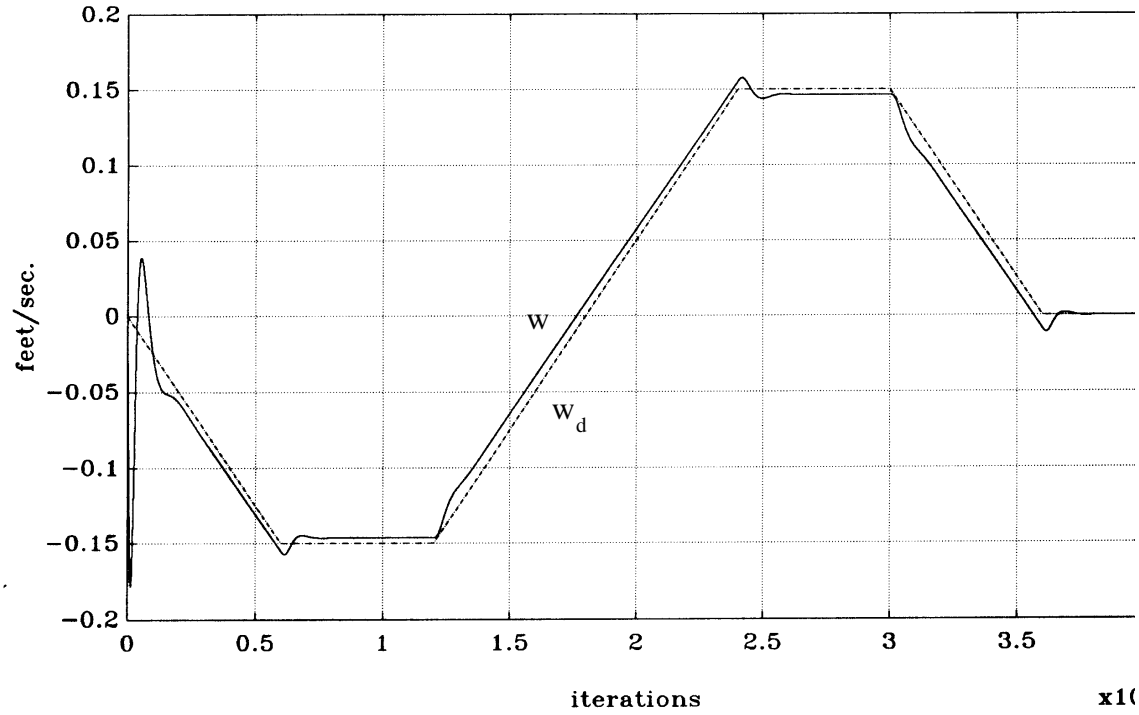

0.2
0.15
0.1
0.05
0
−0.05
−0.1
−0.15
−0.2
feet/sec.
w
$w_d$
0
0.5
1
1.5
2
2.5
3
3.5
iterations
x1(

Fig. 10b.

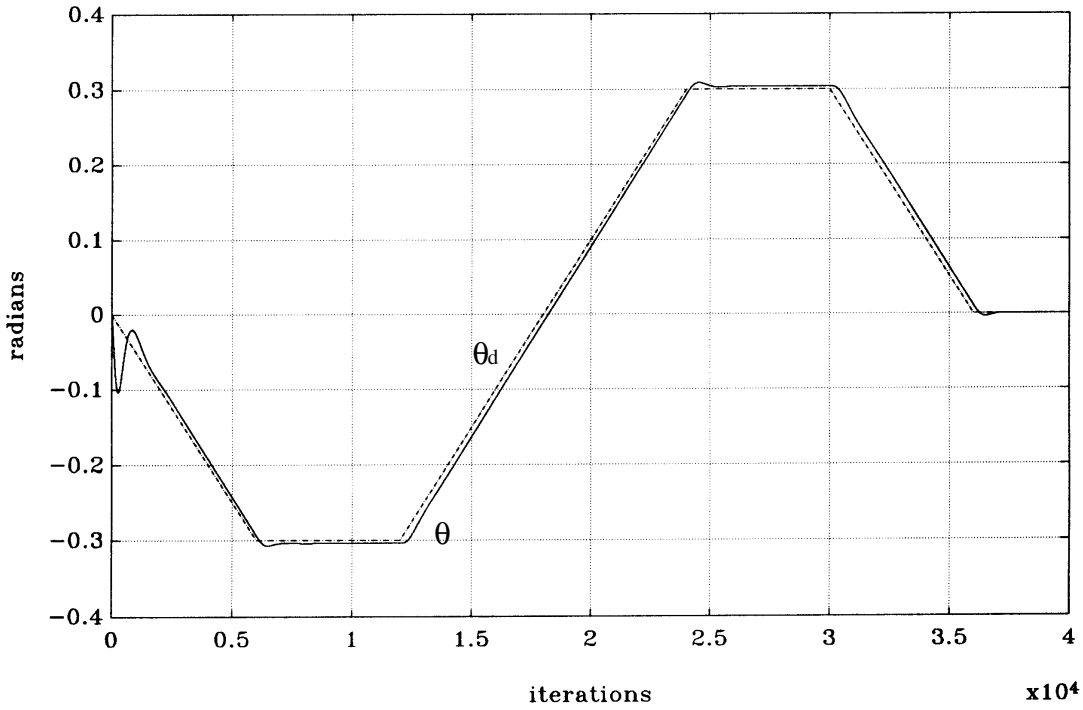

0.4
0.3
0.2
0.1
0
−0.1
−0.2
−0.3
−0.4
radians
θd
θ
0
0.5
1
1.5
2
2.5
3
3.5
4
iterations
x10⁴

Fig. 10a

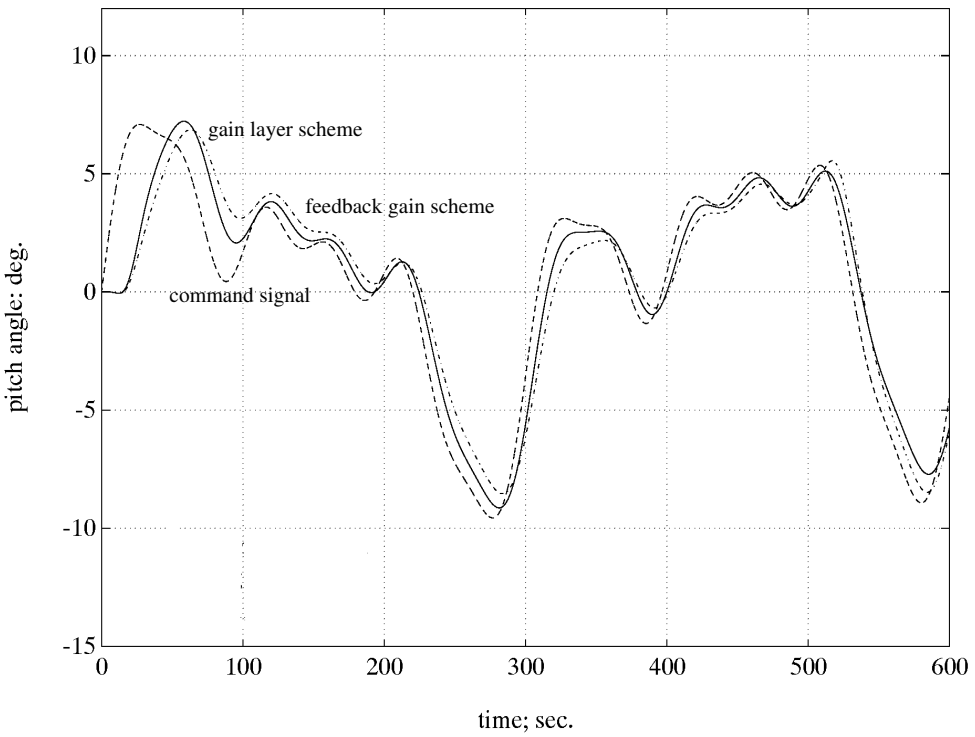

10
5
0
-5
-10
-15
0
100
200
300
400
500
600
pitch angle: deg.
time; sec.
gain layer scheme
feedback gain scheme
command signal

Fig. 9.

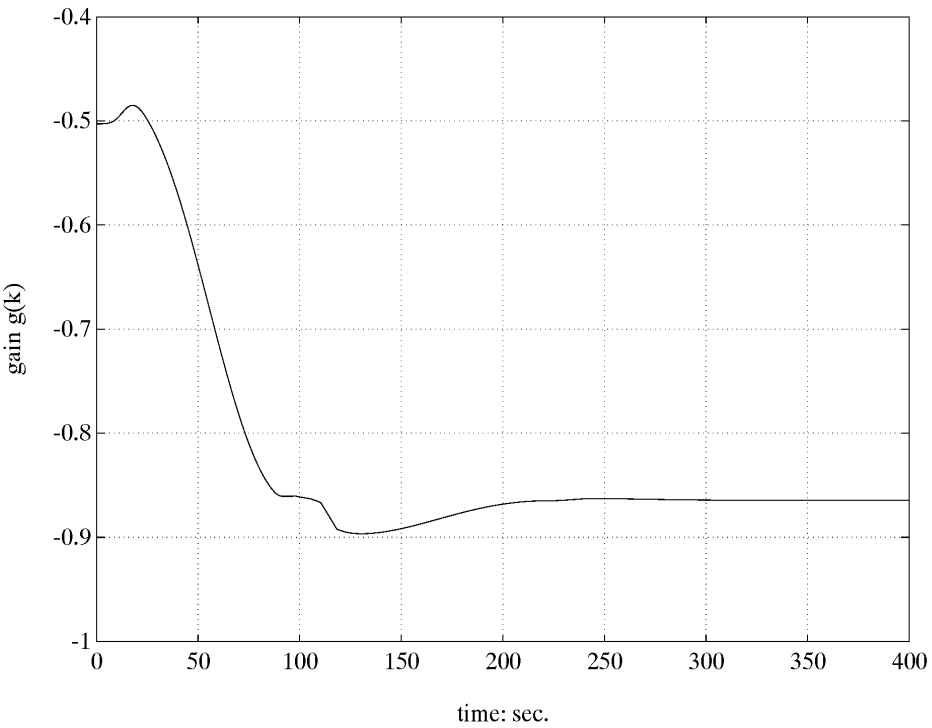

-0.4
-0.5
-0.6
-0.7
-0.8
-0.9
-1
gain g(k)
0
50
100
150
200
250
300
350
400
time: sec.

Fig. 8.

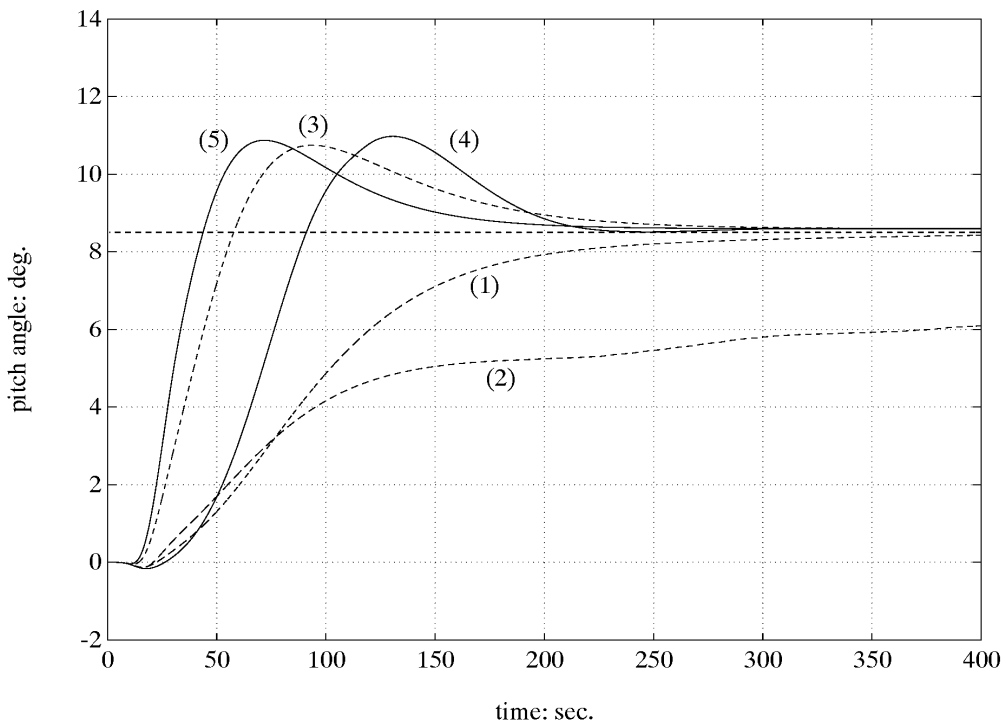

14
12
10
8
6
4
2
0
-2
0
50
100
150
200
250
300
350
400
pitch angle: deg.
time: sec.
(1)
(2)
(3)
(4)
(5)

Fig. 7.

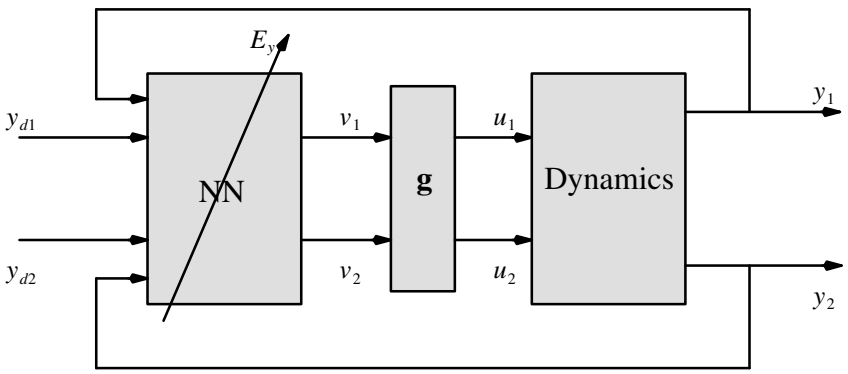



Fig. 6

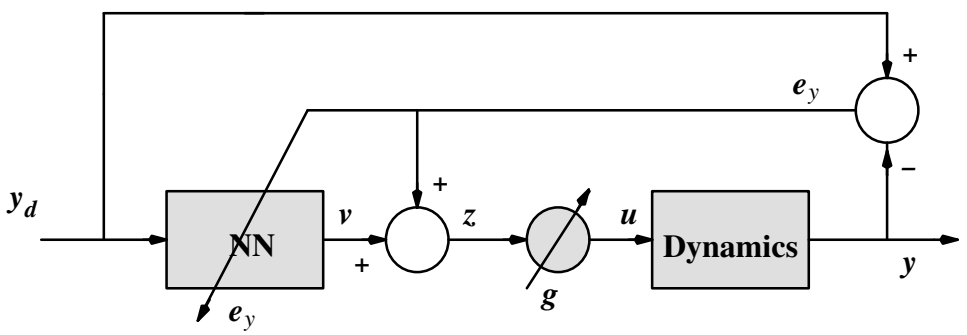



Fig. 5.

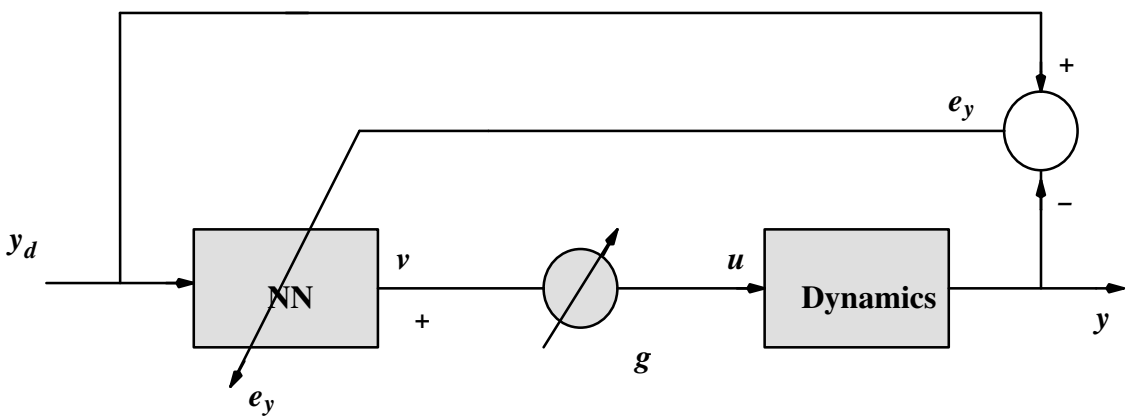



Fig. 4.

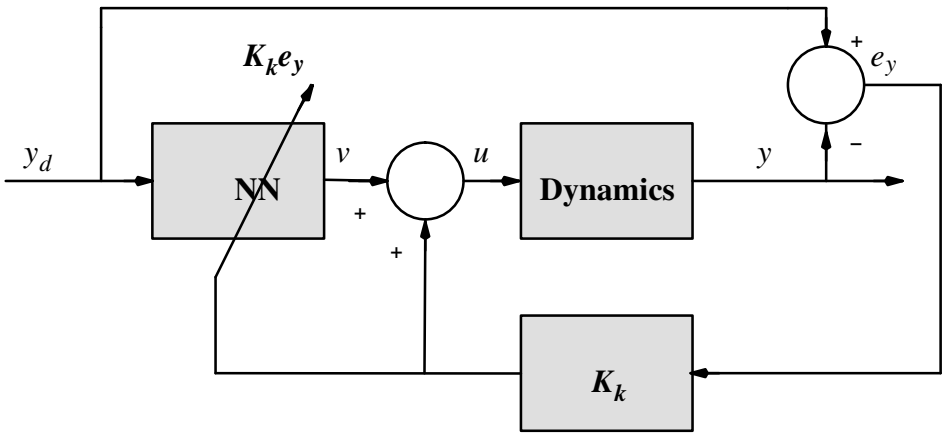



Fig. 3

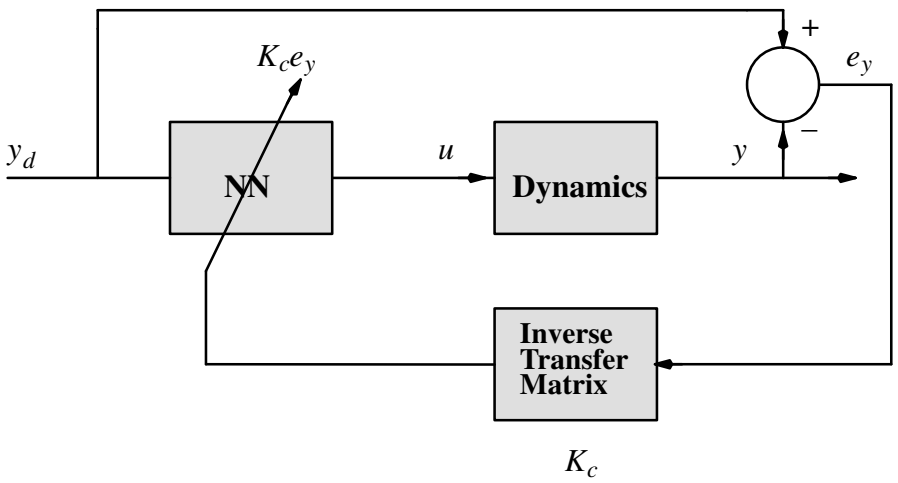
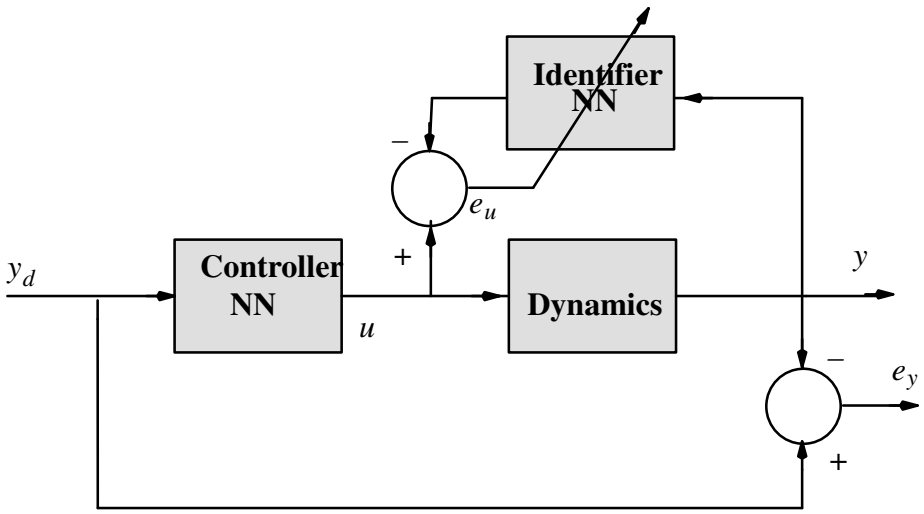
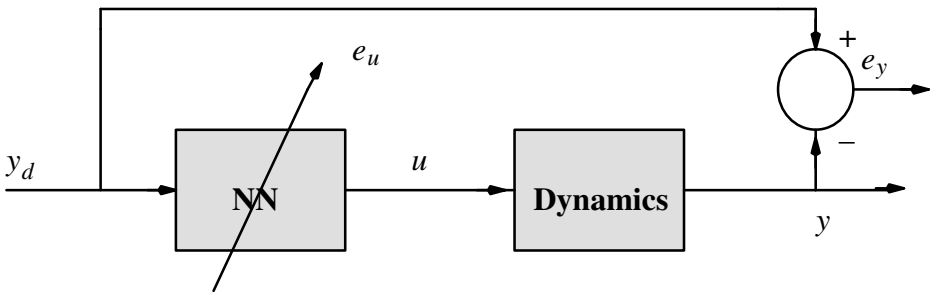



Fig. 2


Identifier NN
−
$e_u$
+
$y_d$
Controller NN
$u$
Dynamics
$y$
−
$e_y$
+

(Fig. 1a)


$e_u$
+
$e_y$
$y_d$
$u$
NN
Dynamics
−
$y$

(Fig. 1b)

Learning rate for gain layer = 0.0001. The dotted line represents the command signal and the solid line corresponds to the vehicle response.

(a) Pitch angle (deg.)

(b) Vertical velocity (ft/sec.)

## FIGURE CAPTIONS

Fig. 1 (a) Indirect Control Scheme

(b) Direct Control Scheme. The error $e_u$ corresponds to the error at the output of the neural network

Fig. 2. Inverse Transfer Matrix Scheme

Fig. 3. Feedback Gain Scheme

Fig. 4. Gain Layer Scheme

Fig. 5. Gain Layer Scheme with error feedback

Fig. 6. Gain Layer Scheme for the control of MIMO dynamics

Fig. 7. System response for a commanded pitch of 8.5 deg.

(1) Specialized learning scheme (2) Inverse Transfer matrix scheme

(3) Feedback gain scheme (4) Gain layer scheme (without error feedback)

(5) Gain layer scheme ( with error feedback )

A learning rate of 0.0001 was used for schemes (1) and (2). For the feedback gain scheme and the gain layer schemes, the learning rate used for the neural network controller was 0.008. The gain layer was updated using a learning rate of 0.001. All the schemes were compared using a network of size 2x20x10x1.

Fig. 8. Variation of the gain factor for the response shown in Fig. 7. The learning rate used for adapting the gain is 0.001. The learning rate for the controller NN is 0.008.

Fig. 9. Comparison of the feedback gain scheme and the gain layer scheme for a time varying command signal. The network learning rates used for both the schemes are 0.005.

Fig. 10. Simultaneous control of Pitch(θ) and Vertical velocity(w). The NN and gain layer are of sizes 4x20x10x2 and 2x1, respectively. Learning rate for the NN = 0.001.

[13] J. Yuh, ‘‘A neural network controller for underwater robotic vehicles,’’ *IEEE Journal of Oceanic Engineering*, vol. 15, no. 3, pp. 161 – 166, July 1990.

[14] R. M. Sanner and J. J. E. Slotine, ‘‘Direct adaptive control using Gaussian networks,’’ *Nonlinear Systems Laboratory, MIT, Tech. Report: NSL–910303*, March 1991.

[15] M. M. Polycarpou and P. A. Ioannaou, ‘‘Identification and control of nonlinear systems using neural network models: design and stability analysis,’’ *Technical Report: TR 91–09–01, University of Southern California, Department of Electrical Engineering–Systems*, California, 1991.

[16] K. P. Venugopal, R. Sudhakar, and A. S. Pandya, ”On–line learning control of autonomous underwater vehicles using feedforward neural networks,” *IEEE Journal of Oceanic Engineering*, vol. 17, no. 4, pp. 308–319, October 1992.

[17] K. P. Venugopal, A. S. Pandya and R. Sudhakar, ‘‘Adaptive neural network controllers for autonomous underwater vehicles,’’ *Proceedings of ROV ’91*, pp. 361–366, 1991.

[18] F. L. Lewis, K. Liu and A. Yesildirek, ‘‘Multilayer neural net robot controller with guaranteed tracking performance,’’ Technical Report, Automation and Robotics Research Institute, University of Texas at Arlington, March, 1993.

[19] A. P. Sage and C. C. White, *Optimum system control*, Prentice Hall, 1977, pp. 382–383.

[20] A. Shein and J. Kloskie, ‘‘Development of simulation facilities for the Autonomous Underwater Vehicle research,’’ *Proceedings of the Summer Simulations Conference (SSC)*, pp. 746–748, 1991.

## REFERENCE


[1] M. Kawato, K. Furukawa and R. Suzuki, ‘‘A hierarchical neural network model for control and learning of voluntary movement,’’ *Biological Cybernetics*, vol. 57, pp. 169–185, 1987.

[2] D. Psaltis, A. Sideris and A. Yamamura, ‘‘A multilayered neural network controller,’’ *IEEE Control Systems Magazine,* vol. 4, pp. 17 – 21, April 1988.

[3] W. T. Miller, R. P. Hewes, F. H. Glanz, and L. G. Kraft, ‘‘Real–time dynamic control of an industrial manipulator using a neural network based learning controller,’’ *IEEE J. Robotics and Automation*, vol. 6, pp. 1–9, February 1990.

[4] K. S. Narendra and K. Parthasarathy, ‘‘Identification and control of dynamical systems using neural networks,’’ *IEEE Trans. Neural Networks*, vol. 1, no. 1, pp. 4–27, March 1990.

[5] W. T. Miller, P. J. Werbos and R. J. Williams (Ed.), *Neural Networks for Control*, MIT Press, Cambridge, MA, 1991.

[6] V. C. Chen and Y. H. Pao, ‘‘Learning control with neural networks,’’ *Proc. Int. Conf. Robotics and Automation,* 1989, vol. 3, pp. 1448–1453.

[7] T. Yabuta and T. Yamada, ’’Neural Network controller characteristics with regard to adaptive control,’’ *IEEE Tran. System, Man and Cybernetics*, vol. 22, no. 1, pp. 170–176, Jan./Feb. 1992.

[8] K. S. Narendra and A. M. Annasamy, *Stable Adaptive Systems*, Prentice Hall, 1991.

[9] J. Feldman, ’’DTNSRDC revised standard submarine equations of motion,’’ David W. Taylor Naval Ship Research and Development Center, *Tech. Report: no. SPD–0393–09*, 1979.

[10] D. E. Rumelhart, J. L. Mc Celland and the PDP research group, *Parallel Distributed Processing, vol. 1: Foundations*, MIT Press, 1986.

[11] T. Troudet, S. Garg, D. Mattern and W. Merril, ‘‘Towards practical control design using neural computation,’’ *NASA Tech. Report TM–103785*, 1991.

[12] M. S. Lan, ‘‘Learning tracking controllers for unknown dynamical systems using neural networks,’’ *Proc. IEEE Conf. Systems, Man and Cybernetics*, pp. 29–31, 1990.

$$+ 0.5\varrho l^2[Z'_{|v_z|}\dot{v}_x|\dot{v}_z| + Z'_{v_z v_z}\sqrt{|v_z(v_y^2 + v_z^2)|}\ ]$$

$$+ 0.5\varrho l^2[Z'_{\delta_s}v_x^2\delta_s + Z'_{\delta_s\eta}v_x^2\delta_s(\eta - \frac{1}{c})c]$$

$$- 0.5C_d\int_l b(x)v_z(x)\sqrt{v_y^2(x) + v_z^2(x)}\,dx$$

$$- 0.5\varrho l\overline{C}_L\int_{x_2}^{x_1} v_y(x)\overline{v}_{FW}(t - \tau[x])dx \qquad \text{(A4)}$$

As before, the states of the vehicle ( $v_x$, $v_y$, $v_z$, $p$, $q$, $r$ ) are obtained at each instant by numerically integrating the equations (A1) = (A2) and (A3) = (A4).

$$\sum M_M = 0.5\varrho l^5[M'_{\dot{q}}\ \dot{q} + M'_{rp}\ rp] + 0.5\varrho l^4[M'_{\dot{v}_z}\ \dot{v}_z + M'_q\ v_x q]$$

$$+ 0.5\varrho l^3[M'_*\dot{v}_x^2 + M'_{v_z}\ v_x v_z + M'_{v_z|v_z|}v_z\sqrt{v_y^2 + v_z^2}\ ]$$

$$+ 0.5\varrho l^3[M'_{|v_z|}\dot{v}_x|\dot{v}_z| + M'_{v_z v_z}\sqrt{|v_z(v_y^2 + v_z^2)|}\ ]$$

$$+ 0.5\varrho l^3[M'_{\delta_s}v_x^2\delta_s + M'_{\delta_s\eta}v_x^2\delta_s(\eta - \frac{1}{c})c]$$

$$+ 0.5C_d\int_l xb(x)v_z(x)\sqrt{v_y^2(x) + v_z^2(x)}\,dx$$

$$- 0.5\varrho l\overline{C}_L\int_{x_2}^{x_1} xv_y(x)\overline{v}_{FW}(t - \tau[x])dx \qquad \text{(A2)}$$

where ρ, the density of sea water; $\delta_s$, the stern plane deflection (input); l, the length of the vehicle; other parameters such as *M'* are various hydrodynamics coefficients determined from the vehicle characteristics. The states of the vehicle ( $v_x$, $v_y$, $v_z$, $p, q, r$ ) are obtained at each instant by numerically integrating the equations (A1) = (A2) and (A3) = (A4) (provided below). The vehicle pitch θ is obtained by integrating the angular velocity $q$. A software package has been developed at the Center for Advanced Marine Systems for performing these computations[20].

The corresponding equations for the vertical velocity relationship are given below:

*(b) Vertical Velocity ($V_z$):*

Summation of all the vertical velocities acting on the vehicle, according to the Euler equations of motion;

$$\sum F_z = m[v_z' - v_x q + v_y p - z_G(p^2 + q_2) + x_G(rp - q') + y_G(rq + p')] \qquad \text{(A3)}$$

where the velocities and coordinates of the center of gravity are defined as before.

The summation of physical velocities is given by,

$$\sum F_z = 0.5\varrho l^4[Z'_{\dot{q}}\ \dot{q}] + 0.5\varrho l^3[Z'_{\dot{v}_z}\ \dot{v}_z + Z'_q\ v_x q + Z'_{v_y p}v_y p]$$

$$+ 0.5\varrho l^2[Z'_*\dot{v}_x^2 + Z'_{v_z}\ v_x v_z]$$

## ACKNOWLEDGEMENT

The authors wish to thank Prof. Stanley Dunn for his keen interest in this work. They are also grateful to the reviewers for their constructive criticisms and suggestions, which helped in improving the clarity of this paper.

## APPENDIX

This section provides the governing equations of the autonomous underwater vehicle dynamics employed in the simulation studies[9][16][20].

*(a) Pitch Angle (θ):*

The pitching moment of the vehicle derived from the conventional Euler equations of motion, is given by,

$$\sum M_M = I_y\dot{q} + (I_x - I_y)rp - (\dot{p} + qr)I_{xy} + (p^2 - r^2)I_{xz} + (qp - \dot{r})I_{yz} + m[z_G(\dot{v}_x - v_y\, r + v_z\, q) - x_G(\dot{v}_z - v_x\, q + v_y\, p)] \quad \text{(A1)}$$

where, $v_x$, $v_y$, $v_z$, $\dot{v}_x$, $\dot{v}_y$ and $\dot{v}_z$ are the linear velocities and their derivatives along the three coordinate axes x, y and z respectively; $p$, $q$, $r$, $\dot{p}$, $\dot{q}$, $\dot{r}$ are the angular velocities and their derivatives; and $x_G$, $y_G$ and $z_G$ are the location of the center of gravity with respect to the global coordinate system. The parameters $I_x$, $I_y$ and $I_z$ are the moments of inertia with respect to the three coordinate axes; $I_{xy}$, $I_{yz}$ and $I_{zx}$ are the moments of inertia with respect to the planes xy, yz and zx, respectively; and $m$ corresponds to the mass of the vehicle.

The summation of all the pitching moments estimated from the vehicle body characteristics and the velocities is given by,

of two weights $g_1$ and $g_2$. The NN and the gain layer are updated at each instant by minimizing the combined system error $E_y$ using equations (42) and (49), where

$$E_y = \frac{1}{2}[(\theta_d - \theta)^2 + (v_{z_d} - v_z)^2] \tag{49}$$

The results are shown in Figs. 10 (a) and (b) for trapezoidally varying command signals $\theta_d$ and $v_{z_d}$. It may be pointed out that these signals can not be chosen independently due to the coupled nature of the vehicle dynamics. The network learning rate was 0.001 and the learning rate for the gain layer is 0.0001. The weights were updated at each time instant (0.05 sec. apart). The NN weights were initialized to small random values between –0.1 and +0.1, as in the previous simulation studies, and the gain layer weights $g_1$ and $g_2$ were initialized to [–0.5, –0.39] corresponding to the inverse of steady state gain determined from the linearized vehicle dynamics (equations 46 and 47). From Fig. 10, it can be seen that the NN controller was able to provide close tracking performance.

## CONCLUSION

The choice of the neural network controller architecture plays an important role in obtaining the desired dynamic response and tracking performance. In this paper, we presented an improved neuro–controller architecture for direct control of dynamical systems which uses the back–propagation algorithm effectively. The ill effects of the controlled dynamics on the on–line updating of the NN weights are moderated by including a compensating gain layer. The presented method avoids the transformation needed in the case of other schemes and improves the system response using error feedback. With some examples, we have shown that this scheme performs well and gives better tracking and dynamical responses. Also,the performance of the gain layer scheme was studied on the problem of controlling two output parameters simultaneously.

this value. The gain layer weight $g$ was also initialized to the same value and updated on–line using a learning rate of 0.001. Equation (47) was used for the MIMO scheme described in the next section. The vehicle dynamics described by the nonlinear equations (A1) and (A2) was used in all the experiments. It can be seen from Fig. 7 that the gain layer scheme using the error feedback provided better transient response and smaller tracking error. The convergence of the gain g with respect to time is shown in Fig. 8.

In Fig. 9. we demonstrate the tracking effectiveness of the gain layer scheme for fast varying command signals. A command signal of the form

$$y_d(k) = 4\sin(\frac{2\pi k}{7000}) + 2\sin(\frac{2\pi k}{2000}) + 0.8\sin(\frac{2\pi k}{1000}) + 2.8\sin(\frac{2\pi k}{3000}) \qquad (48)$$

was given as the input to the network. The results shown are with a network of size 2x20x10x1 trained with a learning rate of 0.008. The controller is able to make the system track the command signal closely. The dotted lines corresponds to the gain feedback scheme with the error feedback. The gain $K_k$ as well as the initial value of $g$ are kept the same as in the previous simulation. As in the other example, the on–line adaptation in the gain layer scheme drives the system towards better tracking. In all the simulation studies, the controller is trained starting from random initial conditions.

### *B. Controlling MIMO dynamics:*

We conducted some preliminary simulation studies on the MIMO control by employing the gain layer scheme (without the error feedback) to control a single input/2–output vehicle dynamics. The controller generates the control signal $\delta_s$ (stern plane angle) based on the desired and actual values of the system outputs ( pitch $\theta$ and vertical velocity $v_z$ ). The scheme is similar to the one in Fig. 6 except that there is only one input to the dynamics. The governing equations of the system dynamics is detailed in Appendix A. The NN controller of size 4x20x10x2 has four input neurons corresponding to the desired and actual outputs of pitch ($\theta$) and vertical velocity ($v_z$). The gain layer $\boldsymbol{g}$ consists

ter experimenting with different number of neurons in the hidden layers, we found empirically that a 2x20x10x1 network provides the best performance.

Simulation studies are conducted for all the five schemes and the performance of the resultant control systems is compared by monitoring the speed of convergence, tracking error and stability of the system. The controller NN weights are initialized to small random values between +0.1 and –0.1 in all the cases. On–line updating of the controller and gain layer weights are continued throughout the experiment. All vehicle states are initialized to zero at the beginning of each simulation study.

Fig. 7. shows the results of comparison for a commanded pitch of 8.5 deg. The results shown are the best ones for each method over a number of trials. Scheme 1, which uses transformation of the error by backpropagating it through the dynamics, results in the slow tracking of the command signal. The inverse transfer matrix scheme also suffers from poor tracking as the high value of the numerically computed inverse differential gain causes large fluctuations in the transformed error. In both cases, the learning rate used was 0.0001, and any attempt to increase it further, tends to create unstable responses. The feedback gain scheme and the gain layer schemes were found to provide good tracking performances. The gain layer scheme with no error feedback performs significantly better than the specialized learning scheme of [2] and the inverse transfer matrix scheme[6]. The gain $K_k$ (see Fig. 3) for the feedback gain scheme was found from a linearized model of the vehicle dynamics.

To this end, the linearized system functions of pitch ($\theta$) and vertical velocity ($v_z$) were determined at the operating forward speed of 10.16 ft/sec. They are:

$$\frac{\theta}{\delta_s} = \frac{-(0.07808s^2 + 0.2311s + 0.0048)}{0.5077s^4 + 0.4576s^3 + 0.0780s^2 + 0.0050s + 0.0024} \tag{46}$$

$$\frac{v_z}{\delta_s} = \frac{-(0.4228s^3 + 0.5420s^2 + + 0.491s + 0.00799)}{0.5077s^4 + 0.4576s^3 + 0.0780s^2 + 0.0050s + 0.0024} \tag{47}$$

From equation (46), the inverse of the steady state gain was calculated (= –0.5) and $K_k$ was set to

and

$$\boldsymbol{g}(k) = \boldsymbol{g}(k-1) + \alpha\, \boldsymbol{J}^T[\boldsymbol{I}_m + \boldsymbol{J}\boldsymbol{g}(k-1)]^{-T}[\boldsymbol{Y}_d - \boldsymbol{Y}][\boldsymbol{V} + \boldsymbol{Y}_d - \boldsymbol{Y}]^T \qquad (45)$$

where $\boldsymbol{I}_m$ is the $m$ x $m$ identity matrix and $-T$ indicates inverse followed by transposition. The simulation results for such a scheme is given in the following section.

## SIMULATION RESULTS

We applied the proposed gain layer scheme for control of the nonlinear dynamics of an autonomous underwater vehicle. Assuming that the vehicle is travelling at a constant speed, the input–output relationship of the dynamics can be modeled using a set of fifth order differential equations having nonlinearities up to the degree of four[9]. The objective of the neural network controller is to provide a correct deflection angle '$\delta_s$' to the vehicle 'stern plane' actuator to maintain a commanded pitch $\theta_d$ [9][20]. The governing equations of the dynamics are given in the Appendix.

The output states of the vehicle (the linear velocities along the three coordinate axes $v_x$, $v_y$, $v_z$ and angular velocities $p$, $q$ and $r$) are obtained at each instant by integrating and solving the equations representing motion by integration, using a time–step of 0.05 sec. The vehicle pitch (system output) is obtained by integrating the angular velocity $q$, using the same time step. The typical operating range of the vehicle pitch is +15 deg. to –15 deg.

### *A. Controlling SISO dynamics:*

The controller NN weights and the gain layer are updated at every iteration which corresponds to a time step of 0.05 sec. The controller NN of size 2x20x10x1 is used in all the experiments. The two inputs to the NN are the desired pitch and the actual pitch of the vehicle. The size of the controller NN is selected based on a rough estimate of the complexity of the dynamics. In earlier studies[16][17], we have found that a single hidden layer was insufficient to model the nonlinearity. Af-

able and the initial value of $g$ need not be estimated precisely, unlike in the feedback gain scheme[1], where the gain is fixed and needs to be estimated precisely for better tracking.

### *C. Extension to MIMO systems:*

The proposed gain layer schemes (Sections *A & B*) can be readily extended to deal with MIMO dynamics. Figure 6 shows the scheme for a 2–input/2–output dynamics. For a general *n*–input/*m*–output dynamics, the NN will have *m* outputs, the *n* x *m* matrix $\mathbf{g}$ will represent a fully connected single layer network with *mn* weights. The Jacobian of the dynamics will be *m* x *n* matrix with elements

$$J_{ij} = \frac{\partial y_i}{\partial u_i} \qquad i = 1, ...., m \text{ and } j = 1, ..., n \tag{41}$$

The NN as well as the gain layer are updated on–line by minimizing the accumulated error $E_y$ over all the outputs of the dynamics. That is,

$$E_y = \frac{1}{2}\sum_{1=1}^{m}(y_{d_i} - y_i)^2 = \frac{1}{2}\| \boldsymbol{Y}_d - \boldsymbol{Y} \|^2 \tag{42}$$

The weight updating equations for the NN controller are the same as in equations (5) and (6) except that the output layer rule becomes,

$$\delta_j^p = (y_{d_j} - y_j) f'(net_j^p) \tag{42a}$$

Proceeding as in Sections *A* and *B*, after some matrix manipulations for the chain rule of differentiation[19], the vectorized version of the gain updating equations can be obtained for the MIMO scheme. Equation (28) for Scheme A (Fig. 4), gets extended to

$$\boldsymbol{g}(k) = \boldsymbol{g}(k-1) + \alpha\ \boldsymbol{J}^T\ [\boldsymbol{Y}_d - \boldsymbol{Y}]\ \boldsymbol{V}^T \tag{43}$$

where all the data vectors are assumed to be column vectors of appropriate dimensions and T indicates matrix transpose.

Similarly for the error feedback scheme, equations (33) and (40) get modified to,

$$\boldsymbol{Y} = [\boldsymbol{I}_m + \boldsymbol{J}\ \boldsymbol{g}]^{-1}\ \boldsymbol{J}\ \boldsymbol{g}\ [\boldsymbol{V} + \boldsymbol{Y}_d] \tag{44}$$

$$\frac{\partial E_y}{\partial g} = \frac{\partial E_y}{\partial y}\frac{\partial y}{\partial g} \tag{34}$$

But

$$\frac{\partial y}{\partial g} = \frac{\partial y}{\partial u}\frac{\partial (zg)}{\partial g} \tag{35}$$

$$\text{As } \frac{\partial (zg)}{\partial g} = z + g\frac{\partial z}{\partial g} \text{ and } \frac{\partial z}{\partial g} = -\frac{\partial y}{\partial g}, \tag{36}$$

$$\frac{\partial y}{\partial g} = \frac{\partial y}{\partial u}[z + (-\frac{\partial y}{\partial g})g] \tag{37}$$

On rewriting, we obtain:

$$\frac{\partial y}{\partial g} = \frac{\frac{\partial y}{\partial u}z}{(1 + g\frac{\partial y}{\partial u})} = \frac{J[v + e_y]}{1 + gJ} \tag{38}$$

Or

$$\frac{\partial E_y}{\partial g} = -\frac{e_y J[v + e_y]}{1 + gJ} \tag{39}$$

In this case, the equation (28) gets modified to:

$$g(k) = g(k-1) + \alpha e_y J\frac{[v + e_y]}{1 + gJ} \tag{40}$$

The important differences between the gain layer scheme presented here and the error transformation schemes[1][[2][6] are as follows. In the gain layer scheme we use the system error rather than a transformed error. This reduces the rapid variations in the NN weights due to the ill effects present in numerical estimation of the Jacobian. In our case the Jacobian affects only the gain layer weights, and their updating is controlled by a different learning rate. This implies that the controller NN can use larger learning rate to obtain faster convergence. Our simulation results support this argument. Also, because of the presence of error feedback the system response is faster. The NN time constant plays only a secondary role in the dynamical response of the system. Further the gain layer is adapt-

where $\alpha$ is the updating rate for *g(k)*. Rewriting the above equation using the chain rule and dropping the index for brevity,

$$\frac{\partial E_y}{\partial g} = \frac{\partial E_y}{\partial u}\frac{\partial u}{\partial g} \tag{26}$$

$$\frac{\partial E_y}{\partial u} = -(y_d - y)\frac{\partial y}{\partial u} \tag{27}$$

Assuming that $v$ is a weak function of $g$ and ignoring the implicit effect of $g$ on $v$ through network feedback, we make the approximation that $\frac{\partial u}{\partial g} = v$. Thus equation (24) reduces to:

$$g(k) = g(k-1) + \alpha(y_d - y)v\frac{\partial y}{\partial u} \tag{28}$$

*B. Effect of the error feedback:*

Equation (16) indicates that the error feedback helps to improve the system response. We introduce an error feedback similar to the scheme by Kawato and Fig. (5) shows the modified scheme.

The input to the gain $g$ is:

$$z = v + e_y = v + (y_d - y) \tag{29}$$

Then, the control input to the dynamics is given by

$$u = zg \tag{30}$$

As before, defining

$$y = J(u)u \tag{31}$$

$$= J(u)g[v + (y_d - y)] \tag{32}$$

which on rewriting gives:

$$y = \frac{J(u)vg}{1 + J(u)g} + \frac{J(u)y_d g}{1 + J(u)g} \tag{33}$$

Considering the updating equations of the gain $g$, we have,

The objective function to be minimized is chosen as $E_y = \frac{1}{2}(y_d - y)^2$. The NN weights can be updated using the total error $e_y$, provided we have $E_y \simeq E_v$, where $E_v$ is the squared error at the output of the NN (see Fig. 4). This condition is achieved when

$$\frac{\partial u}{\partial v} \simeq \frac{\partial u}{\partial y} \tag{21}$$

or,

$$g \simeq [J(u)]^{-1} \tag{22}$$

In practice, the gain layer NN is determined through an adaptation process and will serve as an approximation of the inverse Jacobian of the dynamics. The gain layer has a role similar to the feedback gain constant $K_k$ of Kawato's scheme and the inverse transfer matrix $K_c$ of Chen *et al.* The weight $g$ in SISO case is a scalar value; it will be a matrix in the case of MIMO systems.

The NN updating equations are derived as in equations (5) & (6) and the output layer updating rule becomes

$$\delta_j^p = e_y f'_j(net_j^p) \tag{23}$$

In the following section, we outline the procedure for determining the gain layer weights. For simplicity, we first consider a SISO dynamics where $g$ is a scalar value.

*A. Updating the Gain Layer:*

Let $g(k)$ be the gain layer weight at the instant $k$. The adaptation rule for $g(k)$ can be written as,

$$g(k) = g(k-1) + \Delta g(k) \tag{24}$$

where $\Delta g(k)$ is the change in $g(k)$ at instant $k$. Using the gradient decent rule to minimize the squared error $E_y$ (equation 1), we obtain

$$\Delta g(k) = -\alpha \frac{\partial E_y(k)}{\partial g(k)} \tag{25}$$

Comparing equations (11) and (19) and using the equation (6a), we obtain the weight updating equation for the output layer as:

$$\delta_j^p = K_k \frac{e_y}{2} f'_j(net_j^p) \tag{20}$$

Comparing equations (20) and (12), it can be seen that the feedback error scheme is equivalent to the inverse transfer matrix scheme when $J^{-1}(u) = \frac{K_k}{2}$. With off–line estimation of the Jacobian, both methods provide similar results for SISO dynamics.

It is to be noted that for good tracking performance over a wide range of nonlinear time–varying dynamics, sophisticated estimation of $K_k$ is needed. An important aspect of this scheme is the increased speed of learning due to the error feedback. The network error directly affects the control input resulting in faster learning, while in the other schemes, the error is reflected only through the network. Note that the network learning rate is usually kept very small when using backpropagation algorithm for control applications, to avoid the possibility of divergence[4][7]. Thus when the error is reflected only through the NN, the convergence may be slow.

## THE GAIN LAYER SCHEMES

In this section, we propose two different architectures for the direct control of dynamical systems using neural networks[Figs. 4 & 5]. The schemes use the system error $e_y$ directly for updating the weights. The ill effects due to the numerical estimation of the differential gain of the dynamics is compensated by introducing a gain layer between the NN controller and the system. The gain layer is a separate single layer NN with linear transformation. The controller NN weights are updated at each iteration, according to the backpropagation algorithm using the error $e_y$. A rough estimate of the inverse of the steady state gain of the dynamics is used as the initial value of the gain layer NN and its weights are also updated at each iteration using the gradient descent approach.

where $e_y = y_d - y$, as before and $v$ is the NN output. In terms of the Jacobian of the dynamics $J(u)$, the system response $y$ can be expressed as:

$$y = [v + K_k(y_d - y)]\ J(u) \tag{14}$$

Expanding equation (14) and rewriting:

$$y = \frac{J(u)v}{1 + J(u)K_k} + \frac{K_k J(u) y_d}{1 + J(u)K_k} \tag{15}$$

The scheme in [1] corresponds to the approximation $K_k \simeq [J(u)]^{-1}$. Then the equation (15) reduces to:

$$y = \frac{J(u)v}{2} + \frac{y_d}{2} \tag{16}$$

Note that the command signal $y_d$ is directly influencing the system response because of the error feedback.

The error function to be minimized in this scheme is the same as that of equation (8), and the network updating equations can be derived as before. From equation (14) we get;

$$\frac{\partial y}{\partial w} = J(u)\ [\frac{\partial v}{\partial w} - K_k \frac{\partial y}{\partial w}] \tag{17}$$

i. e.

$$\frac{\partial y}{\partial w} = \frac{J(u)\ \partial v / \partial w}{1 + J(u)K_k} \tag{18}$$

When $K_k \simeq [J(u)]^{-1}$ in equation (18) and with $E_{k_y}$ defined to be $\frac{1}{2}(K_k e_y)^2$, we obtain;

$$\frac{\partial E_{k_y}}{\partial w} = -K_k \frac{e_y}{2} \frac{\partial v}{\partial w} \tag{19}$$

As before $J(u) = \frac{\partial y}{\partial u}$ and equation (9) reduces to:

$$\frac{\partial E_{Ky}}{\partial w} = - K_c^2 e_y \, J(u) \frac{\partial u}{\partial w} \tag{10}$$

When $K_c \simeq [J(u)]^{-1}$, the scheme proposed in [6] is obtained; $E_{Ky}$ corresponds to the squared error at the NN output. Then

$$\frac{\partial E_{Ky}}{\partial w} = - K_c \, e_y \, \frac{\partial u}{\partial w} \tag{11}$$

The updating rules in this scheme remains the same as in Psaltis' scheme, except for the output layer weights. For the output layer weights, equation (6a) gets modified as

$$\delta_j^p = K_c \, e_y \, f'_j(net_j^p) \tag{12}$$

The important difference between the specialized learning scheme and the inverse transfer matrix scheme is that in the first method, the system error $e_y$ is minimized while in the second method the error estimated at the output of the NN, $e_u$ is minimized. Comparing equations (6a) and (12), it can be seen that the differential gain of the dynamics affects the network updating in different ways in the above schemes. In the case of inverse transfer matrix scheme[6], if the inverse differential gain $K_c$ is computed off–line, the convergence may be faster as the dynamics is not treated as an additional layer of the NN. However, the numerical problems associated with the on–line computation of $K_c$ can cause large changes in the network weights. Also, for control signals of constant values ($\partial u \simeq 0$), the on–line estimated values of the inverse differential gain may be inaccurate and unreliable.

### *D. The Feedback Gain Scheme:*

Kawato *et al.*[1] suggested an error feedback scheme for the direct control of dynamical systems, motivated by the voluntary movements in biological systems (Fig. 3). The feedback gain $K_k$ is estimated off–line and is used for transforming the error for updating the network weights as well as for generating the control signal $u$. From figure 3, we obtain the following expressions for the control signal $u$ :

$$u = K_k e_y + v \tag{13}$$

where $f_i'$ denotes the derivative of the sigmoidal transformation function $f_i$ which is given by,

$$f_i\ (net_i^p) = \frac{[1\ -\ \exp(-\ net_i^p\ )]}{[1\ +\ \exp(-\ net_i^p\ )]} \tag{7}$$

The differential gain *J(u)*( or instantaneous forward Jacobian ) can be estimated on–line from the changes in *y* and *u* over two consecutive time instants. Alternately, the dynamics can be modeled using a neural network and can be used for transforming the error $e_y$ to $e_u$ , the equivalent error at the NN output (see Fig. 1b) [5]. Troudet *et al*. discusses such a scheme for flight control application, in [11].

One of the problems in treating the dynamics as a nonmodifiable layer is that, the convergence of the NN may be slow[6]. From equation (6a), it can be seen that the differential gain of the dynamics directly affects the network updating. An on–line numerical estimation procedure using ratio of the incremental changes in the output and input of the system dynamics, may generate large variations in the value of *J(u),* causing correspondingly large changes in weights. This problem is more evident when the reference command signals are rapidly varying. Also, under steady state conditions when the change in inputs and outputs are very small, a numerically estimated Jacobian *J(u)*, will be inaccurate. The inverse transfer matrix scheme suggested by Chen *et al*. [6] is a more suitable method under such conditions.

### *C. The Inverse Transfer Matrix Scheme:*

In the inverse transfer matrix scheme ( Fig. 2 ), the error function to be minimized is chosen as

$$E_{Ky} = \frac{1}{2}(K_c e_y)^2 \tag{8}$$

where $K_c$ represents the inverse transfer relationship of the dynamics. The network updating is done by estimating the derivative of the error function with respect to the weights: i. e.

$$\frac{\partial E_{Ky}}{\partial w} = -\ K_c^2 e_y \frac{\partial y}{\partial w} \tag{9}$$

factor $\frac{\partial y}{\partial u}$. If we define $J(u) = \frac{\partial y}{\partial u}$, equation (3) can be rewritten as

$$\frac{\partial E_y}{\partial w} = - e_y J(u) \frac{\partial u}{\partial w} \tag{4}$$

*B. Backpropagating the error through the dynamics:*

If we consider the system dynamics as an additional layer of the NN controller, having interconnection strength *J(u)* with no updating done, equation (4) represents the backpropagation of the system error $e_y$ through the dynamics. This approach originally proposed by Psaltis *et al.* [2] and termed as 'specialized learning scheme', is an effective way for the direct control of dynamical systems using neural networks. It is used in a number of later studies in different forms[11–13]. We review the network updating equations here for comparing with the other schemes discussed in this paper.

Following the notations in [pp 327–330, 10], let us define $\Delta w_{ji}^p$ as the change in weight $w_{ji}$, connecting neuron *i* in the *(p–1)*th layer to the neuron *j* in the *p*th layer. Let $\delta_j^p$ be the derivative of the error at the layer *p*, say $E_p$, with respect to the summed input to the sigmoid transformation ( $net_j^p$) at neuron *j*, and $o_j^p$ as the output of the neuron. Then the updating equations for any layer can be shown to be[10]:

$$w_{ji}^p(n+1) = w_{ji}^p(n) + \Delta w_{ji}^p(n) \tag{5a}$$

where,

$$\Delta w_{ji}^p = \eta \, \delta_j^p \, o_i^p \tag{5b}$$

and $\eta$ is the learning rate of the NN.

Using equation (4), we can get $\delta_j^p$ for the output layer of the network as

$$\delta_j^p = J(u) \, e_y \, f_j'(net_j^p) \tag{6a}$$

and for any other layer as

$$\delta_i^p = f_i'(net_i^p) \sum_j \delta_j^p w_{ji} \tag{6b}$$

represent the input to the dynamics $\boldsymbol{u}$ as a nonlinear function of $\boldsymbol{x}$ and $t$, i.e. $\boldsymbol{u} = \phi(\boldsymbol{x}, t)$, then $\dot{\boldsymbol{x}} = f[\boldsymbol{x}, \phi(\boldsymbol{x}, t), t]$ represents the closed loop system dynamics. The objective of an on–line learning controller is to generate the control signal to drive the dynamics to a state $\boldsymbol{x}^*$, at each instant in time so that the error $e_y(t)$ between the desired output $\boldsymbol{y}_d(t)$ and the actual output $\boldsymbol{y}(t)$, asymptotically tends to zero.

In a NN based on–line learning control scheme the objective of the neural network is to generate the control input $u$ by mapping the nonlinear function $\phi(\boldsymbol{x},t)$. As indicated before, for using the back-propagation learning algorithm, the error at the output of the network has to be estimated in order to update the neural network interconnection strengths.

### *A. The network error and the system error:*

Fig. 1b shows the basic direct control scheme. The NN learns the inverse characteristics of the dynamics implicitly and generates the correct control signal $u$, to drive the system state $y$ to the desired state $y_d$. Defining the squared error at the system output as the objective function to be minimized by the controller, we have:

$$E_y = \frac{1}{2}(y_d - y)^2 \tag{1}$$

If $(y_d - y)$ is defined as $e_y$, the gradient of the error with respect to the network interconnection strength, $w$, is:

$$\frac{\partial E_y}{\partial w} = -e_y \frac{\partial y}{\partial w} \tag{2}$$

$$= -e_y \frac{\partial y}{\partial u}\frac{\partial u}{\partial w} \tag{3}$$

The above equations can be generalized to the case when $w$ is a vector. The term $\frac{\partial y}{\partial u}$ corresponds to the forward gain of the dynamics and is an important factor determining the stability of the overall system[7]. The term $e_y \frac{\partial y}{\partial u}$ represents the transformation of the total error $e_y$ with a transformation

dynamics. The third method[1] uses a fixed feedback gain stage to generate a transformed error signal which is employed for updating the neural networks. In the first two methods, the differential gain of the dynamics (forward Jacobian) directly affects the network updating. These methods suffer from the fact that an on–line numerical estimation of Jacobian may result in sharp changes in the transformed error and hence in the network weights. On the other hand, the third method may be able to provide robust results only over a particular range, since the feedback gain is fixed[1][6].

In this paper, after a brief review of the above three implementations of the direct control scheme using the backpropagation algorithm, we propose a modified architecture. The major difference in the proposed architecture is that, the network interconnection strengths are updated using the system output error rather than the transformed error used in the other schemes. The ill effects due to the numerical estimation of Jacobian of the controlled dynamics, are moderated by adding an extra gain layer between the neural network controller and the system. The gain layer is a single layer NN with linear transformation. The gain layer NN is initialized to the approximate steady state inverse gain of the dynamics and is adapted on–line. An error feedback, similar to the one in Kawato's scheme[1] is introduced to improve the dynamic response of the control system. Results of simulation studies performed on the nonlinear dynamics of an underwater vehicle, show that the presented schemes give promising results. In the following sections we describe the schemes and the simulation results. In the present study, first we develop and analyze the schemes for Single Input Single Output (SISO) dynamics, for the ease of explanation and comparison. The schemes are extended to deal with Multiple Input Multiple Output (MIMO) dynamics. Simulation results are provided for the control of both SISO and MIMO dynamics. Also, in the subsequent discussions, we have used the words controller, NN or network interchangeably, to refer to the neural network controller.

## DIRECT CONTROL SCHEMES

Consider a time varying dynamical system represented by a set of nonlinear differential equations, $\dot{\boldsymbol{x}} = f(\boldsymbol{x}, \boldsymbol{u}, t)$, $\boldsymbol{y} = h(\boldsymbol{x}, t)$ ; where $\boldsymbol{x}$ denotes the state vector, $\boldsymbol{y}$ the system output and $\boldsymbol{u}$ the input. If we

## INTRODUCTION

The recent interest in highly parallel networks of simple processing elements (neurons) has led to the wide use of such neural networks (NN) for the learning control of dynamical systems[1–5, 11–15, 18]. The capabilities of such networks for efficient generalization and adaptation make them excellent candidates for the learning control of linear as well as nonlinear dynamical systems. The objective of a NN controller is to generate the correct control signal to drive the dynamics from the initial state to a desired final state through an optimum state trajectory. The tracking performance and the ease of implementation of a neural network controller largely depends upon the learning algorithm chosen as well as the control architecture used. The two generally used control architectures are the direct control scheme and the indirect control scheme. In most of the implementations, backpropagation is used as the learning algorithm.

In the indirect control scheme, the parameters of the dynamical system are identified at each instant which are then used to estimate the controller parameters (see Fig. 1a). Thus there is an explicit identification process in this approach[4][8]. The disadvantage of such a scheme is that, the identification and control are solely based on the error $e_u$, and hence minimization of the output error $e_y$ can not be guaranteed[2]. In the direct control scheme, the controller parameters are directly adjusted to reduce the output error (Fig. 1b). Hence it is simpler to implement such a scheme[4][8]. It may be noted that even when the dynamics are linear and time–invariant, the above approaches results in overall nonlinear systems[4][8]. One of the problems in having a direct control scheme using neural networks is that the error at the output of the neural network (network error, $e_u$) which is needed for updating the network weights is not directly available (see Fig. 1b). Only the system error $e_y$ can be measured at the output of the dynamics, and it does not represent the error at the NN output. Three different methods have been proposed in the literature to derive the network error from the system error measurements. In the first method[2], the dynamics is treated as an unmodifiable layer of the NN, and the error is backpropagated through the dynamics. In the method proposed by Chen *et al.*[6], the system error is transformed into the network error using inverse transfer matrix of the

*to appear in Journal of Circuits, Systems and Signal Processing*

# AN IMPROVED SCHEME FOR DIRECT ADAPTIVE CONTROL OF DYNAMICAL SYSTEMS USING BACKPROPAGATION NEURAL NETWORKS

K. P. Venugopal[*#1] R. Sudhakar[#] A. S. Pandya[*+]

* Center for Advanced Marine Systems
\# Department of Electrical Engineering
\+ Department of Computer Science and Engineering
Florida Atlantic University
Boca Raton, FL 33431

**Abstract:**

This paper presents an improved direct control architecture for the on–line learning control of dynamical systems using backpropagation neural networks. The proposed architecture is compared with the other direct control schemes. In the present scheme the neural network interconnection strengths are updated based on the output error of the dynamical system directly rather than using a transformed version of the error employed in other schemes. The ill effects of the controlled dynamics on the on–line updating of the network weights are moderated by including a compensating gain layer. An error feedback is introduced to improve the dynamic response of the control system. Simulation studies are performed using the nonlinear dynamics of an underwater vehicle and the promising results support the effectiveness of the proposed scheme.

---

1 Presently with the Medical Image Processing Group, University of Pennsylvania, Philadelphia, PA 19104.