Proceedings of the 2000 IEEE
International Conference on Robotics & Automation
San Francisco, CA • April 2000

Stabilizing and Robustifying the Error Backpropagation Method in Neurocontrol Applications

M. Onder Efe
Electrical and Electronic Eng. Dept.
Bogazici University
Bebek, 80815, Istanbul, Turkey
efemond@boun.edu.tr

Okyay Kaynak
Electrical and Electronic Eng. Dept.
Bogazici University
Bebek, 80815, Istanbul, Turkey
kaynak@boun.edu.tr

Abstract

This paper discusses the stabilizability of artificial neural networks trained by utilizing the gradient information. The method proposed constructs a dynamic model of the conventional update mechanism and derives the stabilizing values of the learning rate. This is achieved by integrating the Error Backpropagation (EBP) technique with Variable Structure Systems (VSS) methodology, which is well known with its robustness to environmental disturbances. In the simulations, control of a three degrees of freedom anthropoid robot is chosen for the evaluation of the performance. For this purpose, a feedforward neural network structure is utilized as the controller.

1 Introduction

One of the major problems in the training of Artificial Neural Networks (ANNs) is the lack of stabilizing forces, the existence of which prevents the unbounded growth in the adjustable parameters. Another major problem of the training phase is the robustness, i. e. how well the ANN structure, which is trained on-line or off-line, will perform under the existence of strong external disturbances. The methods overcoming the mentioned problems aim to maintain a desired behavior in the face of factors influencing the performance and applicability adversely. It is therefore the concept of stability and robustness constitutes a prime requirement of the design methodologies particularly in systems and control engineering practice. Strictly speaking, a method violating the stability requirements is a potential danger from the safety point of view. However, the innovations in data mining, data fusion, sensor technology, recognition technology and fast microprocessors are ever increasingly encouraging the use of ANN structures, whose operating philosophy is suitable to interaction of an *Expert* and *Machine*.

In most of the applications [1-3], ANN structure is trained by EBP method, whose performance is not found satisfactory in the face of strong external disturbances. In order to robustify the algorithm with the idea of injecting stabilizing forces into the training dynamics, one can come up with a strategy, which prevents unbounded growth in parameters without losing tracking precision. A suitable way of tackling with the uncertainties and observation noise is to introduce VSS based components into the training methodology.

Numerous contributions to VSS theory have been made during the last decade; some of them are as follows. Hung [4] has reviewed the control strategy for linear and nonlinear systems. In [4], the switching schemes, putting the differential equations into canonical forms and generating simple VSS strategies are considered in detail. In [5] and [6], applications of Sliding Mode Control (SMC) scheme to robotic manipulators are studied and the quality of the scheme is discussed from the point of robustness. Another systematic examination of SMC approach is presented in [7]. In this reference, the practical aspects of SMC design are assessed for both continuous time and discrete time cases and a special consideration is given to the finite switching frequency, limited bandwidth actuators and parasitic dynamics. In [8], the design of discrete time SMC is presented with particular emphasis on the system model uncertainties.

This paper is organized as follows. The second section briefly reviews the standard EBP technique, which is responsible for achieving the desired performance specifications. The parameter stabilizing part of the training methodology is derived in the third section. The fourth section gives the global stability proof of the mixed training strategy and discusses the constraints on the design parameters. In the fifth section, the feedforward neural network structure with EBP scheme is introduced and the application of the devised training strategy is presented. The sixth section introduces a plant to be controlled. Simulation results are discussed in the seventh section and the conclusions are presented at the end of the paper.

2 Standard Error Backpropagation Technique

In most applications of ANNs, EBP method constitutes the central part of the learning. In this section, the technique is briefly reviewed for systems in which the outputs are differentiable with respect to the parameter of interest. The method has first been formulated by Rumelhart [9] in 1986. The approach has successfully been applied to a wide variety of optimization problems. The algorithm can be stated as follows.

$$e_j = d_j - f_j(\phi, u) \tag{1}$$

$$J_r = \frac{1}{2}\sum_{j=1}^{outputs} e_j^2 \tag{2}$$

$$\Delta\phi = -\eta_\phi \frac{\partial J_r}{\partial \phi} \tag{3}$$

The observation error in (1) is used to minimize the realization cost in (2) by utilizing the rule described by (3), which is known as gradient descent or EBP in the related literature.

$$\Delta\phi = \eta_\phi \sum_{j=1}^{outputs} e_j \frac{\partial f_j(\phi,u)}{\partial\phi} = \eta_\phi N_\phi \tag{4}$$

The minimization proceeds recursively as given in (4), for which the sensitivity derivative with respect to the generic parameter ϕ is needed. Since the update rule in (4) entails the observation error e, the algorithm is quite sensitive to the noisy observations, which directly influence the value of the adjustable parameter and degrade the learning performance.

3 Stabilization of Training Dynamics by VSS Approach

A continuous-time dynamic model of the parameter update rule prescribed by the EBP technique can be written as in (5).

$$\dot{\Delta\phi} = -\frac{1}{T_s}\Delta\phi + \frac{\eta_\phi}{T_s} N_\phi \tag{5}$$

The above model is composed of the sampling time denoted by T_s, the gradient based non-scaled parameter change denoted by N_ϕ and a scaling factor denoted by η_ϕ, for the selection of which, a detailed analysis is presented in the subsequent discussion. Using Euler's first order approximation for the derivative term, one can easily see the relation between (4) and (5). The synthesis of the parameter stabilizing component is based on the integration of the system in (5) with VSS methodology. In the design of VSS based controllers, one method that can be followed is the reaching law approach [4]. For the use of this theory in the stabilization of the training dynamics, let us define the switching function as in (6) and its dynamics as in (7).

$$s_\phi = \Delta\phi \tag{6}$$

$$\dot{s}_\phi = -\frac{Q_\phi}{T_s}\tanh\left(\frac{s_\phi}{\varepsilon}\right) - \frac{K_\phi}{T_s}s_\phi \tag{7}$$

where, Q_ϕ and K_ϕ are the gains, and ε is the width of the boundary layer. In the derivations presented below, a key point is the fact that the system described by (5) is driven both by the learning rate η_ϕ and by the backpropagated error value N_ϕ. Now we demonstrate that some special selection of this quantity leads to a rule that minimizes the magnitude of parametric displacement. With the quantity defined in (8), equating (7) and (5) and solving for $\Delta\phi$ yields the relation in (9).

$$A_\phi = Q_\phi \tanh\left(\frac{\Delta\phi}{\varepsilon}\right) + K_\phi \Delta\phi \tag{8}$$

$$\Delta\phi = \eta_\phi N_\phi + A_\phi \tag{9}$$

The values of the η_ϕ imposed by (9) might be seen as the desired values at the first glance. However, this selection cancels out the backpropagated error value N_ϕ from (5), consequently the update dynamics exactly behaves as that defined by the adopted switching function (7), which does not necessarily minimize the cost in (2). Therefore the further analysis explores the restrictions on η_ϕ as well as the construction of the mixed training criterion.

Now we have a model described by (5), and an equality to be enforced and formulated by (9). If one chooses a positive definite Lyapunov function as given in (10), the time derivative of this function must be negative definite for stability of parameter change ($\Delta\phi$) dynamics. Clearly the stability in parameter change space implies the convergence in system parameters.

$$V_\phi = \frac{1}{2}s_\phi^2 = \frac{1}{2}(\Delta\phi)^2 \tag{10}$$

$$\dot{V}_\phi = (\Delta\phi)(\dot{\Delta\phi}) \tag{11}$$

If (5) and (9) are substituted into (11), the constraint stated in (12) is obtained for stability in the Lyapunov sense.

$$\left(\eta_\phi + \frac{1}{N_\phi}A_\phi\right)\left(\eta_\phi - \frac{1}{N_\phi}\Delta\phi\right) < 0 \tag{12}$$

Since A_ϕ and $\Delta\phi$ have the same signs, the roots of the expression in (12) clearly have opposite signs. The expression on the left-hand side assumes negative values between the roots. Therefore, in order to satisfy the inequality in (12), the learning rate must satisfy the constraint given in (13). In order to preserve the compatibility between the traditional gradient based approaches and the proposed approach; the interval of learning rate is restricted to positive values as described below. An appropriate selection of η_ϕ can be as in (14).

$$0 < \eta_\phi < \min\left\{\left|\frac{1}{N_\phi}\Delta\phi\right|, \left|-\frac{1}{N_\phi}A_\phi\right|\right\} \tag{13}$$

$$\eta_\phi = \beta \min\left\{\left|\frac{1}{N_\phi}\Delta\phi\right|, \left|-\frac{1}{N_\phi}A_\phi\right|\right\}, 0 < \beta < 1 \tag{14}$$

By substituting η_ϕ in (14) into (9), the stabilizing component $\Delta\phi_{VSS}$ is obtained as;

$$\Delta\phi_{VSS} = \beta \min\left\{|\Delta\phi|, |A_\phi|\right\} \operatorname{sgn}(N_\phi) + A_\phi \tag{15}$$

where, $\Delta\phi$ on the right hand side is the final update value yet to be obtained. The law introduced in (15) stabilizes the system in (5). The question now reduces to the following; can this law minimize the cost defined by (2)? The answer is obviously not, because the stabilizing component in (15) is derived from the displacement of the parameter vector denoted by $\Delta\phi$, whereas the minimization of (2) is achieved when ϕ tends to ϕ^* regardless of what the displacement is. In order to minimize (2), the parameter change anticipated by EBP technique given in (16) should somehow be integrated into the final form of parameter update mechanism.

$$\Delta\phi_{EBP} = \zeta_\phi N_\phi \tag{16}$$

where, ζ_ϕ is the learning rate. It is reasonable to expect that under certain constraints, a combination of the laws formulated in (15) and (16) in a weighted average would minimize both (10) and (2), which means the fulfillment of the design specifications. The parameter update rule will then be as in (17).

$$\Delta\phi = \frac{\alpha_1 \Delta\phi_{VSS} + \alpha_2 \Delta\phi_{EBP}}{\alpha_1 + \alpha_2} \tag{17}$$

The parameter update formula given in (17) carries mixed displacement value containing both the parametric convergence, which is introduced by VSS part, and the cost minimization, which is due to the EBP technique. The balancing in this mixture is left to the designer by an appropriate selection of α_1 and α_2.

4 Extracting the Conditions for the Global Stability of the Mixed Training Dynamics

In this section, the global stability of the mixed training strategy is analyzed. For this purpose, a Lyapunov function given in (18) is defined. In (18), γ_ϕ is a positive constant and its properties are discussed at the end of the section.

$$V_\phi = \frac{1}{2}(\Delta\phi)^2 + \frac{\gamma_\phi}{2}(N_\phi)^2 \tag{18}$$

The time derivative of the Lyapunov function is as given in (19). Since the analysis in this section concerns the stability of the mixed training strategy, the combined form of the learning algorithm, as given below, should be used in the formulation.

$$\dot{V}_\phi = \left(-\frac{1}{T_s}\Delta\phi + \frac{\alpha_1 \Delta\phi_{VSS} + \alpha_2 \Delta\phi_{EBP}}{(\alpha_1 + \alpha_2)T_s}\right)\Delta\phi + \gamma_\phi \dot{N}_\phi N_\phi \tag{19}$$

If the $\Delta\phi_{VSS}$ of (15) and $\Delta\phi_{EBP}$ of (16) are substituted into (19), one obtains the following relation, which can assume two different forms due to the minimum operator.

$$\begin{aligned}\dot{V}_\phi = & -\frac{1}{T_s}\Delta\phi^2 + \frac{\alpha_1 \beta}{(\alpha_1 + \alpha_2)T_s} \min\{|\Delta\phi|, |A_\phi|\} \operatorname{sgn}(N_\phi)\Delta\phi + \\ & \frac{\alpha_1}{(\alpha_1 + \alpha_2)T_s} A_\phi \Delta\phi + \frac{\alpha_2}{(\alpha_1 + \alpha_2)T_s} \zeta_\phi N_\phi \Delta\phi + \gamma_\phi \dot{N}_\phi N_\phi\end{aligned} \tag{20}$$

Case #1: $|\Delta\phi| < |A_\phi|$

$$\begin{aligned}\dot{V}_\phi < & \left(\frac{(\beta + Q_\phi + K_\phi)\alpha_1}{(\alpha_1 + \alpha_2)T_s} - \frac{1}{T_s}\right)\Delta\phi^2 + \\ & \frac{\alpha_2}{(\alpha_1 + \alpha_2)T_s} \zeta_\phi N_\phi \Delta\phi + \gamma_\phi \dot{N}_\phi N_\phi\end{aligned} \tag{21}$$

Case #2: $|A_\phi| < |\Delta\phi|$

$$\begin{aligned}\dot{V}_\phi < & \left(\frac{2\alpha_1(\beta + Q_\phi + K_\phi)}{(\alpha_1 + \alpha_2)T_s} - \frac{1}{T_s}\right)\Delta\phi^2 + \\ & \frac{\alpha_2}{(\alpha_1 + \alpha_2)T_s} \zeta_\phi N_\phi \Delta\phi + \gamma_\phi \dot{N}_\phi N_\phi\end{aligned} \tag{22}$$

If the negativity of the quantity on the right hand side of the inequality (22) is ensured, the negativity of the quantity in (21) becomes trivial. Therefore the two inequalities can be reduced to one inequality, which is given below. The global stability of the mixed training dynamics will clearly require the negativity of the quantity on the right-hand side of (23).

$$\dot{V}_\phi < -\frac{C_\phi}{T_s}\Delta\phi^2 + \frac{\alpha_2}{(\alpha_1+\alpha_2)T_s}\zeta_\phi N_\phi \Delta\phi + \gamma_\phi \dot{N}_\phi N_\phi \quad (23)$$

where,

$$C_\phi = 1 - \frac{2\alpha_1(\beta + Q_\phi + K_\phi)}{\alpha_1+\alpha_2} \quad (24)$$

Set

$$\gamma_\phi = \frac{\sigma_\phi^2}{\sup_t \left|\dot{N}_\phi N_\phi\right|} \quad (25)$$

where, σ_ϕ^2 is the least nonzero value of $\Delta\phi^2$ observed during a training course. It should be noted here that one may not know the numerical value of this number but there exists such a number in the course of each training trial. With this value of γ_ϕ, (23) becomes as follows.

$$\dot{V}_\phi < -\frac{C_\phi}{T_s}\Delta\phi^2 + \frac{\alpha_2}{(\alpha_1+\alpha_2)T_s}\zeta_\phi N_\phi \Delta\phi + \sigma_\phi^2 = B_\phi \quad (26)$$

Inequality in (26) follows from the inequality given below.

$$\frac{\dot{N}_\phi N_\phi}{\sup_t \left|\dot{N}_\phi N_\phi\right|} < 1 \quad (27)$$

Since $\sigma_\phi^2 \le \Delta\phi^2$ for all $t \ge 0$;

$$B_\phi \le -\frac{C_\phi}{T_s}\Delta\phi^2 + \frac{\alpha_2}{(\alpha_1+\alpha_2)T_s}\zeta_\phi N_\phi \Delta\phi + \Delta\phi^2 \quad (28)$$

$$< -\left(\frac{C_\phi}{T_s} - 1\right)\Delta\phi^2 + \frac{\alpha_2}{(\alpha_1+\alpha_2)T_s}\zeta_\phi \left|N_\phi\right|\left|\Delta\phi\right| \quad (29)$$

In order to ensure the negativeness of the right hand side of (29), the following inequality must be satisfied.

$$\zeta_\phi < \frac{\alpha_1+\alpha_2}{\alpha_2}\frac{(C_\phi - T_s)\left|\Delta\phi\right|}{\left|N_\phi\right|} \quad (30)$$

This selection of the learning rate for EBP part ensures the negative definiteness of the time derivative of the Lyapunov function in (18). It is clear that the parameter γ_ϕ exists, nonzero, nonnegative and finite. These facts justify the particular chosen form of the Lyapunov function, and the analysis proves that the suggested form of the parameter update rule given in (17) leads to the stable training of artificial neural networks.

It is clear that the derivation and the analysis presented impose some conditions on the design parameters. In the rest of this section these conditions are discussed.

1. Due to the requirement on the negative definiteness of the time derivative of the Lyapunov function, the inequality given in (31) must be satisfied.

$$C_\phi = 1 - \frac{2\alpha_1(\beta + Q_\phi + K_\phi)}{\alpha_1+\alpha_2} > 0 \quad (31)$$

The selection for the learning rate ζ_ϕ imposes the following condition.

$$1 - \frac{2\alpha_1(\beta + Q_\phi + K_\phi)}{\alpha_1+\alpha_2} > T_s \quad (32)$$

The inequality in (32) ensures the learning rate γ_ϕ to assume positive values. Since the condition in (32) includes the condition in (31), the constraint in (32) is one of the restrictions on the design parameters.

2. If (14) is substituted into (5), after some straightforward calculations, one obtains the second condition, which is given below and is not extracted in detail due to the space limit.

$$\beta(Q_\phi + K_\phi) < 1 \quad (33)$$

5 Training of Artificial Neural Networks

In this section, application of the devised scheme to feedforward neural networks is presented. In [1-3], it is demonstrated that the structure can effectively be used for identification and control purposes. In the conventional EBP technique, propagating the output error back through the neural network minimizes the cost function given in (2). Based on the derivation presented in detail in [9], the delta values for the neurons belonging to the output layer and the hidden layers are evaluated as given by (34) and (35) respectively.

$$\delta_j^{k+1} = (d_j - f_j)\,\Psi'(S_j^{k+1}) \quad (34)$$

$$\delta_j^{k+1} = \left(\sum_{h=1}^{neurons_{k+2}} \delta_h^{k+2} w_{jh}^{k+1}\right)\Psi'(S_j^{k+1}) \quad (35)$$

where, Ψ is the neuronal nonlinearity and S_j is the net summation. Having evaluated the delta values during the backward pass, EBP rule described by (36) is applied for each training pair.

$$\Delta w_{ij\,EBP}^k = \zeta_{w_{ij}^k}\,\delta_j^{k+1} o_i^k \quad (36)$$

The VSS part of the proposed approach estimates the following update value for parametric stability.

$$\Delta w^k_{ij\,VSS} = \beta \min\left(\left|\Delta w^k_{ij}\right|, \left|A_{w^k_{ij}}\right|\right) \mathrm{sgn}\left(N_{w^k_{ij}}\right) + A_{w^k_{ij}} \qquad (37)$$

The two update laws are then combined as a weighted average as given in (17).

6 Plant Model

In the simulations the dynamic model of a three degrees of freedom anthropoid robotic manipulator is used as the test bed. Since the dynamics of such a mechatronic system is modeled by nonlinear and coupled differential equations, precise output tracking becomes a difficult objective due to the strong interdependency between the variables involved and the existence of gravitational forces. Therefore the methodology adopted must have the capability of coping with the stated difficulties.

The general form of the dynamics of a robotic manipulator is described by (40) where $M(q)$, $C(q,\dot{q})$, $g(q)$ and u stand for the state varying inertia matrix, vector of coriolis and centrifugal terms, gravitational forces and applied torque inputs respectively. The nominal values of the plant parameters are given in [1] in standard units.

$$M(q)\ddot{q} + C(q,\dot{q})\dot{q} + g(q) = u \qquad (38)$$

If the angular positions and angular velocities are described as the state variables of the system, six coupled and first order differential equations can define the model. In (39) through (42), the nonzero entries of the state varying inertia matrix are described. The nonzero Cristoffel symbols are given in (43) through (46). The details of the plant model are presented in [1].

$$\begin{aligned} M_{11} &= m_2 l_{c2}^2 \cos^2(q_2) + \\ &m_3 (l_2 \cos(q_2) + l_{c3}\cos(q_2+q_3))^2 + E_1 + \\ &A_2 \sin^2(q_2) + E_2\cos^2(q_2) + A_3 \sin^2(q_2+q_3) + \\ &E_3\cos^2(q_2+q_3) \end{aligned} \qquad (39)$$

$$\begin{aligned} M_{22} &= m_2 l_{c2}^2 \sin^2(q_2) + m_3\left(l_2^2 + l_{c3}^2 + 2 l_2 l_{c3}\cos(q_3)\right) + \\ &I_2 + I_3 \end{aligned} \qquad (40)$$

$$M_{23} = M_{32} = m_3\left(l_{c3}^2 + l_{c3} l_2 \cos(q_3)\right) + I_3 \qquad (41)$$

$$M_{33} = m_3 l_{c3} + I_3 \qquad (42)$$

$$\begin{aligned} hc_1 &= \left(-m_2 l_{c2}^2 + A_2 - E_2\right)\cos(q_2)\sin(q_2) + \\ &(A_3 - E_3)\cos(q_2+q_3)\sin(q_2+q_3) + \\ &m_3(l_2\cos(q_2) + l_{c3}\cos(q_2+q_3)) * \\ &(-l_2 \sin(q_2) - l_{c3}\sin(q_2+q_3)) \end{aligned} \qquad (43)$$

$$hc_2 = \sin(q_2+q_3)\begin{pmatrix} -m_3 l_{c3} l_2 \cos(q_2) + \\ (-m_3 l_{c3}^2 + A_3 - E_3)\cos(q_2+q_3) \end{pmatrix} \qquad (44)$$

$$hc_3 = m_2 l_{c2}^2 \cos(q_2)\sin(q_2) \qquad (45)$$

$$hc_4 = -m_2 l_2 l_{c3}\sin(q_3) \qquad (46)$$

Coriolis and centrifugal terms are formulated as follows:

$$C(q,\dot{q}) = \begin{bmatrix} 2hc_1\dot{q}_1\dot{q}_2 + 2hc_2\dot{q}_1\dot{q}_3 \\ -hc_1\dot{q}_1^2 + 2hc_4(\dot{q}_2\dot{q}_3 + \dot{q}_3^2) + hc_3\dot{q}_2^2 \\ -hc_2\dot{q}_1^2 - hc_4\dot{q}_2^2 \end{bmatrix} \qquad (47)$$

Lastly, the gravity terms are obtained as given in (48) where G represents the gravity constant.

$$g(q_1,q_2,q_3) = \begin{bmatrix} 0 \\ \begin{pmatrix} (m_2 l_{c2} + m_3 l_2) G\cos(q_2) + \\ m_3 l_{c3} G\cos(q_2+q_3) \end{pmatrix} \\ m_3 l_{c3} G \cos(q_2+q_3) \end{bmatrix} \qquad (48)$$

7 Simulation Results

In the simulations, the plant is controlled by a neural network structure. The architecture of the control system is the ordinary feedback loop, in which the neural controller has one hidden layer being comprised of 12 neurons having hyperbolic tangent type neuronal activation functions. The output layer neurons have linear activation functions. During the simulations, all weights and biases of the neural network have been adjusted. The initial values of the parameters of the neural network have been set such that the initial control surfaces for all three links approximately resemble to that of a Proportional plus Derivative (PD) controller having the following parameter set.

$$\begin{bmatrix} K_{p\,BASE} & K_{d\,BASE} \\ K_{p\,SHOULDER} & K_{d\,SHOULDER} \\ K_{p\,ELBOW} & K_{d\,ELBOW} \end{bmatrix} = \begin{bmatrix} 40 & 40 \\ 180 & 260 \\ 150 & 70 \end{bmatrix} \qquad (49)$$

The reference angular velocity trajectories used in the simulations are given in (50) through (52). The simulations are started with initial rest conditions. Apart from the dynamic complexity of the system under control, a considerable difficulty to be alleviated by the algorithm discussed is the existence of the observation noise. It is assumed that the encoders provide noisy measurements to the controller. The noise sequence is Gaussian distributed and has the same statistical properties for all six state variables, namely, each sequence has zero mean and variance equal to 33e-4. For the use of the proposed algorithm, α_1 is set to 2 while α_2 is equal to 1. Due to the space limit, only the state tracking error graphs are illustrated for the proposed technique and pure EBP training strategy in Figs. 1 and 2 respectively.

$$\dot{q}_{1d} = 1.5\sin(2\pi t/5) \tag{50}$$

$$\dot{q}_{2d} = 1.2\sin(2\pi t/5) \tag{51}$$

$$\dot{q}_{3d} = 1.0\sin(2\pi t/5) \tag{52}$$

8 Conclusions

In this paper, we propose a method for creating stabilizing forces in the training dynamics. The method is based on the integration of EBP strategy with VSS technique to benefit from the robustness property of VSS approach as well as the cost minimizing property of the EBP method. The results stipulate that the proposed approach fulfills the task to be accomplished much better that that can be observed with pure EBP technique. The comparison strongly recommends the use of the algorithm for the applications requiring on-line tuning of the parameters, stability in the parameter change space and insensitivity to environmental disturbances.

9 Acknowledgments

This work is supported by Bogazici University Research Fund (Project no: 99A202)

10 References

[1] Efe, M. O., and O. Kaynak, "A Comparative Study of Neural Network Structures in Identification of Nonlinear Systems," *Mechatronics*, v.9, no.3, pp.287-300, 1999

[2] Narendra, K. S., K. Parthasarathy, "Identification and Control of Dynamical Systems Using Neural Networks," *IEEE Trans. on Neural Networks*, v.1, no. 1, pp.4-27, 1990.

[3] Efe, M. O., E. Abadoglu and O. Kaynak, "Analysis and Design of a Neural Network Assisted Nonlinear Controller for a Bioreactor," to appear in *Int. Journal of Robust and Nonlinear Control*, v.9, no.11, 1999.

[4] Hung, J. Y., W. Gao and J. C. Hung, "Variable Structure Control: A Survey," *IEEE Trans. on Industrial Electronics*, v.40, no. 1, pp.2-22, February 1993.

[5] Gao, W., J. C. Hung, "Variable Structure Control of Nonlinear Systems: A New Approach," *IEEE Trans. on Industrial Electronics*, v.40, no. 1, pp.45-55, February 1993.

[6] Erbatur, K., O. Kaynak and A. Sabanovic, "A Study on Robustness Property of Sliding-Mode Controllers: A Novel Design and Experimental Investigations," *IEEE Trans. on Industrial Electronics*, v.46, no. 5, pp.1-7, 1999.

[7] Young, K.D., V. I. Utkin and U. Ozguner, "A Control Engineer's Guide to Sliding Mode Control," *IEEE Trans. on Control Systems Technology*, v.7, no.3, pp.328-342, May 1999.

[8] Kaynak, O., and A. Denker, "Discrete-Time Sliding Mode Control in the Presence of System Uncertainty," *Int. Journal of Control*, v. 57, no.5, pp.1177-1189, 1993.

[9] Rumelhart, D. E., G. E. Hinton and R. J. Williams, "Learning Internal Representations by Error Propagation," in D. E. Rumelhart and J. L. McClelland, eds., *Parallel Distributed Processing*, v.1, pp.318-362, MIT Press, Cambridge, M.A., 1986.

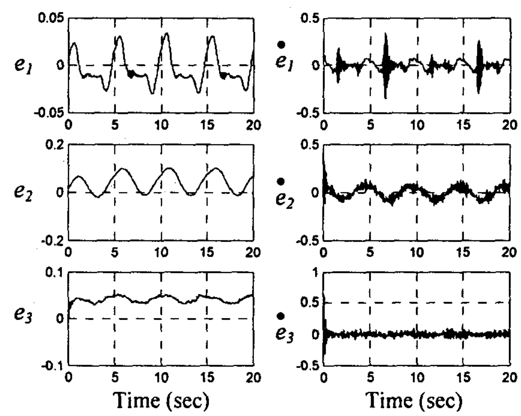

Figure 1. State Tracking Errors Observed with the Proposed Technique

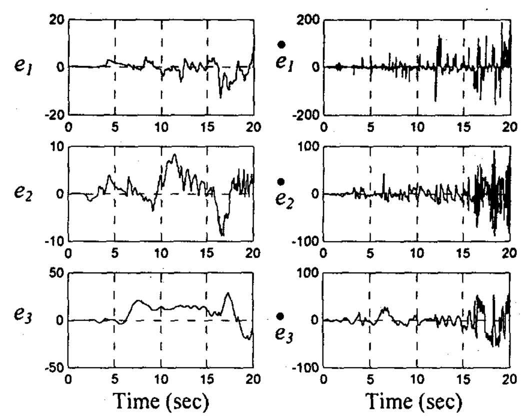

Figure 2. State Tracking Errors Observed with Pure EBP Technique