An Outer Bounding Ellipsoid-Based Algorithm for Identifying Piecewise Affine Output-Error Models

Abdelhak Goudjil[1], Mathieu Pouliquen[2], Eric Pigeon[2], Mostafa-Kamel Smail[1,3]
Abdelwahhab Boudjelal[4], Ali Moradvandi[5]

[1] *IPSA Institute of Polytechnic Science and Aeronautics, 63 Bd de Brandebourg, 94200 Ivry-sur-Seine, France*
[2] *LIS Systems Engineering Laboratory, Normandie University, UNICAEN, ENSICAEN, 14050 Caen, France*
[3] *GeePs Paris, UMR CNRS 8507, CentraleSupélec, Université Paris-Saclay, 91190 Gif-sur-Yvette, France*
[4] *École d'Ingénieurs CESI - LINEACT, 14000 Caen, France*
[5] *Delft Center for Systems and Control, Delft University of Technology, 2628 CD, Delft, The Netherlands*

***Abstract*— The identification of Piecewise Affine Output-Error (PWA-OE) model from input-output data involves estimating a finite set of parameters for the affine output-error submodels and partitioning the regressor space accordingly. Traditional least squares methods fail to provide consistent estimates in the presence of output-error noise, whereas prediction error methods ensure parameter consistency. This paper introduces an enhanced Outer Bounding Ellipsoid (OBE) algorithm tailored for PWA-OE model identification under bounded noise, leveraging the prediction error approach. This class of algorithms is recognized for its computational efficiency. Furthermore, through a numerical example, the proposed method demonstrates excellent performance, achieving accurate parameter estimation with high reliability.**

***Index Terms*— PieceWise Affine system, Output-Error, bounded noise, OBE algorithm.**

I. INTRODUCTION

Hybrid systems represent a class of dynamic systems characterized by the seamless integration of continuous and discrete behaviors. They offer a robust framework for modeling complex real-world phenomena. A distinguished subclass within hybrid systems is the Piecewise Affine (PWA) system. PWA systems are unique in their ability to describe nonlinear dynamics through a set of interconnected linear subsystems, each defined over distinct regions of the state space. These systems combine the accuracy of linear modeling with the flexibility to handle nonlinearities, making PWA systems particularly advantageous for modeling and controlling systems with complex and hybrid characteristics.

The principal difficulty in identifying PWA systems stems from the composite nature of the collected data, which encompasses information from various subsystems. Thus, the identification process entails both assigning each data sample to its corresponding subsystem and estimating the parameter vectors for all the subsystems. Several methodologies have been developed in the literature for the identification of PWA systems utilizing a PWA autoregressive exogenous model (PWARX). Such methodologies include algebraic methods ([1]), clustering-based techniques ([2] and [3]), Bayesian approaches ([4]), optimization-based strategies ([5]), bounded error approaches ([6] and [7]), and recursive methods ([8] and [9]). A comprehensive overview of the majority of these methods is provided in [10]. Despite its widespread application in identifying piecewise affine systems, the PWARX model has several limitations. The autoregressive nature of the PWARX model makes it susceptible to error accumulation, increased sensitivity to noise, and potential stability issues. These limitations undermine their suitability for systems exhibiting abrupt changes or significant disturbances. To mitigate these limitations, we investigate the PWA-OE model, which directly minimizes output error, thus enhancing accuracy, robustness, and stability in managing nonlinear system dynamics.

To the best of our current knowledge, only a limited number of methodologies have been proposed for the identification of PWA-OE model. This specific model presents the advantage of excluding the disturbance model from the process model, as the noise is directly added to the output. Nevertheless, the selection of the PWA-OE model introduces additional complexities to the identification process, since the inputs to the PWA mapping rely on the unobserved 'noise-free' output. The methods discussed in ([11], [12], [13], [14], and [15]) can be employed for the identification of the PWA-OE model. The methodology proposed in [11] presupposes that the discrete mode sequence is known. The method proposed in [14] involves the application of the instrumental variables algorithm for model identification. However, the selection of suitable instrumental variables is crucial, as they must be uncorrelated with the system noise while maintaining a correlation with the inputs. Moreover, the aforementioned methodologies perform optimally under the assumption of typical stochastic noise. The precision of these methods is reliant on this assumption, which necessitates prior knowledge of the noise characteristics influencing the system's output. However, in numerous real-world scenarios, it is often unfeasible to meet the stochastic noise assumption.

To identify the PWA-OE model, we adapt the algorithms outlined in [16] and [9]. The proposed algorithm addresses several previous issues and facilitates the identification of the

PWA-OE model within a bounded noise framework. In real-world applications, the distribution of noise is often challenging to accurately describe. Consequently, a more practical strategy to address the identification problem is to adopt the unknown but bounded noise framework. In this setting, the noise is considered to be unknown, yet is constrained within known bounds. It may follow a uniform distribution, a truncated normal distribution, or any other type, provided that an upper limit on its magnitude is available. This approach allows for the construction of a membership set of model parameters that is consistent with both the measurements and the noise bound. In particular, it facilitates the adaptation of an Outer Bounding Ellipsoid (OBE) algorithm - known for its low computational complexity ([17], [18], [19], and [20]) to identify the PWA-OE model. The proposed algorithm operates in two stages: first, it estimates the active submodel, computes the overall parameter vector and classifies each data point into the appropriate class. Subsequently, the algorithm focuses on estimating the polyhedral partition using Support Vector Machines (SVM).

The rest of this manuscript is structured as follows: Section II delineates the parameterization of the PWA-OE model and formulates the identification problem. Section III introduces the proposed algorithm. To illustrate its efficacy, Section IV presents a numerical example. Lastly, Section V concludes the study.

II. PROBLEM FORMULATION

A. Parametrization of a PWA-OE model

Consider a discrete-time piecewise affine system modeled using the PWA-OE model as follows:

$$\begin{cases} y(k) = y_{\mathrm{o}}(k) + n(k) \\ y_{\mathrm{o}}(k) = f(\varphi_{\mathrm{o}}(k)) \\ n(k) \leq \delta_n \ , \quad \forall\ k \end{cases} \tag{1}$$

In the above equations, $y(k) \in \mathbb{R}$ and $y_{\mathrm{o}}(k) \in \mathbb{R}$ represent the output with noise and the noise-free output, respectively. The unknown but bounded noise is denoted as $n(k) \in \mathbb{R}$, with δ_n representing its known upper bound. The PWA map function $f(\varphi_{\mathrm{o}}(k))$ is articulated as:

$$f(\varphi_{\mathrm{o}}(k)) = \begin{cases} \begin{bmatrix} \varphi_{\mathrm{o}}(k) \\ 1 \end{bmatrix}^\top \theta_1 & \text{if } \varphi_{\mathrm{o}}(k) \in \mathscr{X}_1, \\ \quad\vdots & \\ \begin{bmatrix} \varphi_{\mathrm{o}}(k) \\ 1 \end{bmatrix}^\top \theta_s & \text{if } \varphi_{\mathrm{o}}(k) \in \mathscr{X}_s, \end{cases} \tag{2}$$

s denotes the number of submodels, and $\{\theta_1, \ldots, \theta_s\} \in \mathbb{R}^n$ represent their parameter vectors. The regressor vector $\varphi_{\mathrm{o}}(k)$ is expressed as:

$$\varphi_{\mathrm{o}}(k) = [y_{\mathrm{o}}(k-1) \ \cdots \ y_{\mathrm{o}}(k-n_a) \ \ x(k-1) \ \cdots \ x(k-n_b)]^\top \tag{3}$$

$x(k) \in \mathbb{R}$ signifies the system's input, while n_a and n_b denote, respectively, the orders of the system with respect to the output and the input, as well as $n = n_a + n_b + 1$. It is assumed that the orders of the system are predetermined and identical across all submodels. This assumption is commonly used in the literature. The regions $\{\mathscr{X}_i\}_{i=1}^s$ delineate a polyhedral division of the closed and bounded set $\mathscr{X} \in \mathbb{R}^n$, that is, $\cup_{i=1}^s \mathscr{X}_i = \mathscr{X}$ and $\mathscr{X}_i \cap \mathscr{X}_j = \emptyset$ if $i \neq j$.

At each step k, a submodel is triggered, and its output is denoted as $y(k)$. Consequently, for each step k, there exists a parameter vector θ_i such that the absolute difference between y_k and $\begin{bmatrix} \varphi_{\mathrm{o}}(k) \\ 1 \end{bmatrix}^\top \theta_i$ is less than or equal to δ_n. This parameter vector corresponds to the active submodel at step k.

Objective: *It is desired to estimate $\{\theta_i\}_{i=1}^s$ and $\{\mathscr{X}_i\}_{i=1}^s$ from a set $\{x(k), y(k)\}_{k=1}^N$, which are generated by a system described by (1) and (2).*

B. Reformulation of the identification problem

At each step k, the unidentified active submodel is designated as λ_k, with λ_k taking values from the set $\{1, \cdots, s\}$. At each k, the output is expressed by the parameter vector θ_{λ_k} as:

$$y(k) = \bar{\varphi}_{\mathrm{o}}^\top(k)\theta_{\lambda_k} + n(k)$$

with

$$\bar{\varphi}_{\mathrm{o}}(k) = \begin{bmatrix} \varphi_{\mathrm{o}}(k) \\ 1 \end{bmatrix} \tag{4}$$

The system output, $y(k)$, can also be represented using different parameter vectors as follows

$$\begin{cases} y(k) = \bar{\varphi}_{\mathrm{o}}^\top(k)\theta_1 + n(k) + \bar{\varphi}_{\mathrm{o}}^\top(k)(\theta_{\lambda_k} - \theta_1) \\ \vdots \\ y(k) = \bar{\varphi}_{\mathrm{o}}^\top(k)\theta_{\lambda_k} + n(k) \\ \vdots \\ y(k) = \bar{\varphi}_{\mathrm{o}}^\top(k)\theta_s + n(k) + \bar{\varphi}_{\mathrm{o}}^\top(k)(\theta_{\lambda_k} - \theta_s) \end{cases} \tag{5}$$

Therefore, equation (5) can be restated as:

$$\begin{bmatrix} y(k) \\ \vdots \\ y(k) \\ \vdots \\ y(k) \end{bmatrix} = \begin{bmatrix} \bar{\varphi}_{\mathrm{o}}^\top(k)\theta_1 \\ \vdots \\ \bar{\varphi}_{\mathrm{o}}^\top(k)\theta_{\lambda_k} \\ \vdots \\ \bar{\varphi}_{\mathrm{o}}^\top(k)\theta_s \end{bmatrix} + \begin{bmatrix} n(k) + \bar{\varphi}_{\mathrm{o}}^\top(k)(\theta_{\lambda_k} - \theta_1) \\ \vdots \\ n(k) \\ \vdots \\ n(k) + \bar{\varphi}_{\mathrm{o}}^\top(k)(\theta_{\lambda_k} - \theta_s) \end{bmatrix} \tag{6}$$

Let's define the overall parameter vector $\Theta \in \mathbb{R}^{ns\times 1}$, the extended output vector $Y(k) \in \mathbb{R}^{s\times 1}$ and the extended noise vector $N(k) \in \mathbb{R}^{s\times 1}$ and as:

$$\Theta = \begin{bmatrix} \theta_1 \\ \vdots \\ \theta_s \end{bmatrix}, Y(k) = \begin{bmatrix} y(k) \\ \vdots \\ y(k) \end{bmatrix}, N(k) = \begin{bmatrix} n(k) + \bar{\varphi}_{\mathrm{o}}^\top(k)(\theta_{\lambda_k} - \theta_1) \\ \vdots \\ n(k) \\ \vdots \\ n(k) + \bar{\varphi}_{\mathrm{o}}^\top(k)(\theta_{\lambda_k} - \theta_s) \end{bmatrix}$$

Thus, equation (6) can be restated as:

$$Y(k) = \psi^\top(k)\Theta + N(k) \tag{7}$$

with $\psi(k) = I_s \otimes \bar{\varphi}_o(k)$, where $\otimes$ denotes the Kronecker product.

Redefined Objective: *Considering a dataset of measurements $\{x(k), Y(k)\}_{k=1}^N$ produced by a system as defined in (1) and (2), the task is to estimate Θ and $\{\mathscr{X}_i\}_{i=1}^s$.*

III. IDENTIFICATION ALGORITHM

A. *Notations and definitions*

For the rest of the paper, the following notations are considered:

- $\hat{\lambda}_k$: Estimate of the unknown active submodel λ_k at step k.
- $\widehat{\Theta}(k)$: Estimate of Θ at step k: $\widehat{\Theta}(k) = \begin{bmatrix} \hat{\theta}_1(k) \\ \vdots \\ \hat{\theta}_s(k) \end{bmatrix}$ where, for any $i \in \{1, \ldots, s\}$, $\hat{\theta}_i(k)$ represents an estimate of θ_i.

Definition 1: The a priori predictors, $\hat{Y}_{k/k-1} \in \mathbb{R}^s$, and the a posteriori predictors, $\hat{Y}_{k/k} \in \mathbb{R}^s$, are defined as follows:

$$\begin{cases} \hat{Y}_{k/k-1} = \widehat{\psi}^\top(k)\widehat{\Theta}(k-1) = \begin{bmatrix} \widehat{\bar{\varphi}}^\top(k)\hat{\theta}_1(k-1) \\ \vdots \\ \widehat{\bar{\varphi}}^\top(k)\hat{\theta}_s(k-1) \end{bmatrix} = \begin{bmatrix} \hat{y}^{(1)}_{k/k-1} \\ \vdots \\ \hat{y}^{(s)}_{k/k-1} \end{bmatrix} \\ \hat{Y}_{k/k} = \widehat{\psi}^\top(k)\widehat{\Theta}(k) = \begin{bmatrix} \widehat{\bar{\varphi}}^\top(k)\hat{\theta}_1(k) \\ \vdots \\ \widehat{\bar{\varphi}}\top(k)\hat{\theta}_s(k) \end{bmatrix} = \begin{bmatrix} \hat{y}^{(1)}_{k/k} \\ \vdots \\ \hat{y}^{(s)}_{k/k} \end{bmatrix} \end{cases}$$

where $\{\hat{y}^{(i)}_{k/k-1}\}_{i=1,\cdots,s} \in \mathbb{R}$ and $\{\hat{y}^{(i)}_{k/k}\}_{i=1,\cdots,s} \in \mathbb{R}$ represent the a priori predictors and the posteriori predictors linked with each unique submodel i. $\widehat{\psi}(k)$ is an estimate of the matrix $\psi(k)$ achieved by substituting the elusive and noise-free output $y_o(k-i)$ with its a posteriori calculated predictors, connected to the estimated active submodel $\hat{y}^{(\hat{\lambda}_{k-i})}_{k-i/k-i}$ in the regressor data vector:

$$\widehat{\psi}(k) = I_s \otimes \widehat{\bar{\varphi}}(k), \quad \text{with} \qquad \widehat{\bar{\varphi}}(k) = \begin{bmatrix} \widehat{\varphi}_o(k) \\ 1 \end{bmatrix} \tag{8}$$

and

$$\widehat{\varphi}_o(k) = \begin{bmatrix} \hat{y}^{(\hat{\lambda}_{k-1})}_{k-1/k-1} & \cdots & \hat{y}^{(\hat{\lambda}_{k-n_a})}_{k-n_a/k-n_a} & x(k-1) & \cdots & x(k-n_b) \end{bmatrix}^\top$$

Definition 2: The a priori prediction error $\varepsilon_{k/k-1} \in \mathbb{R}^s$ and the a posteriori prediction error $\varepsilon_{k/k} \in \mathbb{R}^s$ are defined as:

$$\begin{cases} \varepsilon_{k/k-1} = \begin{bmatrix} \varepsilon^{(1)}_{k/k-1} \\ \vdots \\ \varepsilon^{(s)}_{k/k-1} \end{bmatrix} = Y(k) - \widehat{\psi}^\top(k)\widehat{\Theta}(k-1) \\ \varepsilon_{k/k} = \begin{bmatrix} \varepsilon^{(1)}_{k/k} \\ \vdots \\ \varepsilon^{(s)}_{k/k} \end{bmatrix} = Y(k) - \widehat{\psi}^\top(k)\widehat{\Theta}(k) \end{cases}$$

B. *Identification Algorithm*

The proposed algorithm for identifying the PWA-OE model is outlined in Table I. It comprises two distinct stages: the first stage estimates the active submodel, estimates the overall parameter vector, and classifies each data point into the appropriate class. The second stage estimates the regions using SVM.

1) ***Estimation of $\{\lambda_k\}_{k=1}^N$, $\widehat{\Theta}$ and classification of data:*** After initialization, at each step k, the algorithm estimates the active submodel designated by $\hat{\lambda}_k$ by assigning the new data to the submodel that yields the lowest a priori prediction error, as shown by Equation (9). The value of this lowest a priori prediction error is denoted as ζ_k. $\widehat{\Theta}(k)$ is then updated employing an algorithm that utilizes a modified Outer Bounding Ellipsoid approach as specified by the equations (10), (11) and 12). It is important to note that, at each k, only the parameter vector linked to $\hat{\theta}_{\hat{\lambda}_k}(k)$ is updated in $\widehat{\Theta}(k)$. This update is performed using the framework of the symmetric matrix $\sigma_k \in \mathbb{R}^{s\times s}$, which is constructed based on the definition of the matrix $\Delta_k \in \mathbb{R}^{s\times s}$ described in Table I. Based on the values of the lowest a priori prediction error ζ_k and $\widehat{\psi}^\top(k)P_{k-1}\widehat{\psi}(k)$ two cases occur:

- if $\widehat{\psi}^\top(k)P_{k-1}\widehat{\psi}(k) > 0$ and $|\zeta_k| > \delta_e$, then only the parameter vector linked to $\hat{\theta}_{\hat{\lambda}_k}$ is updated in $\widehat{\Theta}(k)$. This update is done using the matrix $\Delta_k \in \mathbb{R}^{s\times s}$.
- Otherwise, if $\widehat{\psi}^\top(k)P_{k-1}\widehat{\psi}(k) = 0$ or $|\zeta_k| \leq \delta_e$, then σ_k is set to 0, which leads to $\Gamma_k = 0$. In this scenario, the adaptation process is static, and the algorithm is unable to determine $\widehat{\Theta}(k)$ that surpasses $\widehat{\Theta}(k-1)$. Therefore, $\widehat{\Theta}(k) = \widehat{\Theta}(k-1)$.

Once the active submodel is estimated and $\widehat{\Theta}(k)$ is updated, $(y(k), \widehat{\bar{\varphi}}(k))$ is linked to the class $D_{\hat{\lambda}_k}$. By forming the classes $\{D_i\}_{i=1}^s$, it becomes possible to estimate the regions $\{\mathscr{X}_i\}_{i=1}^s$ in the second stage.

Remark 1: The condition $\widehat{\psi}^\top(k)P_{k-1}\widehat{\psi}(k) > 0$ for updating $\widehat{\Theta}(k)$ is satisfied by using persistent excitation ([21]). In cases where this condition is not met, indicating insufficient information for an update, the adaptation is halted.

Theorem 1: Consider a discrete-time piecewise affine system modeled using the PWA-OE model as shown in subsection II-A. If persistent excitation is used, then the PWA-OE identification algorithm presented in Tab. I ensures that $\varepsilon_{k/k}$ satisfies:

$$|\varepsilon_{k/k}(\hat{\lambda}_k)| \leq \delta_n \tag{15}$$

■

Proof 1: Based on the definition of $\varepsilon_{k/k}$ and the equation (10), $\varepsilon_{k/k}$ can be rewritten as:

$$\varepsilon_{k/k} = Y(k) - \widehat{\psi}^\top(k)(\widehat{\Theta}(k-1) + \Gamma_k\varepsilon_{k/k-1}) \tag{16}$$

Thus,

$$\varepsilon_{k/k} = \varepsilon_{k/k-1} - \widehat{\psi}^\top(k)\Gamma_k\varepsilon_{k/k-1} = (I_s - \widehat{\psi}^\top(k)\Gamma_k)\varepsilon_{k/k-1}$$

TABLE I

THE PWA-OE IDENTIFICATION ALGORITHM

Initialization:

- $\widehat{\Theta}(0)$ is randomly initialized.
- γ: forgetting factor selected between 0 and 1
- $P_0 = p_0 I_{ns}$ with $p_0 \gg 1$.
- for $k \leq \max\{n_a, n_b\}$, $\hat{Y}_{k/k-1} = \hat{Y}_{k/k} = 0$
- $\{D_i\}_{i=1}^{s} = \emptyset$

I- Estimation of $\{\hat{\lambda}_k\}_{k=1}^{N}$, $\widehat{\Theta}$ and classification of data

For $k = \max\{n_a, n_b\} + 1, \cdots, N$

1- Estimate $\hat{\lambda}_k$ the active submodel:

$$\begin{cases} \hat{\lambda}_k = \underset{i=1,\cdots,s}{\operatorname{argmin}} \mid \varepsilon_{k/k-1}(i) \mid \\ \zeta_k = \min\limits_{i=1,\cdots,s} \varepsilon_{k/k-1}(i) = \varepsilon_{k/k-1}(\hat{\lambda}_k) \end{cases} \tag{9}$$

2- Update the estimate $\widehat{\Theta}(k)$

$$\widehat{\Theta}(k) = \widehat{\Theta}(k-1) + \Gamma_k \varepsilon_{k/k-1} \tag{10}$$

with

$$\begin{cases} \Gamma_k = P_{k-1}\widehat{\psi}(k)\sigma_k(\gamma I_s + \widehat{\psi}^\top(k)P_{k-1}\widehat{\psi}(k)\sigma_k)^{-1} \\ P_k = \dfrac{1}{\gamma}(I_{s(n_a+n_b+1)} - \Gamma_k \widehat{\psi}^\top(k))P_{k-1} \end{cases} \tag{11}$$

and

$$\sigma_k = \begin{cases} \gamma(\widehat{\psi}^\top(k)P_{k-1}\widehat{\psi}(k))^{-1}(\Delta_k - I_s) \\ \quad \text{if} \quad \widehat{\psi}^\top(k)P_{k-1}\widehat{\psi}(k) > 0 \quad \text{and} \quad \mid \zeta_k \mid > \delta_n \\ 0 \quad \text{otherwise} \end{cases} \tag{12}$$

where

$$\Delta_k = \begin{array}{c} \hat{\lambda}_k^{th} \\ \begin{bmatrix} 1 & \cdots & \cdots & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \dfrac{\mid \zeta_k \mid}{\delta_n} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & \cdots & 1 \end{bmatrix} \end{array} \ \hat{\lambda}_k^{th} \tag{13}$$

3- Classify the data into the appropriate class

$$(y(k), \widehat{\varphi}(k)) \in D_{\hat{\lambda}_k} \tag{14}$$

End For

II- Estimation of regions $\{\widehat{\mathscr{X}_i}\}_{i=1}^{s}$

- Use $\{D_i\}_{i=1}^{s}$ and SVM to estimate the regions $\{\widehat{\mathscr{X}_i}\}_{i=1}^{s}$

IV- Return $\widehat{\Theta}_k$ and $\{\widehat{\mathscr{X}_i}\}_{i=1}^{s}$

Using the expression of Γ_k given in the equation (11), this yields:

$$\varepsilon_{k/k} = \gamma(\gamma I_s + \widehat{\psi}^\top(k)P_{k-1}\widehat{\psi}(k)\sigma_k)^{-1}\varepsilon_{k/k-1} \tag{17}$$

If the condition of persistent excitation is fulfilled (that's means: $\widehat{\psi}^\top(k)P_{k-1}\widehat{\psi}(k) > 0$), then the following two scenarios arise:

- If $\zeta_k \leq \delta_n$, then $\sigma_k = 0$, consequently we get:
$$\mid \varepsilon_{k/k}(\hat{\lambda}_k) \mid = \mid \varepsilon_{k/k-1}(\hat{\lambda}_k) \mid$$
This gives $\mid \varepsilon_{k/k}(\hat{\lambda}_k) \mid \leqslant \delta_n$
- If $\zeta_k > \delta_n$, then according to the value of σ_k in (12) for $\sigma_k \neq 0$:
$$\varepsilon_{k/k} = \Delta_k^{-1}\varepsilon_{k/k-1}$$
Therefore
$$\varepsilon_{k/k}(\hat{\lambda}_k) = \frac{\delta_n}{\zeta_k}\ \varepsilon_{k/k-1}(\hat{\lambda}_k)$$
Thus using equation (9) we get $\mid \varepsilon_{k/k}(\hat{\lambda}_k) \mid = \delta_n$.

Taking into account the two cases, it can be concluded that $\mid \varepsilon_{k/k}(\hat{\lambda}_k) \mid \leq \delta_n$.

■

The obtained outcome aligns with the intended goal of the study. The PWA-OE identification algorithm produces an estimation of $\widehat{\Theta}(k)$ that satisfies the condition $\mid \varepsilon_{k/k}(\hat{\lambda}_k) \mid \leq \delta_n$ at each time k. Therefore, δ_n serves as a limit for the a posteriori error of the estimated active submodel.

Remark 2: The successful convergence of the algorithm is fundamentally dependent on the initialization process and the precise estimation of $\hat{\lambda}_k$ at each step k. The negative consequences arising from inaccurate estimates of the active submodel, as well as from improper initialization, can be notably mitigated through the strategic use of the forgetting factor γ. Furthermore, to reduce the impact of initialization and enhance the algorithm's convergence, it is advisable to repeat the first stage of the proposed algorithm multiple times. Additionally, when the noise bound δ_n is unknown, a feasible strategy for selecting the threshold δ_n involves iterating the first stage of the algorithm multiple times with progressively decreasing values of δ_n, using the final estimation of each iteration as the initial condition for the subsequent iteration. This method facilitates the optimal selection of δ_n and improves both the stability and convergence of the algorithm.

*2) **Estimation of regions** $\{\mathscr{X}_i\}_{i=1}^{s}$:* The first stage of the algorithm involves the classification of the data into distinct classes $\{D_i\}_{i=1}^{s}$. The subsequent stage involves the estimation of parameters pertaining to regions $\{\widehat{\mathscr{X}_i}\}_{i=1}^{s}$ that delineate the polyhedral partition, whereby $\widehat{\varphi}_o(k) \in \widehat{\mathscr{X}_i}$ if $(y(k), \widehat{\varphi}(k)) \in D_i$. This process requires the identification of the boundaries that define each region and employing linear classifiers to distinguish the s regions. This topic has been widely studied in the literature, with numerous methodologies that use multi-class SVM ([22], [3], and [7]).

The application of SVM is based on its robust theoretical foundations, its ability to manage substantial data sets with efficiency, and its superior performance in distinguishing the s regions of the polyhedral partition. Consequently, this study also employs multi-class SVM to estimate the parameter vectors of the polyhedral partition. For a comprehensive discussion on the application of SVM and multi-class SVM, readers are directed to the following references: [23], [24], [6], and [25].

IV. Numerical Results

Consider a discrete-time piecewise affine system modeled using the PWA-OE model:

$$y_o(k) = \begin{cases} -0.4y_o(k-1) + x(k-1) + 1.5 + n(k) \\ \qquad \text{if } \; 4y_o(k-1) - x(k-1) + 10 < 0 \\[2ex] 0.5y_o(k-1) - x(k-1) - 1.5 + n(k) \\ \qquad \text{if} \begin{cases} 4y_o(k-1) - x(k-1) + 10 \geq 0 \\ \text{and} \\ 5y_o(k-1) + x(k-1) - 6 \leq 0 \end{cases} \\[2ex] -0.3y_o(k-1) + 0.5x(k-1) - 1.7 + n(k) \\ \qquad \text{if } \; 5y_o(k-1) + x(k-1) - 6 > 0 \end{cases}$$

The system has orders $n_a = 1$ and $n_b = 1$. The excitation sequence $x(k)$ is a random sequence of zero mean uniformly distributed over the interval $[-2, 2]$. Similarly, noise $n(k)$ is a random sequence of zero mean with a uniform distribution in $[-\delta_n, \delta_n]$, where δ_n is tuned to achieve a specified Signal-to-Noise Ratio (SNR).

A sequence of length $N = 2000$ was initially used to assess the performance of the proposed method. The threshold δ_n was chosen to achieve a SNR of 30 dB. The evolution of the posterior prediction error $\varepsilon_{k/k}(\hat{\lambda}_k)$, associated with the estimated active submodel, is illustrated in Fig. 1. It is apparent that the algorithm effectively ensures the boundedness of the posterior prediction error $\varepsilon_{k/k}(\hat{\lambda}_k)$.

In the second case, a Monte Carlo simulation with 100 runs was performed on sequences of length $N = 2000$. The resulting average estimates and standard deviations, obtained for an SNR of 30 dB, are presented in Table II. The results confirm the effectiveness of the proposed approach, as the estimated parameters closely match the true values and show low variability across runs. It is worth noting that the computational time required for this simulation is quite low. Specifically, the Monte Carlo simulation results (100 runs) were obtained in 13.53 s using Matlab R2024a on a 2.3 GHz Intel Core i7 processor.

Furthermore, to assess the impact of noise on estimation accuracy, the noise amplitude δ_n was varied to adjust the SNR. For each SNR level, a Monte Carlo simulation with 100 runs was conducted using sequences of length $N = 2000$. The evaluation was based on the following statistical indicators:

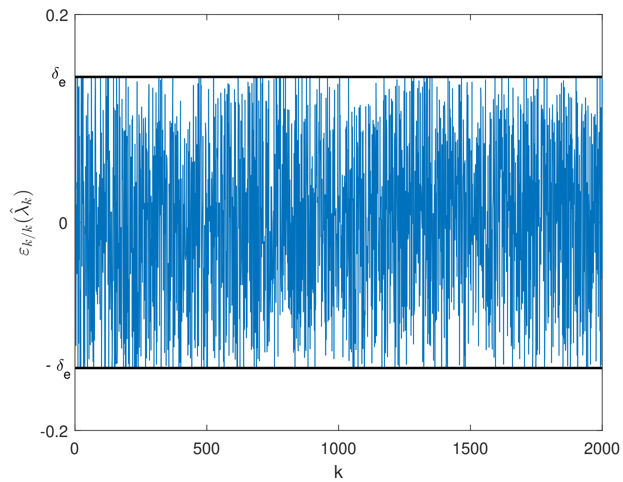

Fig. 1. $\varepsilon_{k/k}(\hat{\lambda}_k)$ as function of time.

$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\theta}_3$
-0.4017 $\pm$ 0.0008	0.4959 $\pm$ 0.0047	-0.3018 $\pm$ 0.0014
0.9985 $\pm$ 0.0009	-0.9998 $\pm$ 0.0019	0.4998 $\pm$ 0.0017
1.4955 $\pm$ 0.0007	-1.5089 $\pm$ 0.0018	-1.6892 $\pm$ 0.0013

TABLE II

Mean and Standard Deviation Analysis of Parameter Vectors for an SNR of 30 dB

- The FIT: $\quad \text{FIT} = 100\left(1 - \frac{\|y-\hat{y}\|}{\|y-\bar{y}\|}\right)\%$.
  The % fit is a measure of how well the estimated output $\hat{y}$ matches the true output y. The mean of y is denoted as $\bar{y}$.
- The Parameter Estimation Error (PEE): $\text{PEE} = \| \Theta^* - \widehat{\Theta} \|$

Figure 2 presents the FIT values obtained for various SNR levels, demonstrating that the proposed algorithm achieves a better FIT as the SNR increases. This result is confirmed in Figure 3, which depicts the relationship between $\| \Theta^* - \widehat{\Theta} \|$ and SNR. It is evident that a higher SNR leads to better estimation quality.

Finally, the polyhedral partition is estimated using multi-class SVM. The resultant average estimation of the hyperplane parameters constituting the polyhedral partition, for SNR$= 30dB$, is given as follows:

$$\begin{cases} \hat{\mathscr{X}}_1 : \begin{bmatrix} 4.0085 & -1.0225 & 10.0132 \end{bmatrix} \\[2ex] \hat{\mathscr{X}}_2 : \begin{cases} \hat{\mathscr{X}}_{21} : \begin{bmatrix} 4.0085 & -1.0225 & 10.0132 \end{bmatrix} \\[1ex] \hat{\mathscr{X}}_{23} : \begin{bmatrix} 5.0019 & 1.0301 & -6.0261 \end{bmatrix} \end{cases} \\[2ex] \hat{\mathscr{X}}_3 : \begin{bmatrix} 5.0019 & 1.0301 & -6.0261 \end{bmatrix} \end{cases}$$

The estimated partition closely resembles the true partition, indicating the precision of SVM.

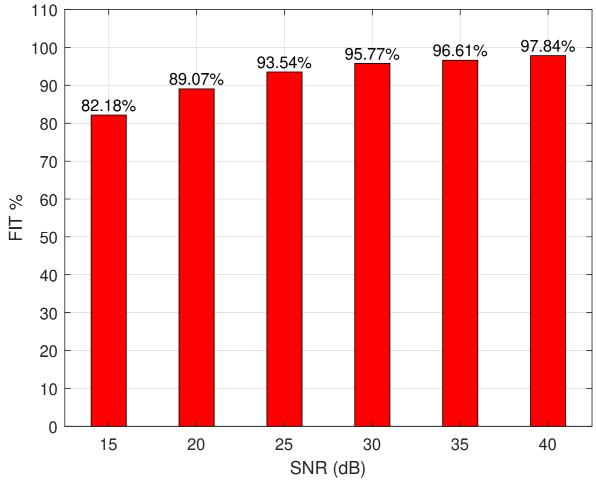

Fig. 2. FIT as function of SNR.

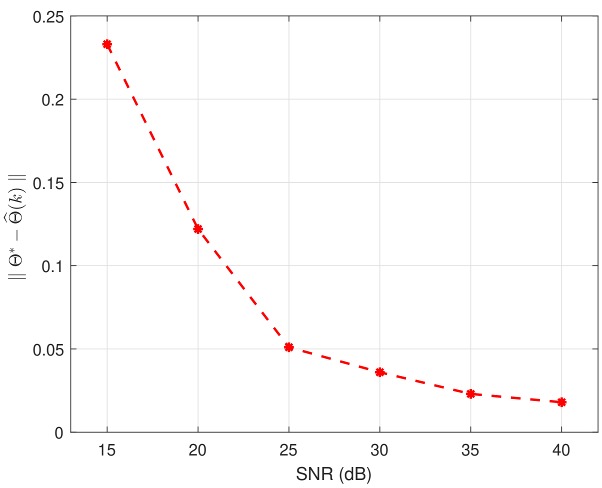

Fig. 3. $\| \Theta^* - \widehat{\Theta} \|$ as function of SNR.

V. CONCLUSION

This paper has introduced an identification algorithm specifically designed for PWA systems described by the PWA-OE model. The numerical results substantiate the precision and efficacy of the algorithm in accurately identifying the PWA systems described by the PWA-OE models. Given that the algorithm depends solely on the boundedness assumption of noises, it is well-suited for identifying real-world applications. Furthermore, the proposed algorithm offers both computational efficiency and ease of implementation. Future research endeavors will focus on conducting convergence analysis and examining stability properties to gain deeper insights into the algorithm's performance under varying conditions and to ensure its reliability in practical scenarios. As a practical application, the algorithm's utility in textured image segmentation is currently being explored.

REFERENCES

[1] R. Vidal, S. Soatto, Y. Ma, and S. Sastry. An algebraic geometric approach to the identification of a class of linear hybrid systems. *IEEE Conference on Decision and Control, Maui, Hawaii, USA*, 2003.

[2] G. Ferrari-Trecate, M. Muselli, D. Liberati, and M. Morari. A clustering technique for the identification of piecewise affine systems. *Automatica*, 39(2):250–217, 2003.

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

[4] A. 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.

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

[6] 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.

[7] A. Goudjil, M. Pouliquen, E. Pigeon, and O. Gehan. Identification algorithm for piecewise affine systems with bounded disturbances. In *2016 24th Mediterranean Conference on Control and Automation (MED)*, pages 628–633, 2016.

[8] L. Bako, K. Boukharouba, E. Duviella, and S. Lecoeuche. A recursive identification algorithm for switched linear/affine models. *Nonlinear Analysis: Hybrid Systems*, 5(2):242–253, 2011.

[9] A. Goudjil, M. Pouliquen, E. Pigeon, and O. Gehan. Convergence analysis of a real-time identification algorithm for switched linear systems with bounded noise. *Conference on Decision and Control, Las Vegas, USA*, 2016.

[10] A. Moradvandi, R. Lindeboom, E. Abraham, and B. De Schutter. Models and methods for hybrid system identification: a systematic survey. *IFAC-PapersOnLine*, 56(2):95–107, 2023.

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

[12] N. Canty, T. O'Mahony, and M.T. Cychowski. An output error algorithm for piecewise affine system identification. *Control Engineering Practice*, 20(4):444–452, 2012.

[13] M. Mejari, V. Breschi, V. Naik, and D. Piga. A bias-correction approach for the identification of piecewise affine output-error models. *IFAC-PapersOnLine*, 53(2):1096–1101, 2020.

[14] N. Canty, T. O'Mahony, and M. Cychowski. An output error algorithm for piecewise affine system identification. *Control Engineering Practice*, 20(4):444–452, 2012.

[15] M. Mejari, V. Breschi, and D. Piga. Recursive bias-correction method for identification of piecewise affine output-error models. *IEEE Control Systems Letters*, 4(4):970–975, 2020.

[16] A. Goudjil, M. Pouliquen, E. Pigeon, O. Gehan, and B. Targui. Recursive output error identification algorithm for switched linear systems with bounded noise. *IFAC-PapersOnLine*, 50(1):14112–14117, 2017.

[17] E. Fogel and Y.F. Huang. On the value of informaton in system identification - bounded noise case. *Automatica*, 18(2):229–238, 1982.

[18] G. Tan, C. Wen, and Y.C. Soh. Identification for systems with bounded noise. *IEEE Transactions on Automatic Control*, 42(7):998–1001, 1997.

[19] M. Pouliquen, E. Pigeon, and O. Gehan. Output Error identification for multi-input multi-output systems with bounded disturbances. *IEEE Conference on Decision and Control - European Control Conference, Orlando, USA, 2011*.

[20] A. Goudjil, M. Pouliquen, E. Pigeon, and O. Gehan. On-line outer bounding ellipsoid algorithm for clustering of hyperplanes in the presence of bounded noise. *Cluster Computing*, pages 1–13, 2023.

[21] L. Ljung. *System Identification - Theory for the User*. Prentice-Hall, 1999.

[22] H. Ohlsson and L. Ljung. Identification of piecewise affine systems using sum-of-norms regularization. *IFAC World Congress, Milan, Italy*, 2011.

[23] V.N. Vapnik and C. Cortes. Support-vector networks. *Machine Learning*, 20(3):273–297, 1995.

[24] V.N. Vapnik. *The Nature of Statistical Learning Theory*. Springer Series in Statistics, 1995.

[25] S. Paoletti, A. L. Juloski, G. Ferrari-Trecate, and R. Vidal. Identification of hybrid systems: a tutorial. *European Journal of Control*, 13(1):242–260, 2007.