Recursive Identification of Piecewise-affine Systems based on Parameter Space Decomposition Theorem

Rui Huang[1], Weiquan Sun[2] and Min Li[3]

[1]System Department, United Technologies Research Center, East Hartford, CT, USA
[2] School of Business Administration, Duquesne University, Pittsburgh, PA, USA
[3] Swanson School of Engineering, University of Pittsburgh, Pittsburgh, PA, USA

li246@iupui.edu

Abstract

The class of Basis Function Piecewise-affine (BPWA) functions is an ideal model structure for nonlinear and hybrid system identification. This paper proposes a parameter space decomposition theorem for BPWA functions. It is proved that any BPWA function has a decomposed parametric representation, in which the algebraic and geometrical parameters can be identified separately. Under the assumption of invariant geometrical structure, a recursive algorithm is proposed to identify the BWPA AutoRegressive Exogenous models from the input-output data. Two benchmark examples are illustrated to show that the proposed algorithm has much higher computational efficiency compared with the competing algorithms with the same identification accuracy.

Keywords: *Piecewise affine techniques, system identification, parameter space decomposition, recursive identification algorithm, hybrid system*

1. Introduction

The class of Piecewise-affine (PWA) functions presents an ideal model structure for nonlinear identification [1, 2]. The PWA functions have a universal approximation capability for nonlinear systems. This provides a systematical way to solve the nonlinear problems using the theorems and techniques developed for linear systems. Using the PWA approximation techniques, the model predictive control (MPC) of nonlinear systems can be formulated as the multi-parametric programs [3]. The control actions can then be obtained as explicit functions of plant measurements. This explicit MPC solution extends the conventional nonlinear MPC into systems with fast dynamics and decreases the hardware implementation cost significantly.

The PWA functions also play an essential role in the modeling and control of hybrid systems. It is shown in [4, 5] that the PWA systems are equivalent (under some mild assumptions) to the mixed logical dynamical (MLD) systems, linear complementarity (LC) systems, extended linear complementarity (ELC) and max-min-plus scaling (MMPS) systems. The control schemes developed for one set of systems can be readily generalized to the other set.

The identification of PWA systems does not have received enough attention until recent years. The known identification approaches can be classified into two categories: the direct method and indirect method.

The direct methods identify the affine functions and domain partitions directly. In these methods, one needs to calculate all the parameter vectors of the local affine functions, and the parameter matrices that define the polyhedral regions. Typical direct methods are the clustering-based method, the bounded-error method, the Bayesian method and the algebraic method [6-9].

The algebraic method formulates the identification of switched AutoRegressive eXogenous (ARX) models as an algebraic geometric problem [6]. This method provides a closed-form solution for deterministic models. For stochastic models, it provides a sub-optimal solution, which can be used as an efficient initial solution to other iterative approaches. The Bayesian method treats the parameters of PWA systems as random variables [7]. The data classification problem is formulated as a problem of Bayes parameter estimation. This problem is combinatorial in natural. An iterative suboptimal algorithm is then derived, whose efficiency can be much improved by utilizing the available prior information on the model parameters. In [8], a bounded-error algorithm is developed, which allows the users to impose a bound on the prediction error. It estimates a PWA ARX (PWARX) model that satisfies the bounded-error condition for all the samples in the estimation data set. In this algorithm, the error bound is used as a tuning parameter to trade off between the quality of fit and model complexity. The clustering-based method exploits the fact that the PWA maps are locally affine. The information about the sub-models can be obtained by clustering the local parameter vectors [9]. This algorithm combines the techniques of clustering, linear identification, and pattern recognition. Measures of confidence on the samples are introduced and exploited, which efficiently improves the performance of both the clustering and the final linear regression procedure. See [10] for a more detailed description of direct approaches. The direct methods can deal with the general PWA systems, including the discontinuous ones. Their performance and computation efficiency have been demonstrated by a wealth of applications. However, a wide application of the direct methods is limited by the complexity of the description and online evaluation of a PWA model. Due to the combinatorial natural, the number of parameters to describe a PWA system can be an exponential function of the number of polyhedral regions [11]. The memory space required is huge and the online evaluation is expensive, when a PWA model consists of many polyhedral regions [12].

The indirect methods are developed based on the concept of PWA basis functions [13]. In an indirect method, the identification parameters are the weights and parameter vectors of the PWA basis functions [14]. Using a suitable model set, a PWA function with many polyhedral regions can be described by a small number of PWA basis functions [15]. This presents a feasible way to deal with the description and online evaluation complexity of continuous PWA systems.

The canonical Piecewise-Affine (CPWA) representation theorem presents a theoretical basis to design a PWA model structure. The class of CPWA functions can approximate a continuous function arbitrarily well with relatively small number of parameters [17-20]. In 1990, Breimann introduces the hinging hyperplane (HH) models. It is proved that the HH models is essentially another form of the CPWA functions [21]. Roll, Bemporad & Ljung proposed the Mixed-Integer Programming (MIP) algorithm for optimal identification of Hinging-Hyperplane ARX (HHARX) models. The identification problem is formulated as a mixed-integer linear/quadratic program, which can be solved for the global optimum. The MIP algorithm is limited to small-medium size problems because of its high computational complexity [1].

In 2005, a PWA basis function (BPWA) representation is proposed to increase the approximation efficiency of the CPWA functions [22]. Based on this representation theorem, Wen, Wang, Jin and Ma developed the PWA Basis Function AutoRegressive eXogenous (BPWARX) models and found successful applications in dynamic system identification and nonlinear function approximation [23]. A modified Gauss-Newton (MGN) algorithm is developed to build the BPWARX models from the input-output data. Although the MGN algorithm is computationally efficient compared with the competing algorithms, its identification is formulated as a non-convex optimization problem. Therefore, the efficiency of this algorithm is limited when dealing with large-scale problems.

This paper proposes a novel parameter space decomposition theorem for the class of BPWA functions. A recursive PWA (RPWA) algorithm is developed using the assumption of invariant geometrical structure. The online identification time of the RPWA algorithm is negligible, and there is no significant loss in precision compared with the MGN algorithm. The performance of the RPWA algorithm is supported by two benchmark problems.

2. BPWA Functions

2.1 BPWA Basis Function

Definition 1 [24] In R^n, a set of points $\{v_i\}_{i=0}^{m}, m \geq n$ are said to be in general position or affinely independent, if no n+1 of them lie in an (n-1)-dimensional hyperplane.

Definition 2 [17] In R^n, a continuous PWA function

$$B(x) = \max\{a_1^T x + b_1, \cdots, a_{n+1}^T x + b_{n+1}\} \tag{1}$$

is defined as a BPWA basis function, if the common intersection of each group of two polyhedral regions forms an (n-1)-dimensional manifold in the domain space, and $a_i \in R^n, b_i \in R$ with $i = 1, \cdots, n+1$.

Fig. 1 shows the domain partition of a BPWA basis function in 2 dimensions. The domain consists of 3 polyhedral regions, and is spanned by a set of 4 points. It is easy to see that the intersection of each couple regions is a 1-dimensional manifold.

In R^n, a $B(x)$ is uniquely specified by n+2 points $\{v_i\}_{i=0}^{n+1}$ in general position. Let $\{x_i, y_i\}$ be the coordinates of $v_i, i = 0, \cdots, n+1$. Assume that the point v_0 be the common intersection of n+1 affine functions, i.e.

$$a_1^T x_0 + b_1 = a_2^T x_0 + b_2 = \cdots = a_{n+1}^T x_0 + b_{n+1} = y_0 \tag{2}$$

Assume further that the i-th affine function is uniquely specified by the set of n+1 points $\{v_0, v_j\}$ with $j = 1, \cdots, n+1, j \neq i$ and $i \in \{1, \cdots, n+1\}$. We can get all the coefficients of affine functions:

$$[a_i^T b_i]^T = A_i^{-1} Y_i \tag{3}$$

where $A_i = [X_1^T, \cdots, X_{i-1}^T, X_{i+1}^T, \cdots, X_{n+1}^T]^T$, $Y_i = [y_0, \cdots, y_{i-1}, y_{i+1}, \cdots, y_{n+1}]^T$ and $X_j = [x_j^T, 1]^T$ with $j = 0, \cdots, n+1$. Meanwhile, we can further get that the equations

$$a_j^T x_k + b_j = y_k, \ j = 1, \cdots, n+1, \ j \neq k \tag{4}$$

hold for any $\{x_k, y_k\}, k \in \{1, \cdots, n+1\}$.

Assume that a PWA basis function $B(x)$ is defined over a compact set $D \subset R^n$ in n-dimensions. We can always find n+1 points $\{\overline{v}_i\}_{i=1}^{n+1}$ at the boundary of D and a point $\overline{v}_0$ at the interior of D, such that Equation (4) holds. The set of n+2 points $\{\overline{v}_i\}_{i=0}^{n+1}$ is defined as the generating points of $B(x)$. They specify the minimum and maximum function values of a PWA basis function [22].

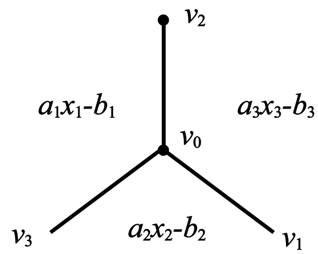

Fig 1. Domain Partition of a BPWA Basis Function in 2 Dimensions

2.2 BPWA Approximation Theorem

Definition 3[20] A function $p(x): R^n \to R$ is defined as a BPWA function, if

$$p(x) = a^T x + b + \sum_{m=1}^{M} \gamma_m B_m(x) \tag{5}$$

where $B_m(x)$ is a BPWA basis function, $a \in R^n$ and $b, \gamma_m \in R$.

Lemma 1[17] Let $f(x)$ be a sufficient smooth function. There must exist a M>0, such that for any positive integer c_B, there exist a BPWA function $p(x)$ with M basis functions, such that

$$\| p(x) - f(x) \| \le \frac{c_B}{M} \tag{6}$$

with $c_B = c_B\left(f(x)\right)$ is a functional of $f(x)$, and $\|\cdot\|$ denotes a norm on Γ.

Lemma 1 provides the theoretical foundation of using the BPWA functions in nonlinear approximation and dynamic identification.

3. Parameter Space Decomposition Theory of BPWA Functions

In this section, the parameter space decomposition theory is first developed for a single BPWA basis function. It is then generalized to the class of BPWA functions.

Lemma 2 Any BPWA basis function $B(x): D \to R$ has a unique decomposed parametric representation

$$B(x) = \alpha^T x + \beta + \gamma \max\left\{0, \alpha_1^T x + \beta_1, \cdots, \alpha_n^T x + \beta_n\right\} \tag{7}$$

where

$$\max_{x \in D}\left\{0, \alpha_1^T x + \beta_1, \cdots, \alpha_n^T x + \beta_n\right\} = 1 \tag{8}$$

with $\alpha \in R^n, \beta \in R, \gamma \in R$ specifies functional values at generating points in the range space, and $\alpha_i \in R^n, \beta_i \in R$ defines the domain partition of $B(x)$ in the domain space.

Proof. Let $B(x)$ be a BPWA basis function generated by a set of n+2 generating points $\{x_i, y_i\}_{i=0}^{n+1}$. Without loss of generalization, we assume that the first n+1 points $\{x_i, y_i\}_{i=0}^{n}$ specify the (n+1)-th local affine function (See Fig. 1 for a 2-dimensional case). This implies that

$$\max\{\alpha_1^T x_i + \beta_1, \cdots, \alpha_{n+1}^T x_i + \beta_{n+1}\} = \begin{cases} \alpha_{n+1}^T x_i + \beta_{n+1} & i=0,\cdots,n \\ \alpha_j^T x_i + \beta_j & i=n+1, j\in\{1,\cdots,n\} \end{cases} \tag{9}$$

Combining (7) and (9), we have

$$\begin{cases} \alpha^T x_0 + \beta + \gamma(\alpha_{n+1}^T x_0 + \beta_{n+1}) & = & y_0 \\ & \vdots & \\ \alpha^T x_n + \beta + \gamma(\alpha_{n+1}^T x_n + \beta_{n+1}) & = & y_n \\ \alpha^T x_{n+1} + \beta + \gamma(\alpha_j^T x_{n+1} + \beta_j) & = & y_{n+1} \end{cases} \tag{10}$$

Rewrite (9) into the matrix form, we can get

$$\begin{bmatrix} x_0^T & 1 & \alpha_{n+1}^T x_0 + \beta_{n+1} \\ & \vdots & \\ x_n^T & 1 & \alpha_{n+1}^T x_n + \beta_{n+1} \\ x_{n+1}^T & 1 & \alpha_j^T x_{n+1} + \beta_j \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix} = \begin{bmatrix} y_0 \\ \vdots \\ y_n \\ y_{n+1} \end{bmatrix} \tag{11}$$

Using elementary row operations, we have

$$\begin{bmatrix} x_0^T & 1 & \alpha_{n+1}^T x_0 + \beta_{n+1} \\ & \vdots & \\ x_n^T & 1 & \alpha_{n+1}^T x_n + \beta_{n+1} \\ x_{n+1}^T & 1 & \alpha_j^T x_{n+1} + \beta_j \end{bmatrix} \sim \begin{bmatrix} x_0^T & 1 & \alpha_{n+1}^T x_0 + \beta_{n+1} \\ & \vdots & \\ \left(x_n - x_0\right)^T & 0 & 0 \\ \left(x_{n+1} - x_0\right)^T & 0 & \left(\alpha_j - \alpha_{n+1}\right)^T x_{n+1} + \left(\beta_j - \beta_{n+1}\right) \end{bmatrix} \tag{12}$$

It follows that

$$\begin{vmatrix} x_0^T & 1 & \alpha_{n+1}^T x_0 + \beta_{n+1} \\ & \vdots & \\ x_n^T & 1 & \alpha_{n+1}^T x_n + \beta_{n+1} \\ x_{n+1}^T & 1 & \alpha_j^T x_{n+1} + \beta_j \end{vmatrix} = \begin{vmatrix} x_0^T & 1 & \alpha_{n+1}^T x_0 + \beta_{n+1} \\ & \vdots & \\ \left(x_n - x_0\right)^T & 0 & 0 \\ \left(x_{n+1} - x_0\right)^T & 0 & \left(\alpha_j - \alpha_{n+1}\right)^T x_{n+1} + \left(\beta_j - \beta_{n+1}\right) \end{vmatrix} \tag{13}$$

$$= 1 \cdot \left[\left(\alpha_j - \alpha_{n+1}\right)^T x_{n+1} + \left(\beta_j - \beta_{n+1}\right)\right] \begin{vmatrix} \left(x_1 - x_0\right)^T \\ \\ \left(x_n - x_0\right)^T \end{vmatrix}$$

Note that the points $\{x_i, y_i\}_{i=0}^{n}$ are in general position. We have $\begin{vmatrix} \left(x_1 - x_0\right)^T \\ \\ \left(x_n - x_0\right)^T \end{vmatrix} \neq 0$. It follows from Definition 3 that $\left(\alpha_j - \alpha_{n+1}\right)^T x_{n+1} + \left(\beta_j - \beta_{n+1}\right) \neq 0$. It immediately leads that the coefficient matrix in (11) is nonsingular. Therefore, if the coefficients of affine functions $\{\alpha_i, \beta_i\}_{i=1}^{n+1}$ of $B(x)$ are known, there always exists one and only one set of $[\alpha^T, \beta, \gamma]^T$, such that the set of linear equations of (9) are satisfied.

Equation (9) can be further decomposed into n+2 sets of equations

$$\begin{cases} \alpha^T x_0 + \beta = y_0 \\ \quad\vdots \\ \alpha^T x_n + \beta = y_n \\ \alpha^T x_{n+1} + \beta + \gamma = y_{n+1} \end{cases} \tag{14}$$

$$\begin{cases} \alpha_1^T x_0 + \beta_1 = 0 \\ \alpha_1^T x_2 + \beta_1 = 0 \\ \quad\vdots \\ \alpha_1^T x_{n+1} + \beta_1 = 1 \end{cases} \quad j = 1 \tag{15}$$

$$\begin{cases} \alpha_j^T x_0 + \beta_j = 0 \\ \quad\vdots \\ \alpha_j^T x_{j-1} + \beta_j = 0 \\ \alpha_j^T x_{j+1} + \beta_j = 0 \\ \quad\vdots \\ \alpha_j^T x_{n+1} + \beta_j = 1 \end{cases} \quad j = 2, \cdots n \tag{16}$$

$$\begin{cases} \alpha_{n+1}^T x_0 + \beta_{n+1} = 0 \\ \quad\vdots \\ \alpha_{n+1}^T x_n + \beta_{n+1} = 0 \end{cases} \quad j = n+1 \tag{17}$$

Equation (15-17) specifies the coefficients $\{\alpha_i, \beta_i\}_{i=1}^{n+1}$. The values of α_i, β_i, are independent from the function values at generating points $\{y_i\}_{i=0}^{n}$. They only define the boundaries of different polyhedral regions in the domain space. For example, $\alpha_1^T x + \beta_1 = \alpha_2^T x + \beta_2$ defines a boundary between polyhedral region R_1 and R_2. It follows from Equation (17) that α_{n+1}=0, β_{n+1}=0. Then $\max\{0, \alpha_1^T x + \beta_1, \cdots, \alpha_n^T x + \beta_n\}$ is defined as the standard form of the BPWA basis function $B(x)$, where $\{\alpha_i, \beta_i\}_{i=1}^{n}$ are defined as the geometrical parameters of $B(x)$. In addition, it is also implicitly defined in Equation (15-17) that $0 \le \max_{x \in D}\{0, \alpha_1^T x + \beta_1, \cdots, \alpha_n^T x + \beta_n\} \le 1$.

Equation (14) specifies the coefficients $[\alpha^T, \beta, \gamma]^T$, which are directly calculated from function values at generating points $\{y_i\}_{i=0}^{n}$. This implies that α, β, γ uniquely specify the function value of a $B(x)$ at the generating point in the range space. Then the coefficients α, β, γ are defined as the algebraic parameters of the BPWA basis function $B(x)$. It should be noted that α, β, γ has no effect on the domain partition of a PWA basis function.

Now we have proved that for any BPWA basis function, there exist one and only one set of algebraic parameters $[\alpha^T,\beta,\gamma]^T$ and geometrical parameters $\{\alpha_i,\beta_i\}_{i=0}^n$, such that Equation (7) holds. This completes the proof of Lemma 2. □

Let $B_1(x)=\max\{a_1^T x+b_1,\cdots,a_{n+1}^T x+b_{n+1}\}$ and $B_2(x)=\max\{\bar{a}_1^T x+\bar{b}_1,\cdots,\bar{a}_{n+1}^T x+\bar{b}_{n+1}\}$ be two BPWA basis functions. Assume that $B_1(x)$, $B_2(x)$ are defined over the same domain partition. According to Lemma 1, we have the decompositions

$$B_1(x)=\alpha^T x+\beta+\gamma\max\{0,\alpha_1^T x+\beta_1,\cdots,\alpha_n^T x+\beta_n\} \tag{18}$$

$$B_2(x)=\bar{\alpha}^T x+\bar{\beta}+\bar{\gamma}\max\{0,\alpha_1^T x+\beta_1,\cdots,\alpha_n^T x+\beta_n\} \tag{19}$$

It follows that the equation

$$\begin{aligned}\bar{B}(x)&=k_1B_1(x)+k_2B_2(x)\\&=(k_1\alpha+k_2\bar{\alpha})^T x+(k_1\beta+k_2\bar{\beta})+(k_1\gamma+k_2\bar{\gamma})\max\{0,\alpha_1^T x+\beta_1,\cdots,\alpha_n^T x+\beta_n\}\end{aligned} \tag{20}$$

holds for any k_1, $k_2\in R$. It is easy to see that the linear combination of two BPWA basis functions can be reduced as the weighted sum of their algebraic parameters, if they share the same set of geometrical parameters.

If $B_1(x)$, $B_2(x)$ are defined over different domain partitions, we can get

$$B_1(x)=\alpha^T x+\beta+\gamma\max\{0,\alpha_1^T x+\beta_1,\cdots,\alpha_n^T x+\beta_n\}+0\max\{0,\bar{\alpha}_1^T x+\bar{\beta}_1,\cdots,\bar{\alpha}_n^T x+\bar{\beta}_n\} \tag{21}$$

$$B_2(x)=\bar{\alpha}^T x+\bar{\beta}+0\max\{0,\alpha_1^T x+\beta_1,\cdots,\alpha_n^T x+\beta_n\}+\bar{\gamma}\max\{0,\bar{\alpha}_1^T x+\bar{\beta}_1,\cdots,\bar{\alpha}_n^T x+\bar{\beta}_n\} \tag{22}$$

It immediately follows that

$$\begin{aligned}\hat{B}(x)&=k_1B_1(x)+k_2B_2(x)\\&=(k_1\alpha+k_2\bar{\alpha})^T x+(k_1\beta+k_2\bar{\beta})\\&\quad+k_1\gamma\max\{0,\alpha_1^T x+\beta_1,\cdots,\alpha_n^T x+\beta_n\}+k_2\bar{\gamma}\max\{0,\bar{\alpha}_1^T x+\bar{\beta}_1,\cdots,\bar{\alpha}_n^T x+\bar{\beta}_n\}\end{aligned} \tag{23}$$

In (23), there are two kinds of calculations. The first one is the weighted sum of the algebraic parameters, which is introduced by the linear combination of $B_1(x)$ and $B_2(x)$. The second is the superposition of the domain partitions, which is produced by two sets of different geometrical parameters. This implies that the two terms $\max\{0,\alpha_1^T x+\beta_1,\cdots,\alpha_n^T x+\beta_n\}$ and $\max\{0,\bar{\alpha}_1^T x+\bar{\beta}_1,\cdots,\bar{\alpha}_n^T x+\bar{\beta}_n\}$ defines the polyhedral partitions of $\hat{B}(x)$ in the domain space. The geometrical domain partition is closed to addition and scalar multiplication. It leads that the domain partition of

$$k_1\gamma\max\{0,\alpha_1^T x+\beta_1,\cdots,\alpha_n^T x+\beta_n\}+k_2\bar{\gamma}\max\{0,\bar{\alpha}_1^T x+\bar{\beta}_1,\cdots,\bar{\alpha}_n^T x+\bar{\beta}_n\} \tag{24}$$

is exactly the same with that of

$$\max\{0,\alpha_1^T x+\beta_1,\cdots,\alpha_n^T x+\beta_n\}+\max\{0,\bar{\alpha}_1^T x+\bar{\beta}_1,\cdots,\bar{\alpha}_n^T x+\bar{\beta}_n\} \tag{25}$$

Here it is shown that the algebraic and geometrical parameters of the BPWA basis functions follow different calculation rules. This conclusion can be generalized to the BPWA functions.

Theorem 1 Any BPWA function has a unique decomposed parametric representation

$$p(x)=\alpha^T x+\beta+\sum_{m=1}^{M}\gamma_m \max\left\{0,\alpha_{m,1}^T x+\beta_{m,1},\cdots,\alpha_{m,n}^T x+\beta_{m,n}\right\} \tag{26}$$

if $p(x): D\to R$ is defined over a compact set $D\in R^n$, where $\max_{x\in D}\left\{0,\alpha_1^T x+\beta_1,\cdots,\alpha_n^T x+\beta_n\right\}=1$, α,β,γ_m are algebraic parameters, and $\alpha_{m,i},\beta_{m,i}$ are geometrical parameters with $i=1,\cdots,n+1$, $m=1,\cdots,M$.

Proof. A BPWA function is formulated as a weighted sum of M BPWA basis functions and an affine function. Using Lemma 1, we have that the decomposition

$$B_m(x)=\alpha_m^T x+\beta_m+\gamma_m \max\{0,\alpha_{m,1}^T x+\beta_{m,1},\cdots,\alpha_{m,n}^T x+\beta_{m,n}\} \tag{27}$$

holds for any $m\in\{1,\cdots,M\}$. It follows that

$$\begin{aligned} p(x)&=a^T x+b+\sum_{m=1}^{M}B_m(x)\\ &=a^T x+b+\sum_{m=1}^{M}\alpha_m^T x+\beta_m+\gamma_m \max\{0,\alpha_{m,1}^T x+\beta_{m,1},\cdots,\alpha_{m,n}^T x+\beta_{m,n}\}\end{aligned} \tag{28}$$

By defining $\alpha=a+\sum_{m=1}^{M}\alpha_m$ and $\beta=b+\sum_{m=1}^{M}\beta_m$, we can obtain

$$p(x)=\alpha^T x+\beta+\sum_{m=1}^{M}\gamma_m \max\{0,\alpha_{m,1}^T x+\beta_{m,1},\cdots,\alpha_{m,n}^T x+\beta_{m,n}\} \tag{29}$$

Note that a,b,α_m,β_m are algebraic parameters. Their affine combinations also specify the function values of $p(x)$ in the range space. Then α,β,γ_m are also algebraic parameters. Meanwhile, the superposition of $\max\{0,\alpha_{m,1}^T x+\beta_{m,1},\cdots,\alpha_{m,n}^T x+\beta_{m,n}\}$ produces the polyhedral partitions in the domain space. Therefore, $\{\alpha_m,\beta_m\}_{m=1}^{M}$ are the geometrical parameters. Here it is proved that any BPWA function has a parametric decomposition representation.

Theorem 1 shows that the parameter space of a BPWA function can be decomposed into two subspaces: the algebraic and geometrical parameter spaces. The parameters in the algebraic subspace have no effect on the geometrical structure of a BPWA function in the domain. Meanwhile, the parameters in the geometrical subspace do not change the function values of a BPWA function in the range space.

Corollary 1 The BPWA basis functions with the addition and scalar multiplication operations define a linear space, if they are defined over the same domain partition.

Proof. According to Theorem 1, the BPWA functions over a fixed domain partition share the same set of geometrical parameters, i.e.

$$p(x)=\alpha^T x+\beta+\sum_{m=1}^{M}\gamma_m \max\{0,\alpha_{m,1}^T x+\beta_{m,1},\cdots,\alpha_{m,n}^T x+\beta_{m,n}\} \tag{30}$$

where $\alpha_{m,i},\beta_{m,i}$ are constants with $m=1,\cdots,M$ and $i=1,\cdots,n+1$.

Denote $p(x)=p(x\mid\Pi)$ with $\Pi=\{\alpha,\beta,\gamma_1,\cdots,\gamma_M\}$. The follow equations

$$\begin{aligned} & p(x \mid k_1\Pi_1 + k_2\Pi_2) \\ & = (k_1\alpha_1 + k_2\alpha_2)^T x + (k_1\beta_1 + k_2\beta_2) + \sum_{m=1}^{M} (k_1\gamma_{m,1} + k_2\gamma_{m,2}) \max\{0, \alpha_{m,1}^T x + \beta_{m,1}, \cdots, \alpha_{m,n}^T x + \beta_{m,n}\} \\ & = k_1 p(x \mid \Pi_1) + k_2 p(x \mid \Pi_2) \end{aligned} \quad (31)$$

hold for any $k_1, k_2 \in R$. In addition, it is trivial to verify the other properties of a linear space, e.g. $p(x \mid 0) = 0$. Therefore, the set of BPWA functions spans a linear space, if their domain partitions are defined by the same set of geometrical parameters. □

It follows from Corollary 1 that the BPWA functions on the same domain partitions present a pseudo-linear model structure. This feature can facilitate the design of an efficient identification algorithm.

4. Recursive Identification of PWA Systems

4.1 BPWARX Model [17]

We define the following BPWA ARX (BPWARX) model

$$y(t) = \alpha^T \varphi(t) + \beta + \sum_{m=1}^{M} \gamma_m B_m(\varphi(t), \theta) + \varepsilon(t) \quad (32)$$

where $B_m(\varphi(t), \theta) = \max\left\{0, \varphi^T(t)\theta_{m,1}, \cdots, \varphi^T(t)\theta_{m,n}\right\}$ and $\theta_{m,i} \in R^{n+1}$ are the coefficients, $y(t) \in R$ is the measured output, $\varepsilon(t) \in R$ is the error term, and t is the sampling time. Here $\phi(t)$ is the regression vector consisting of previous inputs and outputs

$$\varphi(t) = [1, y(t-1), \cdots, y(t-n_a), u(t-1), \cdots, u(t-n_b)]^T \quad (33)$$

4.2 Change of Regression Vector

Given a BPWARX model with M BPWA basis functions. Assume that the geometrical structure is invariant. This implies that the geometrical parameters are kept constant. The set of M BPWA basis functions generates a nonlinear map $g : \varphi(t) \to \phi(t)$, where

$$\phi(t) = [\varphi^T(t), B_1(\varphi(t)), \cdots, B_M(\varphi(t))]^T \quad (34)$$

Although the relationship between $y(t)$ and $\phi(t)$ is nonlinear, the map is linear between $y(t)$ and $\phi(t)$, i.e.

$$y(t) = f(\varphi(t), \theta) = \gamma g(\varphi(t), \theta) = \gamma \phi(t) \quad (35)$$

where θ, γ are the geometrical and algebraic parameters, respectively. Under the assumption of an invariant geometrical structure, the identification of γ is reduced to a least square regression problem.

4.3 Recursive Identification Algorithm

The RPWA algorithm is based on the assumption that the geometrical structure of a BPWARX model can be accurately identified off-line, and it is invariant in the online process. In the off-line stage, the geometrical structure of a BPWARX model is identified using the MGN algorithm [14]. In the online stage, the recursive least square method is utilized for the identification of algebraic parameters. The main steps of this proposed algorithm are summarized as follows:

1) *Off-line stage:*

a) Generate the input-output data $\{u(t), y(t)\}_{t=1}^{T_1}$ for off-line identification;

b) Identify the algebraic and geometrical parameters of a BPWARX model using the MGN algorithm [17], i.e.

$$p(x) = f(\varphi(t), \hat{\theta}, \hat{\gamma})$$

c) Fix the geometrical parameters $\hat{\theta}$, and use (35) to calculate $\phi(t)$;

d) Calculate initial values: $\hat{\gamma}(0) = \hat{\gamma}$, $P(0) = [\Phi^T \Phi]^{-1}$ with $\Phi = [\phi^T(1), \cdots, \phi^T(T_1)]^T$.

2) *On-line stage:*

a) Run the recursive least square method to identify the algebraic parameters $\hat{\gamma}(t)$,

$$\begin{aligned} K(t+1) &= \frac{P(t)\phi(t+1)}{1+\phi^T(t+1)P(t)\phi(t+1)} \\ P(t+1) &= [I - K(t+1)\phi^T(t+1)]P(t) \\ \hat{\gamma}(t+1) &= \hat{\gamma}(t) + K(t+1)[y(t+1) - \phi^T(t+1)\hat{\gamma}(t))] \end{aligned} \tag{36}$$

where I is an identity matrix of proper size.

The main advantage of this algorithm is its high speed of online applications. This facilitates the design of computationally efficient algorithms to identify the time variant systems. The RPWA algorithm can also find promising applications in the adaptive control of nonlinear and hybrid systems.

The accuracy of the RPWA algorithm is also dependent on the geometrical parameters from the off-line calculation. Generally speaking, if the number of BPWA basis functions is big enough, the BPWARX model will consist of many degrees of freedom in its geometrical structure. It can then be flexible enough for general nonlinear systems.

5. Numerical Examples

Example 1: Consider the following nonlinear dynamic system

$$y(t+1) = \frac{y(t)y(t-1)y(t-2)u(t-1)(y(t-2)-1)+u(t)}{1+y(t-1)^2+y(t-2)^2} \tag{37}$$

This system belongs to a set of benchmark problems [25], and is also studied in [26, 27].

As the estimation data, the input $u(t), 1 \le t \le 400$ is generated using a random signal uniformly distributed in the interval [-1, 1]. In the RPWA algorithm, the first 200 data are utilized in the off-line stage, while the rest are for the online identification. By comparison, the MGN algorithm uses the set of 400 data to fit model parameters.

The validation data set is generated with the input signal

$$u(t) = \begin{cases} sin(2\pi t/250) & 401 \le t \le 500 \\ 0.8sin(2\pi t/250) + 0.2sin(2\pi t/25) & 501 \le t \le 800 \end{cases} \quad (38)$$

The algorithm's performance is evaluated using the Variance-Accounted-For (VAF), i.e.

$$VAF = \max\left\{1 - \frac{var(y(t) - \hat{y}(t))}{var(y(t))}, 0\right\} \times 100\% \quad (39)$$

where $var(\cdot)$ denotes the variance of signals, $y(t)$ and $\hat{y}(t)$ are the system and model output, respectively. In this example, the regression vector is $\varphi(t) = [1, y(t-1), y(t-2), y(t-3), u(t-1), u(t-2)]^T$.

Table I summarizes the predicted results of the BPWARX models using different identification methods. Here M and I denote the number of BPWA basis functions and off-line training epochs. τ_{off} and τ_{on} are the online and off-line calculation time, respectively. The one-step-ahead predictions are shown in Fig. 2 and Fig. 3. In this example, these two algorithms have roughly the same accuracy according to the VAF value. However, the online calculation time of the RPWA algorithm is negligible compared with the training time of the MGN algorithm. Therefore, the RPWA algorithm has high online computational efficiency without significant loss in identification precision.

Table I. **Simulation Results of Different Algorithms**

	M	I	τ_{off}	τ_{on}	VAF
RPWA Algorithm	10	2000	159.3	0.08	98.6%
MGN Algorithm	10	2000	316.2	-	99.0%

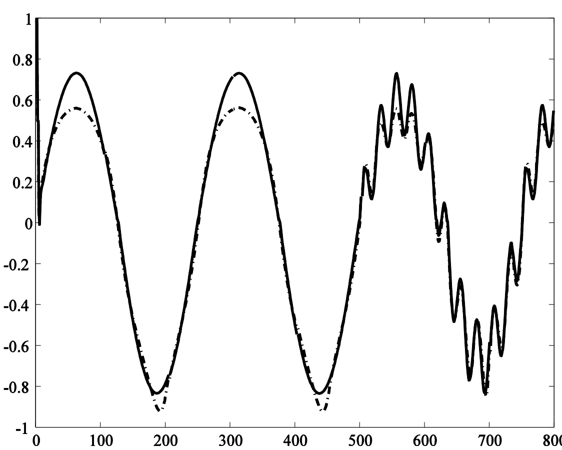

Fig. 2 True (solid) and Simulated (dashed) Outputs for System (37) using RPWA Algorithm

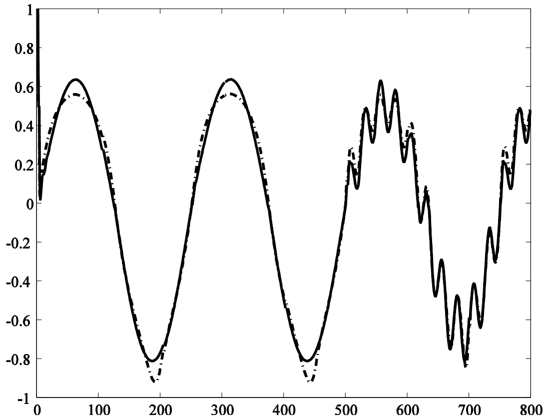

Fig. 3 True (solid) and Simulated (dashed) Outputs for System (37) using MGN Algorithm

Example 2: Consider the Agrawal bioreactor benchmark problem in nonlinear process identification [28, 29]. It is described by the following discrete-time equations

$$\begin{cases} x_1(t+1) = x_1(t) + T\left\{-x_1(t)u(t) + x_1(t)\left[1-x_2(t)\right]e^{x_2(t)/\tau}\right\} \\ x_2(t+1) = x_2(t) + T\left\{-x_2(t)u(t) + x_1(t)\left[1-x_2(t)\right]e^{x_2(t)/\tau} \times \dfrac{1+\beta}{1+\beta-x_2(t)}\right\} \\ \quad y(t) = x_1(t) + u(t) \end{cases} \tag{40}$$

where β = 0.02, τ = 0.48 and T is the sampling time constant. The states $x_1(t)$ and $x_2(t)$ are dimensionless quantities and only $x_1(t)$ is measurable.

The input signal is a multi-step signal with steps of 400 time units and a random magnitude between 0 and 0.6 from a uniform distribution. A set of 800 samples are generated, in which the first 600 data are used for identification and the rest for validation. In the RPWA algorithm, the first half of data are used in off-line training, while the second half in the online updating. The regression vector used is

$$\varphi(t) = [1, y(t-1), y(t-2), u(t-1), u(t-2)]^T$$

Fig. 4 shows the output of the BPWARX model using the RPWA algorithm on the validation data set. For comparison, Fig. 5 visualizes the output of the same model using the MGN algorithm. The simulation results are summarized in Table II. This example also shows that the RPWA algorithm has much higher computational efficiency than the MGN algorithm with the same accuracy.

Table II. Simulation Results of Different Algorithms

	M	I	τ_{off}	τ_{on}	VAF
RPWA Algorithm	15	2000	315.2	0.14	96.6%
MGN Algorithm	15	2000	630.5	-	96.9%

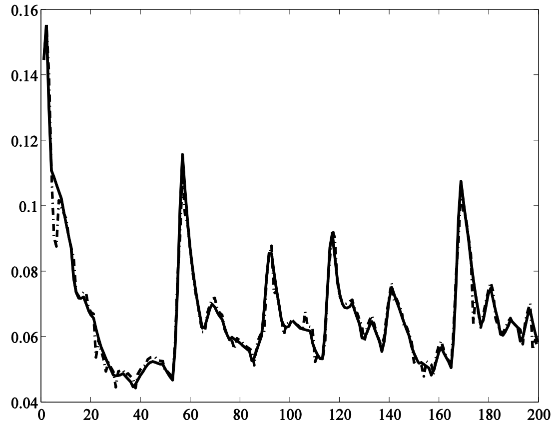

Fig. 4 True (solid) and Simulated (dashed) Outputs for System (40) using RPWA Algorithm

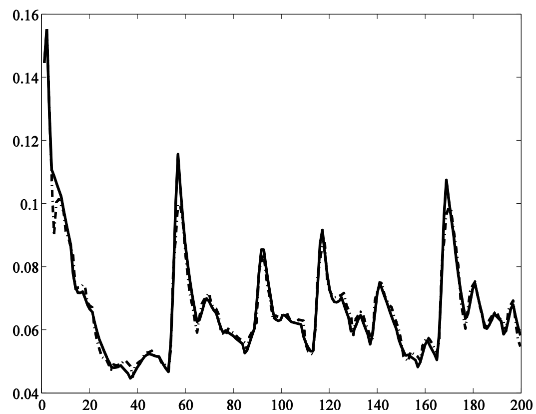

Fig. 5 True (solid) and Simulated (dashed) Outputs for System (40) Using MGN Algorithm

5. Conclusions

This paper proposes a novel parameter space decomposition theorem for the class of BPWA functions. A RPWA algorithm is developed for dynamic system identification. This algorithm identifies the geometrical parameters offline simultaneously and the algebraic parameters online recursively. The RPWA algorithm has much higher computational efficiency and similar identification precision when compared with the modified Gauss-Newton algorithm. Due to the recursive nature, the RPWA algorithm is anticipated to find successful applications in the identification of time-variant systems and the adaptive control of nonlinear systems.

Further researches are necessary to extend the parameter space decomposition theorem into the class of general PWA functions. The generalized theorem is promising to be used in the identification and control of hybrid systems [30-32].

References

[1] J. Roll, A. Bemporad & L. Ljung, "Identification of piecewise affine systems via mixed-integer programming," Automatica, vol. 40, no. 1, **(2004)**, pp. 37-50.

[2] P. Julian & J. M. Desages, “A Canonical piecewise-linear approximation of smooth functions," IEEE Transactions on Circuits Systems I, vol. 45, no. 5, **(1998)**, pp. 567-571.

[3] E. N. Pistikopoulos, V. Dua, N. A. Bozinis, A. Bemporad & M. Morari, “On-line optimization via off-line parametric optimization tools." Computers and Chemical Engineering, vol. 26, **(2001)**, pp. 175-185.

[4] A. Bemporad & M. Morari, “Control of systems integrating logic, dynamics, and constraints." Automatica, vol. 35, no. 3, **(1999)**, pp. 407-427.

[5] W. P. M. H. Heemels, B. De Schutter & A. Bemporad, “Equivalence of hybrid dynamical models," Automatica, vol. 37, no. 7, **(2001)**, pp. 1085-1091.

[6] R. Vidal, “Identification of PWARX hybrid models with unknown and possibly different orders." Proceedings of IEEE American Control Conference, **(2004)**.

[7] A. Juloski, S. Weiland, & W. Heemels, “A Bayesian approach to identification of hybrid systems." In: Proceedings of the 43rd Conference on Decision and Control, Paradise Island, Bahamas, **(2004)**, pp. 13-19.

[8] A. Bemporad, A. Garulli, S. Paoletti & A. Vicino, “A bounded-error approach to piecewise affine system identification," IEEE Transactions on Automatic Control, vol. 50, no. 10, **(2005)**, pp. 1567-1579.

[9] G. Ferrari-Trecate, M. Muselli, D. Liberati & M. Morari, “A clustering technique for the identification of piecewise affine systems." Automatica, vol. 39, no. 2, **(2003)**, pp. 205-217.

[10] A. Lj. Juloski, W. P. M. H. Heemels, G. Ferrari-Trecate, R. Vidal, S. Paoletti & J. H. G. Niessen. “Comparison of four procedures for the identification of hybrid systems". In M. Morari and L. Thiele, editors, Proc. 8th International Workshop on Hybrid Systems: Computation and Control, Lecture Notes in Computer Science, vol. 3414, **(2005)**, pp. 354-369. Springer-Verlag.

[11] C. Wen, X. Ma, X. & B. E. Ydstie, “Analytical expression of explicit MPC solution via lattice piecewise-affine function", Automatica, vol. 45, no. 4, **(2009)**, pp. 910-917.

[12] C. Wen & B. E. Ydstie, “Lattice piecewise-affine representation for explicit model predictive control", AIChE Annual Meeting, Nashville, TN, **(2009)** November.

[13] L. O. Chua & A. C. Deng, “Canonical piecewise-linear representation," IEEE Transactions on Circuits Systems I, vol. 35, no. 1, **(1988)**, pp. 101-111.

[14] C. Wen, S. Wang, H. Zhang & M. J. Khan, “A novel compact piecewise-linear representation," International Journal of Circuit Theory and Applications, vol. 33, no. 1, **(2005)**, pp. 87-97.

[15] C. Wen, S. Wang, X. Sun, H. Zhang, & F. Li. “A global compact representation for three-dimensional continuous piecewise linear functions", Journal of Tsinghua University, vol. 45, no. 4, **(2005)**, pp. 533-536.

[16] P. Pucar and J. Sjöberg, “On the hinge-finding algorithm for hinging hyperplanes." IEEE Transactions on Information Theory, vol. 44, no. 3, **(1998)**, pp. 1310-1319.

[17] C. Wen & X. Ma “A Basis-Function Canonical Piecewise-Linear Approximation." IEEE Transactions on Circuits and Systems I: Regular Papers, vol. 55, no. 5, **(2008)**, pp. 1328-1334.

[18] P. Julian & J. M. Desages, “A Canonical piecewise-linear approximation of smooth functions,” IEEE Transactions on Circuits Systems I, vol. 45, no. 5, **(1998)**, pp. 567-571.

[19] J. N. Lin & R. Unbehauen, “Canonical Piecewise-Linear networks," IEEE Trans. Neural Networks. I, vol. 6, no. 1, **(1995)**, pp. 43-50.

[20] C. Wen & X. Ma, “A max-piecewise-linear neural network for function approximation", Neurocomputing, vol. 71, no. 1-3, **(2008)**, pp. 843-852.

[21] L. Breiman, “Hinging Hyperplanes for regression, classification and function approximation," IEEE Transactions on Information Theory, vol. 39, no. 3, **(1993)**, pp. 999-1013.

[22] C. Wen, S. Wang, F. Li & M. J. Khan, “A compact *f-f* model of high-dimensional piecewise-linear function over a degenerate intersection," IEEE Transactions on Circuits Systems I, vol. 52, no. 4, **(2005)**, pp. 815-821.

[23] C. Wen, S. Wang, X. Jin & X. Ma, “Identification of Dynamic Systems Using Piecewise-Affine Basis Function Models," Automatica, vol. 43, no. 10, **(2007)**, pp. 1824-1831.

[24] M. A. Armstrong, Basic Topology, New York: Springer-Verlag, **(1983)**.

[25] K. S. Narendra & K. Parthasarathy. “Identification and control of dynamical systems using neural networks." IEEE Transactions on Neural Networks, vol. 1, no. 1, **(1990)**, pp. 4-27.

[26] G. P. Liu, Nonlinear identification and control: a neural network approach, London; New York: Springer, **(2001)**.

[27] V. Verdult, L. Ljung & M. Verhaegen, “Identification of composite local linear state-space models using a projected gradient search". International Journal of Control, vol. 75, no. 16, **(2002)**, pp. 1385-1398.

[28] P. Agrawal, C. Lee, H. C. Lim & D. Ramkrishna, "Theoretical Investigations of Dynamic Behaviour of Isothermal Continuous Stirred Tank Biological Reactors," Chemical Engineering Science, vol. 37, no. 3, **(1982)**, pp. 453-462.

[29] L. H. Ungar, "A bioreactor benchmark for adaptive network-based process control," Neural Nefworks for Control, W. T. Miller, 111, R. S. Sutton, and P. J. Werbos, Eds. Cambridge, MA: MIT Press, chap.16, **(1990)**, pp. 387-402 and Appendix A, pp. 476-480.

[30] R. Huang, L.T. Biegler, & S. C. Patwardhan. "Fast offset-free nonlinear model predictive control based on moving horizon estimation." Ind. Eng. Chem. Res., vol. 49, no. 17, **(2010)**, pp. 7882-7890.

[31] R. Huang, V.M. Zavala & L.T. Biegler. "Advanced step nonlinear model predictive control for air separation units." Journal of Process Control, vol. 19, no. 4, **(2009)**, pp. 678-685.

[32] J. Liu, X. Chen, D. Munoz de la Pena & P. D. Christofides, "Sequential and Iterative Architectures for Distributed Model Predictive Control of Nonlinear Process Systems." AIChE Journal, vol. 56, **(2010)**, pp. 2137-2149.

Authors

Rui Huang is currently a senior research scientist in the system department at United Technologies Research Center after receiving his PhD degree from Carnegie Mellon University (2010). He also obtained his bachelor degree and master degree from Zhejiang University, Hangzhou, China. His research interests are in the area of dynamic real time optimization, nonlinear control and estimation theory.

Weiquan Sun is currently pursuing a master degree at Duquesne University, Pittsburgh. His research interest includes system and control, nonlinear optimization and artificial neural network.

Min Li received master and PhD degrees in the years of 2003 and 2007, respectively. He did Postdoc research in 2007 at University of Pittsburgh, USA. He has been serving several journals at various fields. He is the recipient of several overseas fellowships. He works in part of the areas of automation and control. Currently, he is at Indiana University, USA. He is also a member of many professional societies.