NOISE SENSITIVITY OF PROJECTION NEURAL NETWORKS

Seho Oh and Robert J. MarksII

Interactive System Design Laboratory
University of Washington, FT-10
Seattle, WA 98195

I. INTRODUCTION

Projection based artificial neural networks (ANN's) have been discussed extensively elsewhere [1-7]. In this paper, we analyze the noise properties of these networks. Noise is introduced both at the neural and interconnect levels.

Two ANN's will be analyzed. The first is the *alternating projection neural network* (APNN) [1-6] which can effectively be used as a content addressable memory. We show that the performance of such a network improves as the percentage of clamped neurons is increased. The second is the *layered classifier artificial neural network* (L-CANN) [7] whose performance is based on nonlinearly augmented *synthetic discriminant functions* (also called *composite matched filters*) [8-9]. If the signal-to-noise ratio of the interconnect noise remains constant, the performance of a projection based artificial neural network can improve as the number of hidden neurons is increased.

Space does not permit a detailed review of the APNN and L-CANN. The reader is referred to the references (particularly [3] and [7]) for a more detailed explanation. For both ANN's, we will briefly state the algorithm, analyze the noise sensitivity and interpret the results.

II. The APNN

(1) *Algorithm Review*

Consider a set of N continuous level linearly independent library vectors (or patterns) of length L>N : $\{ \mathbf{f}_n \mid 1 \le n \le N \}$. We form the library matrix

$$\underline{F}=[\mathbf{f}_1 \ \mathbf{f}_2 \cdots \mathbf{f}_N]$$

and the neural network interconnection matrix

$$\underline{T}=\underline{F}(\underline{F}^T\underline{F})^{-1}\underline{F}^T$$

where the superscript T denotes transposition. The L neural states are divided into two sets: one in which the states are known (clamped neurons) and the remainder in which the states are unknown (floating neurons). The partition of clamped and floating neurons can change from application to application. We can, however, assume without loss of generality that for a particular application neurons 1 through P are clamped and the remaining neurons are floating. We adopt the vector partition notation

$$\mathbf{i} = \begin{bmatrix} \mathbf{i}^P \\ \hline \mathbf{i}^Q \end{bmatrix}$$

where $\mathbf{i}^P$ is P-tuple of the first P elements of $\mathbf{i}$ and $\mathbf{i}^Q$ is a vector of the remaining Q = L-P. We can thus write, for example,

$$\underline{F}_P = [\ \mathbf{f}^P{}_1 \ \ \mathbf{f}^P{}_2 \cdots \mathbf{f}^P{}_N\ ]$$

Using this partition notation, we define the neural clamping operator by:

$$\underline{\eta}\ \mathbf{i} = \begin{bmatrix} \mathbf{f}^P \\ \hline \mathbf{i}^Q \end{bmatrix}$$

Thus, the first P elements of $\mathbf{i}$ are clamped to $\mathbf{f}^P$. The remaining Q nodes "float"

For synchronous operation, the network iteration can be written as

$$\mathbf{s}(M+1) = \underline{\eta}\ \underline{T}\ \mathbf{s}(M) \tag{1}$$

This operation is the alternating projection between the space which is spanned by columns of matrix $\underline{T}$ and linear variety consisting of the set of all vectors of length L whose first P elements are those of $\mathbf{f}$. If $\underline{F}_P$ is a matrix of full rank, then iteration converges to $\mathbf{f}$.

We can rewrite the matrix $\underline{T}$ in partitioned form as

$$\underline{T} = \left[\begin{array}{c|c} \underline{T}_2 & \underline{T}_1 \\ \hline \underline{T}_3 & \underline{T}_4 \end{array}\right]$$

where $\underline{T}_2$ is a P by P and $\underline{T}_4$ a Q by Q matrix. Eq.(1) can then be written in partitioned form as

$$\begin{bmatrix} \mathbf{f}^P \\ \hline \mathbf{s}^Q(M+1) \end{bmatrix} = \left[\begin{array}{c|c} \underline{T}_2 & \underline{T}_1 \\ \hline \underline{T}_3 & \underline{T}_4 \end{array}\right] \begin{bmatrix} \mathbf{f}^P \\ \hline \mathbf{s}^Q(M) \end{bmatrix}$$

or, equivalently

$$\mathbf{s}^Q(M+1) = \underline{T}_3\mathbf{f}^P + \underline{T}_4\mathbf{s}^Q(M) \tag{2}$$

If the norm of $\underline{T}_4$ is strictly less than one, the steady state solution of Eq.(2) can be written as

$$\mathbf{s}^Q(\infty) = \mathbf{f}^Q = (\underline{I}-\underline{T}_4)^{-1}\underline{T}_3\mathbf{f}^P. \tag{3}$$

(2) *Noise Model for the APNN*

Fig.1 is a block diagram of the APNN iteration in Eq.(2) corrupted with additive noise. The vectors $\mathbf{n}_i$ and $\mathbf{n}_d$ denote the input (data source) and output (detector) noise respectively, and $\mathbf{n}_f$ is the feedback noise vector. These vectors are added appropriately to the vector components (neuron states). The matrices $\underline{N}_3$ and $\underline{N}_4$ denote the system noise which is associated with the interconnects. They may represent the inexactness of analog multiplication or, for digital implementation, round off error. We assume that each neural (vector) noise process consists of elements with identically independent distributions (iid) in spatial domain, and, temporally, are either white or static. For the system, we assume that the noise is spatially iid with temporally white noise. Each noise vector and matrix is assumed to be zero mean and independent from every other one.

Let $\tilde{\underline{T}}_3(i)=\underline{T}_3+\tilde{\underline{N}}_3(i)$ and $\tilde{\underline{T}}_4(i)=\underline{T}_4+\tilde{\underline{N}}_4(i)$.

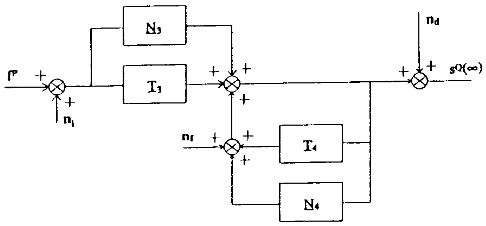

Fig.1 The block diagram of the APNN operation with additive noise. The matrices $\underline{N}_3$ and $\underline{N}_4$ indicate the interconnect noise.

Then the relationship between the noisy $\mathbf{s}^Q(M)$ and $\mathbf{s}^Q(M-1)$ will be

$$\mathbf{s}^Q(M) = \tilde{\underline{T}}_4(M-1)\,\mathbf{s}^Q(M-1) + \tilde{\underline{T}}_3(M-1)\,[\mathbf{f}^P + \tilde{\mathbf{n}}_i(M-1)] + \tilde{\mathbf{n}}_f(M-1)$$

The solution of the above equation is

$$\mathbf{s}^Q(M) = \sum_{k=0}^{M-1} \underline{A}_M(k)\ \tilde{\underline{T}}_3(M-1-k)\ \mathbf{f}^P + \sum_{k=0}^{M-1} \underline{A}_M(k)\ \tilde{\mathbf{n}}_f(M-1-k) + \sum_{k=0}^{M-1} \underline{A}_M(k)\,\tilde{\underline{T}}_3(M-1-k)\ \tilde{\mathbf{n}}_i(M-1-k) \tag{4}$$

where

$$\underline{A}_M(k) = \begin{cases} \underline{I} & ;\ k=0 \\ \tilde{\underline{T}}_4(M-1)\,\tilde{\underline{T}}_4(M-2)\cdots\tilde{\underline{T}}_4(M-k) & ;\ k>0 \end{cases}$$

For notational convenience, we define

$$\underline{T}_3(k) = \tilde{\underline{T}}_3(M-k),\ \underline{T}_4(k) = \tilde{\underline{T}}_4(M-k),$$

$$\mathbf{n}_i(k) = \tilde{\mathbf{n}}_i(M-k-1) \text{ and } \mathbf{n}_f(k) = \tilde{\mathbf{n}}_f(M-k-1).$$

Eq.(4) then become

$$\mathbf{s}^Q(M) = \sum_{k=0}^{M-1} \underline{A}(k)\ \underline{T}_3(k+1)\,\mathbf{f}^P + \sum_{k=0}^{M-1} \underline{A}(k)\ \underline{T}_3(k+1)\ \mathbf{n}_i(k) + \sum_{k=0}^{M-1} \underline{A}(k)\ \mathbf{n}_f(k) \tag{5,a}$$

where

$$\underline{A}(k) = \begin{cases} \underline{I} & ;\ k=0 \\ \underline{T}_4(1)\,\underline{T}_4(2)\cdots\underline{T}_4(k) & ;\ k>0 \end{cases} \tag{5,b}$$

Therefore the noisy steady state result is

$$\mathbf{s}^Q(\infty) = \mathbf{r}_1 + \mathbf{r}_2 + \mathbf{r}_3 + \mathbf{n}_d \tag{6,a}$$

where,

$$\mathbf{r}_1 = \sum_{k=0}^{\infty} \underline{A}(k)\ \underline{T}_3(k+1)\ \mathbf{f}^P \tag{6,b}$$

$$\mathbf{r}_2 = \sum_{k=0}^{\infty} \underline{A}(k)\ \underline{T}_3(k+1)\ \mathbf{n}_i(k) \tag{6,c}$$

$$\mathbf{r}_3 = \sum_{k=0}^{\infty} \underline{A}(k)\ \mathbf{n}_f(k) \tag{6,d}$$

The expectation of $\mathbf{s}^Q(\infty)$ is

$$E[\mathbf{s}^Q(\infty)] = E[\mathbf{r}_1] + E[\mathbf{r}_2] + E[\mathbf{r}_3] + E[\mathbf{n}_d]$$

$$= \sum_{k=0}^{\infty} \underline{T}_4^k\underline{T}_3\mathbf{f}^P = (\underline{I}-\underline{T}_4)^{-1}\mathbf{f}^P = \mathbf{f}^Q$$

Thus, $\mathbf{s}^Q(\infty)$ is an unbiased estimate of our desired steady state result.

(3) *Second Order Analysis*

The second order statistics of the noise are an indicator of the uncertainty of the final result. Indeed, if the noise is Gaussian, the process is uniquely determined by its second order statistics. The covariance of $\mathbf{s}^Q(\infty)$ is

$$\begin{aligned} \mathrm{Cov}[\mathbf{s}^Q(\infty)] &= E[(\mathbf{r}_1+\mathbf{r}_2+\mathbf{r}_3+\mathbf{n}_d)(\mathbf{r}_1+\mathbf{r}_2+\mathbf{r}_3+\mathbf{n}_d)^T]-\mathbf{f}^Q\mathbf{f}^{QT} \\ &= \underline{C}_s + \underline{C}_i + \underline{C}_f + \underline{C}_d \end{aligned} \tag{7}$$

where $\underline{C}_s = E[\mathbf{r}_1\mathbf{r}_1^T]-\mathbf{f}^Q\mathbf{f}^{QT}$, $\underline{C}_i = E[\mathbf{r}_2\mathbf{r}_2^T]$, $\underline{C}_f = E[\mathbf{r}_3\mathbf{r}_3^T]$ and $\underline{C}_d = E[\mathbf{n}_d\mathbf{n}_d^T]$

Let the variances of each of the elements of $\underline{N}_3$ and $\underline{N}_4$ be σ_3^2 and σ_4^2 respectively. We can show [11] that if $\sigma_4^2<(1-||\underline{T}_4||^2)/Q$, we guarantee the convergence of $\mathbf{s}^Q(\infty)$. The covariance of the system noise, $\underline{C}_s$, can be shown [11] to be

$$\underline{C}_s = [\sigma_4^2||\mathbf{f}^Q||^2+\sigma_3^2||\mathbf{f}^P||^2]\gamma(\underline{I}-\underline{T}_4^2)^{-1} \tag{7,a}$$

where $1/\gamma = 1 - \sigma_4^2\mathrm{tr}[(\underline{I}-\underline{T}_4^2)^{-1}]$.

By assumption, $\underline{C}_d = \sigma_d^2\underline{I}$ for the both static and time varying case. The other covariances of the static and time varying cases, however, are different. We will consider each case separately.

STATIC : We can show that, if the feedback and input noise are static, that is, $E[\mathbf{n}_i\mathbf{n}_i^T] = \sigma_i^2\underline{I}$ and $E[\mathbf{n}_f\mathbf{n}_f^T] = \sigma_f^2\underline{I}$, then $\underline{C}_f$ and $\underline{C}_i$ are [11]

$$\underline{C}_f = \sigma_f^2(\underline{I}-\underline{T}_4)^{-2} + \sigma_f^2\sigma_4^2\mathrm{tr}[(\underline{I}-\underline{T}_4)^{-2}]\gamma(\underline{I}-\underline{T}_4^2)^{-1} \tag{7,b}$$

$$\underline{C}_i = \sigma_i^2(\underline{I}-\underline{T}_4)^{-1}\underline{T}_4+\sigma_i^2\{\sigma_4^2\mathrm{tr}[(\underline{I}-\underline{T}_4)^{-1}\underline{T}_4] + P\sigma_3^2\}\gamma(\underline{I}-\underline{T}_4^2)^{-1} \tag{7,c}$$

TIME VARYING : In the time varying case, from the assumption of white noise process, $E[\mathbf{n}_i(k)\mathbf{n}_i^T(l)] = \sigma_i^2\underline{I}\ \delta_{k-l}$ and $E[\mathbf{n}_f(k)\mathbf{n}_f^T(l)] = \sigma_f^2\underline{I}\ \delta_{k-l}$. Expression for $\underline{C}_f$ and $\underline{C}_i$ are follow as [11]

$$\underline{C}_f = \sigma_f^2\ \gamma\ (\underline{I}-\underline{T}_4^2)^{-1} \tag{7,d}$$

$$\underline{C}_i = \sigma_i^2(\underline{I}+\underline{T}_4)^{-1}\underline{T}_4+\sigma_i^2\{\sigma_4^2\mathrm{tr}[(\underline{I}+\underline{T}_4)^{-1}\underline{T}_4] + P\sigma_3^2\}\gamma(\underline{I}-\underline{T}_4^2)^{-1} \tag{7,e}$$

We now have all of the information required to evaluate (7) for both the static and white noise cases. After a review of the L-CANN, we will use this result to analyze the performance of the APNN.

III. THE L-CANN

(1) *Algorithm Review*

The APNN is homogeneous in the sense that any neuron can be clamped or floating. A projection based ANN with fixed stimulus (clamped) and response (floating) neurons is the L-CANN [6,7]. To increase storage capacity and enhance partition diversity, the network is augmented to have R hidden neurons which are generated by some nonlinear operation on the clamped neurons. Let $\mathbf{h}_n$ be the state of the hidden neurons corresponding to a stimulus of $\mathbf{f}^P{}_n$. Then,

$$\mathbf{h}_n = \underline{\vartheta}\ \mathbf{f}^P{}_n \qquad ;\quad 1 \le n \le N.$$

We form the hidden layer matrix

$$\underline{H} = [\mathbf{h}_1\ \mathbf{h}_2 \cdots \mathbf{h}_N]$$

and define that the interconnects from the hidden neurons to output neurons as

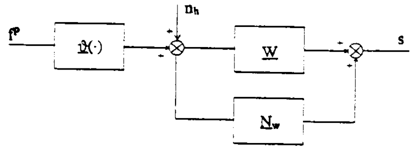

Fig.2 The block diagram of L-CANN with additive noise both at the hidden neurons and the hidden-to-output interconnects.

$$\underline{W} = \underline{F}_Q(\underline{H}^T\underline{H})^{-1}\underline{H}^T. \qquad (8)$$

Note that for a stimulus, $\mathbf{f}^P$, the output at the floating neurons will be

$$\mathbf{s} = \underline{W}\ \underline{\vartheta}\ \mathbf{f}^P = \underline{W}\ \mathbf{h} = \mathbf{f}^Q$$

The operation is one step as opposed to the iterative APNN. If $\underline{F}_Q$ has dimension of 1xN and its elements are bipolar (±1), we can use the L-CANN as a binary classifier.

(2) *Noise Model for the L-CANN*

In this section, we will analyze the effects of additive noise at hidden neurons and hidden-to-output interconnects. We assume that the noise is zero mean and spatially iid. Let $\mathbf{n}_h$ and $\underline{N}_w$ denote the additive noise at the hidden neurons and interconnects respectively as shown in Fig.2. Then the noisy output, $\mathbf{s}$, is

$$\mathbf{s} = (\underline{W}+\underline{N}_w)(\mathbf{h}+\mathbf{n}_h)$$

The expected value of $\mathbf{s}$ is

$$E[\mathbf{s}] = E[(\underline{W}+\underline{N}_w)(\mathbf{h}+\mathbf{n}_h)] = \underline{W}\ \mathbf{h} = \mathbf{f}^Q$$

which is our desired result. The covariance of $\mathbf{s}$ is

$$\mathrm{Cov}[\mathbf{s}] = E[(\mathbf{s}-\mathbf{f}^Q)(\mathbf{s}-\mathbf{f}^Q)^T]$$
$$= E[(\underline{W}\mathbf{n}_h+\underline{N}_w\mathbf{h}+\underline{N}_w\mathbf{n}_h)(\underline{W}\mathbf{n}_h+\underline{N}_w\mathbf{h}+\underline{N}_w\mathbf{n}_h)^T]$$

Because $\underline{N}_w$ and $\mathbf{n}_h$ are iid with variances σ_w^2 and σ_h^2 respectively, the covariance is

$$\mathrm{Cov}[\mathbf{s}] = \underline{C}_h + \underline{C}_w$$

where $\underline{C}_h = \sigma_h^2\underline{W}\underline{W}^T$ and $\underline{C}_w = \sigma_w^2[\,||\mathbf{h}||^2 + R\sigma_h^2]\ \underline{I}$

IV. PROBABILITY OF ERROR FOR BIPOLAR LIBRARY

In this section, we will discuss the probability of error for the bipolar (±1) library vectors for both the APNN and the L-CANN. In both cases, the bipolar response is obtained from noisy output by a *sign* operation. For the APNN, $\mathbf{s}_o = \mathrm{sign}[\mathbf{s}^Q(\infty)]$, and, for the L-CANN, $\mathbf{s}_o = \mathrm{sign}[\mathbf{s}]$. We will assume that all noise is Gaussian.

(1) *Union Bound of Probability of Error*

The probability of error is

$$P_e = \sum_{k=1}^{N} Q_k[1-P_{ck}]$$

where Q_k is the priori probability of the library vector $\mathbf{f}_k$ and P_{ck} is the probability that our classification is correct. We cannot, however, evaluate P_{ck} easily. We can, though, compute the *union bound* [10] for P_e. For Gaussian noise, a probability error bound for k-th library vector can be written as

$$P_{ek} = 1 - P_{ck}$$
$$\le \sum_{i=1}^{Q} \int_0^{\infty} \exp[-(x-1)^2/(2c_i)]/(2\pi c_i)^{1/2}\, dx$$

where c_i is the i-th diagonal element of covariance matrix $\underline{C}$. The probability of error is then represented by following inequality [11]

$$P_e \le \sum_{i=1}^{Q} \mathrm{erfc}(1/c_i^{1/2})$$
$$\le \mathrm{tr}[\underline{C}^{1/2}\exp(-\underline{C}^{-1}/2)]/(2\pi)^{1/2}$$

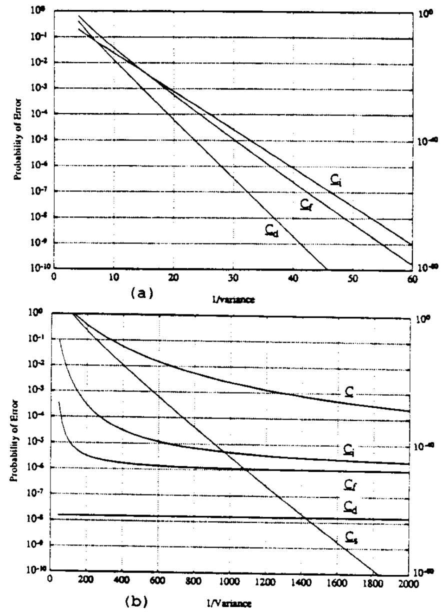

Fig.3 The error probabilities for the APNN with white Gaussian noise. The scale on right side is for $\underline{C}_i$. (a) The error probabilities decrease with the decrease of σ_d, σ_i and σ_f. We assume that σ_3 and σ_4 are all 1/36. (b) The error probabilities from $\underline{C}_i$, $\underline{C}_d$ and $\underline{C}_f$ saturate with the decrease of σ_3 and σ_4, but that from $\underline{C}_s$ decreases. σ_d, σ_i and σ_f are all 1/6.

(2) *Probability of Error for the APNN*

We demonstrate the probability error bounds for the 5 bipolar library of dimension L=30 which are generated randomly. The average value of the diagonal elements of $\underline{T}$ is 1/6. Fig.3a shows the probability of errors corresponding to $\underline{C}_d$, $\underline{C}_i$ and $\underline{C}_f$ vs. σ_d^2, σ_i^2 and σ_f^2 respectively when σ_3 and σ_4 are 1/36. From Fig.3a, the error bound decreases with the decrease of σ_d, σ_i and σ_f. Fig.3b shows the probability errors corresponding to $\underline{C}_d$, $\underline{C}_i$, $\underline{C}_f$ and $\underline{C}_s$ vs. σ_3^2 and σ_4^2 when σ_d, σ_i and σ_f are 1/6. The probability of error from $\underline{C}_d$, $\underline{C}_i$ and $\underline{C}_f$ saturates with the decreases of σ_3 and σ_4, but that from $\underline{C}s$ decreases as shown in Fig.3b. Fig.4 shows that the probability of error from $\underline{C}_d$, $\underline{C}_f$ and $\underline{C}_s$ decrease with the number of clamped neurons.

(3) *Probability of Error for the L-CANN*

In our L-CANN example, we choose the nonlinear function $\underline{\vartheta}$

$$\mathbf{h} = \underline{\vartheta}\ \mathbf{f}^P = \mathrm{sign}[\ \underline{Y}\mathbf{f}^P\ ]$$

where $\underline{Y}$ is generated randomly by a zero mean uniform distribution. [6]. We normalize the noise by the maximum interconnect value in order

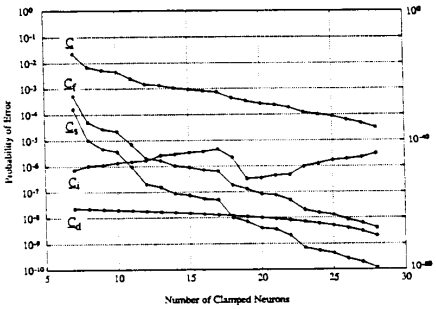

Fig.4 The error probability as a function of the clamped neurons for the APNN. The error probabilities from $\underline{C}_f$, $\underline{C}_d$ and $\underline{C}_3$ decrease with the increase of the percentage of clamped neurons. σ_d, σ_i and σ_f are all 1/6, σ_3 and σ_4 are all 1/36. The scale on right side is for $\underline{C}_i$.

to maintain a constant (maximum) SNR. Fig.5 shows that the probability of error $\underline{C}_w$ increases with the number of hidden neurons, but that from $\underline{C}_h$ decreases. We can show that the probability of error from $\underline{C}_h$ decreases with the number of hidden neurons [11]. Fig.5b shows that the probability of error generally decreases with the increase of the number of hidden neurons.

V. Conclusion

We have presented a preliminary noise analysis of the APNN and L-CANN. For the APNN, the error probability decreases with the increase of the number of clamped neurons. For the L-CANN, the covariances are strongly dependent on the number of hidden neurons. The error resulting from the noise at the hidden neurons decrease with an increase in the number of hidden neurons when the SNR of the interconnects is held constant.

Acknowledgement

This research was supported by the Washington Technology Center and SDIO/IST's ultrahigh speed computing program administered through the U.S. Offices of Naval Research in conjunction with the Optical System Laboratory at Texas Tech University.

References

[1] R.J. Marks II "A class of continuous level associative memory neural nets", **Applied Optics**, vol. 26, pp.2005-2009 (1987).

[2] R.J. Marks II, L.E. Atlas, S. Oh and K.F. Cheung "Optical processor architectures for alternating projection neural networks", **Optics Letters**, vol. 13, pp.533-535 (1988).

[3] R.J. Marks II, S. Oh, L.E. Atlas, and J.A. Ritcey "Alternating projection neural networks", **IEEE Trans. CAS**. (scheduled to appear in the June 1989 issue).

[4] R.J. Marks II, S. Oh, L.E. Atlas and J.A. Ritcey "Homogeneous and layered alternating projection neural networks", *Proceedings of the International Symposium on Optical Engineering and Industrial Sensing for Advanced Manufacturing Technologies,* 26-30 June 1988, Dearborn Hyatt, Michigan.

[5] R.J. Marks II, L.E. Atlas, S. Oh and J.A. Ritcey "The performance of convex set projection based neural networks", **Neural Information Processing Systems,** Dana Z. Anderson, editor, (American Institute of Physics, New York, 1988), pp. 534-543.

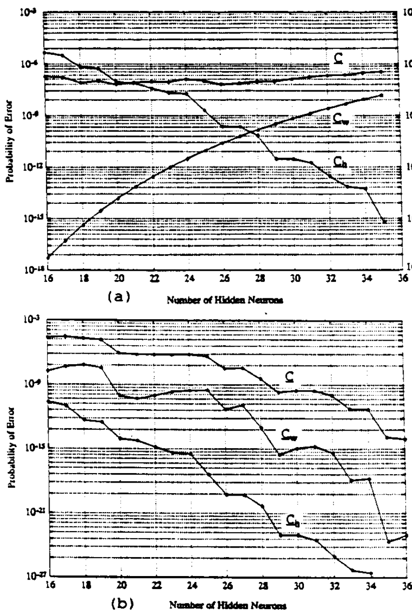

Fig.5 Error probabilities as a function of the number of hidden neurons for the L-CANN. (a) The error probability from $\underline{C}_h$ decreases with the increase of the number of hidden neurons. For $\underline{C}_w$, however, the error probability increases with the square of the norms of hidden states. σ_h and σ_w are 1/6 and 1/36 respectively. The scale on right side is for $\underline{C}_h$. (b) The error probability in the case where the interconnect noise is normalized by the maximum magnitude of the hidden-to-output interconnects. The error probabilities generally decrease with the increase of the number of hidden states. σ_h is 1/6, and the normalized variance of interconnects is $1/12^2$.

[6] R.J. Marks II, L.E. Atlas, D.C. Park and S. Oh "The effect of stochastic interconnects in artificial neural network classification", **Proceedings of the IEEE International Conference on Neural Networks,** San Diego, July 24-27, 1988, vol.II, pp.437-442.

[7] R.J. Marks II, L.E. Atlas and S. Oh, "Generalization in layered classification neural networks". **1988 IEEE International Symposium on Circuits and Systems,** pp. 503-506, Helsinki, 7-9 June, 1988.

[8] R.J. Marks II, J.A. Ritcey, L.E. Atlas and K.F. Cheung "Composite matched filter output partitioning", **Applied Optics,** vol. 26, pp.2274-2278 (1987).

[9] K.F. Cheung, L.E. Atlas, J.A. Ritcey, C.A. Green and R.J. Marks II "A comparison of conventional and composite matched filters with error correction", **Applied Optics,** vol. 26, pp.4235-4239 (1987).

[10] J. M. Wozencraft and I. M. Jacobs, **Principles of communication engineering**. New York: John Wiley and Sons Inc., 1965, pp.83, pp.264-266.

[11] The four page space limitation placed on this paper does not allow inclusion of certain proofs. A copy of the proofs can be obtained from the authors.