A CONTROL STRATEGY USING A CPWL NOE STRUCTURE

L.R. Castro *,1  J. L. Figueroa **,1,2  O. E. Agamennoni ***,1,3

* Dto. de Matemática, Univ. Nac. del Sur, B8000CPB Bahía Blanca, Argentina
** Dto. de Ing. Eléctrica y de Computadoras, CONICET - Univ. Nac. del Sur, B8000CPB Bahía Blanca, Argentina
*** Dto. de Ing. Eléctrica y de Computadoras, CIC - Univ. Nac. del Sur, B8000CPB Bahía Blanca, Argentina

Abstract: In this paper we develop a nonlinear control strategy based on a Nonlinear Output Error (NOE) model structure that uses Canonical High Level Piecewise Linear (HL CPWL) functions to approximate the nonlinear system. This model structure allows the implementation of identification and control algorithms that allows to increase or decrease very easily the complexity of the model during the identification process. This property is very attractive because it allows to find the appropriate NOE model without overfitting. Using this CPWL NOE model structure, we define a simple local linear control scheme.

Keywords: Nonlinear systems, identification, piecewise linear techniques.

1. INTRODUCTION

There exist a set of very well known techniques to design and analyze feedback control strategies for linear systems. If the system under consideration is nonlinear and important performance requirements are imposed, nonlinear control design tools must be used. Canonical Piecewise Linear (CPWL) approximation in the context of Nonlinear Output Error (NOE) model structure allows a systematic multilinear or Linear Parameter Varying (LPV) consideration of a nonlinear dynamical systems. The High Level CPWL (HL CPWL) formulation used in this paper (Julián et al., 1999) is based on a simplicial partition of the input domain such that the system has a linear affine formulation in each simplex which is continuous on the boundaries.

Under slowly varying assumptions, different linear controller scheduling techniques have been proposed in the literature ((Rugh and Shamma, 2000), (Shamma and Athans, 1991), (Shamma and Athans, 1992), (Galán et al., 2004)).

Within the context of fuzzy logic and neural networks, the controller scheduling idea has received the attention of researchers ((Palm and Stutz, 2003), (Chen and Huang, 2004)) as well.

In the framework of these ideas, in this paper we present a new local linear control strategy of a nonlinear system based on a CPWL NOE model description. In ((Castro et al., 2005b), (Castro et al., 2005a)) a CPWL Nonlinear Output Error (NOE) model structure and an identification algorithm were presented. This structure is similar to the one proposed by Narendra and Parthasarathy

---

1 Partially supported by the ANPCyT
2 Partially supported by CONICET
3 Partially supported by CIC
Let the input/output available data of length $L$ corresponding to a given Lipschitz continuous, SISO system. If $\tilde{y}$ is the estimated value corresponding to the input $u$, and
\begin{align}
\mathbf{u}^{k,M+1} &= [u_k, \ldots, u_{k-M}] \\
\mathbf{\tilde{y}}^{k-1,N} &= [\tilde{y}_{k-1}, \ldots, \tilde{y}_{k-N}],
\end{align}
we propose the following black-box identification structure
\begin{align}
\tilde{y}_k &= f_{pwl}(\mathbf{u}_k^{k-M+1}, \mathbf{\tilde{y}}^{k-1,N}) \\
&= c\Lambda(\mathbf{u}^{k,M+1}, \mathbf{\tilde{y}}^{k-1,N}),
\end{align}
where the HL CPWL function $f_{pwl}(\mathbf{x}) = c\Lambda(\mathbf{x})$ is defined as in (Julián et al., 1999), (Julián, 2000) and the model orders $M$ and $N$ are given. This identification structure, pictured in Fig. 1, can be considered a black box model where the regression vector noted by $\varphi^k$ is taken as $\varphi^k = [\mathbf{u}^{k,M+1}, \mathbf{\tilde{y}}^{k-1,N}]$ see (Sjöberg et al., 1995), for example. It is worth to mention that a linear model is a particular case of $f_{pwl}$. Finally, let us note
\begin{align}
z^k &= [\mathbf{u}^{k,M+1}, \mathbf{\tilde{y}}^{k-1,N}],
\end{align}

Fig. 1. CPWL NOE model

The domain of the function $f_{pwl}$ is a compact set $S \subset \mathbb{R}^m, m = M + N + 1$, defined as follows
\begin{align}
S = \{ \mathbf{z} \in \mathbb{R}^m : \bar{z}_i \leq z_i \leq \underline{z}_i; \quad \bar{z}_i = \tilde{z}_i + \delta \cdot n_{div} \\
&= \max z_i \text{ over the entire input/output set with } \mathbf{z} \text{ defined in (4).}
\end{align}

According to ((Julián et al., 1999), (Julián, 1999)), the set defined by (5) is partitioned into polyhedral regions using a simplicial boundary configuration and $f_{pwl}$ is linear on each simplex and continuous on the adjacent boundaries of the simplices.

Taking into account Eq. (5) each dimension is divided into a number of subintervals of equal length $\delta$. Then, when the grid size $\delta$ decreases, the number of divisions $n_{div}$ on each direction increases. As a consequence, using HL CPWL functions for the nonlinear approximation, $n_{div}$ allows to go from a linear model ($n_{div} = 1$) with a coarse partition of the domain to a nonlinear one with a finer partition of $S$. The advantages of using this kind of models is pointed out in (Sjöberg and Ngia, 1998, Ch.1).

According to the proposed methodology we write the identification algorithm using the following notation.
Identification Algorithm

**Notation**

\( n_{div} = 2^d, d \geq 0 \): number of divisions of the region \( S \). Equal number of divisions in each dimension is assumed.

\( V^d \): the set of vertices of the simplicial partition of the set \( S \) with \( n_{div} = 2^d \) number of divisions.

\( \Lambda^d \): The HL CPWL basis defined on \( S \) with vertices belonging to \( V^d \).

\( c \): the row vector of parameters associated with the basis \( \Lambda^d \). The number of parameters is \( (n_{div} + 1)^{M+1+N} \) ([Julián et al., 1999], Julián, 1999).

\( (A)_j \): the j-th row of a matrix \( A \).

\( Niter \in \mathbb{R} \): maximum number of iterations of the optimization algorithm.

\( Maxerror \): maximum allowable approximation error.

\( lr \): learning rate, \( lr_i > 0 \forall i \) (\( lr_i = 0.001 \)).

\( mom \): momentum, \( mom > 0 \) (\( mom = 0.9 \)).

\( lr_{inc} \): learning rate increment, \( lr_{inc} > 1 \) (typically, \( lr_{inc} = 1.05 \)).

\( lr_{dec} \): learning rate decrement, \( 0 < lr_{dec} < 1 \) (typically, \( lr_{dec} = 0.3 \)).

\( \eta \): constant update, \( 0 \leq \eta \leq 1 \).

**Identification Algorithm**

**Step 1.** \( d = 0 \): Linear Approximation.

Compute the set of parameters \( c^{d,*} \) of the linear model solving the following LS problem

\[
e^{d,*} = \min \left( \|y - \hat{y}\|^2 \right)
= \arg \min \left\{ \frac{1}{2} \sum_{i=1}^{L} [y_i - c^{d}A^d ([u^{i,M+1}, \hat{y}^{i-1,N}])^2] \right\}.
\]

**Step 2.** \( d \leftarrow d + 1 \): . Set \( r = 0 \). Evaluate the initial condition \( c^{d,*} \) for the new \( d \) according to the algorithm given in ([Castro et al., 2005a]).

Set \( r = 0 \), \( c^{d,r} = c^{d,*} \), \( \eta = 0 \), \( \Delta c^{d,r} = [0, \ldots, 0] \), \( lr^{d,0} = 0.001 \).

**Step 3.** \( r \leftarrow r + 1 \): Error and gradient evaluation.

\[
E^r = \frac{1}{2} \sum_{i=1}^{L} [y_i - c^{d,r-1}A^d ([u^{i,M+1}, \hat{y}^{i-1,N}])^2]
\]

(6)

\[
\nabla E^r_j = \frac{\partial E^r}{\partial c^{d,r-1}}
= -\sum_{i=1}^{L} [y_i - c^{d,r-1}A^d ([u^{i,M+1}, \hat{y}^{i-1,N}])] \cdot (A^d ([u^{i,M+1}, \hat{y}^{i-1,N}])^T)\]

Step 4. Parameter update. If \( E^r \leq Maxerror \) then STOP; otherwise

\[
\Delta c^{d,r} = \eta \left( -\nabla E^r_j [\Delta c^{d,r}]_{j} + \Delta c^{d,r-1} \right) \text{ mod } \Lambda^d.
\]

(7)

\[
e^{d,r} = c^{d,r-1} + \Delta c^{d,r}.
\]

(8)

where the constant \( \eta \) is precisely defined in Appendix B and the components of the learning rate vector \( lr \) are modified as described below.

\[
lr^{r,j} = \begin{cases} lr^{r,j-1} \times lr_{inc} & \text{if } \text{sign}(\nabla E^r_j) = \text{sign}(\nabla E^{r-1}_j) \\ lr^{r,j-1} \times lr_{dec} & \text{if } \text{sign}(\nabla E^r_j) \neq \text{sign}(\nabla E^{r-1}_j) \end{cases}
\]

If \( r < Niter \) go to Step 3.

else \( c^{d,*} = c^{d,r} \), go to Step 2.

**Remark 1.** In order to improve the algorithm performance, any of the well known stop conditions based on the error evolution, may be applied in Step 4.

**Remark 2.** The described algorithm could be modified to reduce the order of the model. Then, the solution at any level could be found backwards. If only the solution before the last one must be recovered, other possibility would be to retain that vector of parameters and reconstruct the solution using the expression \( \gamma_k = c^{d,r} \Lambda^d ([u^{k,M+1}, \hat{y}^{k-1,N}]) \) for all \( k, 1 \leq k \leq L \).

The advantages and drawbacks of this algorithm have been pointed out in ([Castro et al., 2005b], Castro et al., 2005a).

3. CONTROLLER DESIGN

In order to design the controller, we use the fact that in each sector of the domain, the CPWL NOE structure behaves as a linear model. Then, we will use a linear controller for each one of these sectors.

To design these linear controllers, we adopt the direct synthesis approach (Ogungba and Ray, 1994). The controller specification produces a feedback system with a closed-loop pole in \( a_c \) without offset, when the set point is changed in the form of steps. If the discrete transfer function of the linear model is called \( H(z) \), and \( H^{-1}(z) \) is stable and causal, then the controller can be described as

\[
K(z) = \frac{a_c}{z - 1} \frac{1}{H(z)}.
\]

(9)

Then, the controller algorithm involves two steps, first it is necessary to extract the linear model valid for the current operating point; then, with this model, a controller is designed.

Let us consider that the process is actually in a given operating point described by the vector \([u^{k,M+1}, \hat{y}^{k-1,N}] \). Then it is simple to determine the simplex \( R^{(k)} \) that contains this point, for example using the PWL Toolbox of MATLAB.

IFAC - 337 - ADCHEM 2006
Note that if the dimension of the domain is \( m = N + M + 1 \) each simplex is defined by their \( m + 1 \) vertices. Then the gradient \( J^{(i)} \) in this simplex can be easily evaluated using the values of \( J_{\text{pol}} \) at its vertices (see Appendix A). The parameters of the linear model at the region are the entries of this gradient vector. Then it is possible to implement the control algorithm (see Fig. 2) as follows.

![Control Scheme](image)

**Control Model Algorithm**

**Data.** At time \( k \), the past values of the measured and controlled variables to form \( x = [u^{k,M+1}, y^{k-1,N}] \). The desired closed loop time constant \( (a_c) \).

**Step 1.** Determine the sector \( R^{\text{old}} \) in which is operating the process, the linear model corresponding to this sector (using the algorithm of Appendix B) and the linear controller \( K^{\text{actual}} \) of Eq. (9).

**Step 2.** Determine the actual sector \( R^{\text{actual}} \).

**Step 3.** If \( R^{\text{actual}} \) is not equal to \( R^{\text{old}} \), compute the actual linear model, the new controller \( K^{\text{actual}} \) and make \( R^{\text{old}} = R^{\text{actual}} \).

**Step 4.** Apply the corresponding manipulated variable \( u_k \) to the process, and measure \( y_k \).

**Step 5.** Compute the new manipulated variable \( u_{k+1} \) (computed using \( K^{\text{actual}} \) and \( y_k \)). Make \( k \leftarrow k + 1 \).

**Step 6.** Update \( x = [u^{k-M}, y^{k-N}] \) using \( u_{k+1} \) and \( y_k \) and return to Step 2.

**Remark 3.** It is obvious that the performance of this control algorithm depends on the quality of the model. In this case, we consider that each simplex in the partition of the domain have enough data to allow a good quality model. In this way, it is possible to improve the robustness of the control algorithm by relaxing the control specification i.e. the time constant when the number of data in a given region is small.

4. **EXAMPLE**

In this example we model the neutralization reaction between a strong acid and a strong base in the presence of a buffer agent as described by (Galán, 2001) (for a complete description of this process and a first principles models see (Biagiola et al., 2004)).

The goal is to control the output \( pH \), by manipulating the alkaline solution flow rate \( q_B \). The operating point for the neutralization is \( q_B = 0.5 \) and \( pH = 7.7182 \).

In order to identify the system, \( q_B \) is excited by a random signal with uniform distribution between 0 and 1, the limits of the physical variable. Time simulation is performed for 250 seconds for a sample time of 0.25 sec. The regression vector of the CPWL NOE model is taken as \( \psi_k = [u_k, \tilde{y}_{k-1}, \tilde{y}_{k-2}] \) and the number of divisions for each variable are two, four and eight, giving a total number of parameters equal to 27, 125 and 729, respectively. The parameters of the algorithm for adjusting the vector of parameters are taken as \( lr^0 = 0.0002, mom = 0.9, inc = 1.05 \) and \( dec = 0.3 \).

In Fig. 3 the error is displayed as a function of the iterations; the number of divisions of the simplices increases every 1000 iterations.

![RMS Error for pH Neutralization](image)

The controller is used to follow a set point change. The controller parameter is set at \( a_c = 0.5 \). The simulation for this control is shown in Fig. 4. In this figure, the system response for the controller with \( ndiv = 2, 4 \) and 8, \( ndiv = 4 \) are shown. Again, from these plots it is clear that the controller performance improves when the number of divisions increase.

5. **CONCLUSIONS**

In this paper, a NOE identification algorithm based on HL CPWL functions approximation method is reviewed, and an algorithm to control this model structure is presented. The identification methodology allows to approximate a NOE.
model from a linear one, and the control scheme uses the linear information in each region of operation to design a simple controller. The main feature of this process is that it enables to go from a linear model to a nonlinear one straightforwardly. Finally, the potentials of our approach have been illustrated with two examples.

REFERENCES


Appendix A. LINEAR MODEL FROM CPWL FUNCTIONS

Let us consider a HL CPWL function \( f_{\text{pwl}} : S \subset \mathbb{R}^m \rightarrow \mathbb{R}^4 \) and a vector \( \mathbf{x} \) at a given region \( R^{(k)} \), i.e. \( \mathbf{x} \in R^{(k)} \subset S \). This region \( R^{(k)} \) is uniquely defined by its \( m+1 \) vertices \( V^{(j)}, j = 1, \ldots, m+1 \). Once that these vertices are computed \(^4\), it is possible to obtain the linear model \( J^{(j)} \in \mathbb{R}^m \) that represents the process at this region. An algorithm to perform the linear model can be written as follows.

Linear Model Algorithm

Step 1. Compute the value of \( f_{\text{pwl}} \) at vertices,

\(^4\) This is trivial using the function \text{vertices.m} from (Julián, 2000)
for $i = 1$ to $m + 1$
\begin{align*}
(z)_i &= c_d \Theta (V^{(i)}) \quad ,
\end{align*}
end

**Step 2.** From the vector and matrix, 
for $i = 1$ to $m$
\begin{align*}
(I)^i &= (z)_i - (z)_{m+1}, \\
\Theta(:, i) &= V^{(i)} - V^{(m+1)},
\end{align*}
end

where $\Theta(:, i)$ represents the $i$th column of the $\Theta$ matrix.

**Step 3.** Compute the gain vector $J$ , 
\begin{align*}
J &= \Theta^{-1} \Gamma
\end{align*}

### Appendix B. BIBO STABILITY

Let us suppose that $u \in U \subset \mathbb{R}^{M+1}, y \in O \subset \mathbb{R}^N$, $U$ and $O$ given compact sets, $Q \subset U \times O$. $Q$ compact and $I = [\underline{y}, \overline{y}] \subset \mathbb{R}$, with $\underline{y} = \min y, \overline{y} = \max y$.

**Definition 4.** We say that the model defined by (3) is BIBO stable if $f_{pwl}(Q) \subset I$.

This definition means that the model output remains within the output values when the input is any signal $u \in U$.

The expression (3) defines a mapping $f_{pwl} : Q \rightarrow I$. As $Q$ is a compact set and $f_{pwl}$ is continuous on $Q$, then it attains its maximum and minimum values on $Q$. Moreover, since $f_{pwl}$ is linear on each simplex, the extreme values are attained on $\nabla Q$, the set of vertices of $Q$. Then the NOE identification structure given by (3) will be BIBO stable if the minimization problem
\begin{align*}
\min_{c^d} E^r \quad \text{s.t.} \quad \max_{z^k \in Q} f_{pwl}(z^k) = \max_{v \in V^d} |c^d \Lambda (v)| \subset I,
\end{align*}

where $E^r$ is given by equation Eq. (6) and $V^d$ is defined in Section 2, is solvable. This minimization problem is equivalent to
\begin{align*}
\min_{c^d} E^r \quad \text{s.t.} \quad \begin{cases}
\max_{v \in V^d} (|c^d \Lambda (v)|) \leq \overline{y} \\
\min_{v \in V^d} (|c^d \Lambda (v)|) \geq \underline{y}
\end{cases} \quad \text{(B.1)}
\end{align*}

**Proposition 1.** Let us suppose that, for $d$ and $r - 1$ fixed, the model is BIBO stable. Then the model will be BIBO stable for $d$ and $r$ if the following condition is satisfied
\begin{align*}
\overline{y} - \min_{v \in V^d} \left( c^{d,r-1} \Lambda (v) \right) &\leq \Delta c^{d,r} \Lambda (v) \\
&\leq \overline{y} - \max_{v \in V^d} \left( c^{d,r-1} \Lambda (v) \right), \quad \text{(B.2)}
\end{align*}

where $\Delta c^{d,r}$ is given by equation Eq. (8) and $v \in V^d$.

**Proof.** See (Castro et al., 2005a).

**Corollary 2.** With the hypothesis of Proposition 1, let us note $a = \overline{y} - \min_{v \in V^d} \left( c^{d,r-1} \Lambda (v) \right), b = \overline{y} - \max_{v \in V^d} \left( c^{d,r-1} \Lambda (v) \right)$. Then
\begin{align*}
\eta \geq \min_{v \in V^d} \left[ (\Delta c^{d,r-1}_{\text{mom}}) \Lambda (v) \right] \quad \text{(B.3)}
\end{align*}

\begin{align*}
\eta \leq \max_{v \in V^d} \left[ (\Delta c^{d,r-1}_{\text{mom}}) \Lambda (v) \right] \quad \text{(B.4)}
\end{align*}

simultaneously.

**Proof.** See (Castro et al., 2005a).

**Remark 5.** Considering a 10,000-length input vector, it takes around 16$\mu$s of CPU to solve the minimization problem with restrictions given by (B.1) in the MATLAB environment, with a Pentium IV, 512Mb RAM computer.

Once we have found $c^d$ that satisfies (B.1), we must guarantee that, for any $r$ with $d$ fixed, is $\tilde{y}_k = c^{d,r} \Lambda (z^k) \in I, 1 \leq k \leq L$, where $c^{d,r}$ is obtained from Step 4 of the identification algorithm described in Section 2.

We can state the following sufficient condition.