13th Triennial World Congress, San Francisco, USA

A SLIDING MODE STRATEGY FOR ADAPTIVE LEARNING IN ADALINES

Hebertt Sira-Ramírez* Eliezer Colina-Morles*

* *Departamento Sistemas de Control, Universidad de Los Andes, Mérida 5101, Venezuela.*

Abstract.
A dynamical sliding mode control approach is proposed for robust adaptive learning in analog Adaptive Linear Elements (Adalines), constituting basic building blocks for perceptron-based feedforward neural networks. The zero level set of the learning error variable is regarded as a sliding surface in the space of learning parameters. A sliding mode trajectory can then be induced, in finite time, on such a desired sliding manifold. Neuron weights adaptation trajectories are shown to be of continuous nature, thus avoiding bang–bang weight adaptation procedures.

Keywords. Variable Structure Systems, Sliding Regimes, Adaptive Learning

1. INTRODUCTION

The adjustment of learning parameters in perceptron based feedforward neural networks has been mainly explored form a discrete–time viewpoint. The celebrated Widrow-Hoff *Delta Rule*,(Widrow *et al.*, 1990) constitutes a least mean square learning error minimization algorithm by which an asymptotically stable linear convergence dynamics is imposed on the underlying discrete–time error dynamics. Using *quasi–sliding mode control* ideas ((Sira-Ramirez, 1991b)) a modification of the Delta Rule was proposed by Sira-Ramírez and Żak in (Sira-Ramirez *et al*, 1991a), and in (Żak *et al.*, 1990), whereby a switching weight adaptation strategy is shown to also impose a discrete–time asymptotically stable linear learning error dynamics. This algorithm is at the basis of recently proposed identification and control schemes, based on feedforward neural networks,(Colina-Morles *et al.*, 1993), and (Kuschewski *et al.*, 1993)). To our knowledge, design of learning strategies in adaptive perceptrons, from the viewpoint of sliding mode control in continuous time, has not been addressed in the existing literature. However, the relevance of ordinary differential equations with discontinuous right hand sides was analyzed in the work of (Li *et al.*, 1989), in the context of Analog Neural Networks of the Hopfield type with infinite gain nonlinearities. In that work, it is established under what circumstances sliding mode trajectories do not appear in such a class of neurons.

In this article a *continuous time* sliding mode control approach is proposed for the robust adaptation of variable weights in Adalines, so that its scalar output variable tracks a bounded reference signal with a bounded first order time derivative. The zero level set of the *learning error* variable is regarded as the sliding surface coordinate function and a discontinuous law of adaptive weight variation is proposed which induces, in finte time, a sliding motion which robustly sustains the zero error condition. The sliding mode controlled weight adaptation trajectories are shown to be *continuous*, rather than bang–bang signals. Section 2 contains some definitions,

* This research was supported by the Consejo de Desarrollo Científico, Humanístico and Tecnológico of the Universidad de Los Andes under Research Grants I-455-94-02-C and I-456-94.

assumptions and derivations of the main charcteristics of a sliding mode control approach to weight adaptation in Adalines. In this section, the robustness of the algorithm, with respect to bounded external perturbation inputs, and bounded measurement noises, is also demonstrated along with a derivation of the required matching condition. Section 3 contains some basic examples of relevant significance in the potential applications of the proposed adaptive learning strategy in automatic control applications. The examples include, both, identification of forward and inverse dynamics of unknown, externally perturbed, nonlinear plants. Section 4 contains the conclusions.

2. A SLIDING MODE CONTROL APPROACH TO WEIGHT ADAPTATION IN ADALINES

2.1 *Definitions and basic assumptions*

Consider the perceptron model depicted in Fig.1 where $x(t) = (x_1(t), \ldots, x_n(t))$ represents a vector of bounded time-varying *inputs*, assumed also to exhibit bounded time derivatives, i.e.

$$\begin{aligned} \| x(t) \| &= \sqrt{x_1^2(t) + \ldots + x_n^2(t)} \;\leq\; V_x \;\; \forall t \\ \| \dot{x}(t) \| &= \sqrt{\dot{x}_1^2(t) + \ldots + \dot{x}_n^2(t)} \;\leq\; V_{\dot{x}} \;\; \forall t \end{aligned} \quad (1)$$

where V_x and $V_{\dot{x}}$ are known positive constants. We denote by $\tilde{x}(t)$ the vector of *augmented inputs*, which includes a constant input of value $B > 1$, affecting the *bias*, or *threshold* weight w_{n+1} in the perceptron model, i.e

$$\tilde{x}(t) = \mathrm{col}(x_1(t), \ldots, x_n(t), B) = \mathrm{col}(x(t), B) \quad (2)$$

The scalar signal $y_d(t)$ represents the time-varying *desired output* of the perceptron. It will be assumed that $y_d(t)$ and $\dot{y}_d(t)$ are also bounded signals, i.e.

$$| y_d(t) | \leq V_y \;\; \forall t \qquad | \dot{y}_d(t) | \leq V_{\dot{y}} \;\; \forall t \quad (3)$$

The output signal $y(t)$ is a scalar quantity defined as:

$$y(t) = \sum_{i=1}^{n} \omega_i(t) x_i(t) + \omega_{n+1}(t) = \tilde{\omega}^T(t)\tilde{x}(t) \quad (4)$$

We define the *learning error* $e(t)$ as the scalar quantity obtained from

$$e(t) = y(t) - y_d(t) \quad (5)$$

2.2 *Problem formulation and main results*

Using the theory of *Sliding Mode Control* of *Variable Structure Systems* (see (Utkin, 1992)) we propose to consider the zero value of the learning error coordinate $e(t)$ as a time–varying *sliding surface*, i.e.

$$s(e(t)) = e(t) = 0 \quad (6)$$

Condition (6) guarantees that the perceptron output $y(t)$ coincides with the desired output signal $y_d(t)$ for all time $t > t_h$ where t_h is addressed as the *hitting time*.

Basic Problem Formulation

It is desired to devise a dynamical feedback *adaptation mechanism*, or *adaptation law*, for the augmented vector of variable weights $\tilde{\omega}(t)$ such that the sliding mode condition is enforced.

2.2.1. *Zero adaptive learning error in finite time*

Let " sign $e(t)$ " stand for the *signum* function. We then have the following result

Theorem 2.1 *If the adaptation law for the augmented weight vector $\tilde{\omega}(t)$ is chosen as*

$$\dot{\tilde{\omega}}(t) = -\left(\frac{\tilde{x}(t)}{\tilde{x}^T(t)\tilde{x}(t)}\right) k \text{ sign } e(t) \quad (7)$$

with k being a sufficiently large positive design constant satisfying

$$k > \tilde{W} V_{\dot{x}} + V_{\dot{y}} \quad (8)$$

then, given an arbitrary initial condition $e(0)$, the learning error $e(t)$ converges to zero in finite time, t_h, estimated by

$$t_h \leq \frac{| e(0) |}{k - \tilde{W} V_{\dot{x}} - V_{\dot{y}}} \quad (9)$$

and a sliding motion is sustained on $e = 0$ for all $t > t_h$.

PROOF. The proof follows easily by using a quadratic Lyapunov function candidate. Details of the proof are presented in a paper to be published very soon, (Sira-Ramirez *et al.*, 1995)

Let $p(t) = \frac{\tilde{x}(t)}{\tilde{x}^T(t)\tilde{x}(t)}$ and consider the following theorem.

Theorem 2.2 *If the adaptation law for the augmented weight vector $\tilde{\omega}(t)$ is chosen as*

$$\dot{\tilde{\omega}}(t) = -p(t)\dot{\tilde{x}}^T(t)\tilde{\omega}(t) - p(t)k\text{sign}e(t) \quad (10)$$

with k being a positive design constant satisfying $k > V_{\dot{y}}$, then, given an arbitrary initial condition $e(0)$, the learning error $e(t)$ converges to zero in finite time t_h satisfying

$$t_h \leq \frac{|\,e(0)\,|}{k - V_{\dot{y}}} \tag{11}$$

and a sliding motion is sustained on $e = 0$ for all $t > t_h$.

PROOF. The proof proceeds along similar lines of that of theorem 2.3, (Sira-Ramirez *et al.*, 1995).

2.2.2. *Average features of the adaptation mechanisms*

We proceed, as it is customary in sliding mode control theory, to investigate the *average* behaviour of the involved controlled variables. Such an analysis involves the consideration of the following *invariance conditions*,

$$e(t) = 0\ ;\ \dot{e}(t) = 0 \tag{12}$$

which are ideally satisfied after the sliding motion starts on the sliding surface and is indefinitely sustained thereon. Consideration of such invariance conditions naturally leads to propose the substitution of the discontinuous (bang-bang) input signals by a smooth input signal, known as the *equivalent control input*. This method has been rigorously validated in (Utkin, 1992) as the *Method of the Equivalent Control*. Consider the adaptation law (7) and the associated error equation and substitute the discontinuous signal k sign $e(t)$ by its smooth equivalent value $v_{eq}(t)$.

$$\dot{e}(t) = -v_{eq}(t) + \tilde{\omega}^T(t)\dot{\tilde{x}}(t) - \dot{y}_d(t) \tag{13}$$

The second condition in (12) implies that

$$v_{eq}(t) = \tilde{\omega}^T(t)\dot{\tilde{x}}(t) - \dot{y}_d(t) \ \ \forall\, t > t_h \tag{14}$$

Upon use of (14), a *virtual*, or *equivalent variable weight adaptation law* can also be associated with the actual discontinuous (bang-bang) policy described by (7). We denote such an *equivalent* adaptive weight vector by $\tilde{\omega}_{eq}(t)$. One obtains, for all $t > t_h$,

$$\dot{\tilde{\omega}}_{eq}(t) = -p(t)\dot{\tilde{x}}^T(t)\tilde{\omega}_{eq}(t) + p(t)\dot{y}_d(t) \tag{15}$$

i.e., the average variable weight vector trajectory satisfies a *linear* time–varying vector differential equation with forcing function represented by the bounded function $\dot{y}_d(t)$. Note that $\tilde{\omega}_{eq}(t)$ itself does not, necessarily, lie in the *range* of $\tilde{x}(t)$. The obtained expression (15) describes the projection, along the range of the vector of augmented inputs $\tilde{x}(t)$, of the derivative of the average regulated evolution of $\tilde{\omega}(t)$.

2.2.3. *Requirements for the stability of the average controlled weights dynamics*

Proposition 1 *Suppose the system $\dot{\tilde{\omega}}_{eq}(t) = F(t)\tilde{\omega}_{eq}(t)$ is uniformly stable and let $\dot{y}_d(t)$ be absoultely integrable. Then, the solutions of (15) are bounded*

PROOF. See (Sira-Ramirez *et al.*, 1995).

Proposition 2 *The matrix $F(t)$ is bounded if $\dot{\tilde{x}}(t)$ is bounded*

PROOF. See (Sira-Ramirez *et al.*, 1995).

It is well known (Brockett, 1970) that if the matrix $F(t)$ is bounded, then exponential stability is equivalent to the uniform integrability, over arbitrary interval of times, of the norm of the corresponding transition matrix. The following theorem is proved in (Brockett, 1970). The next result touches upon a special form of the well knwon condition of *persitency of excitation*, of common occurrence in linear and nonlinear adaptive control schemes (see Sastry and Bodson (Sastry *et al.*, 1989)). Let $l(t)$ be defined as

$$l(t) = \frac{\dot{\tilde{x}}(t)\dot{\tilde{x}}^T(t)}{(\dot{\tilde{x}}^T(t)\dot{\tilde{x}}(t))^2} \tag{16}$$

Theorem 2.3 *Let $\dot{\tilde{x}}(t)$ be bounded on $(-\infty, +\infty)$, moreover, assume that the following form of the persistency of excitation condition holds uniformly in t : There exists positive constants δ and ϵ, such that the following matrix condition is satisfied*

$$\int_t^{t+\delta} \Phi(t,\sigma)\ l(\sigma)\Phi^T(t,\sigma)\mathrm{d}\sigma \geq \epsilon I \ \forall t > t_0 \tag{17}$$

Then, the equivalent adaptation law (15) uniformly yields a bounded trajectory for the vector of weights $\tilde{\omega}_{eq}(t)$, for every bounded signal $\dot{y}_d(t)$, if, and only if, the autonomous system $\dot{\tilde{\omega}}_{eq}(t) = F(t)\tilde{\omega}_{eq}(t)$ is exponentially stable.

PROOF. See (Sira-Ramirez *et al.*, 1995)

Condition (17) admits the following scalar form ((Sastry etal., 1989))

$$\int_t^{t+\delta} z^T \Phi(t,\sigma)\, l(\sigma)\Phi^T(t,\sigma)\, z\, \mathrm{d}\sigma \geq \epsilon \;\; \forall t > t_0. \; ; \; \| z \| = 1 \tag{18}$$

which is a condition on the *energy*, averaged over all directions of a unit sphere, of the nonsingularly transformed input vector, $\tilde{\chi}(\tau) = \Phi(\tau,t)\ \tilde{x}(t)/(\tilde{x}^T(t)\tilde{x}(t))$. This means that the vector function $\tilde{\chi}(\tau)$ is quite an "active" time–varying vector, so that the integral of the matrix $\tilde{\chi}(t)\tilde{\chi}^T(t)$ is uniformly positive definite over any interval of finite length δ.

2.3 *Robustness with respect to external perturbations*

2.3.1. *Inputs with bounded additive noise*

Consider a vector–valued norm–bounded *external perturbation input*, denoted by $\xi(t) = (\xi_1(t), \ldots, \xi_n(t))$, which additively affects the values of the input vector $x(t)$ to the perceptron. It is assumed that the perturbation input $\xi(t)$ is not "larger" than the input $x(t)$, i.e.

$$\| \xi(t) \| = \sqrt{\xi_1^2(t) + \ldots + \xi_n^2(t)} \;\leq\; V_\xi < V_x \;\forall t \tag{19}$$

The time derivatives of the components of $\xi(t)$ are assumed to be also bounded

$$\| \dot{\xi}(t) \| = \sqrt{\dot{\xi}_1^2(t) + \ldots + \dot{\xi}_n^2(t)} \;\leq\; V_{\dot{\xi}} \;\; \forall t \tag{20}$$

We define the *augmented external perturbation input* vector as

$$\tilde{\xi}(t) = (\xi_1(t), \ldots, \xi_n(t), 0) \tag{21}$$

This means that it is implicitly assumed that the constant input B to the *bias* weight $\omega_{n+1}(t)$ is a fixed value which does not contain the influence of perturbation signals. The *perturbed learning error* $\hat{e}(t) = y(t) - y_d(t)$ is now given by

$$\hat{e}(t) = [\tilde{x}(t) + \tilde{\xi}(t)]^T \tilde{\omega}(t) - y_d(t) \tag{22}$$

Note that, in spite of the fact that the perturbed input signal $\tilde{x}(t) + \tilde{\xi}(t)$ is actually *available for measurement* , its time derivative $\dot{\tilde{x}}(t) + \dot{\tilde{\xi}}(t)$ *is not*. This means that such time derivatives can not be used in the weight adaptation law. Hence, only an adaptation law of the type proposed in (7) can be actually devised for sliding mode creation on the zero learning error hyperplane. Let $h(t)$ be defined by

$$h(t) = \frac{\tilde{x}(t) + \tilde{\xi}(t)}{[\tilde{x}(t) + \tilde{\xi}(t)]^T [\tilde{x}(t) + \tilde{\xi}(t)]} \tag{23}$$

By virtue of the above considerations, we shall center our attention on the perturbed adaptation law:

$$\dot{\tilde{\omega}}(t) = -h(t)\ k\ \mathrm{sign}\hat{e}(t) \tag{24}$$

The weight adaptation law (24) results, as it easily verified, in the following discontinuous perturbed learning error dynamics

$$\dot{\hat{e}}(t) = -k \text{ sign } \hat{e}(t) + \tilde{\omega}^T(t)[\dot{\tilde{x}}(t) + \dot{\tilde{\xi}}(t)] - \dot{y}_d(t) \tag{25}$$

3. IDENTIFICATION OF FORWARD AND INVERSE DYNAMICS FOR THE KAPITSA PENDULUM

Here we consider a truly nonlinear system of the *non–flat* type, studied by Fliess and coworkers in (Fliess *et al.*, 1993), consisting of a unit mass rod with a suspension point which freely moves only on a vertical direction. The Kapitsa pendulum is, thus, an inverted pendulum where the control actions are constrained to move the suspension point only along a vertical axis. We considered a nonstabilizing open loop control $u(t)$, applied to the plant, and obtained the corresponding output $y_p(t)$ of the nonlinear system, represented by the angular position of the rod with respect to the vertical axis. In the forward dynamics identification problem $y_p(t)$ is regarded as the desired signal, $y_d(t)$, to be followed by the neuron output $y(t)$. In that case, the input function $u(t)$ to the system, is also the input to a stable tranversal filter SF. For the inverse dynamics identification the roles of $u(t)$ and $y_p(t)$ were reversed, with respect to the neuron system. The open loop control function $u(t)$ was chosen, according to ((Fliess *et al.*, 1993)), of the form

$$u(t) = A_1 + A_2 \cos(\frac{t}{2\,\epsilon}) + A_3 \sin(\frac{t}{\epsilon}) \tag{26}$$

where A_1, A_2 and A_3 are constant parameters. The nonlinear system is assumed to be unknown and only its input and output signals are assumed to be measurable for the adaptation process. For simulation purposes, however, the following model was used

$$\begin{aligned}
\dot{\alpha}(t) &= p(t) + \frac{u(t)}{l} \sin\, \alpha(t) \\
\dot{p}(t) &= \left(\frac{g}{l} - \frac{u^2(t)}{l^2} \cos\alpha(t)\right) \sin\alpha(t) - \frac{u(t)}{l}\, p(t)\, \cos\alpha(t) \\
\dot{z}(t) &= u(t) \\
y_p(t) &= \alpha(t)
\end{aligned} \tag{27}$$

where $\alpha(t)$ is the angle of the rod with the vertical axes, $p(t)$ is proportional to the generalized impulsion. The constants g and l represent, respectively, the gravity acceleration and the length of the rod. The velocity of the suspension point acts as the control variable $u(t)$. The variable $z(t)$ is then the vertical position of the suspension point. Numerical values for the parameters of the Kapitsa pendulum model were set to be $g = 9.81[\frac{\text{m}}{\text{sec}^2}]$ and $l = 0.7$[mts]. An open loop control input signal $u(t)$ of the form given in (26) with the following constant parameters

$$A_1 = 0.4 \ ; \ A_2 = 2 \ ; \ A_3 = 3 \ ; \ \epsilon = 0.05$$

was used, for both tasks. The input variables to the neuron are obtained from the state of a transversal stable filter, *SF*, consisting of a string of integrators. The scalar input function $u(t)$ to the stable filter represents the physically available signal to be processed by the neuron -(usually a plant input or output)-. The *SF* module was designed as a stable low pass filter with the following state representation

$$\begin{aligned} \dot{x}_1(t) &= x_2(t) \\ \dot{x}_2(t) &= x_3(t) \\ \dot{x}_3(t) &= -x_1(t) - 3x_2(t) - 3x_3(t) + u(t) \\ \dot{x}_4(t) &= 0 \end{aligned} \tag{28}$$

where the state variable $x_4(t)$ represents the *bias* component with initial condition equals to B. Such a constant parameter is taken, for this example, as $B = 1$. The results of a simulated forward dynamics identification taks, for an Adaline with a total of 4 weights (including one bias variable weight) are shown in Fig.2. In this figure, the desired output trajectory $y_d(t)$ is constituted by the nonlinear pendulum system output i.e. $y_p(t) = \alpha(t) = y_d(t)$ and the input $u(t)$ to the *SF*–module is the same input given to the nonlinear system. The learning (tracking) error response $e(t)$ is shown to converge to zero in approximately 0.02 sec. To alleviate the "chattering" phenomena, present in the neuron output and learning errror responses, as well as to speed up the simulation time for the SIMNON package, the following standard substitution was adopted for the ideal switch function

$$k \text{ sign } e(t) \approx k \frac{e(t)}{|e(t)| + \delta} \tag{29}$$

with $\delta = 0.05$. Highly accurate following is seen to be achieved without chattering around the desired output signal. The open loop unperturbed input signal trajectory $u(t)$, afecting both the pendulum and the *SF*–neuron arrangement, is also shown in this figure. In the simulation no additive noise affecting the input signal $u(t)$ was assumed. The value used for the variable structure gain k was set to $k = 5$. The computer simulation results shown in Fig.3 illustrate the performance of the neuron when the input and desired signals are subjected to noise. The inverse dynamics identification task was also implemented using the same *SF*–module described above. The variable structure control gain used in this case was $k = 90$. The simulation results, without additive noise for the measured output signal $y_p(t)$ of the nonlinear system, are presented, for a 4 weights Adaline, in Fig.4. Fig.5 presents the corresponding results for an additive noise input signal, of the same characteristics as before, affecting the measured signal $y_p(t)$ given as an input to the filter–neuron combination. In this case the value of k was substantially icreased to $k = 1000$ due to the large values of the first order time derivative of the desired output signal $y_d(t)$, represented now by the noisy signal $u(t) + \xi(t)$, with $u(t)$ as given in (26).

4. CONCLUSIONS

In this article a new dynamical discontinous feedback adaptative learning algorithm has been proposed, for linear adaptive combiners, which robustly drives the learning error to zero in finite time. The components of the vector of variable weights are assumed to be provided with continuous time adaptation possibilities. The dynamical adaptive learning scheme is based on sliding mode control ideas and it represents a simple, yet robust, mechanism for guaranteeing finite time reachability of a zero learning error condition. The approach is also highly insensitive to bounded external perturbation inputs, measurement noises and designed input filter parameters. Bounded average weight evolution is guaranteed under several conditions relative to the underlying linear time–varying system describing the average evoultion of the vector of adaptive weights. Some of these conditions are closely related to those of persistency of excitation and thus links our approach with standard adaptive control results. Chattering–free dynamical sliding mode controllers for nonlinear systems have been recently proposed by Sira-Ramírez ((Sira-Ramirez, 1992)) using input-dependent sliding surfaces. The adaline case study presented here represents an instance in which the sliding surface (zero learning error condition) is actually an "input" dependent manifold. The obtained sliding "controller" is thus continuous rather than bang–bang.

Extensions of the results to more general classes of multilayer neuron arrangements is being pursued at the present time, with encouraging results.

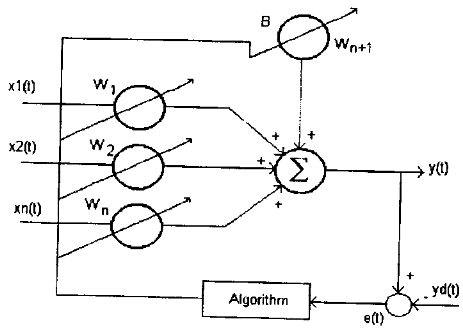

Fig.1 : Perceptron Model

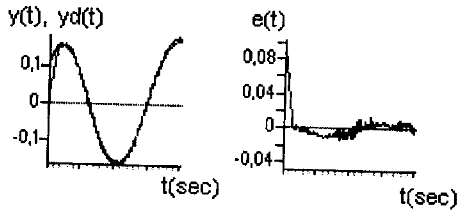

Fig.2 : Noise-free forward dynamics Identification

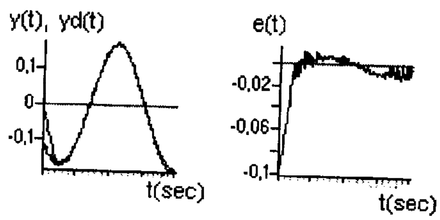

Fig. 3 : Robust Forward dynamics Identification

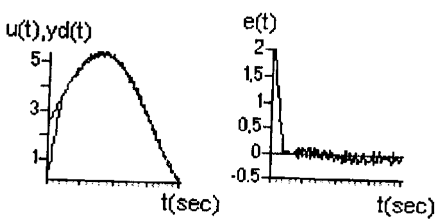

Fig. 4 : Noise-free Inverse dynamics Identification

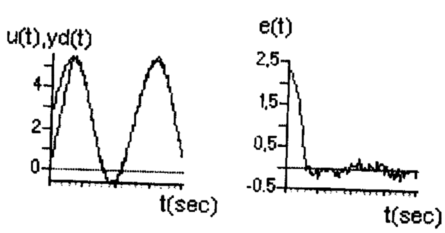

Fig. 5 : Robust Inverse dynamics Identification

5. REFERENCES

Brockett R.W. (1970) *Finite Dimensional Linear Systems* Academic Press, New York.

Colina-Morles E and Mort N (1993) "Neural network-based adaptive control design," *Journal of Systems Engineering* Vol. 1, No. 1, pp. 9-14.

Fliess M., Lévine J., Martin P. and Rouchon P. (1993) "Défaut d'un système non linéaire et commande haute fréquence," *C. R. Acad. Sci. Paris*, Vol. I-316.

Kuschewski J. G., Hui S. and Żak S. H. (1993) "Application of feedforward networks to dynamical system identification and control," *IEEE Transactions on Control Systems Technology*, Vol. 1, No. 1, pp. 37-49.

Li J. H., Michel A. N. and Porod W. (1989) "Analysis and synthesis of a class of neural networks : Variable structure systems with infinite gain," *IEEE Transactions on Circuits and Systmes*, Vol. 36, No. 5, pp. 713–731.

Sastry S. and Bodson M. (1989) *Adaptive Control : Stability, Convergence, and Robustness*, Prentice-Hall, Englewood Cliffs, N.J.

Sira-Ramírez H. and Żak S. H. (1991a) "The adaptation of perceptrons with applications to inverse dynamics identification of unknown dynamic systems," *IEEE Transactions on Systems, Man and Cybernetics*, Vol. 21, No. 3, pp. 634–643.

Sira-Ramírez H. (1991b) "Nonlinear Discrete Variable Structure Systems in Quasi-Sliding Mode, *International Journal of Control*, Vol. 54, No. 5, pp.1171-1187.

Sira-Ramírez H. (1992) "On the Sliding Mode Control of Nonlinear Systems," *Systems and Control Letters*, Vol. 19, pp. 303–312.

Sira-Ramirez H. and Colina-Morles E. (1995) "A sliding mode strategy for adaptive learning in adalines," *Submitted to IEEE Transactions on Circuits and Systems*

Utkin V. I. (1992) *Sliding Modes in Control Optimization* Springer–Verlag, New York.

Widrow B. and Lehr M. A. (1990) "30 years of adaptive neural networks: Perceptron, madaline, and backpropagation," *Proceedings IEEE*, Vol. 78, pp. 1415-1442.

Żak S. H. and Sira-Ramírez H. (1990) "On the adaptation algorithms for feedforward neural networks," in *Theoretical Aspects of Industrial Design*, D.A. Flield and V. Komkov, Eds. Society of Industrial and Applied Mathematics (SIAM), New York, pp. 116-133.