems indexed by a discrete-valued additional variable, called the discrete state. The discrete state can be an exogenous finite-valued input, and in this case one talks of jump affine systems. On the other hand, PieceWise Affine (PWA) systems are a special class of switched affine systems, obtained by partitioning the state-input domain into a finite number of polyhedral regions, and by considering linear/affine dynamics in each region [Sontag, 1981]. The input-output counterpart of switched affine systems is represented by Switched ARX (SARX) systems. A SARX system consists of a collection of ARX systems and a switching signal which selects the ARX dynamics active at each time instant. PieceWise affine ARX (PWARX) systems are a subclass of SARX systems representing the input-output version of PWA systems, with the specification that the switching mechanism is determined by a polyhedral partition of the regressor space.

Switched and PWA models, both in state-space and in input-output form, can be used to describe hybrid and nonlinear phenomena that are frequent in practical situations, e.g., when a system may work in different modes of operation, or the system dynamics changes due to physical limits, dead-zones, switches and thresholds. In addition, since PWA maps have universal approximation properties, PWA models can be used to approximate nonlinear systems that do not exhibit discontinuous or switching behavior. Finally, PWA models are equivalent to several classes of hybrid models, and can therefore be used to describe systems exhibiting hybrid structure. For the aforementioned reasons, identification of both switched and PWA systems has been widely investigated in recent years, starting from the large body of subspace, statistical and set membership tools available in the literature for dealing with linear models. A vast majority of the proposed methods is concerned with the identification of SARX and PWARX models, and this has given raise also to a strong interest for the study of the realization theory of switched affine models, with a specific focus on minimal state-space realizations.

The objective of this paper is to provide an outline of the extremely intense efforts directed towards the development of identification methods for switching models in the last decade. We will make specific reference to results obtained in most recent years, assuming as starting points the literature review in [Roll, 2003] and the tutorial paper [Paoletti et al., 2007], from which there emerge the main approaches developed at the first stages of the research on the topic. An algebraic approach was proposed in [Vidal et al., 2003], providing an exact solution to the identification problem of SARX systems in a noise-free setting. The technique proposed by Roll et al. [2004] relies on the formulation of the identification problem as a mixed integer program (MIP). Although in principle this problem can be solved for the global optimum, it turns out to be tractable only for small data sets. Clustering-based methods were proposed in [Ferrari-Trecate et al., 2003, Nakada et al., 2005], while a Bayesian approach was proposed in [Juloski et al., 2005]. Finally, a technique based on bounded error identification tools was proposed by Bemporad et al. [2005]. The common feature of the last three types of approaches is that they lead to suboptimal solutions to the problem of inferring a PWARX model from data, while keeping an affordable computational burden. Software implementations of the clustering-based method [Ferrari-Trecate et al., 2003], MIP-based method [Roll et al., 2004] and bounded-error method [Bemporad et al., 2005] are available online [Ferrari-Trecate, 2005, Paoletti and Roll, 2007].

From the literature review presented in this survey paper, there emerges a clear research trend towards the use of recently developed, efficient optimization and relaxation tools to tackle the identification problem for switched and PWA systems. In this sense, the tradeoff between computational complexity and quality of the suboptimal solution is still a key issue driving the research in this field. Other emerging trends are the development of efficient identification methods for block-structured systems with PWA nonlinearities, and the application or adaptation of the available methods in specific fields such as computer vision, systems biology, electromechanical and automotive systems.

978-3-902823-06-9/12/$20.00 © 2012 IFAC

10.3182/20120711-3-BE-2027.00332

The paper is organized as follows. Section 2 provides basic notations and definitions on switched systems in state-space and input-output form, with a note on realization theory for these model classes. In Section 3 the identification problems for the different switched model structures are formulated, and some results on identifiability and experiment design are described. Section 4 focuses on the identification methods developed mainly in the last five years, and proposes a classification of these methods based on the model structure and the tools they use. In Section 5 several applications of switched and PWA system identification techniques in different fields are presented. Concluding remarks are reported in Section 6.

2. SYSTEM REPRESENTATIONS

In this section we recall the basic definitions of switched and PWA systems both in state-space and input-output form. We also mention some recent results on realization theory for switched and PWA systems.

2.1 Systems in state-space form

A switched affine system is a collection of affine systems sharing the same continuous state. Switches between affine systems are indexed by a discrete-valued signal called the discrete state. It follows that a discrete-time switched affine system in *state-space* form is described by the equations

$$x(t+1) = A_{\sigma(t)}\, x(t) + B_{\sigma(t)}\, u(t) + f_{\sigma(t)} + w(t) \tag{1a}$$

$$y(t) = C_{\sigma(t)}\, x(t) + D_{\sigma(t)}\, u(t) + g_{\sigma(t)} + v(t), \tag{1b}$$

where $x(t) \in \mathbb{R}^n$, $u(t) \in \mathbb{R}^p$ and $y(t) \in \mathbb{R}^q$ are, respectively, the (*continuous*) state, the input and the output of the system at time $t \in \mathbb{Z}$, and $w(t) \in \mathbb{R}^n$ and $v(t) \in \mathbb{R}^q$ are noise terms. The *discrete* state $\sigma(t) \in \{1, \ldots, s\}$ selects the affine dynamics (or *mode*) which is active at time t, with s denoting the number of modes. The real matrices and vectors A_i, B_i, f_i, C_i, D_i and g_i, $i = 1, \ldots, s$, having appropriate dimensions, describe each affine dynamics.

In general, the discrete state $\sigma(t)$ can be either an exogenous input, or a function of the system continuous state $x(t)$ and input $u(t)$. When referring to general switched (or *jump*) affine models, $\sigma(t)$ is typically intended an unknown, deterministic input. A special case are *jump-Markov* affine models, where the dynamics of $\sigma(t)$ is modeled as an irreducible Markov chain with transition probabilities $p_{i,j}(t) = P\big(\sigma(t+1) = j \,\big|\, \sigma(t) = i\big)$.

In PWA systems [Sontag, 1981], $\sigma(t)$ is determined by a polyhedral partition of the state and input space, according to the rule:

$$\sigma(t) = i \quad \Longleftrightarrow \quad \left[\begin{smallmatrix} x(t) \\ u(t) \end{smallmatrix}\right] \in \Omega_i, \tag{2}$$

where $\{\Omega_i\}_{i=1}^s$ are convex polyhedra defining a complete polyhedral partition of the domain of validity $\Omega \subseteq \mathbb{R}^{n+p}$ of the system.

2.2 Systems in input-output form

A Switched ARX (SARX) system is a collection of ARX systems sharing the same regression vector. As for switched systems in state-space form, the system is equipped with a switching signal selecting the ARX dynamics which is active at each time instant. For fixed model orders n_a and n_b, SARX systems are defined by introducing the regression vector

$$\begin{aligned} r(t) = [\, y(t-1)^T \; \ldots \; y(t-n_a)^T \\ u(t)^T \; u(t-1)^T \; \ldots \; u(t-n_b)^T \,]^T, \end{aligned} \tag{3}$$

and then by expressing the output $y(t)$ as

$$y(t) = \theta_{\sigma(t)}^T \left[\begin{smallmatrix} r(t) \\ 1 \end{smallmatrix}\right] + e(t), \tag{4}$$

where $\sigma(t) \in \{1, \ldots, s\}$ is the discrete state, s is the number of modes, θ_i are the matrices of parameters defining each ARX dynamics, and $e(t) \in \mathbb{R}^q$ is a noise term. In the following, the vector $\varphi(t) = [\, r(t)^T \; 1 \,]^T$ will be referred to as the *extended* regression vector.

PieceWise affine ARX (PWARX) systems are a special case of SARX systems for which the switching mechanism is determined by a polyhedral partition of the regressor domain $\mathcal{R} \subseteq \mathbb{R}^d$, where $d = q \cdot n_a + p \cdot (n_b + 1)$. This means that for these models the discrete state $\sigma(t)$ is given by

$$\sigma(t) = i \quad \Longleftrightarrow \quad r(t) \in \mathcal{R}_i, \tag{5}$$

where $\{\mathcal{R}_i\}_{i=1}^s$ are convex polyhedra defining a complete polyhedral partition of $\mathcal{R}$. In general, the shape of $\mathcal{R}$ reflects the physical constraints on the system inputs and outputs. For instance, typical constraints on the output could be $\|y(t)\|_\infty \leq y_{max}$ or $\|y(t) - y(t-1)\|_\infty \leq \Delta y_{max}$, where $\|\cdot\|_\infty$ is the supremum norm, and y_{max} and Δy_{max} are fixed bounds. Note that, by introducing the PWA map

$$f(r) = \begin{cases} \theta_1^T \varphi & \text{if } r \in \mathcal{R}_1 \\ \vdots & \vdots \\ \theta_s^T \varphi & \text{if } r \in \mathcal{R}_s, \end{cases} \tag{6}$$

with $\varphi = [\, r^T \; 1 \,]^T$, the PWARX model (4) and (5) takes the form of a nonlinear ARX (NARX) system, i.e.:

$$y(t) = f(r(t)) + e(t). \tag{7}$$

Hinging-Hyperplane ARX (HHARX) systems are a subclass of PWARX systems (7) for which the PWA map (6) is continuous. Their definition relies on the concept of hinge functions [Breiman, 1993]. A HH function is defined as:

$$f(r) = \theta_0^T \varphi + \sum_{i=1}^{M} \sigma_i \max\{0, \theta_i^T \varphi\}, \tag{8}$$

where $\sigma_i \in \{-1, 1\}$, θ_i are parameter vectors (here $q = 1$), and M is the number of hinge functions. Though HH functions have universal approximation properties with respect to all continuous functions on a compact subset of $\mathbb{R}^d$, they are not a universal representation of all continuous PWA functions. An extension of the HH functions was proposed by Wen et al. [2005], who defined the concept of PWA basis functions. For continuous PWA functions of d variables, a PWA basis function is either the max or min of $d+1$ affine functions. It is shown that any continuous PWA function of d variables can be expressed as the sum of a suitable number of PWA basis functions. Being an extension of HH functions, the PWA basis function representation enjoys the same universal approximation properties. Based on the PWA basis function representation, Basis function PWARX (BPWARX) systems are presented in [Wen et al., 2007] as systems (7) with

$$f(r) = \theta_0^T \varphi + \sum_{i=1}^{M} \sigma_i \max\{0, \theta_{i,1}^T \varphi, \ldots, \theta_{i,d}^T \varphi\}. \tag{9}$$

Finally, we introduce a switched Output-Error (OE) system as a collection of OE systems sharing the same regression vector. Switched OE systems are defined by the equations:

$$y(t) = z(t) + e(t) \tag{10a}$$

$$z(t) = \theta_{\sigma(t)}^T \begin{bmatrix} r(t) \\ 1 \end{bmatrix}, \tag{10b}$$

where

$$\begin{aligned} r(t) = [\ & z(t-1)^T \ \dots \ z(t-n_a)^T \\ & u(t)^T \ u(t-1)^T \ \dots \ u(t-n_b)^T \ ]^T, \end{aligned} \tag{11}$$

$\sigma(t) \in \{1, \dots, s\}$ is the discrete state, s is the number of modes, θ_i are the matrices of parameters defining each affine dynamics, and $e(t) \in \mathbb{R}^q$ is a noise term.

2.3 Realization theory

Recently there has been a growing interest in realization theory for switched and PWA systems. This is motivated by the fact that algorithms and/or formulae to convert models of a given class into representations in another model class are useful to transfer theoretical properties and tools between classes. For instance, one could identify a PWA model in input-output form, convert it in state-space form, and then apply to the state-space model the tools available for analysis and control design.

Converting a SARX or a PWARX system into an equivalent state-space representation is straightforward. A simple way to construct such an equivalent representation is to define the state as

$$\begin{aligned} x(t) = [\ & y(t-1)^T \ \dots \ y(t-n_a)^T \\ & u(t-1)^T \ \dots \ u(t-n_b)^T \ ]^T, \end{aligned} \tag{12}$$

by eliminating the term $u(t)$ in the regression vector (3). However, the realization thus obtained might not be minimal. In [Petreczky et al., 2011] a characterization of minimality of discrete-time switched linear systems in terms of reachability and observability is presented, and it is shown that minimal realizations are unique up to isomorphisms. Moreover, the authors discuss procedures for converting a switched linear system to a minimal one using results from the theory of rational formal power series in non-commutative variables. Realization theory for switched linear systems with state resets is elaborated in [Petreczky and van Schuppen, 2009].

The inverse problem of converting a switched affine system from state-space to input-output form (the input-output realization problem) has been more thoroughly studied. Preliminary results in [Weiland et al., 2006] show that every observable switched affine system admits a SARX representation. A necessary and sufficient condition for input-output realization of switched affine systems is given in [Paoletti et al., 2008].

Concerning PWA systems, to the best of the authors' knowledge a complete realization theory is still missing. The problem of finding a PWA system which realizes a specified output trajectory has been investigated in [Petreczky, 2006] for autonomous systems. The input-output realization problem for PWA state-space systems was firstly addressed by Rosenqvist and Karlström [2005], who derived sufficient conditions for the existence of an input-output realization of a subclass of PWA systems. Necessary and sufficient conditions for existence of equivalent PWARX representations of a given PWA state-space system were derived in [Paoletti et al., 2010]. The same paper presents a constructive procedure to derive the structure of an equivalent PWARX representation, provided it exists. It is also shown that the number of modes and the number of parameters can grow considerably when a PWA state-space system is converted into a minimum-order equivalent PWARX representation. Since there exist PWA state-space systems which do not admit an equivalent PWARX representation, while the conversion from PWARX to PWA state-space is always possible, one can conclude that the class of PWARX systems is strictly contained in the class of PWA systems. This may have interesting implications from the point of view of the identification of these systems.

3. IDENTIFICATION PROBLEMS

In this section we state the identification problems for switched and PWA systems. We also report some recent contributions on identifiability and experiment design for these classes of systems, since both issues are closely related to identification.

3.1 Identification problem for SARX systems

For SARX models (4), the general identification problem reads as follows.

Problem 3.1. Given a collection of N input-output pairs $\big(y(t), u(t)\big)$, $t = 1, \dots, N$, estimate the model orders n_a and n_b, the number of modes s, and the parameters θ_i, $i = 1, \dots, s$. Moreover, reconstruct the discrete state $\sigma(t)$ for $t > \max\{n_a, n_b\}$.

In practice, the difficulty of Problem 3.1 depends on which quantities are assumed to be known or fixed a priori. If the model orders n_a and n_b are fixed, the problem is to fit the data to s hyperplanes. In this case, the discriminant between different identification methods is whether the number of modes (or hyperplanes) s is estimated from data or fixed a priori.

As will be discussed in Section 4.1, most recent contributions address Problem 3.1 either in an algebraic or an optimization framework. The first category of approaches stems from the original work by Vidal et al. [2003] and its subsequent generalizations, providing an exact algebraic solution to Problem 3.1 for the case of noiseless data. The second category of approaches addresses Problem 3.1 by exploiting recent advances in optimization and relaxation techniques. To better understand and appreciate these approaches, it is worthwhile to recall the following facts (assume n_a and n_b to be fixed).

1. If s is fixed a priori, the classical prediction error approach [Ljung, 1999] can be in principle applied to the identification of SARX models. Given a nonnegative function $\ell(\cdot)$, such as $\ell(\varepsilon) = \varepsilon^2$ or $\ell(\varepsilon) = |\varepsilon|$, this implies to estimate the parameters θ_i's and the discrete state $\sigma(t)$ by solving the following optimization problem:

$$\begin{cases} \min\limits_{\theta_i,\, \chi_{t,i}} \displaystyle\sum_t \sum_{i=1}^{s} \ell\big(y(t) - \varphi(t)^T\theta_i\big)\chi_{t,i} \\ s.t. \quad \displaystyle\sum_{i=1}^{s} \chi_{t,i} = 1 \qquad \forall\, t \\ \qquad\quad \chi_{t,i} \in \{0,1\} \quad \forall\, t, i. \end{cases} \tag{13}$$

In (13), each binary variable $\chi_{t,i}$ determines whether the data point $\big(y(t), r(t)\big)$ is uniquely attributed to the ith submodel. The discrete state $\sigma(t)$ can be finally reconstructed according to the rule:

$$\sigma(t) = i \iff \chi_{t,i} = 1. \tag{14}$$

Unfortunately, the optimization problem (13) is a *mixed integer* program that is computationally intractable, except for small instances. This motivates the investigation of suitable relaxations of (13) that can be solved efficiently and provide solutions close to the original problem.

2. If s is not fixed a priori, and is considered a free variable in (13), one would get perfect fit by associating a submodel to each data point, but this overfit model would have no generalization ability, and therefore would not be very useful. Therefore, when s must be estimated from data, penalties on increasing s should be introduced in order to find a suitable tradeoff between accuracy (in terms of prediction error) and complexity (in terms of number of modes). In this respect, different methods can be distinguished depending on the way the penalties are introduced.

3.2 Identification problem for PWARX systems

For PWARX models defined by (4) and (5), the general identification problem reads as follows.

Problem 3.2. Given a collection of N input-output pairs $\big(y(t), u(t)\big)$, $t = 1, \ldots, N$, estimate the model orders n_a and n_b, the number of modes s, the parameters θ_i and the regions $\mathcal{R}_i$, $i = 1, \ldots, s$.

It is worthwhile to note that, once the partition of the regressor domain is available, the estimation of the discrete state $\sigma(t)$ can be done by applying the rule (5).

As for the identification of SARX models, the difficulty of Problem 3.2 depends on which quantities are assumed to be known or fixed a priori. We note that most techniques specifically developed for the identification of PWARX models, assume fixed model orders n_a and n_b. Nevertheless, also under this assumption, Problem 3.2 still remains very hard. In principle, recalling that PWARX systems are a particular case of SARX systems, one could try to formulate the identification problem for PWARX systems by mimicking (13). However, this optimization problem, which is already prohibitive for SARX systems, shows an additional difficulty in the case of PWARX systems because it is not easy to express efficiently the constraint that the regions $\{\mathcal{R}_i\}_{i=1}^{s}$ must form a complete partition of the regressor domain $\mathcal{R}$. For this reason, most methods for the identification of PWARX systems either discard or soften this constraint during the phase of data classification and parameter estimation, and consider it explicitly only in the final phase of region estimation.

Concerning the estimation of the number of modes s, most methods either assume a fixed s or adjust it iteratively in order to improve the fit. Few techniques allow for the automatic estimation of s from data.

We conclude by recalling that Problem 3.2 becomes easy if the polyhedral partition of the regressor domain is either known or fixed a priori. In this case each regression vector $r(t)$ can be associated to one submodel according to (5). Hence, by introducing the quantities

$$\chi_{t,\,i} = \begin{cases} 1 \text{ if } r(t) \in \mathcal{R}_i \\ 0 \text{ otherwise} \end{cases} \quad \forall\, t, i, \tag{15}$$

under the choice $\ell(\varepsilon) = \varepsilon^2$ the optimization problem (13) boils down to a standard least-squares problem in the unknowns θ_i.

3.3 Identification problem for state space models

For switched affine models defined by (1), or PWA models defined by (1) and (2), the general identification problem reads as follows.

Problem 3.3. Given a collection of N input-output pairs $\big(y(t), u(t)\big)$, $t = 1, \ldots, N$, estimate the model order n, the number of modes s, the 6-tuples $(A_i, B_i, f_i, C_i, D_i, g_i)$, $i = 1, \ldots, s$, the discrete state $\sigma(t)$, $t = 1, \ldots, N$, and, if the model is PWA, the regions Ω_i, $i = 1, \ldots, s$.

As for the models in input-output form, the difficulty of Problem 3.3 depends on which quantities are assumed to be known. Conversely, even when only the model parameters have to be identified (i.e. the model order n, the number of modes s and the switching sequence $\sigma(t)$ are known), Problem 3.3 is still not trivial. As pointed out by Verdult and Verhaegen [2004], though in such a situation the 6-tuples $(A_i, B_i, f_i, C_i, D_i, g_i)$ can be recovered by applying, e.g., subspace identification methods to the data classified in each mode, the 6-tuples of different submodels are isomorphic up to a linear state transformation. To combine the submodels, they need to be transformed into the same state basis, which is an involved task.

The main discriminant between different methods for the identification of switched and PWA models in state-space form is concerned with the requirement of minimum dwell-time, i.e. the minimum sojourn time that the system is required to spend in each mode. This mainly implies that methods constrained with dwell-time are not applicable in case of fast switching physical systems. To the best of the authors' knowledge, the requirement of minimum dwell-time is never adopted in the identification of switched and PWA systems in input-output form.

3.4 Identifibiality

Identifiability is a key issue in system identification, related to the question whether the attempt to infer a given parameterized system from noise-free data is a well-posed problem. The answer to this question has a number of implications in several aspects of system identification, such as experiment design, and the development and analysis of new identification methods. Nevertheless, in spite of the increasing attention deserved by switched and PWA system identification, few results on the identifiability of these classes of systems can be found in the existing literature. Vidal et al. [2002] consider autonomous switched linear systems in state-space form, and study the identifiability of the model parameters by characterizing the set of models that produce the same output measurements. When the data are generated by a model in the class, conditions under which the true model can be identified are given. Recently, Petreczky et al. [2010] presented necessary and sufficient conditions guaranteeing structural identifiability for switched linear systems in state-space form. Compared to [Vidal et al., 2002], systems with inputs are considered, and identifiability with respect to the whole input-output behavior (not only to a finite set of measurements) is investigated.

3.5 Experiment design

Experiment design is another important issue in system identification, which so far has not received much attention for switched and PWA systems. In particular, for PWA systems the choice of the input signal for identification should be such that not only all the affine dynamics are sufficiently excited, but also accurate shaping of the boundaries of the regions is possible. In this context, a very recent contribution has been proposed by Suzuki and Yamakita [2011], who propose a method to generate regression vectors (3) close the boundaries of the regions $\mathcal{R}_i$ by optimizing a suitable cost function. Since the cost function requires information about the boundaries of the regions, an iterative identification algorithm is proposed, which updates the model and the cost function alternatively. For SARX models, a persistence of excitation condition is derived in [Vidal, 2008]. Such a condition guarantees the exponential convergence of an algebraic recursive identification algorithm for SARX systems.

4. METHODS

In this section we present an overview of recent contributions in the framework of switched and PWA system identification. The focus is mainly on what has been proposed in the last five years. For the state-of-the-art prior to 2007, the interested reader is referred to [Roll, 2003] and [Paoletti et al., 2007]. The identification methods are mainly classified based on the model class they are devised for. Therefore we distinguish methods for the identification of SARX, PWARX and switched affine state-space systems. Tables 1–3 list the methods reviewed for the three mentioned model classes. In each table, the first column shows the keywords which characterize the methods listed in the same row. We also present some contributions related to switched output-error systems, and block-structured systems with PWA nonlinearities.

4.1 SARX systems

Most recent contributions related to the identification of SARX systems can be classified as optimization-based, algebraic and recursive. One contribution falling outside these categories is proposed by Porreca and Ferrari-Trecate [2010], who study the partitioning of data sets based on equalities among parameters, and apply their method to the problem of discovering which of the data sets delimited by consecutive switchings (assumed known) have been generated by the same mode of operation of a SARX system.

Table 1. Methods for SARX systems

Optimization-based methods	
Continuous optimization	Lauer et al. [2009, 2011]
Sparse optimization	Ozay et al. [2008], Bako [2011]
Particle swarm optimization	Maruta et al. [2011]
Semi-definite optimization	Ozay et al. [2009]
Sum-of-norms regularization	Ohlsson et al. [2010]
Algebraic methods	
Errors-in-variables	Nazari et al. [2011]
Multi-output systems	Bako and Vidal [2008]
Recursive methods	
Algebraic	Vidal [2008]
Two-step	Bako et al. [2011]
Non-supervised classification	Ayad et al. [2010]
Other methods	
Data set partitioning	Porreca and Ferrari-T. [2010]

Optimization-based methods. Most methods for the identification of SARX systems proposed in the last five years are based either on relaxations of the optimization problem (13) or alternative formulations of it.

In Lauer et al. [2009, 2011] an optimization framework that relies on the minimization of either the minimum or the product of loss functions is proposed. This framework includes the algebraic method [Vidal et al., 2003] and support vector regression method [Lauer and Bloch, 2008a,b] as particular cases. The main feature of the proposed approach is that the identification problem is recast as a continuous (nonlinear and nonconvex) optimization program involving only the model parameters as variables, thus avoiding the use of mixed-integer optimization.

A sparse optimization approach to the identification of SARX models is presented in [Bako, 2011]. This work exploits very recent developments from the community of compressed sensing. The problem of identifying each ARX submodel is formally posed as a combinatorial ℓ_0 optimization problem. This NP-hard problem is then relaxed into a (convex) ℓ_1-norm minimization problem. Sufficient conditions for this relaxation to be exact are presented.

Also Maruta et al. [2011] reduce the SARX system identification problem to an optimization one. Since this problem turns out to be inherently ill-conditioned and nonconvex, they develop a new technique based on particle swarm optimization to avoid being trapped in local minima.

Identification of SARX systems in a set membership framework is addressed by Ozay et al. [2008, 2009]. In this framework, given a data set of noisy input/output data and the available prior information about the feasible system set (e.g. a bound on the number of modes), the goal is to identify a suitable set of ARX models and a switching sequence that are compatible with the available experimental information, while minimizing the number of modes. In [Ozay et al., 2008] the set membership identification problem is reduced to a sparsification form, where the goal is to maximize sparsity of a given vector sequence. Efficient convex relaxations of this NP-hard problem are obtained by exploiting recent results on sparse signal recovery. The contribution in [Ozay et al., 2009] has its starting point in the algebraic method [Vidal et al., 2003] developed for the noise-free case. In the presence of norm-bounded noise, the algebraic procedure leads to a very challenging nonconvex polynomial optimization problem. In [Ozay et al., 2009] this problem is translated into a rank minimization problem which, resorting to suitable convex relaxations, leads to an overall semi-definite optimization problem that can be efficiently solved.

Finally, Ohlsson et al. [2010] study the segmentation of ARX models by formulating this problem as a least-squares problem with sum-of-norms regularization over the model parameter jumps. This formulation turns out to be advantageous with respect to other methods in case of considerable noise and infrequent switches. The proposed method has one tuning parameter, the regularization constant, which is used to trade-off fit and number of modes.

Algebraic methods. Algebraic methods for the identification of SARX systems stem from the original work by Vidal et al. [2003], which has reached a comprehensive treatment in [Vidal, 2008]. The key idea underlying the algebraic approach is to view the identification of multiple ARX models as the identification of a single, "lifted" ARX model which does not depend on the switching sequence. The parameters of the "lifted" ARX model can be identified by applying a polynomial embedding to the input-output data. The parameters of the original ARX submodels are then recovered from the derivatives of this polynomial at different regressors.

The algebraic approach provides the exact solution to the identification of SARX systems in the noise-free case, but may provide poor results in the presence of noise. This problem is considered by Nazari et al. [2011], who propose an improved solution by exploiting the signal-to-noise ratio in the estimation of the highest order of the submodels.

In [Bako and Vidal, 2008] the algebraic approach, originally developed for single-output systems, is extended to the identification of multi-output SARX models with unknown number of modes of unknown and possibly different orders. Preliminary projection of the input-output data onto a low-dimensional linear subspace is proposed in order to cope with the exponential growth of the number of parameters to be identified with the number of outputs and the number of modes.

Recursive methods. Recursive identification of SARX models is addressed by Vidal [2008] in an algebraic framework. In the proposed method, the number of modes, the model orders and the switching sequence are all assumed to be unknown.

In [Bako et al., 2011], a recursive procedure alternating between data assignment to submodels and parameter update is proposed. At each time instant, the discrete state is determined as the index of the submodel that, in terms of the prediction error, appears to have most likely generated the observed regression vector. Then, given the estimated discrete state, the associated parameter vector is updated via recursive least squares.

Online estimation of the number of modes of a SARX system is addressed in [Ayad et al., 2010]. The proposed approach uses a non-supervised classification method to determine online the number of modes as well as the corresponding submodels.

Table 2. Methods for PWARX systems

Optimization-based methods	
Sum-of-norms regularization	Ohlsson and Ljung [2011]
Weighted least-squares	Lai et al. [2010a,b,c]
ℓ_1 optimization	Maruta and Sugie [2011]
Expectation maximization	Jin and Huang [2010]
PrARX models	Taguchi et al. [2009]
BPWARX models	Wen et al. [2007]
Clustering-based methods	
Dempster-Shafer theory	Boukharouba et al. [2009]
Competitive learning	Gegúndez et al. [2008]
Split-and-merge clustering	Baptista et al. [2011]
k-plane clustering	Tabatabaei-Pour et al. [2006b]
Correlation clustering	Ivanescu et al. [2011]
Other methods	
Recursive two-step	Bako et al. [2011]
Bounded-error	Tabatabaei-Pour et al. [2006a]
Adaptive hinging hyperplanes	Xu et al. [2009]

4.2 PWARX systems

Most recent contributions related to the identification of PWARX systems are classified as optimization-based and clustering-based. It is worthwhile to note that most approaches firstly classify the data points and estimate the submodel parameters (either simultaneously or iteratively), and then reconstruct the polyhedral partition of the regressor domain. The latter step is typically addressed by resorting to standard linear, support vector machine or neural network classifiers (see [Paoletti et al., 2007]). Single methods not falling into the aforementioned categories, are reviewed hereafter.

The method by Bako et al. [2011], already mentioned among recursive methods for SARX systems, can be extended to deal with the identification of PWARX systems by appropriately choosing the data assignment criterion.

A modification of the bounded-error procedure [Bemporad et al., 2005] is proposed in [Tabatabaei-Pour et al., 2006a]. In particular, the greedy randomized method adopted in [Bemporad et al., 2005] is substituted with a greedy clustering-based method.

Finally, Xu et al. [2009] propose the model of Adaptive Hinging Hyperplanes (AHH), and show that it is a special case of generalized hinging hyperplanes. A procedure for the identification of AHH models is then derived.

Optimization-based methods. While most methods for the identification of PWARX systems published before 2007 (see, e.g., [Ferrari-Trecate et al., 2003, Nakada et al., 2005, Juloski et al., 2005, Bemporad et al., 2005]) were clever heuristics aimed at finding "good" (suboptimal) solutions to Problem 3.2, in recent years optimization-based methods have received an increasing attention.

In [Ohlsson and Ljung, 2011], the identification of PWARX systems is tackled by over-parameterizing and assigning a submodel to each of the observations. Then, the submodels are forced to be the same if this does not cause a major increase in the fit term. The proposed formulation is a least-squares problem with sum-of-norms regularization over the differences between submodel parameters. The solution to this problem provides simultaneously the number of modes, the data classification and the estimates of the submodel parameters. A regularization constant can be used to trade off fit and number of modes.

In the method proposed by Lai et al. [2010a,b,c], the submodel parameters are estimated via a least-squares based identification method using multiple models, while Maruta and Sugie [2011] present a method for the identification of PWARX systems which exploits the data-based representation of PWA maps and data compression with ℓ_1 optimization.

In Jin and Huang [2010] the identification problem for PWARX systems is formulated and solved within the Expectation-Maximization framework. In particular, contaminated Gaussian distributions are used in the process of constructing the objective function. Solutions are proposed for the problems of sensitivity to initialization and inability to accurately classify the "undecidable" data points (see [Paoletti et al., 2007]).

Probability weighted ARX (PrARX) models are proposed in Taguchi et al. [2009], and their parameter estimation problem is formulated as a single optimization problem. Furthermore, it is shown that the identified PrARX model can be easily transformed into a PWARX model.

Finally, in [Wen et al., 2007], BPWARX models are proposed, and a modified Gauss-Newton algorithm is adopted to identify the models from input-output data.

Clustering-based methods. Clustering-based methods exploit clustering techniques at different stages of the identification process.

Boukharouba et al. [2009] propose a method which is able to solve simultaneously the estimation of the number of modes, the problem of data classification, and the parameter estimation thanks to an evidential procedure based on Dempster-Shafer theory. The proposed approach aims at simultaneously minimizing the error between the measured output and each submodel output, and the distance between data belonging to the same submodel.

In Gegúndez et al. [2008], the submodels of the PWARX system are obtained by means of an algorithm inspired by competitive learning. In particular, the proposed method exploits a process of fuzzy clustering in order to obtain a subset of representatives from the original data set.

Baptista et al. [2011] assume the number of modes of the PWARX system to be unknown, and propose a split-and-merge clustering algorithm to estimate the correct number of modes. The main advantages of this clustering algorithm are that it does not require special care in the initialization step and there is only one tuning parameter to be adjusted.

A k-plane clustering algorithm is employed to provide initial data classification and submodel parameter estimates in the method proposed by Tabatabaei-Pour et al. [2006b]. Then the refinement algorithm proposed in [Bemporad et al., 2005] is repeatedly applied to the estimated clusters in order to improve both the data classification and the parameter estimation.

Finally, Ivanescu et al. [2011] discuss the use of correlation clustering algorithms for robust identification of PWARX models with reduced complexity.

Table 3. Methods for state-space systems

Batch methods	
Change detection	Pekpe et al. [2004]
Switching detection	Borges et al. [2005]
Expectation maximization	Blackmore et al. [2007], Gil and Williams [2009]
Recursive methods	
On-line two-step clustering	Pekpe and Lecœuche [2006]
On-line switching detection	Bako et al. [2007, 2009a]
Continuous state elimination	Bako et al. [2009b]
Sparse optimization	Chen et al. [2011]

4.3 Switched affine state-space systems

Methods for the identification of switched affine state-space systems can be mainly classified in batch and recursive methods. The former category includes the contributions by Pekpe et al. [2004], Borges et al. [2005], Blackmore et al. [2007] and Gil and Williams [2009].

In [Pekpe et al., 2004], a change detection technique is applied to estimate the switching times. Then, the Markov parameters of each submodel are identified via a subspace method. Parameter estimation is alternated with a procedure to merge similar submodels.

Borges et al. [2005] propose a switching detection method which is based on projected subspaces from batches of input-output data.

A method for the identification of switched affine state-space systems based on approximate Expectation-Maximization (EM) is presented in [Blackmore et al., 2007]. This work is continued in [Gil and Williams, 2009]. Since EM-based model estimation techniques may converge to a locally optimal set of model parameters, the authors develop an algorithm for estimating the model parameters that avoids getting stuck in locally optimal solutions. The algorithm is able to guide the search towards better and better maxima of the likelihood function, thus avoiding local convergence.

Recursive methods for the identification of switched affine state-space systems are those proposed by Pekpe and Lecœuche [2006], Bako et al. [2007, 2009a,b] and Chen et al. [2011].

In [Pekpe and Lecœuche, 2006] an on-line clustering approach organized in two steps is proposed. In the first step, the Markov parameter matrices of the submodels are estimated through an on-line estimation method. In the second step, a dynamic classification algorithm is applied to classify these parameters.

A structured subspace identification method for linear systems, which can be implemented on-line, is presented in [Bako et al., 2007, 2009a]. This method, combined with an on-line switching times detection strategy, is then applied to blindly identify switched systems and to classify the obtained submodels. The proposed method requires a minimum dwell-time in each discrete state to achieve good performances.

The assumption of minimum dwell-time is removed in [Bako et al., 2009b], where a technique for eliminating the unknown continuous state vector from the model equations is proposed under an appropriate assumption of observability. This leads to the introduction of suitable structured intermediary matrices, which can be estimated via a randomly initialized algorithm that alternates between data classification and parameter update via recursive least squares. Given these matrices, the submodel parameters associated to each mode can be recovered.

Finally, Chen et al. [2011] consider switched affine state-space systems corrupted by an unknown, non-centered, and sparse error sequence. By taking advantage of some recent developments in sparse optimization theory, a recursive approach for the identification of these systems is presented.

4.4 Switched output-error systems

In [Wang and Chen, 2011], a modified recursive least squares algorithm for the on-line identification of switched OE systems systems is presented. Given a finite input-output data set and a bound on the noise, Feng et al. [2010b] consider the problem of identifying a switched OE system with the smallest number of modes that is compatible with the prior information. This problem is recast into a polynomial optimization form, from which computationally tractable problems are obtained by exploiting its inherent sparse structure and a randomized Hit and Run type approach. A similar problem is considered by Feng et al. [2010a] in a set-membership framework. By recasting the identification problem as polynomial optimization, deterministic algorithms are developed, in which the inherent sparse structure is exploited. A finite dimensional semi-definite problem, equivalent to the original identification problem, is obtained. Moreover, an equivalent rank minimization problem subject to LMI constraints is formulated, in order to deal with computational complexity issues.

4.5 Block-structured systems

Block-structured systems such as Hammerstein, Wiener and Linear Fractional Representation (LFR) systems, composed by the interconnection of a linear time-invariant block and a PWA nonlinearity, can be equivalently described as an overall PWA system. Therefore their identification can be tackled by applying black-box PWA system identification techniques to the overall system. On the other hand, in the identification of interconnected systems it is crucial to exploit the available knowledge about the interconnection structure in order to obtain enhanced identification results. For this reason, in this section we mention some contributions addressing the joint identification of both the linear and the PWA components of block-structured systems with PWA nonlinearity.

Dolanc and Strmčnik [2005] propose a recursive least-squares technique to tackle identification of Hammerstein models in which the nonlinear block is described by a PWA static function. Both Hammerstein and Wiener systems with PWA nonlinearity are considered in [Zimmerschied and Isermann, 2009], where the simultaneous estimation of the linear block and the PWA function is performed via a Levenberg-Marquardt algorithm.

Another block-structured system which has been thoroughly analyzed in recent years is that consisting of a linear system cascaded with a backlash nonlinearity (a special case of PWA nonlinearity). Bounded error estimation techniques have been employed in [Cerone and Regruto, 2007] and [Cerone et al., 2009] to deal with input and output backlash, respectively. An input backlash has been considered also in [Vörös, 2010], where an iterative procedure is proposed to alternatively estimate the system parameters and the unmeasured inner signals related to the backlash.

Identification of interconnected systems in LFR form has been takled in [Pepona et al., 2007, Paoletti et al., 2009, Pepona et al., 2011]. The considered model class consists of the interconnection of a linear time-invariant block and a static nonlinearity. Iterative identification procedures have been proposed which alternate the estimation of the linear and the nonlinear components, the latter being modeled as PWA functions. The main outcome of these works is that remarkable advantages are encountered when a priori knowledge about the structure of the interconnection is explicitly accounted for in the identification process. This is mainly due to the fact that identification of black-box PWA models for interconnected systems usually leads to an explosion in the number of modes of the identified model.

5. APPLICATIONS

In the last decade, PWA identification techniques have been successfully applied to a wide variety of real-world problems, involving systems or phenomena which are inherently nonlinear, but can be modeled as the combination of a finite number of linear behaviors. Table 4 lists problems and applications reviewed in this section, and the corresponding references.

A benchmark application that has been considered in several papers, by using different classes of models and identification techniques, is the electronic component placement process in a pick-and-place machine [Juloski et al., 2004, Bemporad et al., 2005, Juloski et al., 2005, Sepasi and Sadrnia, 2008, Boukharouba et al., 2009]. Another popular benchmark in nonlinear system identification, the *silverbox* model, has been tackled in [Pepona et al., 2011] by using block-structured PWA models.

Computer vision is a field which has seen a number of application of PWA identification techniques. Challenging problems such as dynamic texture classification and video segmentation have been tackled in [Vidal and Ma, 2006, Vidal et al., 2007, Vidal, 2008, Ozay et al., 2008, Boukharouba et al., 2010a]. These works usually adopt SARX models in which each submodel corresponds to a coherent video sequence, while the switching between two submodels describes the transition between two different scenes or video segments.

The use of PWA identification tools for electromechanical systems and automotive applications is also becoming quite popular. Hybrid model identification of DC motors has been considered in [Canty and O'Mahony, 2009] and [Maruta et al., 2011]. A PWARX model of the electronic throttle regulating the air inflow of a car engine has been estimated in [Vašak et al., 2005]. PWA approximations of Wiener/Hammerstein systems have been exploited in

Table 4. Applications

Benchmark applications	
Pick-and-place machine	Juloski et al. [2004], Bemporad et al. [2005], Juloski et al. [2005], Sepasi and Sadrnia [2008], Boukharouba et al. [2009]
Silverbox data set	Pepona et al. [2011]
Computer vision	
Dynamic textures/ video segmentation	Vidal and Ma [2006], Vidal et al. [2007], Vidal [2008], Ozay et al. [2008], Boukharouba et al. [2010a]
Electromechanical/automotive systems	
DC motors	Canty and O'Mahony [2009], Maruta et al. [2011]
Electronic throttle	Vašak et al. [2005]
Diesel engine	Zimmerschied and Isermann [2009]
Automatic driving system	Sato et al. [2008], Taguchi et al. [2009]
Hybrid powertrain system	Ripaccioli et al. [2009]
Wind turbine	Vašak et al. [2011]
Sealing mechanism	Verspecht et al. [2011]
Systems with backlash	Cerone and Regruto [2007], Cerone et al. [2009], Vörös [2010]
Systems biology	
Cortical neuroprostheses	Hudson and Burdick [2007]
Hormone pulses	Liberati [2009a]
Sleep apneas	Liberati [2009a]
Dialysis process	Liberati [2009b]
Genetic networks	Porreca and Ferrari-T. [2008], Porreca et al. [2009]
Cell biology	Vries et al. [2009]
Environmental systems	
Open-channel systems	Boukharouba et al. [2010b]

[Zimmerschied and Isermann, 2009] to model the air- and exhaust-manifold of a modern common-rail diesel engine. Identification of PWARX models for automatic driving systems has been addressed in [Sato et al., 2008] and [Taguchi et al., 2009]. PWA models for an advanced hybrid powertrain system and for a pitch-controlled wind turbine have been proposed in [Ripaccioli et al., 2009] and [Vašak et al., 2011], respectively. Verspecht et al. [2011] address PWARX modeling of a sealing mechanism used in minimally invasive surgery. Identification of systems with backlash, which is one of the most common nonlinearities in mechanical systems, has been treated in [Cerone and Regruto, 2007], [Cerone et al., 2009] and [Vörös, 2010].

Systems biology is another research area that has attracted the attention of researchers, due to the fact that PWA models provide good approximations of complex phenomena. Cortical neuroprostheses, used to restore motor function in individuals with high level spinal cord injuries or severe motor disorders, can be modeled as PWA systems in which each mode represents a different brain state with its associated neural dynamics, as shown in [Hudson and Burdick, 2007]. The presence of switching in several physiological and pathological processes, such as hormone pulses or sleep apneas, are described by PWARX models in [Liberati, 2009a]. A PWA approximation of the continuous process of dialysis is proposed in [Liberati, 2009b]. Genetic regulatory networks are well suited to be modeled by PWA systems, as gene expression profiles can be split into segments, each one generated by a single affine mode. PWA identification of genetic networks has been addressed in [Porreca and Ferrari-Trecate, 2008] and [Porreca et al., 2009]. Challenges in PWA modeling and identification problems relevant to systems biology are discussed in [Vries et al., 2009].

Finally, PWA modeling of open-channel systems, which are useful in the management of water resources, has been treated in [Boukharouba et al., 2010b].

6. CONCLUSIONS

In this paper, a review of the most recent contributions on switched and PWA system identification has been presented. From this review, there emerges a clear research trend towards the use of novel optimization and relaxation tools to tackle the identification problems for switched and PWA systems. On the other hand, algebraic and clustering-based methods have received a steady attention from researchers in the field. Other emerging trends are the development of identification methods for block-structured systems with PWA nonlinearities, and for switched nonlinear systems, though the latter have not been considered in this paper (see, e.g., [Lauer and Bloch, 2008a, Bako et al., 2010, Lauer et al., 2010]).

The review of the literature has highlighted a continuously increasing interest towards the identification of switched and PWA systems. Nevertheless, to construct a complete identification theory for these classes of systems, much work still remains to be done. To this end, research directions to be further pursued include the study of identifiability issues, the development of experiment design techniques and of a complete realization theory, and the full exploitation of a priori information about the system structure within the identification procedures.

REFERENCES

O. Ayad, M. Sayed-Mouchweh, and P. Billaudel. Switched hybrid dynamic systems identification based on pattern recognition approach. In *2011 IEEE International Conference on Fuzzy Systems*, Taipei, Taiwan, 2010.

L. Bako. Identification of switched linear systems via sparse optimization. *Automatica*, 47(4):668–677, 2011.

L. Bako and R. Vidal. Algebraic identification of MIMO SARX models. In M. Egerstedt and B. Mishra, editors, *Hybrid Systems: Computation and Control*, volume 4981 of *Lecture Notes in Computer Science*, pages 43–57. Springer, 2008.

L. Bako, G. Mercère, and S. Lecœuche. On-line structured identification of switching systems with possibly varying orders. In *Proc. 2007 European Control Conference*, pages 4065–4072, Kos, Greece, 2007.

L. Bako, G. Mercère, and S. Lecœuche. On-line structured subspace identification with application to switched linear system. *International Journal of Control*, 82(8): 1496–1515, 2009a.

L. Bako, G. Mercère, R. Vidal, and S. Lecœuche. Identification of switched linear state space models without minimum dwell time. In *Proc. 15th IFAC Symposium* *on System Identification*, pages 569–574, Saint-Malo, France, 2009b.

L. Bako, K. Boukharouba, and S. Lecœuche. An ℓ_0-ℓ_1 norm based optimization procedure for the identification of switched nonlinear systems. In *Proc. 49th IEEE Conference on Decision and Control*, pages 4467–4472, Atlanta, GA, 2010.

L. Bako, K. Boukharouba, E. Duviella, and S. Lecœuche. A recursive identification algorithm for switched linear/affine models. *Nonlinear Analysis: Hybrid Systems*, 5:242–253, 2011.

R.S. Baptista, J.Y. Ishihara, and G.A. Borges. Split and merge algorithm for identification of piecewise affine systems. In *2011 American Control Conference*, pages 2018–2023, San Francisco, CA, 2011.

A. Bemporad, A. Garulli, S. Paoletti, and A. Vicino. A bounded-error approach to piecewise affine system identification. *IEEE Transactions on Automatic Control*, 50 (10):1567–1580, 2005.

L. Blackmore, S. Gil, S. Chung, and B. Williams. Model learning for switching linear systems with autonomous mode transitions. In *Proc. 46th IEEE Conference on Decision and Control*, pages 4648–4655, New Orleans, LA, 2007.

J. Borges, V. Verdult, M. Verhaegen, and M.A. Botto. A switching detection method based on projected subspace classification. In *Proc. Joint 44th IEEE Conference on Decision and Control and 2005 European Control Conference*, pages 344–349, Seville, Spain, 2005.

K. Boukharouba, L. Bako, and S. Lecœuche. Identification of piecewise affine systems based on Dempster-Shafer theory. In *Proc. 15th IFAC Symposium on System Identification*, pages 1662–1667, Saint-Malo, France, 2009.

K. Boukharouba, L. Bako, and S. Lecœuche. Temporal video segmentation using a switched affine models identification technique. In *Proc. 2nd International Conference on Image Processing Theory, Tools and Applications*, pages 157–160, Paris, France, 2010a.

K. Boukharouba, E. Duviella, L. Bako, and S. Lecœuche. Multimodeling vs piecewise affine modeling for the identification of open channel systems. In *Proc. 12th IFAC Symposium on Large Scale Systems: Theory and Applications*, Villeneuve d'Ascq, France, 2010b.

L. Breiman. Hinging hyperplanes for regression, classification, and function approximation. *IEEE Transanctions on Information Theory*, 39(3):999–1013, 1993.

N. Canty and T. O'Mahony. Design considerations for piecewise affine system identification of nonlinear systems. In *Proc. 17th Mediterranean Conference on Control and Automation*, pages 157–162, Thessaloniki, Greece, 2009.

V. Cerone and D. Regruto. Bounding the parameters of linear systems with input backlash. *IEEE Transactions on Automatic Control*, 52(3):531–536, 2007.

V. Cerone, D. Piga, and D. Regruto. Parameter bounds evaluation for linear system with output backlash. In *Proc. 15th IFAC Symposium on System Identification*, pages 575–580, Saint-Malo, France, 2009.

D. Chen, L. Bako, and S. Lecœuche. A recursive sparse learning method: Application to jump Markov linear systems. In *Proc. 18th IFAC World Congress*, pages 3198–3203, Milan, Italy, 2011.

G. Dolanc and S. Strmčnik. Identification of nonlinear systems using a piecewise-linear Hammerstein model. *Systems & Control Letters*, 54(2):145–158, 2005.

C. Feng, C.M. Lagoa, N. Ozay, and M. Sznaier. Hybrid system identification: An SDP approach. In *Proc. 49th IEEE Conference on Decision and Control*, pages 1546–1552, Atlanta, GA, 2010a.

C. Feng, C.M. Lagoa, and M. Sznaier. Hybrid system identification via sparse polynomial optimization. In *Proc. 2010 American Control Conference*, pages 160–165, Baltimore, MD, 2010b.

G. Ferrari-Trecate. Hybrid identification toolbox (HIT), 2005. http://www-rocq.inria.fr/who/Giancarlo.Ferrari-Trecate/HIT_toolbox.html.

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.

M.E. Gegúndez, J. Aroba, and J.M. Bravo. Identification of piecewise affine systems by means of fuzzy clustering and competitive learning. *Engineering Applications of Artificial Intelligence*, 21(8):1321–1329, 2008.

S. Gil and B. Williams. Beyond local optimality: An improved approach to hybrid model learning. In *Proc. Joint 48th IEEE Conference on Decision and Control and 28th Chinese Control Conference*, pages 3938–3945, Shangai, China, 2009.

N. Hudson and J. Burdick. A stochastic framework for hybrid system identification with application to neurophysiological systems. In A. Bemporad, A. Bicchi, and G. Buttazzo, editors, *Hybrid Systems: Computation and Control*, volume 4416 of *Lecture Notes in Computer Science*, pages 273–286. Springer, 2007.

A.M. Ivanescu, T. Albin, D. Abel, and T. Seidl. Employing correlation clustering for the identification of piecewise affine models. In *Proc. 2011 Workshop on Knowledge Discovery, Modeling and Simulation*, pages 7–14, San Diego, CA, 2011.

X. Jin and B. Huang. Robust identification of piecewise/switching autoregressive exogenous process. *AIChE Journal*, 56(7):1829–1844, 2010.

A.Lj. Juloski, W.P.M.H. Heemels, and G. Ferrari-Trecate. Data-based hybrid modelling of the component placement process in pick-and-place machines. *Control Engineering Practice*, 12(10):1241–1252, 2004.

A.Lj. Juloski, S. Weiland, and W.P.M.H. Heemels. A Bayesian approach to identification of hybrid systems. *IEEE Transactions on Automatic Control*, 50(10):1520–1533, 2005.

C.Y. Lai, C. Xiang, and T.H. Lee. Identification and control of nonlinear systems via piecewise affine approximation. In *Proc. 49th IEEE Conference on Decision and Control*, pages 6395 –6402, Atlanta, GA, 2010a.

C.Y. Lai, C. Xiang, and T.H. Lee. Identification and control of nonlinear systems using piecewise affine models. In *Proc. 11th International Conference on Control, Automation, Robotics & Vision*, pages 2259–2265, Singapore, 2010b.

C.Y. Lai, C. Xiang, and T.H. Lee. Identification of piecewise affine systems and nonlinear systems using multiple models. In *Proc. 8th IEEE International Conference on Control and Automation*, pages 2005 – 2012, Xiamen, China, 2010c.

F. Lauer and G. Bloch. Switched and piecewise nonlinear hybrid system identification. In M. Egerstedt and B. Mishra, editors, *Hybrid Systems: Computation and Control*, volume 4981 of *Lecture Notes in Computer Science*, pages 330–343. Springer, 2008a.

F. Lauer and G. Bloch. A new hybrid system identification algorithm with automatic tuning. In *Proc. 17th IFAC World Congress*, pages 10207–10212, Seoul, Korea, 2008b.

F. Lauer, R. Vidal, and G. Bloch. A product-of-errors framework for linear hybrid system identification. In *Proc. 15th IFAC Symposium on System Identification*, pages 563–568, Saint-Malo, France, 2009.

F. Lauer, G. Bloch, and R. Vidal. Nonlinear hybrid system identification with kernel models. In *Proc. 49th IEEE Conference on Decision and Control*, pages 696–701, Atlanta, GA, 2010.

F. Lauer, G. Bloch, and R. Vidal. A continuous optimization framework for hybrid system identification. *Automatica*, 47(3):608–613, 2011.

D. Liberati. Biomedical applications of piece-wise affine identification for hybrid systems. *Annals of biomedical engineering*, 37(9):1871–1876, 2009a.

D. Liberati. Piece-wise affine identification in dialysis. *Nonlinear Analysis: Hybrid Systems*, 3(4):708–712, 2009b.

L. Ljung. *System Identification: Theory for the User.* Prentice-Hall, Upper Saddle River, NJ, 1999.

I. Maruta and T. Sugie. Identification of PWA models via data compression based on ℓ_1 optimization. In *Proc. Joint 50th IEEE Conference on Decision and Control and 2011 European Control Conference*, pages 2800–2805, Orlando, FL, 2011.

I. Maruta, T. Sugie, and T.H. Kim. Identification of multiple mode models via distributed particle swarm optimization. In *Proc. 18th IFAC World Congress*, pages 7743–7748, Milan, Italy, 2011.

H. Nakada, K. Takaba, and T. Katayama. Identification of piecewise affine systems based on statistical clustering technique. *Automatica*, 41(5):905–913, 2005.

S. Nazari, Q. Zhao, and B. Huang. An improved algebraic geometric solution to the identification of switched ARX models with noise. In *2011 American Control Conference*, pages 1230–1235, San Francisco, CA, 2011.

H. Ohlsson and L. Ljung. Identification of piecewise affine systems using sum-of-norms regularization. In *Proc. 18th IFAC World Congress*, pages 6640–6645, Milan, Italy, 2011.

H. Ohlsson, L. Ljung, and S. Boyd. Segmentation of ARX-models using sum-of-norms regularization. *Automatica*, 46(6):1107–1111, 2010.

N. Ozay, M. Sznaier, C. Lagoa, and O. Camps. A sparsification approach to set membership identification of a class of affine hybrid systems. In *Proc. 47th IEEE Conference on Decision and Control*, pages 123–130, Cancun, Mexico, 2008.

N. Ozay, C. Lagoa, and M. Sznaier. Robust identification of switched affine systems via moments-based convex optimization. In *Proc. Joint 48th IEEE Conference on Decision and Control and 28th Chinese Control Conference*, pages 4686–4691, Shanghai, China, 2009.

S. Paoletti and J. Roll. Piecewise affine system identification toolbox (PWAID), 2007. http://www.control.isy.liu.se/∼roll/PWAID/.

S. Paoletti, A. Lj. Juloski, G. Ferrari-Trecate, and R. Vidal. Identification of hybrid systems: A tutorial. *European Journal of Control*, 13(2–3):242–260, 2007.

S. Paoletti, A. Garulli, J. Roll, and A. Vicino. A necessary and sufficient condition for input-output realization of switched affine state space models. In *Proc. 47th IEEE Conference on Decision and Control*, pages 935–940, Cancun, Mexico, 2008.

S. Paoletti, A. Garulli, E. Pepona, and P. Date. Exploiting structure in piecewise affine identification of LFT systems. In *Proc. 15th IFAC Symposium on System Identification*, pages 581–586, Saint-Malo, France, 2009.

S. Paoletti, J. Roll, A. Garulli, and A. Vicino. On the input-output representation of piecewise state space models. *IEEE Transactions on Automatic Control*, 55 (1):60–73, 2010.

K.M. Pekpe and S. Lecœuche. Online classification of switching models based on subspace framework. In *Proc. 2nd IFAC Conference on Analysis and Design of Hybrid Systems*, pages 154–159, Alghero, Italy, 2006.

K.M. Pekpe, G. Mourot, K. Gasso, and J. Ragot. Identification of switching systems using change detection technique in the subspace framework. In *Proc. 43rd IEEE Conference on Decision and Control*, pages 3720–3725, Paradise Island, Bahamas, 2004.

E. Pepona, S. Paoletti, A. Garulli, and P. Date. An iterative procedure for piecewise affine identification of nonlinear interconnected systems. In *Proc. 46th IEEE conference on Decision and control*, pages 5098–5103, New Orleans, LA, 2007.

E. Pepona, S. Paoletti, A. Garulli, and P. Date. Identification of piecewise affine LFR models of interconnected systems. *IEEE Transactions on Control Systems Technology*, 19(1):148–155, 2011.

M. Petreczky. Realization theory for discrete-time piecewise-affine hybrid systems. In *Proc. 17th International Symposium on Mathematical Theory of Networks and Systems*, pages 1275–1294, Kyoto, Japan, 2006.

M. Petreczky and J.H. van Schuppen. Realization theory of discrete-time linear hybrid system. In *Proc. 15th IFAC Symposium on System Identification*, pages 593–598, Saint-Malo, France, 2009.

M. Petreczky, L. Bako, and J.H. van Schuppen. Identifiability of discrete-time linear switched systems. In *Proc. 13th ACM International Conference on Hybrid systems: Computation and Control*, pages 141–150, 2010.

M. Petreczky, L. Bako, and J.H. van Schuppen. Realization theory of discrete-time linear switched systems. Technical Report arXiv:1103.1343, Arxiv, 2011.

R. Porreca and G. Ferrari-Trecate. Identification of piecewise affine models of genetic regulatory networks: the data classification problem. In *Proc. 17th IFAC World Congress*, pages 307–312, Seoul, Korea, 2008.

R. Porreca and G. Ferrari-Trecate. Partitioning datasets based on equalities among parameters. *Automatica*, 46 (2):460–465, 2010.

R. Porreca, S. Drulhe, H. de Jong, and G. Ferrari-Trecate. Identification of parameters and structure of piecewise affine models of genetic networks. In *Proc. 15th IFAC Symposium on System Identification*, pages 587–592, Saint-Malo, France, 2009.

G. Ripaccioli, A. Bemporad, F. Assadian, C. Dextreit, S. Di Cairano, and I. Kolmanovsky. Hybrid modeling, identification, and predictive control: An application to hybrid electric vehicle energy management. In R. Majumdar and P. Tabuada, editors, *Hybrid Systems: Computation and Control*, volume 5469 of *Lecture Notes in Computer Science*, pages 321–335. Springer, 2009.

J. Roll. *Local and Piecewise Affine Approaches to System Identification*. PhD thesis, Department of Electrical Engineering, Linköping University, Linköping, Sweden, 2003. http://www.control.isy.liu.se/publications/.

J. Roll, A. Bemporad, and L. Ljung. Identification of piecewise affine systems via mixed-integer programming. *Automatica*, 40(1):37–50, 2004.

F. Rosenqvist and A. Karlström. Realisation and estimation of piecewise-linear output-error models. *Automatica*, 41(3):545–551, 2005.

H. Sato, T. Zanma, and M. Ishlda. Identification of human operation using data clustering and its application to automated system. In *Proc. 10th IEEE International Workshop on Advanced Motion Control*, pages 411–416, Trento, Italy, 2008.

S. Sepasi and M.A. Sadrnia. On-line identification of an electronic component placement process using a potential fuzzy clustering scheme. In *Proc. 2nd International Conference on Electrical Engineering*, Lahore, Pakistan, 2008.

E. D. Sontag. Nonlinear regulation: The piecewise linear approach. *IEEE Transactions on Automatic Control*, 26 (2):346–358, 1981.

H. Suzuki and M. Yamakita. Input design for hybrid system identification for accurate estimation of submodel regions. In *2011 American Control Conference*, pages 1236 –1241, San Francisco, CA, 2011.

M. Tabatabaei-Pour, M. Gholami, K. Salahshoor, and H.R. Shaker. A clustering-based bounded-error approach for identification of PWA hybrid systems. In *Proc. 9th International Conference on Control, Automation, Robotics and Vision*, Singapore, 2006a.

M. Tabatabaei-Pour, K. Salahshoor, and B. Moshiri. A modified k-plane clustering algorithm for identification of hybrid systems. In *Proc. 6th World Congress on Intelligent Control and Automation*, volume 1, pages 1333–1337, Dalian, China, 2006b.

S. Taguchi, T. Suzuki, S. Hayakawa, and S. Inagaki. Identification of probability weighted multiple ARX models and its application to behavior analysis. In *Proc. Joint 48th IEEE Conference on Decision and Control and 28th Chinese Control Conference*, pages 3952–3957, Shanghai, China, 2009.

M. Vašak, N. Hure, and N. Perić. Identification of a discrete-time piecewise affine model of a pitch-controlled wind turbine. In *Proc. 34th International Convention on Information and Communication Technology, Electronics and Microelectronics–Computers in Technical Systems*, Opatija, Croatia, 2011.

M. Vašak, L. Mladenović, and N. Perić. Clustering-based identification of a piecewise affine electronic throttle model. In *Proc. 31st Annual Conference of the IEEE Industrial Electronics Society*, pages 177–182, Raleigh, NC, 2005.

V. Verdult and M. Verhaegen. Subspace identification of piecewise linear systems. In *Proc. 43rd IEEE Conference on Decision and Control*, pages 3838–3843, Paradise Island, Bahamas, 2004.

J. Verspecht, L. Catoire, S. Torfs, and M. Kinnaert. Identification of a hybrid model for simulation of the instrument/trocar interaction force. In *Proc. 2011 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 4477–4482, San Francisco, CA, 2011.

R. Vidal. Recursive identification of switched ARX systems. *Automatica*, 44(9):2274–2287, 2008.

R. Vidal and Y. Ma. A unified algebraic approach to 2-D and 3-D motion segmentation and estimation. *Journal of Mathematical Imaging and Vision*, 25(3):403–421, 2006.

R. Vidal, A. Chiuso, and S. Soatto. Observability and identifiability of jump linear systems. In *Proc. 41st IEEE Conference on Decision and Control*, pages 3614–3619, Las Vegas, NV, 2002.

R. Vidal, S. Soatto, Y. Ma, and S. Sastry. An algebraic geometric approach to the identification of a class of linear hybrid systems. In *Proc. 42nd IEEE Conference on Decision and Control*, pages 167–172, Maui, HI, 2003.

R. Vidal, S. Soatto, and A. Chiuso. Applications of hybrid system identification in computer vision. In *Proc. 2007 European Control Conference*, pages 4853–4860, Kos, Greece, 2007.

J. Vörös. Modeling and identification of systems with backlash. *Automatica*, 46(2):369–374, 2010.

D. Vries, P.J.T. Verheijen, and A.J. den Dekker. Hybrid system modeling and identification of cell biology systems: perspectives and challenges. In *Proc. 15th IFAC Symposium on System Identification*, pages 227–232, Saint-Malo, France, 2009.

J. Wang and T. Chen. Online identification of switched linear output error models. In *Proc. 2011 IEEE International Symposium on Computer-Aided Control System Design*, pages 1379 –1384, Denver, CO, 2011.

S. Weiland, A.Lj. Juloski, and B. Vet. On the equivalence of switched affine models and switched ARX models. In *Proc. 45th IEEE Conference on Decision and Control*, pages 2614–2618, San Diego, CA, 2006.

C. Wen, S. Wang, F. Li, and M.J. Khan. A compact f-f model of high-dimensional piecewise-linear function over a degenerate intersection. *IEEE Transactions on Circuits and Systems I*, 52(4):815–821, 2005.

C. Wen, S. Wang, X. Jin, and X. Ma. Identification of dynamic systems using piecewise-affine basis function models. *Automatica*, 43(10):1824–1831, 2007.

J. Xu, X. Huang, and S. Wang. Adaptive hinging hyperplanes and its applications in dynamic system identification. *Automatica*, 45(10):2325–2332, 2009.

R. Zimmerschied and R. Isermann. Nonlinear system identification of block-oriented systems using local affine models. In *Proc. 15th IFAC Symposium on System Identification*, pages 658–663, Saint-Malo, France, 2009.