Exploiting structure in piecewise affine identification of LFT systems

Simone Paoletti * **Andrea Garulli** * **Eleni Pepona** ** **Paresh Date** ***

* *Dipartimento di Ingegneria dell'Informazione, Università di Siena, Via Roma 56, 53100 Siena, Italy*
** *Dept of Mathematical Sciences, Brunel University, Uxbridge, Middlesex, UB8 3PH, United Kingdom*
*** *CARISMA, Dept. of Mathematical Sciences, Brunel University, Uxbridge, Middlesex UB8 3PH, United Kingdom*

Abstract: Identification of interconnected systems is a challenging problem in which it is crucial to exploit the available knowledge about the interconnection structure. In this paper, identification of discrete-time nonlinear systems composed by interconnected linear and nonlinear systems, is addressed. An iterative identification procedure is proposed, which alternates the estimation of the linear and the nonlinear components. Standard identification techniques are applied to the linear subsystem, whereas recently developed piecewise affine (PWA) identification techniques are employed for modelling the nonlinearity. A numerical example analyzes the benefits of the proposed structure-exploiting identification algorithm compared to applying black-box PWA identification techniques to the overall system.

1. INTRODUCTION

Real world systems often consist of interconnected linear and nonlinear components. Models for interconnected systems range from the well-known Hammerstein and Wiener systems, to more complex networked structures. A quite general framework for representing system interconnections is based on the linear fractional transformation (LFT) modeling formalism (see, e.g., Hsu et al. [2005a]). Applying black-box identification techniques to LFT systems has the drawback that the resulting model does not reflect the internal structure of the system, and may need a large number of parameters to reach the required accuracy. On the other hand, exploiting the knowledge about the structure of the interconnection is not only expected to improve the accuracy of the estimated model, but also to reduce the computational burden of the identification procedure. In spite of this, identification of interconnected systems has received little attention so far. Previdi and Lovera [2003] set up a parameter estimation procedure for models in LFT form where the forward part is represented by a classical linear regression and the feedback part is given by a nonlinear dynamic map parameterized by a neural network. A new paradigm for identification of interconnected systems has been introduced in Hsu et al. [2005a,b, 2006], where the authors consider LFT interconnections of linear dynamic systems and static nonlinear maps. Under the assumption that the linear part is known, several nonparametric estimation algorithms for the static nonlinear maps are proposed.

This paper addresses the identification of an LFT interconnection composed by a linear and a nonlinear system, both of which are unknown. The proposed solution relies on an iterative scheme that alternates the identification of the linear and the nonlinear part. Iterative procedures have been widely used in nonlinear system identification, starting from the seminal work by Narendra and Gallman [1966]. In our approach, standard linear identification techniques are employed to identify the linear part of the system. Moreover, we choose to represent the system nonlinearity by a piecewise affine (PWA) model. This is mainly motivated by the universal approximation properties of PWA maps [Lin and Unbehauen, 1992, Breiman, 1993]. PWA system identification is a challenging problem that has received an increasing attention in recent years, and a number of identification techniques have been proposed (see Roll [2003] and Paoletti et al. [2007] for an extensive overview). Here we adopt the bounded-error procedure for identification of piecewise affine ARX (PWARX) models presented in Bemporad et al. [2005], which does not require to fix a priori the number of modes of the PWA map, but estimates such a number from data. Numerical examples have shown that the proposed iterative scheme is able to successfully profit from the knowledge of the system interconnection structure. The framework considered in this paper contains the framework in Pepona et al. [2007] as a particular case.

The paper is organized as follows. The considered identification problem is introduced in Section 2, while Section 3 describes the proposed identification algorithm. In Section 4, a detailed numerical example both clarifies several aspects of the identification problem and illustrates the effectiveness of the identification algorithm. Conclusions and future research directions are summarized in Section 5.

2. PROBLEM FORMULATION

We consider discrete-time networked dynamical systems formed by the interconnection of linear and nonlinear components. By a suitable rearrangement, these intercon-

nected systems can be always recast into the LFT form of Fig. 1, where the block $\mathcal{L} = \begin{bmatrix} \mathcal{L}_{yu} & \mathcal{L}_{ye} & \mathcal{L}_{yw} \\ \mathcal{L}_{zu} & \mathcal{L}_{ze} & \mathcal{L}_{zw} \end{bmatrix}$ is a linear time-invariant dynamical system (partitioned as usual in control theory), while the block $\mathcal{N}$ takes into account the system nonlinearities. Signals u_t, y_t and e_t are the system input, output and noise at time $t \in \mathbb{Z}$, while $\boldsymbol{z}_t$ and w_t are internal signals representing the input and the output of the nonlinear block. In this preliminary work, we restrict our attention to u_t, y_t and w_t being scalar signals.

The question we address is how to exploit the structure information for identifying both the linear part $\mathcal{L}$ and the nonlinear part $\mathcal{N}$ of the system in Fig. 1. A careful analysis of the problem highlights that input-output data $\{u_t, y_t\}$ and prior information about the interconnection structure are still not sufficient to allow for the estimation of both $\mathcal{L}$ and $\mathcal{N}$. Additional information concerning the internal signals of the system is needed to tackle somehow the problem. In this respect, we distinguish three situations:

i) Both $\boldsymbol{z}_t$ and w_t are known.

ii) w_t is known, while $\boldsymbol{z}_t$ is unknown.

iii) $\boldsymbol{z}_t$ is known, while w_t is unknown.

In case *i*), the identification problem does not present particular difficulties, since it can be decomposed into the separate identification of the linear system $\mathcal{L}$ and of the nonlinear system $\mathcal{N}$, which are well-studied problems. On the other hand, only the part $[\ \mathcal{L}_{yu}\ \mathcal{L}_{ye}\ \mathcal{L}_{yw}\ ]$ of $\mathcal{L}$ can be identified in case *ii*). We thus argue that case *iii*) is the most general situation worth of investigation if one wants to identify both $\mathcal{L}$ and $\mathcal{N}$.

Based on the above discussion, we tackle the identification problem for the system in Fig. 1 under the following main assumption.

Assumption 1. The signal $\boldsymbol{z}_t$ is either measured or can be inferred from past inputs u_t and outputs y_t, i.e. there exists a known stable linear system $\mathcal{M}$ such that $\boldsymbol{z}_t = \mathcal{M}\left[\begin{smallmatrix} u_t \\ y_t \end{smallmatrix}\right]$.

Assumption 1 is similar to the "measurability" assumption on z_t adopted by Hsu et al. [2005a,b, 2006] in a context where $\mathcal{L}$ is known and only $\mathcal{N}$ needs to be identified. Stability of $\mathcal{M}$ is required to cope with uncertainty on its initial condition. We also consider the following assumption.

Assumption 2. The nonlinear part $\mathcal{N}$ is static.

Assumption 2 is only motivated by the tests we have carried out so far, and will be possibly removed in the final version of this paper, as discussed in Remark 3.1.

We are now ready to state the identification problem addressed in this paper.

Problem 1. Given the data set $\{u_t, y_t, \boldsymbol{z}_t\}_{t=1}^{N}$, identify a model of the LFT system in Fig. 1 of the following type:

$$A(q)y_t = B(q)u_{t-n_k} + G(q)w_t + \varepsilon_t \tag{1a}$$
$$w_t = f(\boldsymbol{z}_t), \tag{1b}$$

where the ARX model (1a) describes the linear part $\mathcal{L}$, and the static map $f(\cdot)$ in (1b) is a PWA approximation of the static nonlinearity $\mathcal{N}$. □

In (1a), $\varepsilon_t \in \mathbb{R}$ is the error term, and $A(q), B(q), G(q)$ are finite polynomials in the delay operator q^{-1}, namely:

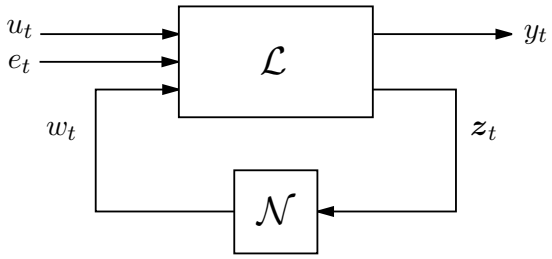

Fig. 1. LFT model structure.

$$A(q) = 1 + a_1 q^{-1} + \ldots + a_{n_a} q^{-n_a} \tag{2a}$$
$$B(q) = b_0 + b_1 q^{-1} + \ldots + b_{n_b} q^{-n_b} \tag{2b}$$
$$G(q) = 1 + g_1 q^{-1} + \ldots + g_{n_g} q^{-n_g}. \tag{2c}$$

Without loss of generality, the first coefficient in (2c) is taken equal to 1, since it is always possible to scale the nonlinear map $f(\cdot)$. In the following, it will be useful to rewrite (1a) in regression form as follows:

$$y_t = \vartheta^T \varphi_t + w_t + \varepsilon_t, \tag{3}$$

where $\vartheta = [\ \vartheta_\zeta^T\ \vartheta_\omega^T\ ]^T$ is the unknown parameter vector, formed by

$$\vartheta_\zeta = [\ a_1\ \ldots\ a_{n_a}\ b_0\ b_1\ \ldots\ b_{n_b}\ ]^T \tag{4a}$$
$$\vartheta_\omega = [\ g_1\ \ldots\ g_{n_g}\ ]^T, \tag{4b}$$

and $\varphi_t = [\ \zeta_t^T\ \omega_t^T\ ]^T$ is the (partially unknown) regression vector, formed by

$$\zeta_t = [\ -y_{t-1}\ \ldots\ -y_{t-n_a}\ u_{t-n_k}\ \ldots\ u_{t-n_k-n_b+1}\ ]^T \tag{5a}$$
$$\omega_t = [\ w_{t-1}\ \ldots\ w_{t-n_g}\ ]^T. \tag{5b}$$

Remark 2.1. The linear part $\mathcal{L}$ of the LFT system can be described in a variety of ways. State space models or pseudolinear regression models (like ARMAX, OE, BJ, etc.), can be adopted. An ARX model has been chosen in (1a) in order to simplify the presentation. □

The PWA map $f(\cdot)$ in (1b) has the form:

$$f(\boldsymbol{z}) = \theta_i^T \begin{bmatrix} \boldsymbol{z} \\ 1 \end{bmatrix} \quad \text{if } \boldsymbol{z} \in \mathcal{Z}_i, \quad i = 1, \ldots, s, \tag{6}$$

where s is the number of modes, $\theta_i \in \mathbb{R}^{n_z+1}$, $i = 1, ..., s$, are the parameters of each mode, and $\{\mathcal{Z}_i\}_{i=1}^{s}$ is a complete polyhedral partition of the domain $\mathcal{Z} \subseteq \mathbb{R}^{n_z}$ where $f(\cdot)$ is defined.

3. IDENTIFICATION ALGORITHM

The most challenging issue in Problem 1 is the fact that the internal signal w_t is not measured, and hence it must be estimated along with the polynomials $A(q)$, $B(q)$, and $G(q)$. To this aim, we propose an iterative identification procedure which alternates the estimation of the linear and the nonlinear part. Assuming the linear part known, it is possible to recover w_t by means of smoothing techniques, as suggested in Claassen [2001]. Then, a PWA map $f(\cdot)$ can be fitted to the pairs $(w_t, \boldsymbol{z}_t)$. On the other hand, if w_t is known, the identification of the linear part can be easily carried out by means of standard linear identification techniques. The above discussion suggests the formulation of the iterative identification algorithm summarized in Table 1. After an initialization step, each algorithm iteration consists of two stages, namely the PWA approximation of

Table 1. Iterative identification algorithm

GIVEN: ϑ^0, γ
SET: $j = 0$
REPEAT
  SET: $j = j + 1$
  % PWA approximation of the nonlinear part
    Estimate $\{v_t^j\}$ according to (8)
    Fit a PWA map $f^j(\cdot)$ to the samples $(\boldsymbol{z}_t, v_t^j)$
    Compute w_t^j as in (9)
  % Identification of the linear part
    Compute ϑ^j from the linear regression (10)
UNTIL $\|\vartheta^j - \vartheta^{j-1}\| \leq \gamma \|\vartheta^j\|$
RETURN: $\vartheta^j, f^j(\cdot)$

the nonlinear part and the identification of the linear part. These steps are described below.

Remark 3.1. The ideas of the proposed identification algorithm can be straightforwardly extended to the case of dynamical nonlinearity $\mathcal{N}$ modelled by a PWARX model of the type $w_t = f(\phi_t)$, with $f(\cdot)$ a PWA map and ϕ_t a finite-length regression vector composed by past values of w_t and $\boldsymbol{z}_t$. At present time, we have only devised to test our algorithm in such a case, and plan to include results of our tests (including real applications), if any, in the final version of this paper. □

3.1 Initialization

An estimate ϑ^0 of the parameter vector ϑ in (3) must be provided for initialization. A rule of thumb is to estimate the part ϑ_ζ^0 of ϑ^0 by fitting the ARX model

$$A(q)y_t = B(q)u_{t-n_k} + \bar{w} + \varepsilon_t \tag{7}$$

to the data through standard linear identification techniques, and setting all the components of ϑ_ω^0 equal to zero. The constant term $\bar{w}$, to be estimated in (7), accounts for the effects of the neglected terms depending on the unknown signal w_t (i.e. for the presence of the nonlinearity).

In the following, let $j = 1, 2, \ldots$ be the iteration counter, and ϑ^{j-1} be the estimate of ϑ in (3) computed at iteration $j-1$. Moreover, let $\bar{n} = \max\{n_a, n_b + n_k - 1, n_g\}$.

3.2 PWA approximation of the nonlinear part

Given the estimate ϑ^{j-1} of the ARX coefficients, standard smoothing techniques are used to recover a suitable signal $\{v_t^j\}_{t=\bar{n}-n_g+1}^N$ according to the estimated linear dynamics

$$y_t = (\vartheta_\zeta^{j-1})^T \zeta_t + v_t^j + (\vartheta_\omega^{j-1})^T \begin{bmatrix} v_{t-1}^j \\ \vdots \\ v_{t-n_g}^j \end{bmatrix} + \varepsilon_t, \tag{8}$$

with $t = \bar{n}+1, ..., N$. Note that (8) is obtained by replacing ϑ with ϑ^{j-1} in (3). In general, the role of the error term ε_t depends on the particular smoothing algorithm chosen. A very simple choice for the considered ARX model is to set $\varepsilon_t = 0$ for all t, and then to recover the signal v_t^j as a particular solution of the set of linear equations resulting from (8).

The next step is to fit a PWA map $f^j(\cdot)$ to the samples $(\boldsymbol{z}_t, v_t^j)$, $t = \bar{n} - n_g + 1, \ldots, N$. To this aim we adopt the bounded-error procedure proposed in Bemporad et al. [2005]. An attractive feature of this method is that the number of modes of the PWA map is automatically estimated from data. given the maximum allowable approximation error.

After reconstructing the PWA map $f^j(\cdot)$, an estimate of the unknown sequence $\{w_t\}_{t=\bar{n}-n_g+1}^N$ can be obtained as

$$w_t^j = f^j(\boldsymbol{z}_t), \quad t = \bar{n} - n_g + 1, \ldots, N. \tag{9}$$

It is stressed that, in the proposed scheme, v_t^j is seen as a "noisy" version of w_t^j, where by "noisy" it is meant that v_t^j will typically include the effects of both the system noise and the model error. Hence, the role of the PWA approximation stage is not only to provide an analytic expression of the static nonlinearity, but also to improve the smoothing of the unknown signal w_t. This feature of the proposed procedure is illustrated through examples in the bottom part of Fig. 4.

3.3 Identification of the linear part

Given the estimated signal $\{w_t^j\}_{t=\bar{n}-n_g+1}^N$, one can form the estimated regression vectors $\{\varphi_t^j\}_{t=\bar{n}+1}^N$ by replacing w_t with w_t^j in the definition (5b) of ω_t. Then, an estimate ϑ^j of the ARX coefficients can be computed through standard linear identification techniques from the model equation

$$y_t - w_t^j = (\vartheta^j)^T \varphi_t^j + \varepsilon_t, \quad t = \bar{n}+1, \ldots, N. \tag{10}$$

The procedure terminates when no significant changes occur to the estimated ARX coefficients between two consecutive iterations. This criterion is implemented by checking if

$$\|\vartheta^j - \vartheta^{j-1}\| \leq \gamma \|\vartheta^j\|, \tag{11}$$

where γ is a positive threshold defined by the user, and $\|\cdot\|$ denotes the Euclidean norm of a vector.

4. NUMERICAL EXAMPLE

In this section, we apply the proposed identification algorithm to an interconnected system with a sigmoid static nonlinearity. Hence, the true system does not belong to the considered model class. Identification results for different noise levels are compared with those obtained by fitting directly to the input-output data $\{u_t, y_t\}$ a black-box PWARX model of the type

$$y_t = g(\boldsymbol{r}_t) + \epsilon_t, \tag{12}$$

where $g(\cdot)$ is a PWA map, $\boldsymbol{r}_t$ is a finite-length regression vector composed by past inputs and outputs, and ϵ_t is an error term. Given $\delta > 0$, the adopted bounded-error identification procedure [Bemporad et al., 2005] estimates a PWARX model with minimum number of modes such that the condition $|\epsilon_t| \leq \delta$ is satisfied for all data points in the estimation data set.

Consider the Lur'e system in Fig. 2. The linear system $\widetilde{\mathcal{L}}$ is described by the ARX model:

$$\begin{aligned} y_t = &- \widetilde{a}_1\, y_{t-1} - \widetilde{a}_2\, y_{t-2} + b_1\,(u_{t-1} - \widetilde{w}_{t-1}) \\ &+ b_2\,(u_{t-2} - \widetilde{w}_{t-2}) + e_t, \end{aligned} \tag{13}$$

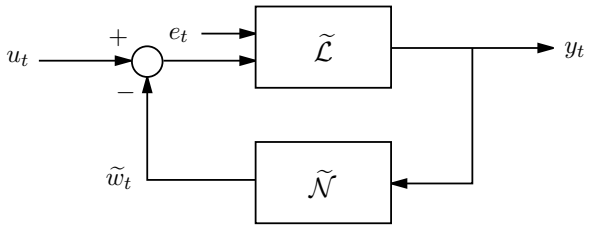

Fig. 2. Lur'e system of the example of Section 4.

where $\widetilde{a}_1 = -1.20$, $\widetilde{a}_2 = 0.85$, $b_1 = 1.00$, $b_2 = -0.10$, while the static block $\widetilde{\mathcal{N}}$ is the sigmoid function

$$\widetilde{\mathcal{N}}(z) = \frac{1 - e^{-5z}}{1 + e^{-5z}} \,. \tag{14}$$

An equivalent representation of the Lur'e system in the LFT form of Fig. 1 can be obtained by defining

$$w_t = -\widetilde{a}_1 \, y_{t-1} - b_1 \, \widetilde{w}_{t-1}. \tag{15}$$

Note that $\widetilde{w}_{t-1} = \widetilde{\mathcal{N}}(y_{t-1})$, and hence we can write $w_t = \mathcal{N}(z_t)$ with $z_t = y_{t-1}$ and

$$\mathcal{N}(z) = -\widetilde{a}_1 \, z - b_1 \widetilde{\mathcal{N}}(z). \tag{16}$$

Moreover, by substituting (15) into (13), the linear system $\mathcal{L}$ of the LFT form is described by the equations

$$\begin{aligned} y_t &= -a_2 \, y_{t-2} + b_1 \, u_{t-1} + b_2 \, u_{t-2} \\ &\quad + w_t + g_1 \, w_{t-1} + e_t \end{aligned} \tag{17a}$$

$$z_t = y_{t-1}, \tag{17b}$$

where $a_2 = \widetilde{a}_2 - \dfrac{\widetilde{a}_1 b_2}{b_1} = 0.73$ and $g_1 = \dfrac{b_2}{b_1} = -0.10$. In the following, we apply the iterative algorithm of Section 3 to identify an LFT model (1) of the Lur'e system with the linear part $\mathcal{L}$ having the structure (17).

The identification experiment is carried out by exciting the Lur'e system with input u_t uniformly distributed in $[-1.5, 1.5]$, while the noise e_t is assumed to be uniformly distributed in $[-\eta_e, \eta_e]$, $\eta_e > 0$. Two data sets of $N = 1000$ data points (u_t, y_t, z_t) are generated for different noise levels: $\eta_e = 0.05$ in the first data set, and $\eta_e = 0.25$ in the second data set.

4.1 Dataset with $\eta_e = 0.05$

As described in Section 3.1, the iterative algorithm is initialized with estimates $a_2^0 = 0.4936$, $b_1^0 = 0.9984$ and $b_2^0 = 0.6572$ obtained by fitting the ARX model

$$y_t = -a_2^0 \, y_{t-2} + b_1^0 \, u_{t-1} + b_2^0 \, u_{t-2} + \bar{w}^{\,0} + \varepsilon_t \tag{18}$$

to the data by least squares, while g_1^0 is set equal to 0. At each iteration, the bounded-error procedure [Bemporad et al., 2005] is applied to approximate the static nonlinearity, while least squares are used for estimation of the ARX coefficients. By setting $\gamma = 0.001$, the algorithm terminates after 13 iterations. The sequence of the estimates of the ARX coefficients is plotted in Fig. 3, showing the convergence of the iterative procedure. Note that, while the parameter b_1 is well estimated already from the beginning, the parameter b_2 is poorly initialized with even the wrong sign. Nevertheless, its estimates converge to the true value quite rapidly. Concerning the estimation of the static nonlinearity, the nonlinear map $\mathcal{N}(z)$ in (16) is finally approximated by a PWA map $f(z)$ with 5 modes and

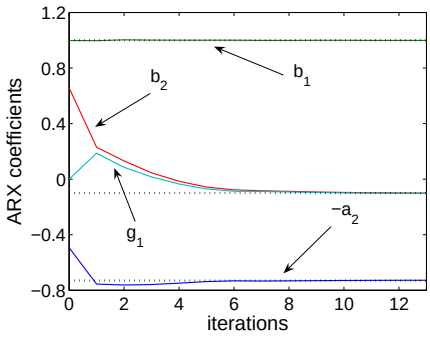

Fig. 3. Sequence of the estimates of the ARX coefficients $-a_2$, b_1, b_2 and g_1 in the case $\eta_e = 0.05$ of Section 4.1 (dotted lines are true values).

maximum absolute error 0.0503 over the interval $|z| \leq 5.5$ where measured data are available. The nonlinear map $\mathcal{N}(z)$ and its PWA approximation $f(z)$ are plotted in Fig. 4 (top left) for $|z| \leq 2$. Fig. 4 (top right) shows how the bound δ of the bounded-error algorithm is selected at each iteration to fit a suitable PWA map to the samples (z_t, v_t^j). At the first iteration, the samples (z_t, v_t^1) are quite scattered (bottom left in Fig. 4), and a large δ (namely, $\delta = 1.5$) is chosen to fit the points within the allowable error δ while keeping the number of modes low. Note that some samples fall outside the error bands: these are points discarded as outliers by the PWA approximation algorithm to obtain linearly separable clusters. As long as the iterations proceed, the samples are less and less scattered, and δ can be reduced. For instance, $\delta = 0.4$ is chosen at the fourth iteration (bottom right).

We stress that the original Lur'e system (13)-(14) is not identifiable from input-output measurements only, because there exist infinite pairs $(\widetilde{\mathcal{L}}, \widetilde{\mathcal{N}})$ giving origin to the same input-output behaviors. Nevertheless, a model of the system (13)-(14) can be recovered from the identified LFT model if additional prior knowledge about the nonlinearity is assumed. For instance, if one knows that the nonlinearity is a saturation, i.e. $\lim_{|z|\to\infty} \frac{\partial \widetilde{\mathcal{N}}(z)}{\partial z} = 0$, and thus, from (16), $\widetilde{a}_1 = -\lim_{|z|\to\infty} \frac{\partial \mathcal{N}(z)}{\partial z}$, the PWA approximation $f(z)$ of $\mathcal{N}(z)$ can be used to obtain an estimate of $\widetilde{a}_1$ as

$$\hat{\widetilde{a}}_1 = -\frac{1}{2}\left(\lim_{z\to-\infty} \frac{\partial f(z)}{\partial z} + \lim_{z\to+\infty} \frac{\partial f(z)}{\partial z} \right) = -1.1911. \tag{19}$$

Then, $\widetilde{a}_2$ is estimated as $\hat{\widetilde{a}}_2 = \hat{a}_2 + \frac{\hat{\widetilde{a}}_1 \hat{b}_2}{\hat{b}_1} = 0.8474$. Both estimates $\hat{\widetilde{a}}_1$ and $\hat{\widetilde{a}}_2$ are very close to their true values. A PWA approximation $\widetilde{f}(z)$ of $\widetilde{\mathcal{N}}(z)$ is finally obtained as $\widetilde{f}(z) = -\frac{1}{\hat{b}_1}\big(f(z) + \hat{\widetilde{a}}_1 \, z\big)$. The nonlinear map $\widetilde{\mathcal{N}}(z)$ and its 5-mode PWA approximation $\widetilde{f}(z)$ are plotted in Fig. 5. The maximum absolute approximation error is 0.0438 over the interval $|z| \leq 5.5$.

For comparison purposes, a black-box PWARX model (12) of the overall Lur'e system is identified by applying the bounded-error identification algorithm [Bemporad et al., 2005] to the same estimation data set. In order to guarantee a fair comparison, the bound $\delta = 0.0870$ is chosen equal to the maximum absolute identification error of the iden-

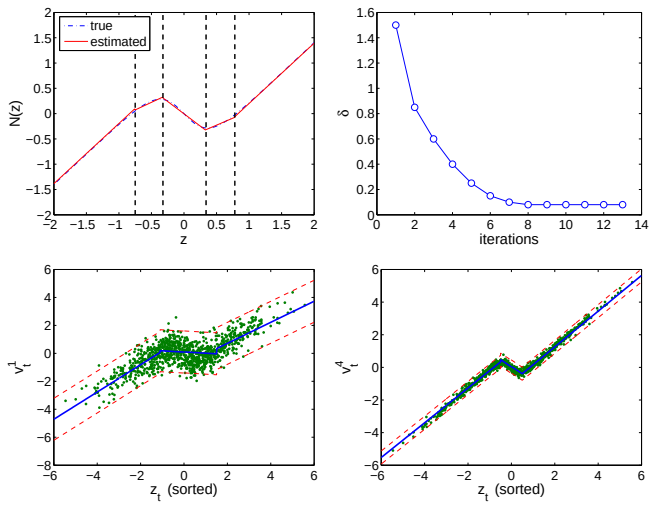

Fig. 4. (Top left) Nonlinear map $\mathcal{N}(z)$ (dash-dotted) and its PWA approximation $f(z)$ (solid) in the case $\eta_e = 0.05$ of Section 4.1. Vertical dashed lines represent the z-domain partition of the PWA map. (Top right) Choices of δ in the PWA approximation step at each iteration of the iterative identification process. (Bottom left) Samples (z_t, v_t^1) and PWA map fitted to these samples for $\delta = 1.5$ (solid). (Bottom right) Samples (z_t, v_t^4) and PWA map fitted to these samples for $\delta = 0.4$ (solid). In the bottom figures, dashed lines limit the error band $\pm\delta$ around the PWA map.

tified LFT model (i.e. the maximum $|\varepsilon_t|$ in (1) evaluated on estimation data). Note that, since $w_t = f(y_{t-1})$ with $f(\cdot)$ a 5-mode PWA map, the identified LFT model admits an equivalent PWARX representation (12) with regression vector $\boldsymbol{r}_t = [\ y_{t-1}\ \ y_{t-2}\ \ u_{t-1}\ \ u_{t-2}\ ]^T$ and up to $5^2 = 25$ modes, corresponding to all possible mode combinations for the pair (w_t, w_{t-1}) in (17). In the considered data set, there exist regression vectors $\boldsymbol{r}_t$ in each of the 25 regions of such a PWARX model. However, 21 modes out of 25 contain a number of points between 7 and 50 in a 4-dimensional space, which makes us expect possible difficulties in the classification step. Indeed, the procedure fails in partitioning the whole data set into linearly separable clusters. A 6-mode PWARX model satisfying the bound δ is obtained by discarding 124 data points, but the corresponding maximum absolute identification error computed on the whole data set is 0.2511. Note that, in this data set where the signal-to-noise ratio is quite large ($\frac{\sigma_y}{\sigma_e} = 58.92$) and the true system is not PWA, the identification error is mainly due to the model error. We thus argue that the PWARX identification procedure experiences difficulties in the classification step due to a deficient distribution of the data points in the zones of the 4-dimensional space where more points would be needed to get the required accuracy. This problem does not show up in the PWA approximation step of the proposed LFT identification algorithm, as can be readily seen, e.g., in Fig. 4 (bottom right).

The identified LFT and PWARX models are compared using $M = 1000$ validation data sets composed by $N = 1000$ data points (u_t, y_t, z_t) each. The validation data sets are generated by feeding the Lur'e system with different realizations of the input and noise signals. Simulated

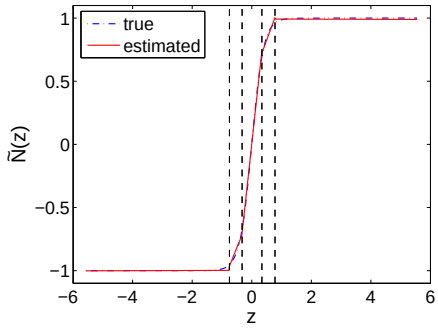

Fig. 5. Sigmoid map $\widetilde{\mathcal{N}}(z)$ (dash-dotted) and its PWA approximation $\widetilde{f}(z)$ (solid) in the case $\eta_e = 0.05$ of Section 4.1. Vertical dashed lines represent the z-domain partition of the PWA map.

outputs $\hat{y}_t$ are obtained by simulating the identified models with inputs u_t only. Also the Lur'e system is simulated without noise, to serve as a reference for evaluating the performance of the two identified models. Validation is carried out by calculating the fit between the system and the simulated outputs. Fig. 6 shows the histogram obtained by grouping the M values of FIT for each model into bins with centers $\{30, 40, 50, 60, 70\}$. The better performance of the LFT model compared to the PWARX model is apparent.

4.2 Dataset with $\eta_e = 0.25$

This data set is more challenging, because the signal-to-noise ratio is $\frac{\sigma_y}{\sigma_e} = 11.25$. The iterative identification procedure is initialized with values $a_2^0 = 0.5010$, $b_1^0 = 0.9571$ and $b_2^0 = 0.5368$ obtained by fitting the ARX model (18) to the data, while g_1^0 is set equal to 0. By setting $\gamma = 0.001$, the algorithm terminates after 9 iterations. The final values $\hat{a}_2 = 0.7357$, $\hat{b}_1 = 0.9915$, $\hat{b}_2 = -0.0982$ and $\hat{g}_1 = -0.0884$ provide very good estimates of the corresponding true parameters. The nonlinear map $\mathcal{N}(z)$ in (16) is approximated by a PWA map $f(z)$ with 3 modes and maximum absolute error 0.1460 over the interval $|z| \leq 7.3$ where measured data are available. Note that the PWA approximation of $\mathcal{N}(z)$ is less accurate compared to the previous identification experiment. This is due to the higher noise level, which results into more scattered samples (z_t, v_t^j), so that the bound δ cannot be reduced below 0.4 in the PWA approximation step of the last algorithm iterations (compare this value with the plot at top right in Fig. 4).

A black-box PWARX model (12) of the overall Lur'e system is also identified by applying the bounded-error identification algorithm to the same estimation data set. The bound $\delta = 0.3777$ is taken equal to the maximum absolute identification error of the identified LFT model. In this case, the identified LFT model admits an equivalent PWARX representation with $3^2 = 9$ modes, but the bounded-error procedure is able to fit to the data a simpler PWARX model with 3 modes.

The identified LFT and PWARX models of the Lur'e system present the same worst-case error on estimation data, so that no clear preference can be accorded to one

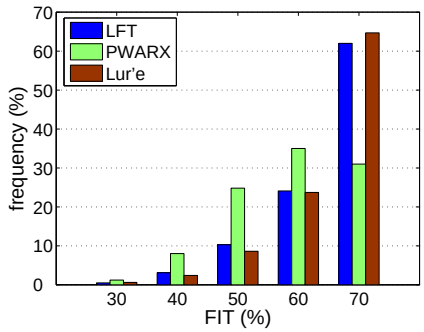

Fig. 6. Histogram obtained in the case $\eta_e = 0.05$ of Section 4.1 by grouping the $M = 1000$ values of FIT for the LFT, PWARX and noiseless Lur'e models into bins with centers $\{30, 40, 50, 60, 70\}$.

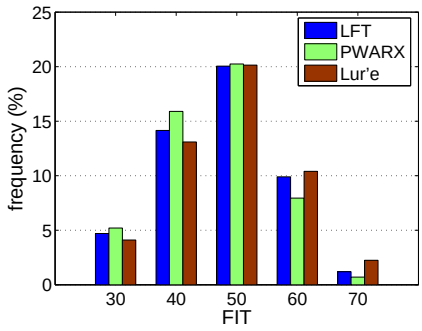

Fig. 7. Histogram obtained in the case $\eta_e = 0.25$ of Section 4.2 by grouping the $M = 1000$ values of FIT for the LFT, PWARX and noiseless Lur'e models into bins with centers $\{30, 40, 50, 60, 70\}$.

model or to the other based on this criterion. The number of parameters is also similar: 10 for the LFT model, and 12 for the PWARX model. As for the previous data set, the two models are further compared using $M = 1000$ validation data sets composed by $N = 1000$ data points (u_t, y_t, z_t) each. Fig. 7 shows the histogram obtained by grouping the M values of FIT for each model into bins with centers $\{30, 40, 50, 60, 70\}$. The performance of the two models is similar, as confirmed by the median of the corresponding distributions, being 47.9141% for the LFT model, and 46.6157% for the PWARX model. We may argue that, in this case, the structure information contained in the data is somehow hidden by the high noise level, so that exploiting the system structure does not lead to significant improvements as in the identification experiment with $\eta_e = 0.05$.

5. CONCLUSIONS AND FUTURE RESEARCH

In this paper an iterative algorithm for PWA identification of complex systems composed by interconnected linear and nonlinear systems has been presented. The proposed approach provides explicit models for both the linear and the nonlinear part of the system. Numerical examples have shown that the LFT models provided by the proposed structure-exploiting algorithm are typically more accurate than the PWARX models obtained from black-box identification of the whole system. This is particularly evident when the noise level is not sufficient to mask the information about the inner interconnection structure contained in the data.

Several open problems are worth to be addressed in the considered identification framework. Identifiability issues related to LFT structures with PWA nonlinearities need to be investigated. Moreover, the proposed identification algorithm should be extended to the case of MIMO systems and multi-output nonlinearities.

REFERENCES

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

M. S. Claassen. *System identification for structured nonlinear systems*. PhD thesis, Department of Mechanical Engineering, University of California at Berkeley, USA, 2001.

K. Hsu, M. Claassen, C. Novara, P. Khargonekar, M. Milanese, and K. Poola. Nonparametric identification of static nonlinearities in a general interconnected system. In *Proc. 16th IFAC World Congress*, Prague, Czech Republic, 2005a.

K. Hsu, T. Vincent, C. Novara, M. Milanese, and K. Poola. Identification of nonlinear maps in interconnected systems. In *Proc. 44th IEEE Conference on Decision and Control and European Control Conference 2005*, pages 6430–6435, Seville, Spain, 2005b.

K. Hsu, T. Vincent, and K. Poola. A kernel-based approach to structured nonlinear system identification Part I: Algorithms. In *Proc. 14th IFAC Symposium on System Identification*, pages 1198–1203, Newcastle, Australia, 2006.

J.-N. Lin and R. Unbehauen. Canonical piecewise-linear approximations. *IEEE Transactions on Circuits and Systems – I: Fundamental Theory and Applications*, 39 (8):697–699, 1992.

K. S. Narendra and P. G. Gallman. An iterative method for the identification of nonlinear systems using a Hammerstein model. *IEEE Transaction on Automatic Control*, 11(3):546–550, 1966.

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.

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

F. Previdi and M. Lovera. Identification of a class of non-linear parametrically varying models. *International Journal of Adaptive Control and Signal Processing*, 17 (1):33–50, 2003.

J. Roll. *Local and piecewise affine approaches to system identification*. PhD thesis, Department of Electrical Engineering, Linköping University, Linköping, Sweden, 2003.