Geometrical Structure Identification of General Piecewise-affine Functions for Nonlinear Dynamic Systems

Weiquan Sun[1] and Min Li[2]

[1]*School of Business Administration, Duquesne University, Pittsburgh, PA, USA*
[2]*School of Medicine, Indiana University,*
975 West Walnut Street, Indianapolis, IN, 46202

sunw@duq.edu, li246@iupui.edu

Abstract

This paper proposes a class of General Piecewise-Affine (GPWA) AutoRegressive eXogenous (GPWARX) models for nonlinear black-box identification. A GPWARX model is a weighted sum of GPWA basis functions. In n-dimensions, a GPWA basis function is defined over n+2 polyhedral regions, which are pairwise directly adjacent in the domain space. The geometrical structures of GPWA basis functions provide the GPWARX models with stronger representation capability and higher approximation efficiency in estimation of nonlinear systems than the other canonical PWA functions. Benchmark nonlinear identification examples are illustrated to show that GPWARX models have better estimation accuracy than HHARX and BPWARX models with the same number of parameters and using the same parameter identification algorithm.

Keywords: *Piecewise-affine systems, system identification, model structure, hybrid system, general PWA basis function, canonical PWA representation*

1. Introduction

Piecewise-affine (PWA) functions provide an attractive black-box model structure for system identification. The class of PWA models has universal approximation capability for nonlinear dynamic systems [1, 2]. This suggests a systematic way of using well-developed linear technologies to solve nonlinear problems. PWA functions can describe hybrid systems, which consists of both continuous and discrete dynamics [3-5]. Therefore, many analysis, synthesis and control techniques for PWA systems can be generalized to hybrid systems [6-12].

It is a challenging task to identify PWA models from a set of input-output data. PWA models are defined by local affine functions and polyhedral domain partitions. Therefore, PWA identification algorithms require fitting two sets of parameters: coefficient vectors of local affine functions and parameter matrices of affine inequalities, which define polyhedral regions in the domain.

PWA model identification can be formulated as a mixed integer nonlinear program (MINLP) problem. Global optimal solutions are feasible for small-scale or some medium-scale problems. Problem scale is defined by the number of sampling data and the number of polyhedral regions in a PWA model. A typical medium-scale problem may have 20-30 sampling data scatter into several polyhedral regions. The main idea of most PWA identification algorithms is developing computationally efficient approaches to build suboptimal PWA models of high quality. PWA identification approaches can be classified into two categories: direct and indirect method.

Direct method identifies local affine functions and domain partitions directly. This method requires fitting all the parameter vectors of local hyperplanes and parameter matrices of polyhedral regions. Typical direct methods include the clustering-based method [13], the bounded-error method [14], the Bayesian method [15] and the algebraic method [16]. Direct methods can deal with general PWA systems, including discontinuous systems. Their high performances in computational efficiency have been shown in many applications [17].

Wider applications of direct methods are limited by descriptional and computational complexities of PWA models. Due to the combinatorial natural, the number of parameters to describe a PWA system may be an exponential function of the number of independent variables. In addition, functional evaluation of PWA models defined over polyhedral regions might be too computationally intensive to run in real time.

Indirect methods are developed based on the concept of PWA basis function expansion [18-26]. A PWA function is represented with a weighted sum of PWA basis functions. Note that a PWA function over many polyhedral regions may be described by a small number of PWA basis functions. Indirect method is an efficient way to deal with description and evaluation complexities of continuous PWA systems.

Indirect methods require identifying the weights and parameter vectors of PWA basis functions. This problem is addressed by formulating a criterion function and minimizing it through numerical optimization methods. Then the performances of indirect PWA identification algorithm are implicitly specified by the number of PWA basis functions, which is denoted as model size.

Canonical Piecewise-Affine (CPWA) representation theorem provides a theoretical basis to design PWA basis functions and determine the size of PWA models. The CPWA representation theorem is developed based on the concept of minimal degenerate intersection (MDI). The MDIs are geometrical objects, which are elementary “building blocks” of PWA functions’ polyhedral partitions in the domain. The superposition of MDIs can generate all types of domain partitions of continuous PWA functions. Mathematically, by defining a PWA function over a MDI as basis function, a weighted sum of these basis functions can describe any continuous PWA system.

In [26], it is proved that a PWA function over a MDI can be mathematically represented by an absolute-value function of n nesting levels. A PWA function with n variables can be expressed by a weighted sum of no more than n-nesting absolute-value functions. Unfortunately, the nested absolute-value functions usually have implicit functional forms. This limits the use of nested absolute-value functions in building a general PWA model structure.

In [2, 21, 27], a class of CPWA functions are developed using absolute-value functions as basis functions. A CPWA function is a sum of an affine function and one or more absolute values of affine functions. The CPWA functions can approximate continuous functions arbitrarily well. Hinging Hyperplane (HH) functions is essentially another form of the CPWA functions [8, 18, 21]. The class of absolute-value functions does not have a universal representation capability for general PWA functions with n variables. This limits their approximation efficiency in estimation of nonlinear systems. It implies that a nonlinear system may require an approximate CWPA function with a big number of absolute-values functions.

In [30], a PWA basis function (BPWA) representation is proposed, which describes a subset of continuous PWA function with a linear combination of BPWA basis functions. In n dimensions, a BPWA basis function is the maximum or minimum of n+1 affine functions. In [21, 27], Wen, Wang, Jin and Ma developed the PWA Basis

Function AutoRegressive eXogenous (BPWARX) models and found successful applications in nonlinear identifications and function approximation. In [31], a parameter space decomposition theorem is developed 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. A recursive algorithm is developed to identify the BWPARX models from the input-output data.

The class of BPWA functions has a universal representation capability with a simple and explicit functional form. However, they may have low representation efficiency. A BPWA basis function is defined over a single MDI. According to BPWA representation theorem, the number of MDIs in a PWA function specifies the size of its BPWA representation, i.e. the number of basis functions. Note that the approximation accuracy of BPWA functions is implicitly determined by the number of MDIs. A BPWA function defined over larger number of MIDs gives a smaller approximation error. Therefore, BPWARX models may have low approximate efficiency, because they might require a big model size to estimate nonlinear systems.

In 2011, Wen and Ma develop a General PWA (GPWA) representation theorem [32]. This theorem provides a uniform theoretical framework for many known CPWA models. It is proved that the geometrical structures of domain partitions determine the representation capability of different classes of PWA basis functions. The GPWA representation theorem shows that an ideal identification algorithm should be developed based on identifying the geometrical structures of PWA functions in the domain space. This paper proposes a GPWARX model using the GPWA representation theorem. A class of GPWA basis functions is proposed, which are defined over n+2 MDIs in n dimensions. A modified Guass-Newton algorithm [27] is used to build a GPWARX model from input-output data. The performance of GPWARX models is supported by two benchmark examples of nonlinear dynamic identification.

The organization of this paper is as follows. The preliminary knowledge on CPWA representation is given in Section 2. Section 3 reviews CPWA models and CPWA approximation theorems. In Section 4, GPWARX models are introduced, and numerical simulation results are illustrated in Section 5. Section 6 gives a brief concluding remark.

2. Preliminaries

This section provides the preliminary knowledge on the PWA functions and the CPWA representation theorem.

Definition 1. Assume that $\Gamma \in R^n$ be a compact set. Assume further that there is a finite family of nonempty polyhedral regions R_i, such that

$$\Gamma = \cup_{i=1}^{m} R_i, \hat{R}_p \cap \hat{R}_q = \Phi, \forall p \neq q \tag{1}$$

where $\hat{R}_i$ is the interior of region R_i, Φ denotes an empty set, and p, $q \in \{1,\cdots,m\}$. A function $P(x)$:$\Gamma \rightarrow R$ is said to be piecewise affine if

$$P(x) = \varphi^T \theta_j \qquad \forall x \in R_j \tag{2}$$

where $\varphi = [1 \;\; x^T]^T, \theta_j \in R^{n+1}, H_j \in R^{M_j \times (n+1)}$, and "$\leq$" denotes component-wise inequality. The sets

$$R_j = \{x \mid H_j \varphi \leq 0\}, j = 1, \cdots, s \tag{3}$$

form a polyhedral partition of the x-space. Here R_j is called as a region and $\varphi^T\theta_j$ the corresponding local affine function.

In order to measure the complexity of a PWA function, the CPWA representation theorem introduces the concept of MDIs.

Definition 2 [25] In R^n, an intersection of two regions, which is an (n-1)-dimensional affine manifold, is said to be a first-order intersection $S^{(1)}$. An affine manifold of dimension n-k in R^n is a kth-order intersection $S^{(k)}$, if it is the intersection of two or more affine manifolds of $S^{(k-1)}$

$$S^{(k)} = \cap_{i=1}^{\geq 2} S_i^{(k-1)} \tag{4}$$

An affine manifold $\bar{S}^{(k)}$ is called a minimal degenerate intersection (MDI), if it is formed by the smallest possible number of $\bar{S}^{(l)}$'s for all $1 \leq l < k$.

The MDIs are the basic "building blocks" of general CPWL functions. They can be used to define the geometrical relationships of different polyhedral regions in the domain space.

Definition 3 Two regions are directly adjacent if the intersection forms a first-order MDI.

Definition 3 can be clarified by the domain partition in 2 dimensions as shown in Figure 1. Figure 1(a-c) consists of directly adjacent regions only, because each couple of two regions intersects with a first-order MDI. In Figure 1 (d), the regions OAB and OCD are not directly adjacent, because OAB∩OCD=O is a second-order degenerate intersection. There are three regions in Figure 1 (c) and four ones in Figure 1 (d). Both domain partitions are defined a second-order degenerate intersection, i.e. point O. Figure 1 (c) defines a MDI, which is the simplest domain partition that defines a second-order degenerate intersection. Here the simplest domain partition implies that it has the minimum number of regions.

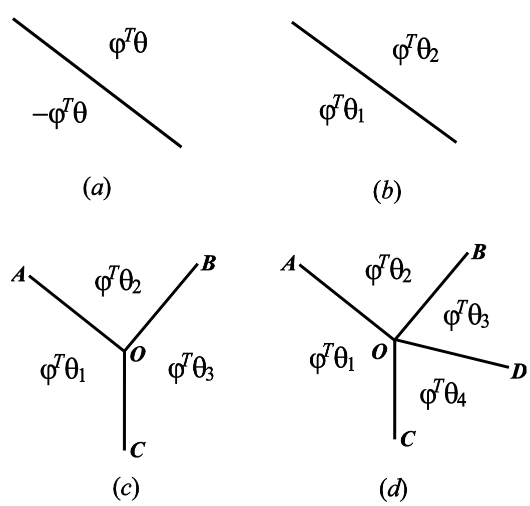

Figure 1. Domain partition of: (a) absolute-value function; (b) hinge function; (c) PWA basis function; (d) PWA function with dis-adjacent regions

3. PWA Approximation Theorem

3.1. PWA Models

Consider the nonlinear system

$$y(t) = P(x(t)) + \varepsilon(t) \tag{5}$$

where $x(t)\in R^n$ is the regression vector, $y(t)\in R$ is the measured output, $\varepsilon(t)\in R$ is the error term, and $P(x)$ is a PWA function. When the regression vector $\varphi(t)$ consists of previous inputs and outputs

$$\varphi(t) = [1, y(t-1),\cdots, y(t-n_a), u(t-1),\cdots, u(t-n_b)]^T \tag{6}$$

the system is defined as a PWARX system [20].

Rewrite $P(x)$ into a basis function expansion, we have

$$P(x) = \varphi^T\theta + \sum_{m=1}^{M}\lambda_m B_m(x) \tag{7}$$

where $B_m(x)$ is a basis function with $\lambda_m\in R$, $m=1,\cdots,M$. Typical examples of basis functions are absolute-value function $a(x)$, hinge functions $h(x)$ and PWA basis function $b(x)$. In R^n, analytical expressions of basis functions are formulated as follows:

$$a(x) = |\varphi^T\theta| \tag{8}$$

$$h(x) = \max\left\{\varphi^T\theta_1, \varphi^T\theta_2\right\} \tag{9}$$

$$b(x) = \max\left\{\varphi^T\theta_1, \cdots, \varphi^T\theta_{n+1}\right\} \tag{10}$$

It should be noted that the number of affine function in a BPWA basis function is specified by the dimensions of domain space. By comparison, the absolute-value functions and hinge functions are constructed by 2 affine functions. Figure 1(a-c) visualizes the domain partitions of the absolute-value functions, hinge functions and BPWA basis functions in 2 dimensions. The domain space is partitioned into two directly adjacent regions for an absolute-value functions and a hinge function. Note that two regions only define one first-order MDI. This limits the representation capability of the absolute-value functions and hinge functions in high dimensional space.

The domain of a BPWA basis function contains n+1 pairwise directly adjacent regions in n dimensions. It defines a single n-dimensional MDI. This is the simplest domain partition with universal representation capability. Figure 1(c) shows the domain partition of a 2-dimensional BPWA basis function. Point O defines a second-order MDI. The order of O will be reduced by one, if any one region is removed from the domain partition as shown in Figure 1(c). Therefore, a 2-dimensional BPWA basis function can be used as an elementary "building block" to represent any continuous PWA function with two variables.

Figure 1(d) consists of regions that are not directly adjacent. It implies that this domain partition can be further decomposed into a superposition of simpler domain partitions defined by BPWA basis functions. Denote $P(x)=P(x|\theta_1,\theta_2,\theta_3,\theta_4)$ as a PWA function defined over a domain partition shown as in Figure 1(d). Using the decomposition algorithm in [30], we can get

$$P(x) = P_1(x\,|\,\theta_{11},\theta_{12},\theta_{13}) + P_2(x\,|\,\theta_{21},\theta_{22},\theta_{23})$$

where $P_1(x|\theta_{11},\theta_{12},\theta_{13})$ and $P_2(x|\theta_{21},\theta_{22},\theta_{23})$ are BPWA basis functions defined over domain partitions shown in Figure 1(c).

3.2 PWA Approximation Theorem

Definition 1[18]**:** A function $f(x)$, $x\in C$ is sufficiently smooth if the following integral is finite:

$$\int \|\omega\|^2 |\hat{f}(\omega)|\, d\omega < \infty \tag{11}$$

where $\hat{f}(\omega)$ is the Fourier transform of $f(x)$ and C is a compact set in R^n.

Lemma 1[18,21]**:** Let $f(x)$ be a sufficiently smooth function. Then there must exist a $c_B>0$, such that for any positive integer M, there exist M basis functions $B_m(x)$ and coefficients $\lambda_m\in R$, such that

$$\|\varphi^T\theta + \sum_{m=1}^{M}\lambda_m B_m(x) - f(x)\|^2 \le \frac{c_B}{M} \tag{12}$$

where $B_m(x)$ denotes an absolute-value function $a_m(x)$, a hinge function $h_m(x)$ or a PWA basis function $b_m(x)$ with $m=1,\cdots,M$.

Lemma 1 shows that the absolute-values functions, hinge functions and PWA basis functions can approximate sufficiently smooth nonlinear functions arbitrarily well. However, these three classes of basis functions have different approximation efficiency, because they have different representation capabilities. Here the approximation accuracy denotes the number of parameters in a PWARX model required for a given approximation accuracy.

The hinge functions have the same representation capability with the absolute-value functions. Those two classes of basis functions can be formulated as the minimum or maximum of two affine functions. The absolute-value functions and hinge functions only cover a small subset of continuous PWA functions, i.e. the functions that fulfill the so-called consistent variation property [22]. Therefore, they can not efficiently approximate the PWA functions with the boundary configurations that break the consistent variation property. Accordingly, a huge number of hinge functions are required to achieve desired estimation accuracy. The resulting HHARX models may involve very large number of parameters in real identification settings. This conclusion is also valid for the class of absolute-value functions.

The class of BPWA functions covers all continuous PWA functions defined over a single degenerate intersection. Note that PWA functions are generally defined over many degenerate intersections and a PWA basis function generates a single MDI. Theoretically, a lot of BPWA basis functions are required to describe a PWA function over many degenerate intersections. Therefore, the resulting BPWARX model may have low representation efficiency, because they might contain a large number of parameters.

4. GPWA Representation Model

4.1 GPWA Basis Function

Definition 4 In R^n, a continuous PWA function

$$g(x) = \min\left\{0, \max\left\{\varphi^T\theta_{m,1},\cdots,\varphi^T\theta_{m,n+1}\right\}\right\} \tag{13}$$

is defined as a GPWA basis function, if regions R_i, i=1,⋯,n+2 are pairwise directly adjacent in the domain space.

From a geometrical point of view, a GPWA basis function $g(x)$ is defined over a compact domain structure, in which each couple of its regions is directly adjacent, and each group of n regions forms an nth-order MDI. Note that (n+1) affinely independent points spans an n-dimensional hyperplane. A $g(x)$ is a direct extension of an affine function, because its domain partition is spanned by n+2 points in general position. A set of n+2 points in general position means that any set of n+1 points span an n-dimensional hyperplane.

From a mathematical point of view, a $g(x)$ realizes the maximum parametric degree of freedom in n+2 affinely independent points. In R^n, n+2 points $\{x_i, y_i\}_{i=1}^{n+2}$ at most offer (n+1)×(n+2) degree of freedom. This is equal to the number of coefficient parameters in n+2 affine functions of a $g(x)$. Therefore, a GPWA basis $g(x)$ realizes the upper bound of the degree of freedom in n+2 points. It is the most complex PWA function specified by n+2 points in general position. Therefore, the class of the GPWA basis functions is the most complicated CPWA functions, which can be used as the elementary "building blocks" of general PWA functions.

The class of GPWA basis functions is defined over a particular geometrical structure, which enables them with universal representation capability and high approximation efficiency in nonlinear black-box modeling.

Figure 2 visualizes the surface plot of a GPWA basis function with 2 variables. The corresponding domain partitions are shown in Figure 3. It is easy to see that any two polyhedral regions are directly adjacent to form a first-order MDI, e.g. $OAC \cap ABC = AC$, and each group of three regions intersect at a second-order MDI, e.g. $OAC \cap ABC \cap OAB = O$.

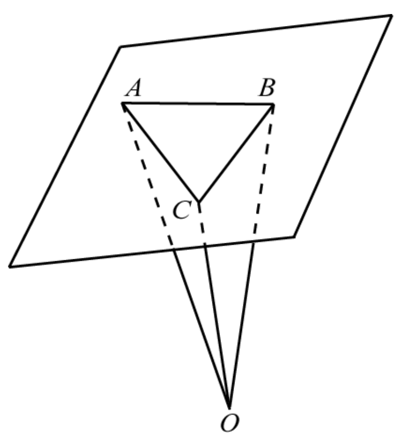

Figure 2. Surface Plot of a GPWA Basis Function in 2 Dimensions

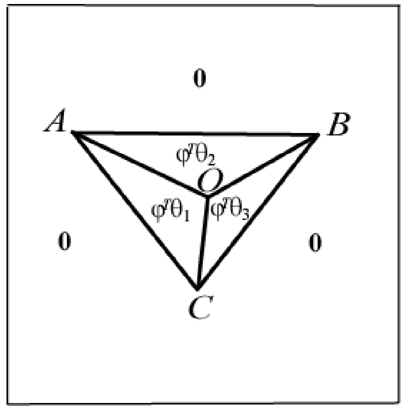

Figure 3. Domain Partition of GPWA Basis Function in 2 Dimensions

4.2 GPWA Approximation

Lemma 2[32] Let $P(x)$ be a continuous PWA function. Then there must exist a GPWA function

$$p(x) = \varphi^T\theta + \sum_{m=1}^{M} \lambda_m g_m(x) \tag{14}$$

such that

$$P(x) = p(x), \quad \forall x \in R^n \tag{15}$$

where $g_m(x)$ is a GPWA basis function.

Theorem 1: In R^n, any continuous function on a compact domain can be approximated to an arbitrary precision by a GPWA function.

Proof. Because of the theory of spline functions, any continuous nonlinear function can be approximated arbitrarily well by a PWA function. It is stated in Lemma 2 that any continuous PWA function has a GPWA representation. The class of GPWA functions is a universal model set for all continuous functions.

Theorem 1 presents the theoretical basis to use the class of GPWA functions to approximate continuous nonlinear functions. If the nonlinear functions satisfy the sufficient smooth condition, the approximation accuracy can be estimated by the number of GPWA basis functions.

Theorem 2: Let $f(x)$ be a sufficiently smooth function. Then there must exist a positive real c_G>0, such that for any positive integer M, there exist M basis functions $g_m(x)$ and coefficients $\lambda_m \in R$, such that

$$\| \varphi^T\theta + \sum_{m=1}^{M} \lambda_m g_m(x) - f(x) \|^2 \le \frac{c_G}{M} \tag{16}$$

and

$$c_G \le c_B \tag{17}$$

where $g_m(x)$ is a GPWA basis function with $m=1,\cdots,M$.

Proof. A BPWA basis is defined over a single MDI, while a GPWA basis is over n+1 MDIs. It is easy to see that a BPWA basis function is essentially a special case of

GPWA basis function. Then similar error bound as (12) is obtained for a sufficiently smooth function i.e.

$$\| \varphi^T \theta_0 + \sum_{m=1}^{M} \lambda_m g_m(x) - f(x) \|^2 \leq \frac{c_G}{M} \tag{18}$$

Note that the BPWA functions with M basis functions are a proper subset of the GPWA functions with M basis functions. The error bound c_B/M in (12) is also a looser upper bound for the BWPA functions in (17), i.e. $c_G \leq c_B$.

The GPWARX models have a stronger approximation capability than both the HHARX and BPWARX models. They can get higher estimation accuracy than the other two models with the same number of parameters in nonlinear identifications. Therefore, the GPWARX models present a more suitable model structure, because they can efficiently approximate the geometrical structures in the domain space of general PWA functions.

5. Numerical Examples

Example 1: Consider the following Agrawal bioreactor benchmark problem in nonlinear identification and control [33, 34]. This problem is described by the following discrete-time equations

$$\begin{cases} x_1(t+1) = x_1(t) + \left\{ -x_1(t)u(t) + x_1(t)[1 - x_2(t)]e^{x_2(t)/\tau} \right\} \\ x_2(t+1) = x_2(t) + \left\{ -x_2(t)u(t) + x_1(t)[1 - x_2(t)]e^{x_2(t)/\tau} \times \dfrac{1+\beta}{1+\beta - x_2(t)} \right\} \\ \quad y(t) = x_1(t) + u(t) \end{cases} \tag{19}$$

with β = 0.02 and τ = 0.48. The states $x_1(t)$ and $x_2(t)$ are dimensionless quantities and only $x_1(t)$ is measurable.

The input 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 30000 data samples is generated, where the first 20000 samples are used for identification and the rest for validation. The input-output data sets are re-sampled at 1/50. Then 400 data samples are available for identification. The regression vector is $\varphi(t) = [1, y(t-1), y(t-2), u(t-1), u(t-2)]^T$.

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

$$VAF = \max\left\{ 1 - \frac{var(y(t) - \hat{y}(t))}{var(y(t))}, 0 \right\} \times 100\% \tag{20}$$

where $var(\cdot)$ denotes the variance of signals, $y(t)$ and $\hat{y}(t)$ are the system and model output, respectively.

Three models are used to generate the one-step-ahead predictions of the system. Figure 4 shows the outputs of the real system and the GPWARX model after 2000 circles of iterations. By comparison, Figure 5 demonstrates the simulation results of the HHARX and BPWARX model with roughly the same number of parameters.

The modified Gauss-Newton algorithm [27] is used to build GPWARX, BPWARX and HHARX models. Table I summarizes the predicted results of different models. Here M and I denote the number of basis functions and training epochs. N_B is the number of parameters in each "basis function", and N denotes the total number of model parameters. In this example, the GWARX model has a VAF value of 99.4%, which is larger than the HHARX model (97.6%) and BPWARX model (98.8%). The former model has better prediction accuracy than the latter two models.

Table 1. Comparison of Approximation Accuracies

	M	N_B	*N*	*I*	*VAF*
GPWARX Model	10	42	430	2000	99.4%
BPWARX Model	14	31	448	2000	98.8%
HHARX Model	34	12	442	2000	97.6%

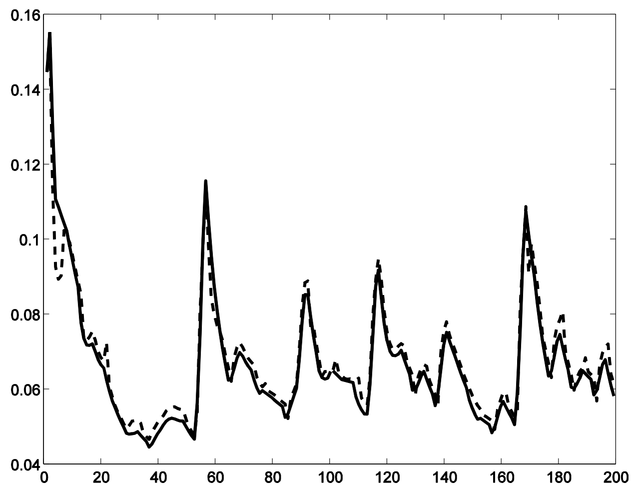

Figure 4. True (solid) and Predicted (dashed) Outputs using GPWARX Model

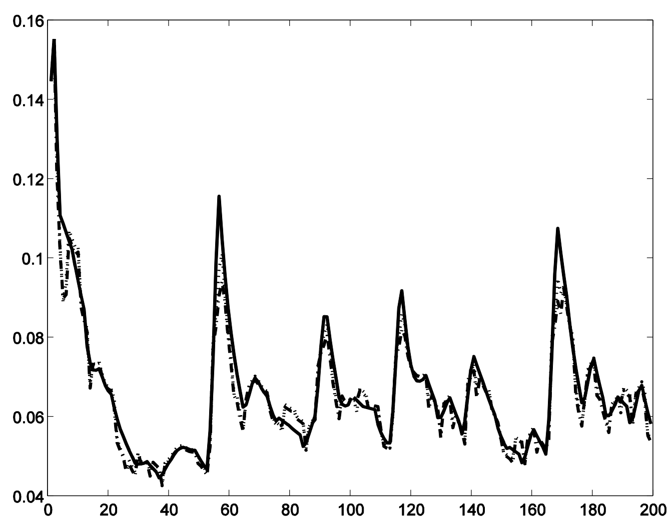

Figure 5. True (solid) and Predicted Outputs using HHARX Models (dash dotted) and BPWARX Model (dotted)

Example 2: Consider another nonlinear benchmark system described in the state-space form [35, 36]

$$\begin{cases} x_1(t+1) = \left(\dfrac{x_1(t)}{1+x_1^2(t)}+1\right)\sin\left(x_2(t)\right) \\ x_2(t+1) = x_2(t)\cos\left(x_2(t)\right)+x_1(t)e^{-\frac{x_1^2(t)+x_2^2(t)}{8}}+\dfrac{u^3(t)}{1+u^2(t)+0.5\cos\left(x_1(t)+x_2(t)\right)} \\ y(t) = \dfrac{x_1(t)}{1+0.5\sin\left(x_2(t)\right)}+\dfrac{x_2(t)}{1+0.5\sin(x_1(t))} \end{cases} \tag{21}$$

To generate identification data, the system is excited with a random input signal $u(t)$, uniformly distributed on the interval [-2, 2] with $1 \le t \le 200$. The validation data set is generated with the input

$$u(t) = \sin(2\pi t/10)+\sin(2\pi t/5), 1 \le t \le 200$$

The regression vector used is

$$\varphi(t) = [1, y(t-1), y(t-2), y(t-3), u(t-1), u(t-2), u(t-3)]^T$$

Figure 6 and Figure 7 show one-step-ahead predictions of the real system and three PWA models. Table II lists the simulation results and the models' parameters. It is easy to see that the GPWARX models obtain much better prediction accuracy than the BPWARX and HHARX models.

Table 2. Comparison of Approximation Accuracies

	M	N_B	N	I	VAF
GPWARX Model	20	56	1140	3000	93.4%
BPWARX Model	26	43	1144	3000	89.4%
HHARX Model	76	14	1144	3000	87.5%

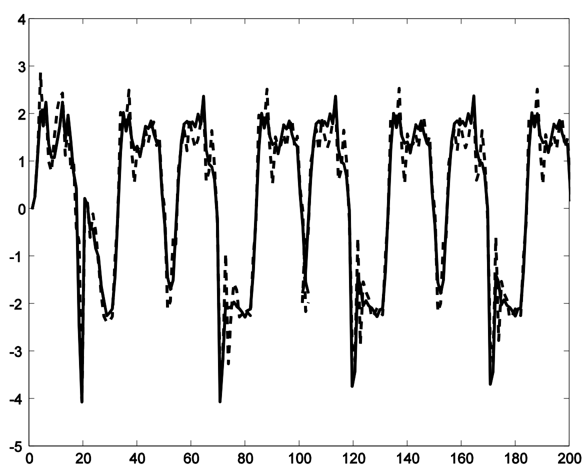

Figure 6. True (solid) and Predicted Outputs using GPWARX Model

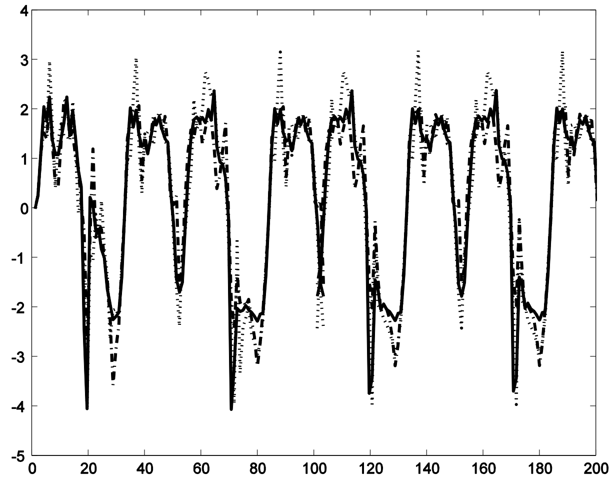

Figure 7. True (solid) and Predicted Outputs using HHARX Models (dash dotted) and BPWARX Model (dotted)

6. Conclusions

This paper proposes a class of GPWARX models for nonlinear system identification. Geometrical structures of GPWA basis functions provide the GPWARX models with stronger representation capability and higher approximation efficiency than the other canonical PWA models. Numerical simulation results demonstrate that GPWARX models obtain higher precision in the estimation of nonlinear systems than HHARX and BPWARX models with the same number of parameters and using the same parameter identification algorithm.

The class of GPWA functions has a global and compact functional form. They offer an alternative model structure to conventional nonlinear models. GPWA models can find many applications in nonlinear/hybrid system identification and control. For example, GPWA functions can describe any continuous PWA system dynamics, which are used in the framework of explicit model predictive control (eMPC). The eMPC reformulates the online optimization in a model predictive control (MPC) into a multi-parametric program problem. The optimal control action is calculated off-line as explicit functions of the state and reference vectors. The online computational time of conventional MPC can be reduced to the microsecond-millisecond range. The eMPC can extend the application domain of MPC into fast sampling systems.

It is shown in [37] that redundant parameters may be introduced into GPWA models, which are identified using a modified Gauss-Newton algorithm from a set of input-output data. Therefore, it is a promising research direction to develop a model reduction algorithm for the class of GPWA functions. This algorithm can generate a reduced GPWA model with smaller number of parameters for a given approximation accuracy. Further researches in this direction are being performed.

References

[1] J. Roll, “Local and piecewise affine approaches to system identification", Ph.D. thesis, Department of Electrical Engineering, Linkoping University, **(2003)**.

[2] P. Julian and 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] A. Bemporad, G. Ferrari-Trecate and M. Morari, “Observability and controllability of piecewise affine and

hybrid systems", IEEE Transactions on Automatic Control, vol. 45, no. 10, **(2000)**, pp. 1864-1876.
[4] W. P. M. H. Heemels, B. De Schutter and A. Bemporad, "Equivalence of hybrid dynamical models", Automatica, vol. 37, no. 7, **(2001)**, pp. 1085-1091.
[5] A. Bemporad and M. Morari, "Control of systems integrating logic, dynamics, and constraints", Automatica vol. 35, no. 3, **(1999)**, pp. 407-427.
[6] E. D. Sontag, "Nonlinear regulation: The piecewise linear approach", IEEE Transactions on Automatic Control, vol. 26, no. 2, **(1981)**, pp. 346-358.
[7] B. De Schutter and T. J. J. van den Boom, "MPC for continuous piecewise affine systems", Systems and Control Letters, vol. 52, **(2004)**, pp. 179-192.
[8] J. Roll, A. Bemporad and L. Ljung, "Identification of piecewise affine systems via mixed-integer programming", Automatica, vol. 40, no. 1, **(2004)**, pp. 37-50.
[9] A. Rantzer and M. Johansson, "Piecewise linear quadratic optimal control", IEEE Transactions on Automatic Control, vol. 45, no. 4, **(2000)**, pp. 629-637.
[10] Chengtao Wen, X. Ma & B. E. Ydstie. "Analytical Expression of Explicit MPC Solution via Lattice Piecewise-affine Function", Automatica, 45(4): 910-917, 2009.
[11] C. Wen and B. E. Ydstie, "Lattice Piecewise-affine Representation for Explicit Model Predictive Control", AIChE Annual Meeting, Nashville, TN, **(2009)** November.
[12] G. Schullerus, V. Krebs, B. De Schutter and T. J. J. van den Boom, "Input signal design for identification of max-plus-linear systems", Automatica, vol. 42, no. 6, **(2006)**, pp. 937-943.
[13] G. Ferrari-Trecate, M. Muselli, D. Liberati and M. Morari, "A clustering technique for the identification of piecewise affine systems", Automatica, vol. 39, no. 2, **(2003)**, pp. 205-217.
[14] A. Bemporad, A. Garulli, S. Paoletti and A. Vicino, "A bounded-error approach to piecewise affine system identification", IEEE Transactions on Automatic Control, vol. 50, no. 10, **(2005)**, pp. 1567-1579.
[15] A. Juloski, S. Weiland and 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.
[16] R. Vidal, S. Soatto, Y. Ma and S. Sastry, "An algebraic geometric approach to the identification of a class of linear hybrid systems", Proceedings of IEEE Conference on Decision and Control, **(2003)**.
[17] A. Lj. Juloski, W. P. M. H. Heemels, G. Ferrari-Trecate, R. Vidal, S. Paoletti and 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.
[18] L. Breiman, "Hinging hyperplanes for regression, classification and function approximation", IEEE Transactions on Information Theory, vol. 39, no. 3, **(1993)**, pp. 999-1013.
[19] 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.
[20] J. Roll, A. Bemporad and L. Ljung, "Identification of piecewise affine systems via mixed-integer programming", Automatica, vol. 40, no. 1, **(2004)**.
[21] Chengtao Wen, S. Wang, X. Jin and X. Ma, "Identification of dynamic systems using piecewise-affine basis function models", Automatica, vol. 43, no. 10, **(2007)**, pp. 1824-1831.
[22] L. O. Chua and A. C. Deng, "Canonical piecewise-linear representation", IEEE Transactions on Circuits Systems I, vol. 35, no. 1, **(1988)**, pp. 101-111.
[23] C. Wen, S. Wang, H. Zhang and M. J. Khan, "A novel compact piecewise-linear representation", International Journal of Circuit Theory and Applications, vol. 33, no. 1, **(2005)**, pp. 87-97.
[24] J. N. Lin and R. Unbehauen, "Canonical Piecewise-Linear networks", IEEE Trans. Neural Networks. I, vol. 6, no. 1, **(1995)**, pp. 43-50.
[25] C. Kahlert and L. O. Chua, "The complete canonical piecewise-linear representation– Part I: The geometry of the domain space", IEEE Transactions on Circuits Systems I, vol. 39, no. 3, **(1992)**, pp. 222-236.
[26] J. N. Lin and R. Unbehauen, "Canonical piecewise-linear approximation", IEEE Transactions on Circuits Systems I, vol. 39, no. 8, **(1992)**, pp. 697-699.
[27] C. Wen, & X. Ma, "A max-piecewise-linear neural network for function approximation", Neurocomputing, vol. 71, no. 1-3, **(2008)**, pp. 843-852.
[28] C. Wen, S. Wang, X. Sun, H. Zhang and 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.
[29] C. Wen and S. Wang, "A Complete 3-D Canonical Piecewise-linear Representation", International Conference on Technology and Automation- ICTA'05, Thessaloniki, Greece, **(2005)**.
[30] C. Wen, S. Wang, F. Li and M. J. Khan, "A compact f-f model of high-dimensional piecewise-linear function over a degenerate intersection," IEEE Transactions on Circuits Systems I, vol. 52, no. 4, **(2005)**, pp. 815-821.
[31] R. Huang, W. Sun and M. Li, "Recursive Identification of Piecewise-affine Systems based on Parameter Space Decomposition Theorem", International Journal of Control and Automation, vol. 4, no. 4, **(2011)**, pp.23-37.

[32] C. Wen and X. Ma, "A general canonical piecewise-linear function theorem: geometrical structures determines representation capability", IEEE Transactions on Circuits Systems II, vol. 58, no. 2, **(2011)**, pp. 936-940.
[33] P. Agrawal, C. Lee, H. C. Lim and 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.
[34] 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, pp. 387-402 and Appendix A, **(1990)**, pp. 476-480.
[35] K. S. Narendra and S. M. Li, "Neural networks in control systems", In P. Smolensky, M. C. Mozer, & D. E. Rumelhart (Eds.), Mathematical perspectives on neural networks. London: Lawrence Erlbaum Associates, **(1996)**.
[36] J. Roll, A. Nazin and L. Ljung, "Nonlinear system identification via direct weight optimization," Automatica, vol. 41, no. 3, **(2005)**, pp. 475-490.
[37] C. Wen and X. Ma, "Polyhedral Partition Simplification of General Canonical Piecewise-affine Functions for Dynamic System Identification", Automatica, 2nd revision, submitted in **(2011)**.

Authors

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. He has published three papers in international academic journals in the area of modeling and cotnrol of hybrid systems.

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.