

# Zonotope parameter identification for piecewise affine system

WANG JIAN-HONG

School of Electronic Engineering and Automation, Jiangxi University of Science and Technology,  
Ganzhou, 343100, CHINA

*Abstract:* This paper studies the identification problem for piecewise affine system, which is a special nonlinear system. As the difficulty in identifying piecewise affine system is to determine each separated region and each unknown parameter vector simultaneously, we propose a multi class classification process to determine each separated region. This multi class classification process is similar to the classical data clustering process, and the merit of our strategy is that the first order algorithm of convex optimization can be applied to achieve this classification process. Furthermore to relax the strict probabilistic description on external noise in identifying each unknown parameter vector, zonotope parameter identification algorithm is proposed to compute a set that contains the parameter vector, consistent with the measured output and the given bound of the noise. To guarantee our derived zonotope not growing unbounded with iterations, a sufficient condition for this requirement to hold may be formulated as one linear matrix inequality. Finally a numerical example confirms our theoretical results

*Keywords:* Piecewise affine system, Multi class classification, Zonotope, Parameter identification

Received: September 23, 2019. Revised: December 29, 2019. Accepted: January 23, 2020. Published: March 16, 2020.

## 1. Introduction

Piecewise affine system considered in this paper is one of hybrid dynamical systems, as piecewise affine system represents switching dynamics among a collection of linear differential or difference equations with state space being partitioned by a finite number of linear hyperplanes. Hybrid dynamical systems are a class of complex systems that involve interacting discrete event and continuous variable dynamics. They are important in applications in embedded systems, cyber physical systems, robotics, manufacturing systems, traffic management, biomolecular networks, and have recently been at the center of intense research activity in the control theory, computer aided verification, and artificial intelligence communities. But in the control theory when to control a system, one needs to know at least something about how it behaves and reacts to different actions taken on it. Hence we need a model of the system. A system can informally be defined as an entity which interacts with the rest of the world through more or less well defined input and output data. A model is then an approximate description of the system, and an ideal model may be simple, accurate and general. This approximate

description of the system can be constructed by system identification strategy, as the goal of system identification is to build a mathematical model of a dynamic system based on some initial information about the system and the measurement data collected from the system. According to (Ljung Lennart, 1999), the process of system identification consists of designing and conducting the identification experiment in order to collect the measurement data, selecting the structure of the model and specifying the parameters to be identified and eventually fitting the model parameters to the obtained data. Finally the quality of the obtained model is evaluated through model validation process. Generally system identification is an iterative process and if the quality of the obtained model is not satisfactory, some or all of the listed phases can be repeated in order to obtain one satisfied model for that considered system. Because of our complex world, all phenomenon are described as nonlinear systems. But nonlinear systems can not be convenient for other applications such as controller design, filter and prediction etc. So during system identification procedure, most common models are linear difference equation descriptions, such as ARX and ARMAX models, as well as linear state

space models. When linear models are not sufficient for describing accurately the dynamics of a system, nonlinear identification can be employed. A large number of nonlinear model structures have been constructed to investigate their properties, see (S Boyd, 2004), where some real time fast convex algorithms are proposed to identify model parameters. Many tools for identification, as well as for control, stability analysis, have emerged in recent years. To be able to use these proposed tools, a mathematical model of the system is needed.

Identification of hybrid systems (for example, piecewise affine systems) is an area that is related to many other research fields within nonlinear system identification, as such hybrid systems are sufficiently expressive to model a large number of physical process, and can approximate nonlinear dynamics with arbitrary accuracy. In addition, given the equivalence between piecewise affine systems and several classes of hybrid systems (R Pintelon, 2001), piecewise affine system identification techniques can be used to obtain hybrid models. The identification of piecewise affine system is a challenging problem, as it involves the estimation of both the parameters of the affine sub-models, and the coefficients of the hyper-planes defining the partition of the state-input set. This issue clearly underlies a classification problem such that each data point corresponds to one sub-model. In particular, one can find several different approaches which are applicable, or related to piecewise affine system identification problem. Some examples of approaches for piecewise affine systems are neural networks with piecewise affine perceptions. In (J Roll, 2004), mixed integer programming is used to solve piecewise affine system identification problem. As the number of integer variables increases with the number of training samples, mixed integer programming is limited to problems with a small number of observations. To be able to reconstruct a possible discontinuous piecewise affine map with a multi-dimensional domain, Giancarlo F Trecate (2013) proposes to exploit the combined use of clustering, linear identification, and pattern recognition techniques, and allows to identify both the affine

sub-models and the polyhedral partition of the domain. In (A Lj Jouloski, 2013) the sub-model parameters are described through probability density functions, which are iteratively updated through particle filtering algorithms. The sum of norms regularization strategy in (Henrik Ohlsson, 2013) can be computationally heavy in case of appropriate step size. Piecewise affine system identification problem amounts to learning from a set of training data and the parameters defining each affine sub-model (Fabien Lauer, 2011). This piecewise affine system identification problem is an NP hard problem in general, see (Fabien Lauer, 2015, 2016), for a detailed explanation on the complexity of piecewise affine system identification. For the sake of simplicity, sparse property is imposed in piecewise affine systems (Laurent Bako, 2011), then the sparse optimization can greatly improve computational efficiency. The strengths of piecewise affine system identification problem of (Valentina Breschi, 2016) are the computational efficiency and the ability to be run both in a batch and in a recursive way, where the combined use of recursive multiple least squares and linear multi-category discrimination is chosen for computing a solution of unconstrained optimization problems. When the identification error in piecewise affine system is bounded by a quantity, a three stage procedure of a bounded error approach for parametric identification of piecewise affine autoregressive exogenous models is proposed in (Alberto Bemporad, 2005). But the performance of that bounded error approach is dependent of noise, overestimated model orders and classification accuracy greatly. The conversion of piecewise affine models from state space input-output form was addressed by deriving necessary and sufficient conditions for a given piecewise affine state space model to admit equivalent representation (Simone Paoletti, 2010). From modern control theory, this state space input-output form can not guaranteed to be a minimal realization, so piecewise affine models are widely used in nonlinear system identification or control. A convex relaxation, based on L1 regulation is proposed in (Laurent Bako, 2011) to approximate the underlying combinatorial problem appearing from piecewise affine regression.

The statistical clustering technique in (Nakada H, 2005) first computes the parameters of the affine local models, then partition of the regressor space. The greedy algorithm of (Simone Paoletti, 2008) to partition in feasible sets of linear inequalities can be computationally heavy in case of large training sets. A main limitation of a above approach is that the polyhedral partition of the regressor space is given by Voronoi diagram, which will limit flexible capability (Alberto Bemporad, 2000). The problem of finding a lower complexity estimation of piecewise affine models from noise corrupted input-output data is dealt with in (Giuseppe C Calafiore, 2017), where an identification criterion formed by the average of a standard prediction error cost is combined with an L1 regularization term to promote sparse solution (Xin Xu, 2017). The piecewise linear Hammerstein model can be identified in the presence of a special excitation signal in (Gregor Dolanc, 2005), where it is convenient for the description of the processes with highly nonlinear or discontinuous memoryless static functions. Naoko Miyashita (2007) employs identification method using pseudo random binary sequences input for decoupling the identification of nonlinear static block with that of piecewise affine dynamic block. New iterative algorithms to identify Hammerstein system with piecewise linear nonlinearities are proposed in (Yun Lin, 2006), further that normalized iterative method produces convergent result with smooth nonlinear part and finite impulse linear part. A new form of the Hammerstein model with modified parameterization, which will eliminate the main practical limitations, is introduced in (Gregor Dolanc, 2013), a parameter estimation algorithm and a novel pole placement controller are presented to tune the identified model. By means of a sparse overparameterization, the identification of nonlinear systems using piecewise linear models is turned into a convex optimization problem in [Per Mattsson, 2016], where a recursive likelihood based methodology is proposed to penalize model complexity.

As our considered piecewise affine system is one popular modeling framework for hybrid systems proposed in control theory, furthermore piecewise

affine system can be seen as a special case of switched systems with linear dynamics in each mode and mutual exclusive partitions of the state space. Based on above descriptions on piecewise affine system, in this paper we continue to study the detailed identification strategy for piecewise affine system. It is well known that in such piecewise affine system, the space is partitioned into many separate regions and a local linear form is used for each separate region. So the first step in identifying piecewise affine system is to determine these separate regions. After the separate regions are given, the second identification problem is reduced to identify the linear submodels for each region. To deal with above mentioned steps, we reformulate the problem of determining the separate regions as a multi class classification problem, which can be solved by classical first order algorithm of convex optimization theory, such as mirror descent algorithm or Nesterov's optimal algorithm. As a multi class classification problem coincides with a data clustering process into the separate regions. When to identify the unknown parameter in each separate region, many classical identification algorithms can be used directly here, for example, least squares algorithm, maximal likelihood algorithm and Bayesian algorithm, etc. But all the classical identification algorithms hold in case that the considered noise may be a zero mean random signal. This condition corresponds to the classical probabilistic description on noise. To relax this strict probabilistic description on noise, we investigate the zonotope parameter identification algorithm in the presence of bounded noise. This bounded noise is considered in set membership identification field widely, and it is a new deterministic identification algorithm. The zonotope parameter identification algorithm computes a set that contains the parameters consistent with the measured output and the given bound of the disturbance. To guarantee our derived zonotope not growing unbounded with iterations, some contracting properties can be imposed. In this work, a sufficient condition for these contracting properties to hold may be formulated as one linear matrix inequality. By solving one optimization problem with linear matrix

inequality constraint, one approximate feasible solution set is obtained to contain the parameters.

The rest of the paper is organized as follows. In section 2, the problem setting and the piecewise affine system are presented. In section 3, a multi class classification problem based on the first order algorithm of convex optimization theory, which can be used to determine the separate regions, is introduced. In section 4, The zonotope parameter identification algorithm is proposed to identify the unknown parameters in the presence of bounded noise for each separate region. In section 5, a very simple numerical example is used to illustrate the proposed algorithm. Finally, conclusions and comments about future research are presented in section 6.

## 2. Piecewise affine system

Consider an affine model as follows.

$$y(t) = -\sum_{i=1}^{n_a} a_i y(t-i) + \sum_{j=1}^{n_b} b_j u(t-j) + e(t) \quad (1)$$

where  $u(t)$  and  $y(t)$  are input and output respectively,  $\{a_i\}$  and  $\{b_j\}$  are the unknown model parameters,  $e(t)$  is an external noise. Two orders  $n_a$  and  $n_b$  are priori known. This affine model can be rewritten as a linear regression form, after a regression vector  $\phi(t)$  is introduced.

$$\begin{cases} \phi(t) = [-y(t-1) \ \cdots \ -y(t-n_a) \ u(t-1) \ \cdots \ u(t-n_b)]^T \\ y(t) = \phi^T(t)\theta + e(t) \end{cases} \quad (2)$$

where the unknown parameter vector is stacked as.

$$\theta = [a_1 \ \cdots \ a_{n_a} \ b_1 \ \cdots \ b_{n_b}]^T \quad (3)$$

For large enough orders  $n_a$  and  $n_b$ , that affine model can be approximated any linear system. Although affine model (1) is beneficial for local approximation of some nonlinear systems, but it can not capture any nonlinear properties, so in order to introduce affine model into nonlinear system, piecewise affine system is obtained. It means that the parameter vector  $\theta$  is dependent of the region in the regression space, where regression vector  $\phi(t)$  lies. Then the regression space is divided into  $n$  separate regions  $R_1 \cdots R_n$ , our considered piecewise affine system can be defined as.

$$y(t) = \phi^T(t)\theta_i + e(t) \quad \text{if } \phi(t) \in R_i \quad (4)$$

where the parameter vector  $\theta_i$  depends on its separate region  $R_i$ . The problem of identifying piecewise affine system is reformulated as that, after output and input  $\{u(t), y(t)\}$  are collected, how to identify those unknown parameter vectors  $\{\theta_i\}_{i=1}^n$ ? Due to the fact that regression vector  $\phi(t)$  is constituted by output and input  $\{u(t), y(t)\}$ , so the first step is to judge which region the regression vector belongs to.

## 3. Multi-class classification process

As there exist  $n$  separate regions  $R_1 \cdots R_n$ , the problem of determining which region the regression vector lies is in conjunction with a multi class classification process. We observe  $N$  data points  $\{z(t)\}_{t=1}^N$ .

$$\begin{aligned} z(t) &= [y(t) \ \phi(t)]^T \\ &= [y(t) \ -y(t-1) \ \cdots \ -y(t-n_a) \ u(t-1) \ \cdots \ u(t-n_b)]^T \end{aligned} \quad (5)$$

where  $N$  denotes the number of observed data points, and each data point  $z(t)$  belongs to one of  $n$  non-overlapping classes, along with labels  $\lambda_t \in R^n$  which are basic orthonormal vectors in  $R^n$ ; the index of the only nonzero entry in  $\lambda_t$  is the number of class  $t$  to which  $z(t)$  belongs. We want to build a multi class analogy of the standard linear classifier as follows: a multi class classifier is specified by a matrix  $A$  and a vector  $a \in R^n$ . Given a data point  $z(t)$ , we compute the  $n$  dimensional vector  $Az(t) + a$ , identify its maximal component, and treat the index of this component as our guess for the serial number of the class to which  $z(t)$  belongs.

Let  $\bar{\lambda}_t = 1 - \lambda_t$  be the component of  $\lambda_t$ . Given a data point  $z$  and the corresponding label  $\lambda$ , let us set.

$$h = h(A, a, z, \lambda) = [Az + a] - [\lambda(Az + a)] + \bar{\lambda} \quad (6)$$

If  $i_*$  is the index of the only nonzero entry in  $\lambda$ , then the  $i_*$ th entry in  $h$  is zero. And  $h$  is nonpositive if and only if the classifier, given by  $A, a$  and evaluated at  $z$ , recovers the class  $i_*$  of  $z$  with margin 1, i.e. we have.

$$[Az + a]_j \leq [Az + a]_{i_*} - 1 \quad \text{for } j \neq i_* \quad (7)$$

On the other hand if the classifier fails to classify  $z$

correctly, that is.

$$[Az + a]_j \geq [Az + a]_{i_*} \quad \text{for some } j \neq i_* \quad (8)$$

Then the maximal entry in  $h$  is  $\geq 1$ . So we set.

$$\eta(A, a, z, \lambda) = \max_{1 \leq j \leq n} [h(A, a, z, \lambda)]_j \quad (9)$$

We get a nonnegative function which vanishes for the pairs  $(z, \lambda)$ , which are quite reliably-with margin  $\geq 1$ -classified by  $(A, a)$ , and is  $\geq 1$  for the pairs  $(z, \lambda)$  with  $z$  not classified correctly. Thus the function is that.

$$H(A, a) = E\{\eta(A, a, z, \lambda)\} \quad (10)$$

The expectation being taken over the distribution of the pairs  $(z, \lambda)$ , is an upper bound on the probability for classifier  $(A, a)$  to misclassify a data point. What we would do is to minimize  $H(A, a)$  over  $A$  and  $a$ . To do this, since  $H(A, a)$  is not observable, we replace the expectation by its empirical counterpart.

$$H_N(A, a) = \frac{1}{N} \sum_{t=1}^N \eta(A, a, z(t), \lambda_t) \quad (11)$$

For the sake of simplicity, imposing an upper bound on some norm  $\|A\|$  of  $A$ , one optimization problem is obtained.

$$\min_{A, a} \frac{1}{N} \sum_{t=1}^N \max_{1 \leq n} [Az(t) + a - \lambda_t (Az(t) + a) + \bar{\lambda}_t]_i \quad (12)$$

subject to  $\|A\| \leq 1$

A natural choice of the norm  $\|A\|$  is the maximum of the  $\|A\|_2$  norm. The classical first order algorithm of convex optimization theory can be used to solve that optimization problem (12), such as mirror descent algorithm or Nesterov's optimal algorithm. Once optimization variables  $A$  and  $a$  are obtained, then the linear classifier  $Az(t) + a$  is get. From above multi class classification process, we see that once one data point is collected, we can cluster it with a linear classifier. So based on above linear classifier, all data points can be clustered together, then those data points clustering together as one class can be used in the second identification problem for unknown parameter.

#### 4. Zonotope parameter identification

### algorithm

After all collected data points are clustered as  $n$  classes, then those data points belonging to one same class can be used to identify one unknown parameter. Here we only rewrite the following piecewise affine system in the  $i$ th separate region.

$$y(t) = \phi^T(t)\theta_i + e(t) \quad \phi(t) \in R_i \quad (13)$$

The mission of this section is to identify the unknown parameter vector  $\theta_i$  in case of unknown but bounded noise. It is well known that in equation (13), there are two ways to represent uncertainties: the statistical approach and the deterministic approach. In the statistical approach, the uncertainty or disturbance is modeled by a random process with a known statistical property, when estimates of the probability distributing of the uncertainty or disturbance are available. But in many applications, there are situations when then probability distributing of the uncertainty or disturbance is not known and only bound of the uncertain domain can be fixed, then the probabilistic assumptions on the uncertainty are no longer valid. In the deterministic approach, disturbance is assumed to belong to a set. Different families of classical sets are used depending on their properties. The main advantage of the deterministic approach is that disturbance is assumed to be unknown but bounded and this is often simpler to verify than the criterion on the probability disturbance. This is the main reason why we choose the deterministic approach to model the disturbance affecting the system behavior. Based on this remark, one of the deterministic approach-zonotope parameter identification algorithms has been chosen in this section to identify two unknown parameter vector  $\theta_i$ . This identification algorithm computes a set that contains the parameters consistent with the measured output and the given bound of the disturbance. This set is represented by a zonotope, that is an affine map of an unitary hypercube.

Observing equation (13) again, as  $e(t)$  represents the considered noise, this noise belongs to a bounded set, i.e.

$$e(t) \in \{e \in R : |e| \leq \sigma\} \quad (14)$$

where  $\sigma \in R$  is an upper bound and  $e(t)$  is unknown, but has known bound.

From set membership identification theory (J M Bravo & T Alamo & E F Ca macho,2006), given a set of measured outputs, the feasible solution set is defined as the set of parameters that are consistent with measured outputs and the given bounds. More precisely, the following definitions are given through this section.

**Definition 1 (Feasible Solution Set):**

Suppose that the pairs  $\{y(t), \phi(t)\}, t=1, 2 \dots N$  are given. The vector  $\theta_i$  is said to belong to the feasible solution set if there exists  $\theta_i$  such that..

$$|y(t) - \phi^T(t)\theta_i| \leq \sigma, t=1, 2 \dots N \quad (15)$$

**Definition 2 (Information set):**

Given the pairs  $\{y(t), \phi(t)\}, t=1, 2 \dots N$  at time instant  $t$ , the information set  $I_t$  is a set of all feasible parameters, that are consistent with the model (13), the measured output  $y(t)$  and the known bound at time instant  $t$ , namely:

$$I_t = \{\theta_i \in R^{n_a+n_b} : -\sigma \leq y(t) - \phi^T(t)\theta_i \leq \sigma\} \quad (16)$$

Geometrically  $I_t$  represents a strip, that is consisted with  $\{y(t), \phi(t)\}, t=1, 2 \dots N$ . Feasible solution set at time instant  $t+1$ , denoted as  $FSS_{t+1}$ , can be computed exactly from the one corresponding to time instant  $t$  by the following recursion.

$$FSS_{t+1} = FSS_t \cap I_t \quad (17)$$

It is difficult to compute the feasible solution set (J M Bravo & A Suarez & M V asallo,2016), so an outer bound of the feasible solution set can be used.

**Definition 3 (Approximated Feasible Solution Set):**

An approximated feasible solution set denoted AFSS, is a set that satisfies that FSS. The intersection  $FSS_t \cap I_t$  is approximated by means of the intersection between a zonotope and a strip at time instant  $t$ .

**Definition 4 (Zonotope of order  $m$ ):**

Given a vector  $p \in R^{n_a+n_b}$  and a matrix  $H \in R^{(n_a+n_b) \times m}$ , a zonotope of order  $m$  is a set of  $n_1 = n_a + n_b$  dimensional vectors defined by (J M Bravo T Alamo M Vasallo,2017).

$$Z = \{\theta_i \in R^{n_1} : \theta_i \in p \oplus HB^m\} \quad (18)$$

where  $HB^m$  is a linear projection of  $B^m$  into  $n_1 = n_a + n_b$  dimensional parameter space,  $B^m$  is a unit hypercube of order  $m$ , and  $\oplus$  denotes Miniowski sum.

Using the approximated feasible solution set on the intersection (17), then

$$FSS_{t+1} = FSS_t \cap I_t \subseteq AFSS_{t+1} \quad (19)$$

If in equation (41) feasible solution set  $FSS_t$  is denoted by a zonotope and information set  $I_t$  is a strip, then a family of zonotopes which over bound the intersection between a zonotope and a strip, are derived as the following Theorem 1.

**Theorem 1:**

Suppose FSS at time instant  $t$ , denoted as a zonotope

$$FSS_t = \hat{p}_t \oplus \hat{H}_t B^r \subset R^{n_1} \quad (20)$$

The information set or a strip

$$I_t = \{\theta_i \in R^{n_a+n_b} : -\sigma \leq y(t) - \phi^T(t)\theta_i \leq \sigma\} \quad (21)$$

and a scalar  $\gamma$ , define

$$\begin{cases} \hat{p}_t(\gamma) = \hat{p}_t + \gamma(y(t) - \phi^T(t)\hat{p}_t) \\ \hat{H}_t(\gamma) = [(I - \gamma\phi(t))\hat{H}_t \quad \sigma\gamma] \end{cases} \quad (22)$$

Thus we have

$$\begin{aligned} FSS_{t+1} &= FSS_t \cap I_t \subseteq AFSS_{t+1} \\ &= \hat{p}_t(\gamma) \oplus \hat{H}_t(\gamma) B^{r+1} \end{aligned} \quad (23)$$

where  $I$  is an identity matrix.

In (Marko Tanaskovic & Lorenzo Fagiano & Carlo Novara,2017), the optimization-based method is used to choose a scalar  $\gamma \in R$ , through minimizing the volume of a zonotope. Now the minimization of the P-radius of a zonotope is applied, as the P-radius criterion allows to guarantee the non-increasing property of the guaranteed zonotope at each time instant. It tells us that to guarantee the approximated feasible solution set not growing unbounded with iteration steps, the following inequality relation between two neighboring zonotope is imposed.

$$l_t \leq \beta l_{t-1} + \sigma^2 \quad (24)$$

where  $\beta \in (0,1]$  is a contraction rate, and  $l_t$  is the P-radius of zonotope parameter estimation set at time

instant  $t$ , which is defined by.

$$l_i = \max_{\theta_i \in FSS_i} \left( \|\theta_i - \hat{p}_i\|_P^2 \right) \quad (25)$$

where  $P$  is an  $n_1$ -dimensional positive definite matrix.

Substituting the definition (25) into the inequality relation (24), we have

$$\max_{\hat{z} \in B^{r+1}} \left( \|\hat{H}_i(\gamma) \hat{z}\|_P^2 \right) \leq \max_{z \in B^r} \beta \left( \|\hat{H}_{i-1} z\|_P^2 \right) + \max_{\eta \in B^1} \|\sigma \eta\|_2^2 \quad (26)$$

Expanding equation (26) to obtain

$$\hat{z}^T \hat{H}_i(\gamma) P \hat{H}_i(\gamma) \hat{z} - \beta z^T \hat{H}_{i-1} P \hat{H}_{i-1} z - \eta^2 \sigma^2 \leq 0 \quad (27)$$

Due to the recursion of  $\hat{H}_i(\gamma)$  in equation (22), we compute.

$$\begin{aligned} \hat{H}_i(\gamma) \hat{z} &= (I - \gamma \phi(t)) \hat{H}_{i-1} z + \sigma \gamma \eta \\ &= (I - \gamma \phi(t)) \bar{z} + \sigma \gamma \eta \end{aligned} \quad (28)$$

where we set

$$\begin{cases} \bar{z} = \hat{H}_{i-1} z \\ \hat{z} = [z \quad \eta]^T \end{cases} \quad (29)$$

Applying equation (28) in (27), we get.

$$\begin{aligned} &\left[ \bar{z}^T (I - \gamma \phi(t))^T + \sigma \gamma \eta \right] P \left[ (I - \gamma \phi(t)) \bar{z} + \sigma \gamma \eta \right] \\ &- \beta \bar{z}^T P \bar{z} - \eta^2 \sigma^2 \leq 0 \end{aligned} \quad (30)$$

Formulation above inequality as that.

$$\begin{aligned} &\bar{z}^T (I - \gamma \phi(t))^T P (I - \gamma \phi(t)) \bar{z} + \bar{z}^T (I - \gamma \phi(t))^T P \sigma \gamma \eta \\ &+ \sigma \gamma \eta P (I - \gamma \phi(t)) \bar{z} + \sigma^2 \gamma^2 \eta^2 - \beta \bar{z}^T P \bar{z} - \eta^2 \sigma^2 \leq 0 \end{aligned} \quad (31)$$

A sufficient condition for equation (31) to hold can be rewritten as one linear matrix inequality.

$$\begin{aligned} &\begin{bmatrix} \bar{z} \\ \eta \end{bmatrix}^T \\ &\times \begin{bmatrix} (I - \gamma \phi(t))^T P (I - \gamma \phi(t)) - \beta P & (I - \gamma \phi(t))^T P \sigma \gamma \\ P \sigma \gamma (I - \gamma \phi(t)) & \sigma^2 (\gamma^2 - 1) \end{bmatrix} \begin{bmatrix} \bar{z} \\ \eta \end{bmatrix} \leq 0 \\ &\forall \bar{z}, \eta \end{aligned} \quad (32)$$

Using the definition of positive definite matrix allows us to rewrite as that.

$$\begin{aligned} &\begin{bmatrix} (I - \gamma \phi(t))^T P (I - \gamma \phi(t)) - \beta P & (I - \gamma \phi(t))^T P \sigma \gamma \\ P \sigma \gamma (I - \gamma \phi(t)) & \sigma^2 (\gamma^2 - 1) \end{bmatrix} \leq 0 \\ &\forall \begin{bmatrix} \bar{z} \\ \eta \end{bmatrix} \neq 0 \end{aligned} \quad (33)$$

The linear matrix inequality in (33) defines the feasible solution for scalar  $\gamma$ , i.e.  $\gamma$  can be computed by solving the following Eigenvalue problem.

$$\begin{aligned} &\max_{\tau, \gamma} \tau \\ &\text{subject to } \frac{(1 - \beta) P}{\max_{\eta \in B} \|\gamma \eta\|_2^2} \geq \tau I, \quad \tau > 0 \\ &\begin{bmatrix} (I - \gamma \phi(t))^T P (I - \gamma \phi(t)) - \beta P & (I - \gamma \phi(t))^T P \sigma \gamma \\ P \sigma \gamma (I - \gamma \phi(t)) & \sigma^2 (\gamma^2 - 1) \end{bmatrix} \leq 0 \end{aligned} \quad (34)$$

The above Eigenvalue problem can be solved by using convex optimization algorithm, then based on this optimal scalar  $\gamma \in R$ , a zonotopic outer approximation of the intersection between a zonotope and a strip is obtained by using matrix inequality optimization strategy. Finally our zonotope parameter identification algorithm is formulated as follows.

#### Algorithm 1: Zonotope Parameter Identification Algorithm

- (1) Obtain measured input-output data and construct regressor vector  $\phi(t)$ ;
- (2) Build a strip that bounds the consistent parameters, i.e. information set;

$$I_t = \{ \theta_i \in R^{n_a + n_b} : -\sigma \leq y(t) - \phi^T(t) \theta_i \leq \sigma \}$$

- (3) Construct a zonotope  $FSS_t = \hat{p}_t \oplus \hat{H}_t B^r \subset R^{n_1}$  to denote the feasible solution set at time instant  $t$ ;
- (4) Compute the intersection between a zonotope and a strip at time instant  $t$  and obtain a new zonotope

$$FSS_{t+1} = FSS_t \cap I_t \subseteq AFSS_{t+1} = \hat{p}_t(\gamma) \oplus \hat{H}_t(\gamma) B^{r+1}$$

to denote the approximated feasible solution set at time instant  $t+1$ ;

- (5) Choose one optimal scalar  $\gamma$  through solving one matrix inequality optimization strategy;
- (6) Repeat the above steps and terminate the recursive algorithm when the P-radius  $l_i$  is zero, then denote  $\hat{p}^*$  as the vector in the last zonotope, so the unknown parameter vector  $\theta_i$  is given by

$$\hat{\theta}_i = \hat{p}^* \quad (35)$$

It is similar to applying above six steps to identify another unknown parameter vector.

### 5. Numerical example

In this section, one simple piecewise affine system is used to prove our strategies, such as two class classification process and zonotope parameter identification algorithm. This simple piecewise affine system is given as follows.

$$y(t) = \begin{cases} \phi^T(t)\theta_1 + e(t) & \text{if } \phi(t) > 0 \\ \phi^T(t)\theta_2 + e(t) & \text{if } \phi(t) \leq 0 \end{cases} \quad (36)$$

where regression vector  $\phi(t)$  and two unknown parameter vectors  $(\theta_1, \theta_2)$  are described as that.

$$\begin{cases} \phi(t) = [-y(t) \quad -u(t)]^T \\ \theta_1 = [7 \quad 2]^T \\ \theta_2 = [2 \quad 0.5]^T \end{cases} \quad (37)$$

we exert the input signal  $u(t)$  in piecewise affine system. The actual input signal is given in Figure.1 (a), but this actual input signal is not suited for simulation. So we use its approximated input signal to replace the actual input signal in our simulation, where the approximated input signal is similar to sinusoidal signal-Figure.1 (b). Then we measure the output signal  $y(t)$  by using some measuring devices, the observed output signal is plotted in Figure.2.

Firstly our mentioned multi class classification process is reduced to two class classification problem in this simulation. Given one data point  $(y(t), \phi(t))$ , we need to determine which region this data point belongs to. Here the number of given data points is  $N = 500$ , i.e. these 500 data points belong to one of two classes. The clustering process can be seen Figure.3, where data points are clustering around two ellipsoids. As three points deviate away these two ellipsoids, then they are regarded as outliers and we delete them. From Figure.3, we see that all data points are classified correctly, except three data points.

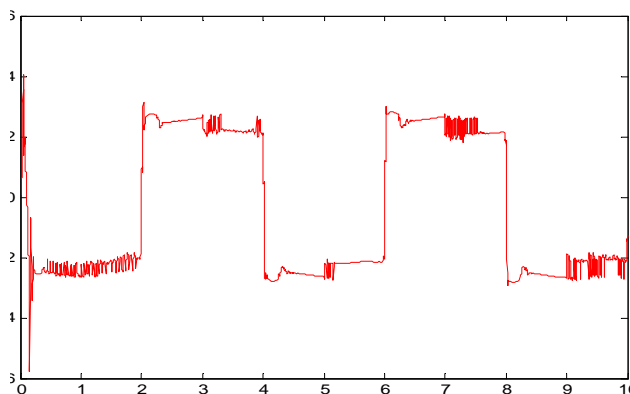
Secondly in the presence of bounded noise, choose upper bound  $|e(t)| \leq \sigma = 0.5$ , and all initial parameter

values  $\hat{\theta}_0 = \frac{1}{p_0} I$ . Zonotope parameter identification

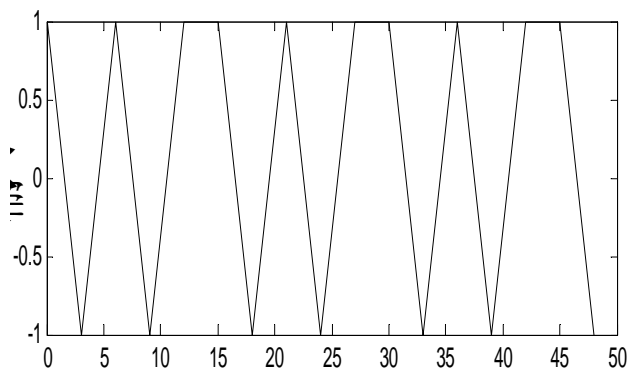
algorithm is applied to identify those two unknown parameter vectors. Applying above six steps to

construct a sequence of candidate zonotopes, and after 20 iterations, these candidate zonotopes are given in Figure.4 and Figure.5.

In Figure.4, the black star denotes the optimal parameter vector as  $\theta_1 = [7 \quad 2]^T$ , and a sequence of candidate zonotopes generated by zonotope parameter identification algorithm include  $\theta_1 = [7 \quad 2]^T$  as their interior point. as these candidate zonotopes have decreasing volumes with it erations, i.e. certain contracting properties hold. Generally the other unknown parameter vector corresponding to  $\theta_1 = [7 \quad 2]^T$  can be chosen as the center of the smallest zonotope. Further the black star is the optimal parameter vector as  $\theta_2 = [2 \quad 0.5]^T$  in Figure.5, and results are similar to them in Figure.4



(a) Actual input signal



(b) Approximated input signal

Figure.1: The applied input signal



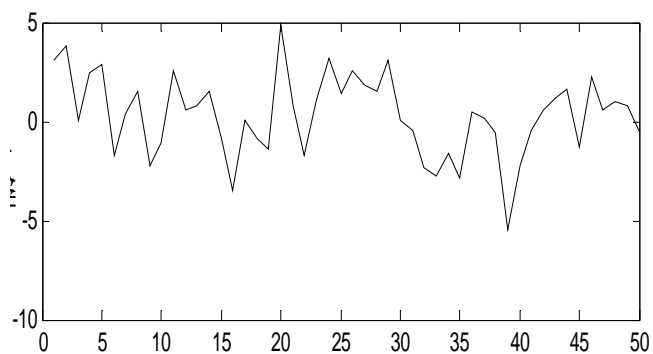


Figure.2: The observed output signal

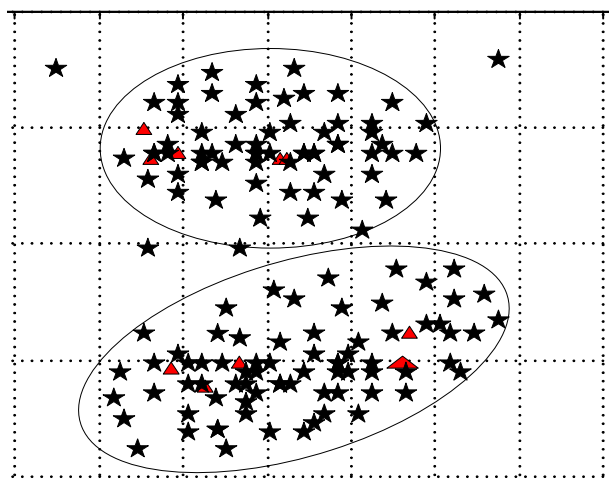


Figure.3: The observed output signal

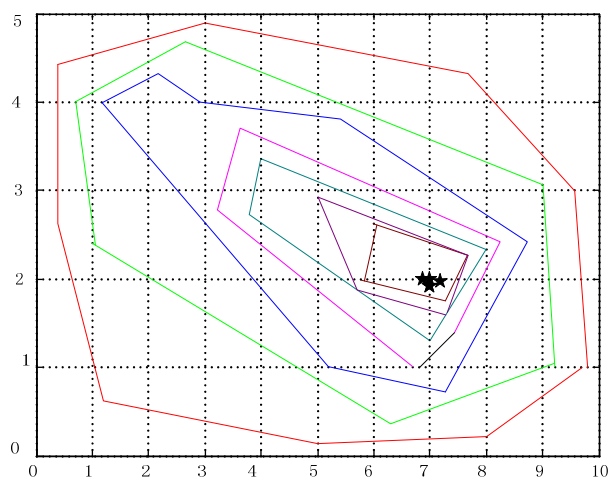


Figure 4: The observed output signal

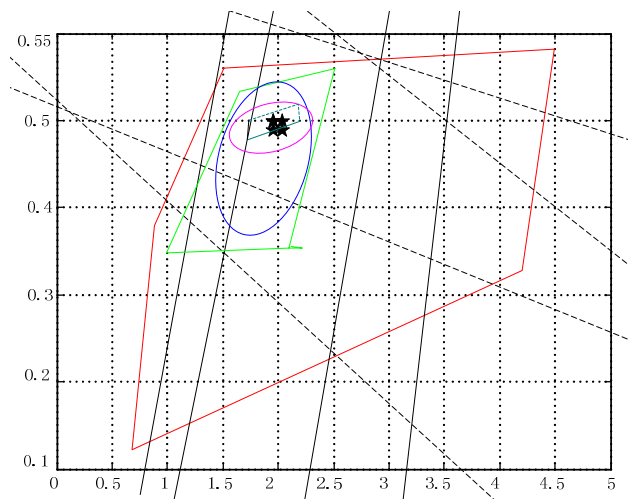


Figure.5: The observed output signal

## 6. Conclusion

In this paper, we study the problem of identifying piecewise affine system, which combines the linear and nonlinear properties. As it is a nonlinear system that is piecewise affine in the regression space, so the parameter vector depends on the region in the regression space. The separate regions are determined as a multi class classification problem, which is solved by classical first order algorithm of convex optimization theory. In the presence of unknown but bounded noise, zonotope parameter identification algorithm is proposed to identify unknown parameter vector in each separated region. Generally finite sample property of zonotope parameter identification algorithm is our ongoing work.

## Acknowledgements

The author is grateful to Professor Eduardo F Camacho for his warm invitation in his control lab at the University of Seville, Seville, Spain. Thanks for his assistance and advice on zonotopes in guaranteed state estimation and model predictive control. This work was supported by the Grants from the National Science Foundation of China (no. 61364014).

## References

- [1] Ljung, Lennart, (1999), "System identification: Theory for the user," Prentice Hall Press, Upper Saddle River, New Jersey, USA

- [2] S Boyd and L Vandenberghe, (2004), "Convex optimization," Cambridge University Press, Cambridge, UK.
- [3] R Pintelon and J Schoukens, (2001), "System identification: a frequency domain approach," IEEE Press, New York, NJ, USA.
- [4] J Roll, A Bemporad and Lennart Ljung, (2004), "Identification of piecewise affine systems via mixed integer programming," *Automatica*, Vol 40, No. 1, pp. 37-50.
- [5] Giancarlo Ferrari Trecate, Marco Muselli, (2003), "A clustering technique for the identification of piecewise affine systems," *Automatica*, Vol 39, No.2, pp. 205-217.
- [6] A Lj Joulouski, S Weiland, (2013), "A Bayesian approach to identification of hybrid systems," *IEEE Transactions on Automatic control*, Vol 50, No.10, pp. 1520-1532.
- [7] Henrik Ohlsson, Lennart Ljung, (2013), "Identification of switched linear regression models using sum of norms regularization," *Automatica*, Vol 49, No.4, pp. 1045-1050.
- [8] Fabien Lauer, Gerard Bloch, (2011), "A continuous optimization framework for hybrid system identification," *Automatica*, Vol 47, No.3, pp. 608-613.
- [9] Fabien Lauer, (2016), "On the complexity of switching linear regression," *Automatica*, Vol 74, No.12, pp. 80-83.
- [10] Fabien Lauer, (2015), "On the complexity of piecewise affine system identification," *Automatica*, Vol 62, No.12, pp. 148-153.
- [11] Laurent Bako, (2011), "Identification of switched linear systems via sparse optimization," *Automatica*, Vol 47, No.4, pp. 668-677.
- [12] Valentina Breschi, Dario Piga, (2016), "Piecewise affine regression via recursive multiple least squares and multi-category discrimination," *Automatica*, Vol 73, No.11, pp. 155-162.
- [13] Alberto Bemporad, Andrea Garulli, (2005), "A bounded error approach to piecewise affine system identification," *IEEE Transactions on Automatic control*, Vol 50, No.10, pp. 1567-1580.
- [14] Simone Paoletti, Jacob Roll, (2010), "On the input output representation of piecewise affine state space models," *IEEE Transactions on Automatic control*, Vol 55, No.1, pp. 60-72, 2010.
- [15] Laurent Bako, (2011), "A recursive identification algorithm for switched linear/affine models," *Nonlinear Analysis-Hybrid Systems*, Vol 5, No.2, pp. 242-253.
- [16] Nakada H, Takaka K, (2005), "Identification of piecewise affine systems based on statistical clustering technique," *Automatica*, Vol 41, No.5, pp. 905-913.
- [17] Paoletti S, A Lj Joulouski, (2008), "Identification of hybrid systems a tutorial," *European Journal of Control*, Vol 13, No.2, pp. 242-260.
- [18] Alberto Bemporad, Giancarlo Ferrari Trecate, (2000), "Observability and controllability of piecewise affine and hybrid systems," *IEEE Transactions on Automatic control*, Vol 45, No.10, pp. 1864-1876.
- [19] Giuseppe C Calafiore, (2017), "Leading impulse response identification via the elastic net criterion," *Automatica*, Vol 80, No.4, pp. 75-87.
- [20] Xin Xu, (2017), "Distributed emitter parameter refinement based on maximum likelihood," *International Journal of Intelligent Computing and Cybernetics*, Vol 10, No.1, pp. 2-11.
- [21] Gregor Dolanc, Stanko Strmcnik, (2005), "Identification of nonlinear systems using a piecewise-linear Hammerstein model," *Systems & Control Letters*, Vol.54, No.2, pp. 145-158.
- [22] Naoko Miyashita and Masaki Yamakita, (2007), "Identification of Hammerstein Systems with Piecewise-Affine Nonlinearities," *Proceedings of the 2007 American Control Conference*, New York, USA, July 11-13.
- [23] Yun Liu and Er-Wei Bai, (2006), "Iterative Identification of Hammerstein Systems with Piecewise-linear Nonlinearities," *Proceedings of the 2006 American Control Conference*, Minneapolis, Minnesota, USA, June 14-16.
- [24] Gregor Dolanc, Stanko Strmcnik, (2013), "Identification and Control of Nonlinear Systems Using a Piecewise-Linear Hammerstein Model, in Case Studies in Control," Springer-Verlag London.
- [25] Per Mattsson, Dave Zacharian, Peter Stoica, (2016), "Recursive identification method for piecewise ARX models: a sparse estimation approach," *IEEE Transactions on Signal Processing*, Vol 64, No.19, pp. 5082-5093.
- [26] JM Bravo, T Alamo, EFCamacho (2006), "Bounded error identification of systems with time varying parameters," *IEEE Transactions on Automatic control*, Vol 51, No.7, pp. 1144-1150.

- [27] J M Bravo, A Suarez, M Vasallo (2016), "Slide window bounded error time varying system identification," *IEEE Transactions on Automatic control*, Vol 61, No.8, pp. 2282-2287.
- [28] Marko T anaskovic, Lorenzo Fagiano, Carlo Novara, (2017), "Data driven control of nonlinear systems: an on line direct approach," *Automatica*, Vol 75, No.1, pp. 1-10.