

# The influence of limited presynaptic growth and synapse removal on adaptive synaptogenesis

Dawn M. Adelsberger-Mangan, William B Levy

Department of Neurological Surgery, Box 420, University of Virginia Health Sciences Center, Charlottesville, VA 22908, USA

Received: 20 August 1993 / Accepted in revised form: 16 April 1994

**Abstract.** This report continues our research into the effectiveness of adaptive synaptogenesis in constructing feed-forward networks which perform good transformations on their inputs. Good transformations are characterized by the maintenance of input information and the removal of statistical dependence. Adaptive synaptogenesis stochastically builds and sculpts a synaptic connectivity in initially unconnected networks using two mechanisms. The first, synaptogenesis, creates new, excitatory, feed-forward connections. The second, associative modification, adjusts the strength of existing synapses. Our previous implementations of synaptogenesis only incorporated a postsynaptic regulatory process, receptivity to new innervation (Adelsberger-Mangan and Levy 1993a, b). In the present study, a presynaptic regulatory process, presynaptic avidity, which regulates the tendency of a presynaptic neuron to participate in a new synaptic connection as a function of its total synaptic weight, is incorporated into the synaptogenesis process. In addition, we investigate a third mechanism, selective synapse removal. This process removes synapses between neurons whose firing is poorly correlated. Networks that are constructed with the presynaptic regulatory process maintain more information and remove more statistical dependence than networks constructed with postsynaptic receptivity and associative modification alone. Selective synapse removal also improves network performance, but only when implemented in conjunction with the presynaptic regulatory process.

## 1 Introduction

In order to make the predictions, particularly the behaviorally relevant decisions, necessary for survival, an animal must create an efficient working model of its world. Creation of this working model requires that the animal interpret and store the associative and causal relationships of its environment. Finding these associations and

causal relationships is not easy from a computational complexity perspective because of the constantly changing, high-dimensional nature of the sensory input (Levy 1989). At the level of sensory processing, Barlow (1959) suggests combating this complexity by transforming the raw sensory input into a form which can be more easily stored and utilized by the brain. More specifically, sensory systems perform a series of recodings designed to remove the redundancy of the input and thereby increase the relative entropy of the neuronal representations (Barlow 1959, 1961a, b, 1985). In fact, relative entropy is a natural measure of statistical dependence (Watanabe 1960, 1969), or equivalently, from a coding theory viewpoint, of complexity.

Therefore, neural recoding may increase the efficiency of neural representations by transforming redundant, nonindependent patterns of neural firing into patterns of neural firing with less statistical dependence. However, any transformation of a neural representation for the purpose of removing statistical dependence risks information loss (Richards and Levy 1990). Therefore, when considering the quality of a transformation that reduces statistical dependence, it is important also to consider the information loss that is incurred with such a recoding.

In recent years, various groups have taken advantage of information theory measures to describe different neural recodings that remove statistical dependence (Adelsberger-Mangan and Levy 1992, 1993a, b; Atick 1992; Atick and Redlich 1992; Barlow and Földiák 1989; Földiák 1990; Levy 1985, 1989; Linsker 1988, 1990, 1992; Rolls 1989; van Hateren 1992). Of these, two groups, our own and the Földiák and Barlow collaboration, have concentrated on biologically plausible mechanisms for the construction of networks which perform such recodings. The networks of Földiák and Barlow contain full feed-forward connectivity between the input and output layer and full lateral connectivity between the neurons of the output layer. Three types of adaptive modification underlie the ability of these networks to reduce statistical dependence: synaptic modification of the excitatory feed-forward synapses, synaptic modification of the inhibitory lateral synapses, and individual modification of

the output neuron firing threshold. Taken together, these processes tend to remove all pairwise correlations between the output neurons.

Except for employing an associative modification rule for the modification of excitatory synapse weights, our networks are different. First, because we are inspired by the dentate gyrus region of the hippocampus, our networks are sparsely connected with excitatory, feed-forward connections. This contrasts with the fully connected architecture of Barlow and Földiák. Additionally, we (Adelsberger-Mangan and Levy 1993a, b; Levy and Colbert 1991; Levy and Desmond 1985) employ synaptogenesis to create the synaptic connectivity and have begun to use synapse removal to arbitrate the ultimate synaptic connectivity.

Previously, we investigated adaptive synaptogenesis (Adelsberger-Mangan and Levy 1993b) that only consisted of a postsynaptic regulatory process (receptivity to new innervation) and associative modification. Here we build on our previous investigations (Adelsberger-Mangan and Levy 1992, 1993b) by adding a presynaptic regulatory process, called avidity (Levy and Desmond 1985), and a rule for synapse removal, dropweight (Levy and Colbert 1991). As shown in the simulations, the addition of the avidity principle alone or, even better, the addition of avidity plus synapse removal produces significant improvements in the information measures that define network performance.

## 2 Simulations

### 2.1 Architecture

The neural networks have two neuronal layers, an input and an output layer. The number of neurons in the input layer is always 64. The number of output neurons is always 6.

Initially, there are no synaptic connections between the input and output layers. Synaptogenesis creates feed-forward excitatory synaptic connections between the input and output layers. No restriction is placed on the number of synaptic connections that may develop between an input/output neuron pair.

### 2.2 Neuronal properties

All neurons produce binary outputs  $\{0, 1\}$  corresponding to  $\{\text{not firing, firing}\}$  at a particular time step. At each time step  $t$ , the activity,  $Y_j(t)$ , of an output neuron  $j$  is determined by the input layer activities,  $X_i(t)$ , and by the synaptic strengths  $W_{i,j}(t)$  according to:

$$Y_j(t) = f\left(\sum_{i=1}^N \sum_{k=1}^K X_i(t) \cdot W_{i,j}(t)\right)$$

where  $k$  indexes multiple synapses between input neuron  $i$  and output neuron  $j$  and where:

$$f(s) = 1 \quad \text{if } s \geq \text{the output neuron firing threshold} \\ = 0 \quad \text{otherwise.}$$

### 2.3 Inputs

There are 25 unique environments that are used as the inputs to the network. Each environment contains 64 different 64-bit patterns; therefore each input environment contains 6.0 bits of representational information. The mean statistical dependence of the 25 environments is  $44.65 \pm 2.7$  bits (mean  $\pm$  standard deviation). All reported values are the means of the 25 environments.

### 2.4 Network development

All networks are initially unconnected, i.e., there are no synaptic connections between the input and output layers at time zero. The synaptic connectivity between the input layer and the output layer develops adaptively under the control of three local processes: synaptogenesis (the creation of new synapses), associative synaptic modification of existing synapses, and synapse removal. Synaptogenesis and synapse removal alter the synaptic connectivity between the input and output layers; associative modification adjusts the strengths of existing synapses. Changes in synaptic connectivity occur on a time scale 640 times longer than that of synapse weight modification (see the synapse construction section).

### 2.5 Associative synaptic modification

At each time step  $t$ , the synaptic weight of synapse  $k$  between input neuron  $i$  and output neuron  $j$  modifies according to:

$$W_{i,j}(t+1) = W_{i,j}(t) + \Delta W_{i,j}(t, t+1)$$

where:

$$\Delta W_{i,j}(t, t+1) = 0.05 \cdot Y_j(t) \cdot (X_i(t) - W_{i,j}(t))$$

A natural consequence of this synaptic modification rule and of the maximum and minimum values of each  $X_i$  is that individual synaptic strengths remain bounded within the range 0 to 1. Thus, excitatory synapses are not converted into inhibitory synapses. Nevertheless, the sum of the synaptic weights, that is, the connection strength between input neuron  $i$  and output neuron  $j$ , is from naive considerations unbounded because multiple synapses can exist between a given input/output neuron pair.

### 2.6 Changing synaptic connectivity

At the time steps where changes in the synaptic connectivity are permitted, two mechanisms are implemented: (1) construction of new synaptic connections and (2) removal of synaptic connections.

**2.6.1 Synapse construction.** During each synaptogenesis step, the creation of a new synapse between input neuron  $i$  and output neuron  $j$  occurs randomly with probability:

$$P_{ij}(t) = \gamma \cdot A_i(t) \cdot R_j(t)$$

where  $\gamma$  equals 0.005 and where  $A_i$  equals the presynaptic avidity for  $i$  reaching any postsynaptic site, and  $R_j$  equals

the postsynaptic receptivity of  $j$  for any presynaptic innervation. Because  $A_i$  and  $R_j$  each range from 0 to 1,  $P_{ij}$  ranges from 0 to 0.005. The particular value of  $\gamma$  is not important as long as it is small enough to guarantee that synapse construction is a rare process. After  $P_{ij}(t)$  is computed for an input/output neuron pair, a pseudo-random number generator determines whether a new synapse is placed between the pair. All newly created synapses have an initial weight of 0.50.

The avidity of an input neuron quantifies its tendency to participate in a new synaptic connection. For each input neuron, avidity is a monotonically decreasing function of the total synaptic weight supported by that input neuron. By restricting the participation of an input neuron in new synapse formation as a function of its total synaptic weight, we prevent the domination of the output layer, and thereby the output firing patterns, by a small number of input neurons.

Specifically, for input neuron  $i$ , the avidity at time step  $t$  is given by:

$$A_i(t) = \frac{a}{a + \sum_{j \neq k} W_{i,j}(t)^{10}}$$

The constant  $a$  is a variable of the study and is equal to  $1.0 \times 10^{33}$ , 1.0, or  $1.0 \times 10^{-3}$ . When  $a$  equals  $1.0 \times 10^{33}$ , the avidity of the input neurons remains, for all practical purposes, equal to 1.0 regardless of the number and weight of the neuron's synaptic connections; therefore, when this constant is employed, input neurons are 'unlimited by avidity'. When  $a$  equals 1.0, a total synaptic weight of 1.0 results in an avidity of 0.50. Under these conditions, if an input neuron makes four synapses, each of which remains at its initial synapse weight of 0.50, the input neuron has less than a  $3.0 \times 10^{-5}$  probability of making a new synapse with the output layer at each synaptogenesis step. This value of the avidity constant is referred to as 'moderate avidity'. Finally, we test an avidity constant of  $1.0 \times 10^{-3}$ : in this case, a total synaptic weight of 1.0 yields an input neuron avidity of only  $1.0 \times 10^{-3}$ . With this avidity constant, if a neuron maintains only two synapses at their initial weight of 0.50, it has less than a  $3.0 \times 10^{-5}$  probability of making a new synapse with the output layer at each synaptogenesis step. Therefore, this avidity constant is referred to as 'limiting avidity'. Note that avidity only affects the input neuron's ability to participate in a new synaptic connection; it exerts no effect on the strengthening of existing synapses.

The receptivity of a postsynaptic neuron to new synaptic innervation is a monotonically decreasing function of the neuron's recent firing history. Specifically, for postsynaptic neuron  $j$ , the receptivity at time step  $t$  is given by:

$$R_j(t) = \frac{r_1}{r_1 + \bar{y}_j(t)^2}$$

where  $\bar{y}_j(t)$  is the running average of the firing level of output neuron  $j$  and is calculated as:

$$\bar{y}_j(t) = 0.99\bar{y}_j(t-1) + 0.01 Y_j(t)$$

Before the simulations, the constants  $r_1$  and  $r_2$  were determined so that a receptivity of 1.00 results when the firing level of the output neuron is 0.40, and a receptivity of 0.001 results when the firing level of the output neuron is 0.50 (specifically,  $r_1 = 2.1977 \times 10^{-15}$ ,  $r_2 = 38.7286$ ). The specific shape of the receptivity function (and, correspondingly, specific values of  $r_1$  and  $r_2$ ) is of little importance; the only requirement is that  $R_j(t)$  equals 1.0 when the output neuron never fires and decreases to essentially zero as the firing level of the output neuron approaches 0.50. In fact, nearly identical results are obtained with a simple biphasic receptivity function which allows only two values, 1 when  $\bar{y}_j(t) < 0.5$ , and 0 when  $\bar{y}_j(t) \geq 0.5$  (unpublished observations).

Synaptogenesis begins at time step 1 by running the 384 ( $64 \times 6$ ) individual Bernoulli processes,  $P_{ij}$ , that control the creation of new synaptic connections. After the first synaptogenesis step, the process recurs periodically with a period length equal to 640 time steps. The rate of synapse formation is purposely kept very slow; this allows approximate convergence of the strengths of existing synaptic connections and  $\bar{y}_j(t)$  before the next bout of synaptogenesis (Levy and Desmond 1985; Levy and Colbert 1991). Not surprisingly, many opportunities for synapse creation are necessary to ensure output firing reaches the receptivity stabilizing value of 0.5. (The number of synaptogenesis opportunities required to attain output firing levels of 0.5 depends on the firing threshold, avidity constant, and whether synapse removal is employed; the number of opportunities required in this report ranged from 40 to 95 000.)

*2.6.2 Synapse removal.* Synapse removal is based on a minimum acceptable synaptic strength, the synaptic drop weight. Clearly, the drop weight must be less than 0.50, or all new synapses will be lost. More specifically, if the strength of a synapse between input neuron  $i$  and output neuron  $j$  falls below the firing probability of the input neuron, the synapse is removed. The computational rationale for this is as follows. Associative modification results in a synaptic weight which converges, approximately, to the probability that the presynaptic neuron is firing, given that the postsynaptic neuron is firing. Therefore, when a synaptic strength falls to a value which is less than the probability of presynaptic neuron firing, the firing of the postsynaptic neuron suggests that the presynaptic neuron is not firing, that is, the pre- and postsynaptic neuron are poorly correlated. Because we want to produce convergence between highly correlated inputs, such a presynaptic neuron is a poor choice for innervation to that postsynaptic neuron.

## 2.7 Simulation duration

The simulations continue until it is determined that the network connectivity has stabilized. Specifically, at each synaptogenesis step, the number of output neurons with a firing level greater than or equal to 0.50 is determined. If this number is 6 (that is, all the output neurons have sufficient input innervation to drive the receptivity of the output neurons to essentially zero), and there were no

alterations to the network connectivity at that step (that is, no synapses were removed or constructed), the simulation ends.

Between the synaptogenesis steps, the input layer patterns are presented randomly to the input layer for 640 time steps.

### 2.8 Measures

After the network is fully developed (using the mechanisms described above), the full input environment is presented to the network and the corresponding output patterns are archived to calculate the representational information and statistical dependence of the transformed environment.

Calling the input neuron layer space  $X$ , and the output neuron layer space  $Y$ , the information lost in the transformation of the input patterns into output patterns is:

$$H(X|Y) = H(X) - H(Y).$$

A measure of the effect of network transformation on the statistical dependence (Stat. Depend.) of the input layer firing pattern is determined from:

$$\Delta \text{Stat. Depend.} = \text{Stat. Depend. } X - \text{Stat. Depend. } Y$$

The statistical dependence of input space  $X$  is defined as  $\sum_i H(X_i) - H(X)$  (Watanabe 1969). For more details explaining these calculations, see Adelsberger-Mangan and Levy (1992, 1993b).

## 3 Results

The simulations demonstrate that decreasing the avidity constant, thereby restricting the ability of the input neurons to participate in a large number of synaptic connections, improves network performance. The simulations also illustrate that removal of synapses, when implemented in conjunction with avidity constants that restrict input neuron participation in synaptogenesis, further increases maintenance of the input information and reduction of input statistical dependence. Without limiting input neuron avidity, however, synapse removal is not useful and can instead be harmful to network performance.

### 3.1 Effect of avidity constant

**3.1.1 Output representational information.** Restricting the input neurons' participation in synaptogenesis produces networks that maintain more information than networks formed without this constraint. Over all thresholds tested, as seen in Fig. 1, the networks constructed with limiting input neuron avidity (avidity constant of  $1.0 \times 10^{-3}$ ) have higher levels of output representational information than those constructed with moderate input avidity (avidity constant 1.0). Similarly, over all thresholds tested, the networks constructed with moderate avidity have higher levels of output layer representa-

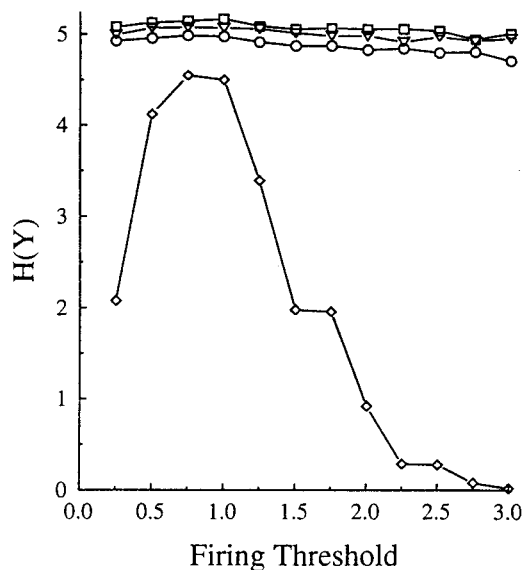
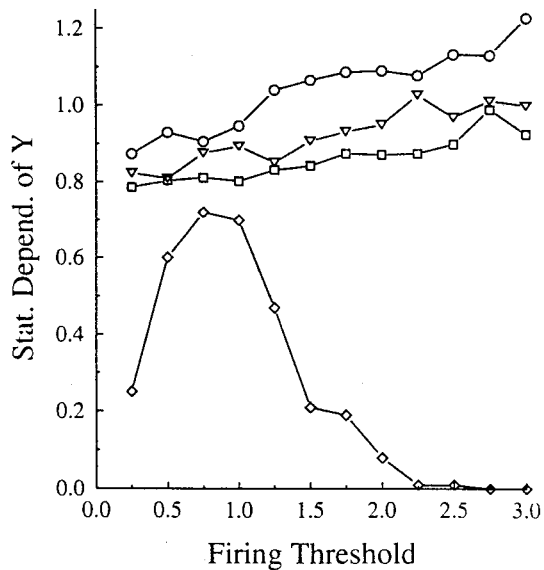


Fig. 1. Effect of avidity constant on output representational information as a function of firing threshold. The networks constructed with limiting avidity slightly, but consistently, produce more output representational information than networks constructed with greater levels of input neuron avidity. However, each of the three levels of input avidity, when incorporated into our adaptive synaptogenesis mechanism, yields networks which maintain high levels of output representational information over a wide range of output firing threshold. By contrast, networks with fixed, random connectivity fail to produce high levels of output representational information over most firing thresholds.  $H(Y)$ , Output representational information:  $\square$ , limiting avidity;  $\nabla$ , moderate avidity;  $\circ$ , unlimited avidity;  $\diamond$ , nonadaptively constructed networks with a fixed connectivity of 15.2% and a fixed connection strength of 0.367

tional information than those constructed with unlimited input avidity (avidity constant of  $1.0 \times 10^{33}$ ). Over the threshold range 0.25 to 3.0, the networks created with unlimited avidity produce a mean output representational information content of  $4.88 \pm 0.20$  bits ( $n = 300$ ). By comparison, over the same threshold range, the networks constructed with limiting avidity produce a mean output representational information content of  $5.07 \pm 0.13$  bits ( $n = 300$ ), an increase of 3.9%. While this increase is small, it is significant ( $P < 0.001$ , two-tailed, large sample difference of means test). Additionally, if we consider that the maximum level of output representational information that can be achieved is 6.0 bits, this average increase of 0.19 bits of representational information is 17% of what can be attained. The highest output representational information levels result with a firing threshold of 1.0 and limiting avidity. Under these conditions, the networks average 5.171 bits, 86.2% of the input information.

**3.1.2 Output statistical dependence.** Along with the increase of output layer representational information produced by limiting avidity, there is a corresponding decrease of output layer statistical dependence. As is seen in Fig. 2, over all thresholds tested, the networks constructed with limiting avidity have lower levels of output

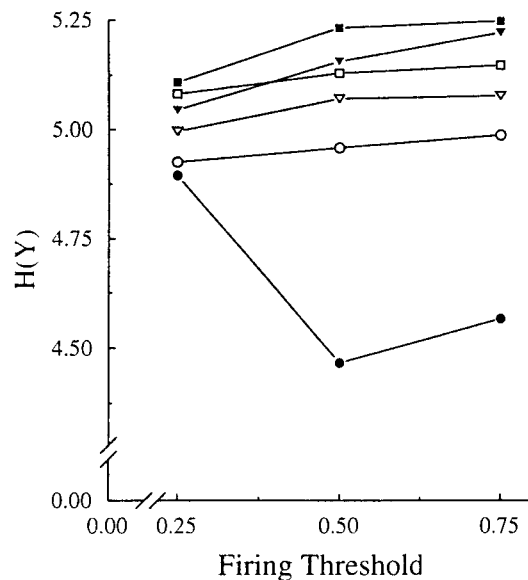


**Fig. 2.** Effect of avidity constant on output statistical dependence as a function of firing threshold. The networks constructed with limiting avidity consistently produce lower levels of output statistical dependence than networks constructed with greater levels of input neuron avidity. However, each of the three levels of input avidity, when incorporated into our adaptive synaptogenesis mechanism, yields networks which produce low levels of output statistical dependence over a wide range of output thresholds. Although the networks with fixed, random connectivity produce the lowest levels of statistical dependence, it should be remembered that these networks lose large amounts of input representational information over most firing thresholds.  $\circ$ , unlimited avidity;  $\square$ , moderate avidity;  $\triangle$ , limiting avidity;  $\diamond$ , nonadaptively constructed networks with a fixed connectivity of 15.2% and a fixed connection strength of 0.367

statistical dependence than those constructed with moderate avidity. Similarly, over all thresholds tested, the networks constructed with moderate avidity have lower levels of output statistical dependence than those constructed with unlimited avidity. Over the threshold range 0.25 to 3.0, the networks constructed with unlimited avidity produce a mean output statistical dependence level of  $1.04 \pm 0.21$  bits ( $n = 300$ ), and those constructed with limiting avidity produce  $0.86 \pm 0.13$  bits ( $n = 300$ ), a 17.3% decrease ( $P < 0.001$ ). The lowest levels of output statistical dependence result in networks constructed with a threshold of 0.25 and limiting avidity. Under these conditions, the networks average  $0.79 \pm 0.12$  bits of output statistical dependence, a reduction of over 98.2% of the input statistical dependence.

### 3.2 Interaction of input neuron avidity and synapse removal

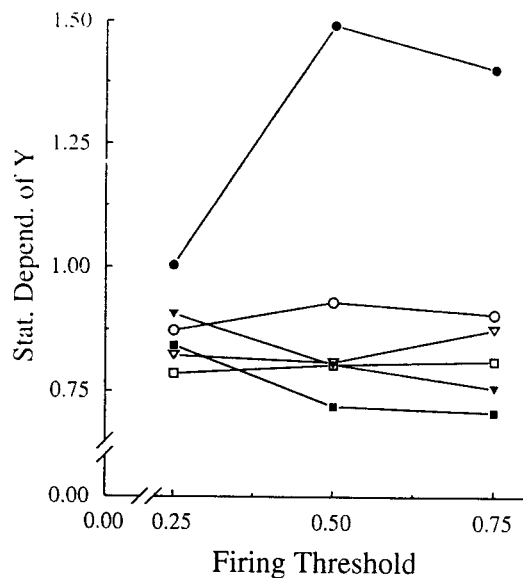
**3.2.1 Output representational information.** As can be seen in Fig. 3, synapse removal can increase or decrease the maintenance of the input information depending on the constant controlling avidity values. When the avidity constant restricts the ability of the presynaptic neuron to participate in a large number of synaptic connections (avidity constants 1.0 and  $1.0 \times 10^{-3}$ ), synapse removal increases input information maintenance. However,



**Fig. 3.** Effect of avidity constant and synapse removal on output representational information as a function of firing threshold. Synapse removal combined with limiting or moderate input avidity yields networks which consistently produce higher levels of output representational information than networks constructed without synapse removal. In contrast, synapse removal combined with unlimited input avidity yields networks which produce less output representational information than networks constructed without synapse removal.  $H(Y)$ . Output representational information:  $\circ$ , unlimited avidity;  $\square$ , moderate avidity;  $\triangle$ , limiting avidity;  $\diamond$ , nonadaptively constructed networks with a fixed connectivity of 15.2% and a fixed connection strength of 0.367

when input neurons are unlimited by avidity, synapse removal decreases input information maintenance. Specifically, at a threshold of 0.75 and with input neuron avidity unlimited, synapse removal reduces output representational information by 8.2% ( $4.58 \pm 0.61$  bits versus  $4.99 \pm 0.15$  bits,  $n = 25$ ,  $P < 0.002$ ). By contrast, when input avidity is restricted, synapse removal increases output representational information by 2.7% ( $5.22 \pm 0.11$  bits versus  $5.08 \pm 0.13$  bits with moderate avidity,  $n = 25$ ,  $P < 0.001$ ;  $5.25 \pm 0.09$  bits versus  $5.15 \pm 0.10$  bits with limiting avidity,  $n = 25$ ,  $P < 0.001$ ). The largest levels of output layer representational information result with a firing threshold of 0.75, with limiting avidity and synapse removal. Under these conditions, the networks average 5.25 bits of output representational information, that is 87.5% of the information presented at the input layer and 5.2% more of the input information maintained in networks constructed with unlimited avidity and without synapse removal ( $5.25 \pm 0.09$  bits versus  $4.99 \pm 0.15$  bits,  $n = 25$ ,  $P < 0.001$ ). Considering that the maximum level of output representational information that can be achieved is 6.0 bits, a 0.26 bit increase in information maintenance represents a gain of 25% of the total improvement possible.

**3.2.2 Output statistical dependence.** As can be seen in Fig. 4, synapse removal can increase or decrease output



**Fig. 4.** Effect of avidity constant and synapse removal on output statistical dependence as a function of firing threshold. Synapse removal combined with limiting or moderate input avidity yields networks which typically produce lower levels of output statistical dependence than networks constructed without synapse removal. In contrast, synapse removal combined with unlimited avidity yields networks which produce more output statistical dependence than networks constructed without synapse removal.  $\circ$ , unlimited avidity; no synapse removal;  $\bullet$ , unlimited avidity; synapse removal;  $\square$ , moderate avidity; no synapse removal;  $\blacksquare$ , moderate avidity; synapse removal;  $\triangle$ , limiting avidity; no synapse removal;  $\blacktriangle$ , limiting avidity; synapse removal

statistical dependence depending on the constant controlling avidity values. Synapse removal, when implemented with avidity constants which restrict the ability of an input neuron to participate in synaptogenesis (constants 1.0 and  $1.0 \times 10^{-3}$ ), reduces output layer statistical dependence at thresholds 0.50 and 0.75. Conversely, synapse removal, when input neurons are unlimited by avidity, increases output layer statistical dependence over all firing thresholds tested (0.25, 0.50, and 0.75). At a firing threshold of 0.75 and with unlimited input avidity, synapse removal increases output statistical dependence by 54.8% ( $1.40 \pm 0.60$  bits versus  $0.91 \pm 0.16$  bits,  $n = 25$ ,  $P < 0.001$ ). However, at the same threshold and with avidity values that restrict synaptogenesis, synapse removal decreases output statistical dependence by 13.0% ( $0.76 \pm 0.12$  bits versus  $0.88 \pm 0.13$  bits with moderate avidity,  $n = 25$ ,  $P < 0.002$ ;  $0.71 \pm 0.08$  bits versus  $0.81 \pm 0.10$  bits with limiting avidity,  $n = 25$ ,  $P < 0.001$ ).

### 3.3 Networks with fixed architecture

Networks constructed without any adaptive processes, that is, ones in which the synapses are placed by a random process and the synapse strengths are fixed, do not perform as well as networks constructed adaptively (see Figs. 1 and 2). The connectivity of these nonadaptively constructed networks is 15.2%; all synapses have strength 0.367. [These values are the averages from networks constructed with unlimited input neuron avidity,

without synapse removal, and with the threshold (0.75) that maximizes such networks' performance.] By comparing the results obtained in such nonadaptively formed networks with those obtained in adaptively constructed networks, we get a better idea of the improvements produced by adaptive synaptogenesis. As is clearly evident in Figs. 1–4 (as well as Figs. 1 and 2 in Adelsberger-Mangan and Levy 1993b), adaptive synaptogenesis is robust across threshold settings. By contrast, networks constructed with fixed synaptic connectivity fail to maintain the input information at nearly all firing thresholds tested (Fig. 1).

## 4 Discussion

The simulations of this and our previous report (Adelsberger-Mangan and Levy 1993b) demonstrate that adaptive synaptogenesis provides a robust mechanism for creating networks that, as defined by the information measures used here, improve neural representations. Our previous report illustrates the utility of the postsynaptic receptivity mechanism, which forces output neuronal firing to a preset value, and of associative weight modification, which adjusts arbitrary and probably suboptimal choices of initial synapse strength. The present simulations illustrate the utility of two additional mechanisms for improving network performance, input neuron avidity and synapse removal. In particular, the simulations of this report demonstrate that networks adaptively constructed under the additional constraint of limited input avidity maintain more information and produce less output statistical dependence than networks constructed with only the adaptive processes of receptivity and associative modification. Synapse removal is also helpful but only when used in concert with an avidity constant which limits the ability of the input neurons to participate in synaptogenesis. Specifically, the networks constructed with limiting input neuron avidity (avidity constant of  $1.0 \times 10^{-3}$ ) and employing synapse removal, maintained 5.2% more of the input information and removed 21.8% more input statistical dependence than networks with unlimited input avidity (avidity constant of  $1.0 \times 10^{33}$ ) and without synapse removal (when output neuron firing threshold was optimized).

It must be noted that the avidity constant that produces the best performance here may not be optimal. That is, we make no claim that the avidity constant of  $1.0 \times 10^{-3}$ , the constant which resulted in the best performance in the networks constructed for this report, is the one which maximizes network performance, and it may be that less avidity produces better performance. However, there is a cost to increasing network performance by reducing the avidity constant: the rate of network convergence. Indeed, an avidity constant which is too small may result in networks which do not converge in terms of postsynaptic activity. For an extreme example, suppose the avidity of the input neurons is zero, while the receptivity of the output neurons is greater than zero. Under these conditions, it is impossible for the output neurons to attain the desired firing level, and as a result input

information will be lost. A less extreme version of this argument specifies an avidity constant allowing only one synapse per presynaptic neuron. In fact, the smallest avidity used here essentially prohibits more than two synapses of initial strength per presynaptic neuron. The problem of setting the avidity constant so that it exerts a strong effect on the network construction without being ultimately limiting is a matter of trial and error for us, but this procedure should not be viewed as unbiological because natural selection could use this same process of trial and error.

The networks created by the adaptive processes, specifically, synaptogenesis, associative modification of synapse weights, and synapse removal, are distinct in several ways when compared with other networks which perform information maintaining recodings. The first distinction of our adaptively constructed networks is their very simple architecture; they are composed only of excitatory feed-forward connections. This is in contrast to the majority of networks designed to perform information maintaining transformations on their inputs. Typically, such networks are fully connected between the input and output layers and additionally have lateral connectivity between the output neurons (Linsker 1992; Oja 1989; Rubner and Tavan 1989; Sanger 1989a, b). However, such fully connected networks are not typical of cortical networks. Additionally, networks with only feed-forward connectivity provide for a fast input-output transformation (Adelsberger-Mangan and Levy 1993a). In contrast, the lateral connections of Földiák's networks force a series of recurrent operations to achieve output representation convergence (Földiák 1990). From a biological perspective, the ability to perform transformations quickly, is clearly desirable in facilitating an animal's successful interaction with its environment.

Our adaptively constructed networks also provide an alternative for control of output firing levels. That is, the adaptively constructed networks adjust network connectivity to create a desired level of output firing. This is in contrast to firing threshold adaptation, the more commonly employed mechanism of controlling neuronal activity (Adelsberger-Mangan and Levy 1992; Buckingham and Willshaw 1993; DeSieno 1988; Földiák 1990; Grossberg 1976). In Földiák's network, output neuron firing levels are regulated through a combination of adjustable firing threshold and lateral interactions. However, this mechanism for regulating firing threshold does not always guarantee stable output codes (Földiák 1990). By contrast, our research demonstrates that regulating firing levels by adjusting feed-forward connectivity provides a robust mechanism for transforming a wide variety of input environments over a range of firing thresholds (see Figs. 1 and 2 and also Adelsberger-Mangan and Levy 1992). Additionally, while the receptivity constants chosen in this research resulted in output firing levels of 50%, alternate receptivity constants can be chosen to create output neurons with other firing levels (unpublished observations).

The major distinction of the adaptive mechanisms explored in this research is that they employ biologically plausible mechanisms for network construction. The

implementation of synaptogenesis, as well as the rule for synapse weight modification and synapse removal, is rather simple and only requires information available locally at the neuron or synapse. Indeed, it would be surprising if Nature did not use this or some related adaptive mechanisms in wiring the networks of the brain.

*Acknowledgements.* This work was supported by NIH MH48161, MH00622 and RR07864 to W.B.L., by NIH MH09967 to D.M.A., by NSF MSS-9216372 and EPRI RP8030-08 to P. Papanicolaou-Kazakos, and by the Department of Neurosurgery, Dr. John A. Jane, Chairman.

## References

- Adelsberger-Mangan DM, Levy WB (1992) Information maintenance and statistical dependence reduction in simple neural networks. *Biol Cybern* 67:469-477
- Adelsberger-Mangan DM, Levy WB (1993a) Character recognition using adaptively constructed feed-forward networks. *Proc World Congress Neural Networks II*:423-426
- Adelsberger-Mangan DM, Levy WB (1993b) Adaptive synaptogenesis constructs networks that maintain information and reduce statistical dependence. *Biol Cybern* 70:81-87
- Atick JJ (1992) Could information theory provide an ecological theory of sensory processing? *Network* 3:213-251
- Atick JJ, Redlich AN (1992) What does the retina know about natural scenes? *Neural Comput* 4:196-210
- Barlow HB (1959) Sensory mechanisms, the reduction of redundancy, and intelligence. (National Physical Laboratory Symposium no. 10. The mechanisation of thought processes) Her Majesty's Stationary Office, London, pp 537-559
- Barlow HB (1961a) The coding of sensory messages. In: Thorpe WH, Zangwill OL (eds) *Current problems in animal behavior*. Cambridge University Press, Cambridge, UK, pp 331-360
- Barlow HB (1961b) Possible principles underlying the transformations of sensory messages. In: Rosenblith W (ed) *Sensory communication*. MIT Press, Cambridge, Ma., pp 217-234
- Barlow HB (1985) Perception: what quantitative laws govern the acquisition of knowledge from the senses? In: Coen CW (ed) *Functions of the brain*. Clarendon Press, Oxford, pp 11-43
- Barlow HB, Földiák P (1989) Adaptation and decorrelation in the cortex. In: Durbin RM, Miall C, Mitchison GJ (eds) *The computing neuron*. Addison-Wesley, Wokingham, pp 54-72
- Buckingham J, Willshaw D (1993) On setting unit thresholds in an incompletely connected associative net. *Network* 4:441-459
- DeSieno D (1988) Adding a conscience to competitive learning. *IEEE Int Conf Neural Networks* 1:117-124
- Földiák P (1990) Forming sparse representations by local anti-Hebbian learning. *Biol Cybern* 64:165-170
- Grossberg S (1976) Adaptive pattern classification and universal recoding. 1. Parallel development and coding of neural feature detectors. *Biol Cybern* 23:121-134
- Hateren van J (1992) A theory of maximizing sensory information. *Biol Cybern* 68:23-29
- Levy WB (1985) An information/computation theory of hippocampal function. *Soc Neurosci Abst* 11:493
- Levy WB (1989) A computational approach to hippocampal function. In: Hawkins RD, Bower GH (eds) *Computational models of learning in simple neural systems*. Academic, San Diego, pp 243-305
- Levy WB, Colbert CM (1991) Adaptive synaptogenesis can complement associative modification/depression. In: Commons ML (ed) *Neural network models of conditioning: quantitative analyses of behavior*, 13. Lawrence Erlbaum, Hillsdale, pp 53-68
- Levy WB, Desmond NL (1985) The rules of elemental synaptic plasticity. In: Levy WB, Anderson JA, Lehmkuhle S (eds) *Synaptic modification, neuron selectivity, and nervous system organization*. Lawrence Erlbaum, Hillsdale, pp 105-121

- Linsker R (1988) Self-organization in a perceptual network. *Computer* 21:105-117
- Linsker R (1990) Perceptual neural organization: some approaches based on network models and information theory. *Annu Rev Neurosci* 13:257-281
- Linsker R (1992) Local synaptic learning rules suffice to maximize mutual information in a linear network. *Neural Comput.* 4:672-683
- Oja E (1989) Neural networks, principle components, and subspaces. *Int J Neural Syst* 1:61-68
- Richards DSP, Levy WB (1990) Optimum preprocessing networks and a data processing theorem. *Proc Int J Conf Neural Networks* 1:19-22
- Rolls ET (1989) Information processing in the taste system of primates. *J Exp Biol* 146:141-164
- Rubner J, Tavan P (1989) A self-organizing network for principle component analysis. *Europhys Lett* 10:693-698
- Sanger TD (1989a) Optimal unsupervised learning in a single-layer linear feed forward neural network. *Neural Networks* 2:459-473
- Sanger TD (1989b) An optimality principle for unsupervised learning. In: Touretzky DS (ed) *Advances in Neural Information Processing Systems, I*. Morgan Kaufmann, San Mateo, pp 11-19
- Watanabe S (1960) Information theoretical analysis of multivariate correlation. *IBM J Res Dev* 4:66
- Watanabe S (1969) *Knowing and guessing: a quantitative study of inference and information*. Wiley, New York