

The statistical relationship between connectivity and neural activity in fractionally connected feed-forward networks

William B. Levy¹, Hakan Deliç², Dawn M. Adelsberger-Mangan¹

¹ Department of Neurological Surgery, Health Sciences Center Box 420, University of Virginia, Charlottesville, VA 22908, USA

² Signal and Image Processing Laboratory, Department of Electrical and Electronics Engineering, Boğaziçi University, Bebek 80815 Istanbul, Turkey

Received: 18 October 1996 / Accepted in revised form: 6 October 1998

Abstract. It is desirable to have a statistical description of neuronal connectivity in developing tractable theories on the development of biological neural networks and in designing artificial neural networks. In this paper, we bring out a relationship between the statistics of the input environment, the degree of network connectivity, and the average postsynaptic activity. These relationships are derived using simple neurons whose inputs are only feed-forward, excitatory and whose activity is a linear function of its inputs. In particular, we show that only the empirical mean of the pairwise input correlations, rather than the full matrix of all such correlations, is needed to produce an accurate estimate of the number of inputs necessary to attain a prespecified average postsynaptic activity level. Predictions from this work also include distributional aspects of connectivity and activity as shown by a combination of analysis and simulations.

1. Introduction

This present report explores the relationship between the statistics of connectivity and the statistics of activity of a simple feed-forward network. The analyses here grow out of two concerns. First, there are fundamental limits on the number of connections that exist in the brain or in artificial neural networks (Amaral et al. 1990; Mitchison 1991). Second, activity levels profoundly interact with the ability of feed-forward networks to reduce statistical dependence while maintaining information (Adelsberger-Mangan and Levy 1993).

Here we investigate a relationship between these two limiting processes for a specific and widely used synaptic modification rule (Földiák 1990; Grossberg 1976; Rolls 1989). As a starting point, we introduce the idea that there is a limit on the connectivity of the postsynaptic

neuron. In our analyses, we will individually hypothesize one of the following three types of postsynaptic constraints on connectivity: (A) the postsynaptic neuron can only maintain, or only has room for, a total synaptic weight of some fixed preset value, or (B) the postsynaptic neuron only has room for a fixed number of inputs, or (C) the postsynaptic neuron dynamically regulates synaptogenesis (that is, the creation of new synapses between the input and output neurons) in order to achieve a preset level of average activity.

Each of these limits has its appeal. Hypotheses A and B seem most relevant to an engineer with design constraints due to limitations on the size of an integrated circuit and its components, including the interconnections between neuron-like elements (Mead 1989; Mitchison 1991). Hypothesis A seems biologically sensible in terms of neuronal stability as well as the capability for plasticity in connections that maintain functionality when some of the neurons are lost (Hillman and Chen 1984). Hypothesis C is supported by neurobiological data that indicate a dynamic, activity-based regulation of connectivity (its neural relevance is discussed in Levy et al. 1990; Levy and Desmond 1985, and the developmental perspective is explored in Levy and Colbert 1991). In this regulatory theory, decreased postsynaptic activity produces more synaptogenesis. In essence, Hypothesis C introduces an additional adaptive process that allows us to consider – and Nature to use – stochastically driven synaptogenesis to achieve desirable neuronal architectures and network characteristics (e.g., redundancy reduction that largely preserves information).

The relationships between network connectivity and neural activity are important to the computational neuroscientist seeking to understand brain function. In our research, we are investigating a developmental model of synapse creation that is sensible and compatible with associative synaptic modification. Using a biologically inspired synaptogenesis rule, synapse creation continues until each output neuron has met a specific requirement, e.g., a desired firing level (Adelsberger-Mangan and Levy 1993) or decoding accuracy (unpublished observations). Indeed, it is generally assumed that

the brain tends to have set point levels of activity. Furthermore, energy-efficient transmission of information puts a set point on activity levels (Levy and Baxter 1996). The networks created using synapse creation and weight modification end up fractionally connected; indeed, we have found that networks where the output neurons are innervated by a small subset of the input neurons outperform (as determined by information maintenance and decoding accuracy) networks that are more densely connected (Adelsberger-Mangan and Levy 1993). With a better understanding of the statistical relationships between network connectivity and neural activities, we can refine our developmental model and increase our understanding of hippocampal and neocortical functions. Further, as regulation of neural activity may play a role in other cognitive functions (for a discussion of the role of neural activity in sensory processing, see for example Barlow 1959; Field 1994; Földiák 1990), an understanding of the interaction between neural activity and synaptic connectivity is of general interest.

The organization of the paper is as follows. In Section 2, we investigate the relationship between different neuronal parameters for a particular synaptic modification rule. Section 3 connects the number of input neurons to average postsynaptic activity via an approximation of the input correlation matrix. In Section 4, a recursive method for the solution of the eigenvalue problem leads to an expression for the distribution of the average postsynaptic activity. Simulation results evaluating the estimates developed in the paper are presented in Section 5, and a discussion on extending the results to recurrent nonlinear networks is provided in Section 6. Some concluding remarks can be found in Section 7.

2 Single neuron statistics

In this paper, we only consider the simplest feed-forward neuron or network of neurons whose inputs are only excitatory and whose activity is a linear function of its inputs. (When neurons are considered as linear devices, neural output equals activity; alternatively, when neurons are considered nonlinear devices, neural output is a nonlinear function of activity.) The most basic statistic of such a network is the average activity of a postsynaptic neuron, which is essential in determining the information capacity of the network (Adelsberger-Mangan and Levy 1993).

As a first-order approximation, we assume a linear activation expression:

$$\mathbf{Y}_j = c_y \mathbf{X}^T \mathbf{W}_j = c_y \sum_{i \in S_j} x_i w_{ij} \quad (1)$$

where \mathbf{Y}_j is the activation of the j th postsynaptic neuron, c_y is a constant applicable to all neurons of this class, and $\mathbf{X}^T \mathbf{W}_j$ is the inner product between the vector of afferent activities \mathbf{X} , made up of presynaptic activities $x_i \in \{0, 1\}$, and the vector of synaptic weights \mathbf{W}_j , made up of synapse strengths w_{ij} . S_j is the subset of inputs that

synapse on (i.e., connect to) neuron j . The vectors \mathbf{Y}_j , \mathbf{X} , \mathbf{W}_j and their elements are all implicit functions of time t . Then, if synaptic weights are fixed, the average activity of a postsynaptic neuron is

$$E[\mathbf{Y}_j] = c_y E_j[\mathbf{X}^T \mathbf{W}_j] = c_y \sum_{i \in S_j} E_j[x_i] w_{ij}$$

where $E_j[x_i] = E[x_i | i \in S_j]$ and $E_j[\mathbf{X}] = E[\mathbf{X} | i \in S_j, \forall i]$ are the conditional expectation operators.

Even a crude approximation that answers our initial question about the interrelationship between connectivity and the statistics of the input excitation and output activity is useful for a start. Suppose that the average presynaptic activity is nearly the same across all afferents; that is, $E_j[x_i] \approx E_j[x_k], \forall i, k \in S_j$. Then,

$$E[\mathbf{Y}_j] \approx c_y E_j[x_i] \sum_{i \in S_j} w_{ij} \quad (2)$$

and

$$\frac{E[\mathbf{Y}_j]}{E_j[x_i]} \approx c_y \sum_{i \in S_j} w_{ij} \quad (3)$$

This is the form which is most interesting when we consider a neuron governed by Hypothesis A that fixes $\sum_i w_{ij}$. In this crude approximation, we see a proportionality between the total synaptic weight on a postsynaptic neuron and the ratio of average input activities and output activities.

Suppose synaptic modification of the existing vector of synapses \mathbf{W}_j on neuron j occurs in a manner that is governed by a particular associative modification rule that receives experimental support both in the visual cortex (Rauschecker and Singer 1979, 1981) and in the hippocampus (Levy and Steward 1983).

$$w_{ij}(t+1) = w_{ij}(t) + \Delta w_{ij}(t)$$

$$\Delta w_{ij}(t) = \varepsilon \mathbf{Y}_j(t) [c_w x_i(t) - w_{ij}(t)]$$

where t is time, ε is a small positive constant controlling the modification rate, and c_w is a positive constant of suitable units so that the indicated subtraction makes sense.

As per the theorem in Appendix I, for small ε and a stationary, strongly mixing stochastic environment, this modification rule leads to a globally, asymptotically ($t \rightarrow \infty$) stable state of synaptic strengths:

$$\mathbf{W}_j = \frac{c_w \mathbf{e}_j^1 \lambda_j^1}{E\left[\sum_{i \in S_j} x_i e_{ij}^1\right]} \quad (4)$$

where the pair $(\lambda_j^1, \mathbf{e}_j^1)$ is, respectively, the dominant eigenvalue and the associated eigenvector of the correlation matrix $E_j[\mathbf{X}\mathbf{X}^T]$ of the subset of inputs S_j which connect to neuron j , and e_{ij}^1 is the i th element of the vector \mathbf{e}_j^1 . That is, this eigenpair is defined by

$$E_j[\mathbf{X}\mathbf{X}^T] \mathbf{e}_j^1 = \lambda_j^1 \mathbf{e}_j^1$$

where $\mathbf{e}_j^T \mathbf{e}_j^1 = 1$, and the absolute value of λ_j^1 is the largest eigenvalue of $E_j[\mathbf{X}\mathbf{X}^T]$.

With this asymptotic result, we can improve upon our first approximation by dispensing with the assumption that the average presynaptic activities are all the same. Taking synaptic strengths to be at their asymptotic values, we again have average postsynaptic activity as in (2). This time, however, we do not assume identical average activities; rather, we note that multiplication of (4) by $E_j[\mathbf{X}^T]$ gives

$$E_j[\mathbf{X}^T] \mathbf{W}_j = c_w \lambda_j^1$$

Combining this equation with (1) shows that the average activity of neuron j is solely a function of the pairwise correlations of the inputs; that is,

$$E[\mathbf{Y}_j] = c_y c_w \lambda_j^1 \quad (5)$$

This relationship might be of use when engineering a network for a desired level of output activity; that is, the constants c_y and c_w can be adjusted to achieve this activity level.

In addition, from (4) also develops a more precise expression that relates the total synaptic strength onto j as a function of the input statistics

$$\sum_{i \in S_j} w_{ij} \equiv \mathbf{1}^T \mathbf{W}_j = \frac{c_w \lambda_j^1 \mathbf{1}^T \mathbf{e}_j^1}{E_j[\mathbf{X}^T] \mathbf{e}_j^1} \quad (6)$$

$$= \frac{E[\mathbf{Y}_j] \mathbf{1}^T \mathbf{e}_j^1}{c_y E_j[\mathbf{X}^T] \mathbf{e}_j^1} \quad (7)$$

where $\mathbf{1}$ is the vector of all ones of appropriate dimension. Equation (6) shows that the total synaptic strength depends, via the dominant eigenpair, on these same correlational statistics and, moreover, on the average activity of the afferents synapsing on j . In (7), we see the precise relationship that (3) only approximated. That is, the crude assumption of identical average activity used to produce (3) is replaced by a weighted average over all the relevant afferents where the weighting $\mathbf{1}^T \mathbf{e}_j^1 / E_j[\mathbf{X}^T] \mathbf{e}_j^1$ depends on the correlation matrix $E_j[\mathbf{X}\mathbf{X}^T]$.

3 Network statistics

So far, we have been discussing only the statistics of a single postsynaptic neuron. Now we turn to a more challenging issue: the distribution of these statistics across a population of neurons. In order to take on this challenge, we need to consider the empirical distributions of subsets of the correlation matrix $E[\mathbf{X}\mathbf{X}^T]$. The setting for these empirical distributions seems quite natural for any of the three hypotheses mentioned in the Introduction. Specifically, we assume a synaptogenesis process such that the synapses between the set of inputs $i \in \{1, 2, \dots, n\}$ are made equiprobably with one synapse added at a time. Furthermore, we assume that synaptogenesis stops before all afferents are connected

to all postsynaptic neurons. In the case of Hypothesis B, synaptogenesis continues until each neuron has a complement of m_j synapses, where $m_j < n$. In the case of Hypotheses A and C, we allow each postsynaptic neuron to make one new synaptic connection and then run the input environment often enough so that a steady state is achieved for all synapse weights w_{ij} . This cycle of synaptogenesis followed by asymptotic associative modification continues until the individual criterion of the appropriate hypothesis is achieved. [This procedure can be sped up by allowing a lot more synaptogenesis to occur early on (Levy and Colbert 1991). The important point is that synaptogenesis must ultimately be a much slower process than associative modification, and this requirement fits biological intuition.]

Each postsynaptic neuron is likely to receive a different set of inputs. To indicate this more precisely, we introduce some notation to designate the subset of the inputs that synapse on any one particular postsynaptic neuron. Call such a subset of inputs to the j th postsynaptic neuron $S_j \subset \{1, 2, \dots, n\}$. Then, consider the related function that just selects a lower dimensional subspace of the random vector \mathbf{X} : $S_j(\mathbf{X}) \in [0, 1]^{m_j}$ where $m_j = |S_j|$ is the number of synapses. Because we want to arrive at distributional statements about the population of postsynaptic neurons, we must turn our attention to items such as the correlation matrix of the inputs to neuron j , $E_j[\mathbf{X}\mathbf{X}^T]$. Of course, the elements in such a correlation matrix appear in the larger correlation matrix $E[\mathbf{X}\mathbf{X}^T]$ as well. In fact, we are repeatedly sampling from $E[\mathbf{X}\mathbf{X}^T]$ to create the various matrices $E_j[\mathbf{X}\mathbf{X}^T]$.

Consider the following empirical expectations of the elements of the matrix $E_j[\mathbf{X}\mathbf{X}^T]$:

$$\xi_j = \frac{1}{m_j(m_j - 1)} \sum_{i=1}^{m_j} \sum_{\substack{k=1 \\ k \neq i}}^{m_j} E_j[x_i x_k]$$

and

$$\zeta_j = \frac{1}{m_j} \sum_{i=1}^{m_j} E_j[x_i^2]$$

where $E_j[x_i x_k]$, $i, k \in S_j$, denote the elements of $E_j[\mathbf{X}\mathbf{X}^T]$.

Next, suppose that the number of synapses, m_j , received by each postsynaptic neuron j is asymptotically large. Then,

$$\xi_j \longrightarrow \xi = \frac{1}{n(n-1)} \sum_{i=1}^n \sum_{\substack{k=1 \\ k \neq i}}^n E[x_i x_k], \quad \forall j \quad (8)$$

and

$$\zeta_j \longrightarrow \zeta = \frac{1}{n} \sum_{i=1}^n E[x_i^2], \quad \forall j \quad (9)$$

Furthermore, under the assumption that m_j is asymptotically large, the row sums of the $m_j \times m_j$ symmetric matrix $E_j[\mathbf{X}\mathbf{X}^T]$ will be nearly the same (rows differ only

in diagonal elements), and each row sum can be approximated by $m_j \xi$. If we factor this product out of $E_j[\mathbf{X}\mathbf{X}^T]$, the resulting matrix has a dominant eigenvalue of 1 since it is a stochastic matrix. The associated eigenvector aligns with $\mathbf{1}$. This means that, in the limit of m_j , the dominant eigenvalue of the original unfactored matrix $E_j[\mathbf{X}\mathbf{X}^T]$ is

$$\lambda_j^1 \approx m_j \xi \quad (10)$$

Because most neurons in the cortical systems of the forebrain have so many synapses, considering asymptotic results is certainly sensible. In particular, (10) is a good approximation to such limit calculations, and the following rearrangement produces a relationship between the number of synapses and the input statistics:

$$m_j \approx \frac{\lambda_j^1}{\xi} \quad (11)$$

More to the point, using (5) and (11), we see the approximate asymptotic relationship between the number of synapses and the average postsynaptic activity as

$$m_j \approx \frac{E[\mathbf{Y}_j]}{c_y c_w \xi}$$

which, to be more precise and without ignoring the value of the diagonal elements of $E_j[\mathbf{X}\mathbf{X}^T]$ while assuming they are all equal to ζ , implies (see Appendix B for the proof)

$$m_j \approx \frac{1}{\xi} \left(\frac{E[\mathbf{Y}_j]}{c_y c_w} - \zeta \right) + 1 \quad (12)$$

Thus, we now have another prescription for the biologist who entertains Hypothesis C. Equation (12) says that for a system with many synapses per neuron, the number of synapses per neuron can be estimated from the average postsynaptic activity and the average of the input correlations. This type of relationship is a quantitative reflection of a possible activity dependence of synaptogenesis. That is, postsynaptic activity may regulate synaptogenesis in order to maintain a set point level of average activity (Adelsberger-Mangan and Levy 1993, 1994). Moreover, the biologist must consider both modification of existing synapses and formation of new synapses as changing when the correlational input statistics are altered.

For the engineer who works with the constraints of Hypothesis A, it may be more useful to rewrite the relationship as $E[\mathbf{Y}_j] \sim c_y c_w m_j \xi$. Hence, if the number of synapses are limited to a specific value of m_j and if ξ can be estimated, then the constants c_y and c_w can be adjusted to produce the desired value of average postsynaptic activity.

4 Distribution of the dominant eigenvalue

We have shown in (5) that for a small modification rate and a stationary, strongly mixing stochastic environment, the distribution of the average postsynaptic

activity is asymptotically (in time) the same as that of the dominant eigenvalue scaled by a constant. Unfortunately, the dominant eigenvalue of a matrix of dimension greater than four cannot be algebraically solved for in an exact manner. However, it is possible to obtain an approximation for such an eigenvalue by applying several iterations of the power method (Isaacson and Keller 1966). One version of the power method consists of the recursion

$$\frac{M\mathbf{V}(k)}{\|\mathbf{V}(k)\|_\infty} = \mathbf{V}(k+1)$$

where $\mathbf{V}(k)$ is the eigenvector associated with the dominant eigenvalue of the positive-definite matrix M at step k , and $\|\cdot\|_\infty$ is the infinity norm.

A single iteration that starts with $\mathbf{V}(1) = \mathbf{1}$ can produce a vector that yields an estimate of λ^1 . Specifically, the one iteration is

$$E_j[\mathbf{X}\mathbf{X}^T]\mathbf{1} = \mathbf{V}_j \quad (13)$$

where the new eigenvector estimate, \mathbf{V}_j , is just the vector of row sums of the matrix. Using this estimate of the eigenvector, we can estimate the associated eigenvalue m_j times (once for each element in \mathbf{V}_j) and then average. That is, we can start with an approximate eigenequation

$$E_j[\mathbf{X}\mathbf{X}^T]\mathbf{V}_j = \lambda_j^1 \mathbf{V}_j \quad (14)$$

and multiply both sides of (14) by $\mathbf{1}$ to find

$$\mathbf{1}^T E_j[\mathbf{X}\mathbf{X}^T]\mathbf{V}_j = \lambda_j^1 \mathbf{1}^T \mathbf{V}_j \quad (15)$$

From (13) and (15), we finally produce the following estimate of the dominant eigenvalue λ_j^1 :

$$\hat{\lambda}_j^1 = \frac{\mathbf{1}^T E_j[\mathbf{X}\mathbf{X}^T]\mathbf{V}_j}{\mathbf{1}^T \mathbf{V}_j} = \frac{\mathbf{V}_j^T \mathbf{V}_j}{\sum_{i=1}^{m_j} v_{ij}} = \frac{\sum_{i=1}^{m_j} v_{ij}^2}{\sum_{i=1}^{m_j} v_{ij}} \quad (16)$$

where v_{ij} is i th entry in \mathbf{V}_j .

Then from this estimate of the dominant eigenvalue λ_j^1 , we determine $E[\mathbf{Y}_j]$ [directly from equation (5)]:

$$E[\mathbf{Y}_j] \approx \frac{c_y c_w \sum_{i=1}^{m_j} v_{ij}^2}{\sum_{i=1}^{m_j} v_{ij}} \quad (17)$$

It is hard to conclude a distributional form for $\hat{\lambda}_j^1$ from (17). The following rearrangement is more revealing:

$$\frac{\sum_{i=1}^{m_j} v_{ij}^2}{\sum_{i=1}^{m_j} v_{ij}} = \frac{\frac{\sum_{i=1}^{m_j} v_{ij}^2}{m_j} - \left(\frac{\sum_{i=1}^{m_j} v_{ij}}{m_j} \right)^2}{\frac{\sum_{i=1}^{m_j} v_{ij}}{m_j}} + \frac{\sum_{i=1}^{m_j} v_{ij}}{m_j} \quad (18)$$

Note that the first term on the right-hand side of (18) is the empirical variance of the i th row sum v_{ij} , divided by the empirical mean of v_{ij} . If this variance is much smaller than the row sum mean, (16) simplifies to

$$\hat{\lambda}_j^1 \approx \frac{\sum_{i=1}^{m_j} v_{ij}}{m_j} \quad (19)$$

Thus, as a corollary to (19),

$$E[\mathbf{Y}_j] \approx \frac{c_y c_w \sum_{i=1}^{m_j} v_{ij}}{m_j} \quad (20)$$

which is obtained from (5).

The row sums v_{ij} , $i = 1, \dots, m_j$, in (19) and (20) are statistically dependent if we consider the sampling process for the matrix $E_j[\mathbf{X}\mathbf{X}^T]$. While we hesitate to evoke a specific central limit theorem without formulating the nature of this dependence, heuristics indicate that the random variables λ_j^i and $E[\mathbf{Y}_j]$ may be normally distributed (Aldous 1989). Furthermore, there is experimental evidence, as presented in the next section, that normal distribution is a satisfactory approximation of the distribution of the dominant eigenvalue.

5 Simulations

In this section, we present some computer simulations that test the accuracy of the approximations made in the previous sections. In the first set of computer simulations, we examine the distribution of average postsynaptic activity, as well as the accuracy of our prediction of postsynaptic activity when output neurons receive innervation from a fixed number of input neurons (Hypothesis B). In the second set of simulations, we examine the distribution of the number of input neurons needed to achieve a desired level of postsynaptic activity (Hypothesis C), as well as the accuracy of our prediction of the number of input neurons needed to achieve a desired level of average postsynaptic activity.

In all simulations, the input layer consists of 1024 binary neurons. The firing state of the input neurons is determined by a uniform pseudo-random number generator biased for thresholds of 0.25. (Lower firing levels require a larger network.) There are five different input environments, consisting of 8, 16, 32, 100, or 400 neuronal firing patterns. After the input patterns of a particular environment are generated, the average firing level of the input neurons is determined. If an input neuron fires at an average level that is less than 0.23 or greater than 0.27, the firing for that input neuron is re-determined via the random-number generator. This process continues until the average firing level of all 1024 input neurons is between 0.23 and 0.27. Although the firings of the input neurons are independent in terms of the generator, by virtue of finite or sparse sampling, there is statistical dependence between input layer neurons.

Based on the patterns of input neuron firing, there is an input correlation matrix, $E[\mathbf{X}\mathbf{X}^T]$, of dimension 1024×1024 . Therefore, because we create five different input environments (differing in the number of firing patterns), we generate five input correlation matrices. The average entry of the five input correlation matrices (corresponding to the input environments consisting of 8, 16, 32, 100, or 400 neuronal firing patterns) is 6.27×10^{-2} . The variance of the entries of the correlation matrices ranges from a high of 5.06×10^{-3} (8 patterns) to a low of 1.23×10^{-4} (400 patterns).

The output layer consists of 500 or 2000 neurons. The synaptic connections between the input and output layers are created in one of two fashions: either each output neuron receives input from a randomly selected but fixed number of input neurons (Hypothesis B), or synapses are formed at random until the average postsynaptic activity of each output neuron reaches a preset value (Hypothesis C). In all simulations, input neurons are chosen randomly, and multiple inputs from a single input neuron to an output neuron are permitted. For experiments simulating the problem within the framework of Hypothesis A see Levy (1992).

In all simulations, the constants c_w and c_y are set to 1. The constant ε controlling the rate of synaptic modification is $\varepsilon = 0.05$.

5.1 Hypothesis B

The first series of simulations is run in accordance with Hypothesis B; that is, the input and output neuronal layers are connected such that each output neuron receives a fixed number of inputs. In this set of simulations, we illustrate that: (1) the distribution of average postsynaptic activity given a fixed number of inputs is roughly normal, and (2) the accuracy of the estimate of postsynaptic activity (20) over a wide range number of inputs and input variance.

Figure 1 shows the distribution of average postsynaptic activity when $m_j = 100$ input neurons. The data are obtained from 2000 output neurons. The input environment consists of 32 neuronal firing patterns (the average entry of the input correlation matrix is 6.27×10^{-2} with variance 1.17×10^{-3}).

Superimposed upon the histogram of $E[\mathbf{Y}_j]$ values is the appropriately scaled normal distribution with identical mean and variance (Fig. 1). Although a strong theoretical conclusion regarding the distribution of $E[\mathbf{Y}_j]$

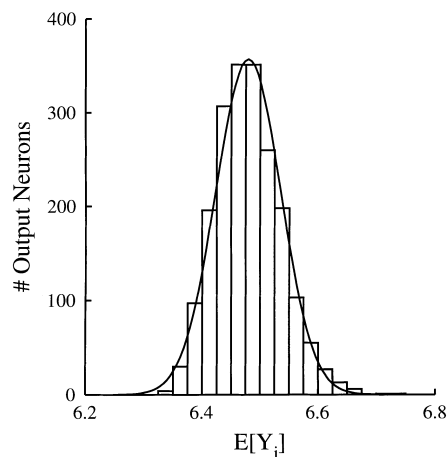


Fig. 1. Histogram of postsynaptic activation ($E[\mathbf{Y}_j]$) for $m_j = 100$. The data are from 2000 output neurons. The input environment consists of 32 neuronal firing patterns. The average entry of the input correlation matrix is 6.27×10^{-2} with variance 1.17×10^{-3} . Superimposed upon the histogram is an appropriately scaled normal distribution with a mean and variance identical to the histogram data set

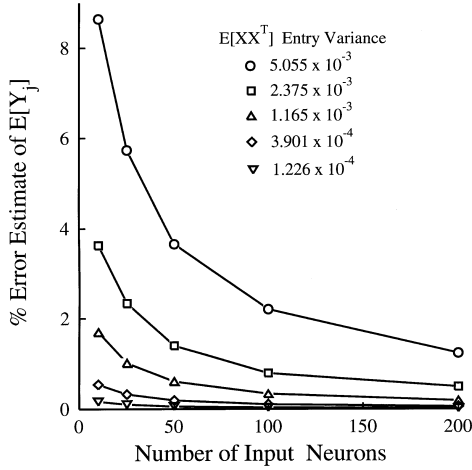


Fig. 2. Percent error estimate of $E[\mathbf{Y}_j]$ as a function of the number of input neurons (m_j) and the variance of the elements of $E[\mathbf{X}\mathbf{X}^T]$. Each plotted value is the average of 500 output neurons. The input environments consist of 8 (\circ), 16 (\square), 32 (\triangle), 100 (\diamond), or 400 (∇) neuronal firing patterns. The percentage error estimate of $E[\mathbf{Y}_j]$ is calculated as $|\lambda_j^1 - \hat{\lambda}_j^1|/\lambda_j^1$

has not been presented, the distribution of $E[\mathbf{Y}_j]$ is apparently close to normal.

Figure 2 demonstrates the accuracy of the estimate of postsynaptic activity (20) as a function of the number of input neurons synapsing on each output neuron and the variance of the entries of the input correlation matrix ($E[\mathbf{X}\mathbf{X}^T]$). The number of input neurons chosen, 10, 25, 50, 100, and 200, correspond to 0.98%, 2.4%, 4.9%, 9.8%, and 19.5% of the input layer neurons, respectively. Thus, these networks are only fractionally connected. Given the number of input neurons, $E[\mathbf{Y}_j]$ is determined using (5) and then estimated using (20) (as c_y and c_w equal 1, $E[\mathbf{Y}_j]$ equals λ_j). The percent error estimate of $E[\mathbf{Y}_j]$ is calculated as $|\lambda_j^1 - \hat{\lambda}_j^1|/\lambda_j^1$. Averages are obtained from 500 output neurons and plotted against the number of input neurons. The figure indicates that as the output neurons make synapses with more input neurons, the average error in the estimate of postsynaptic activity provided by (20) decreases. In fact, when the output neurons accrue at least 50 synapses (4.9% of the input neurons), the average percent error is less than 2% for each of the five input environments tested. Moreover, Fig. 2 indicates that (20) yields acceptable estimates even when the variance of the entries of $E[\mathbf{X}\mathbf{X}^T]$ approaches the maximum possible value (a variance of 1.17×10^{-2} , obtained when one-quarter of the entries of $E[\mathbf{X}\mathbf{X}^T]$ equal 0.25 and the rest equal zero, and the mean entry equals 6.25×10^{-2}) and only 10 input neurons innervate each output neuron. Considering that neurons make many more synapses than the numbers in these simulations, it is clear that the estimate of postsynaptic activity given in (20) is quite robust.

5.2 Hypothesis C

For the next series of simulations, the input and output layers of the network are connected in accordance with

Hypothesis C, that is, each postsynaptic neuron accrues input neuron innervation until a desired level of average postsynaptic activity is obtained. This set of simulations illustrates that: (1) the distribution of the number of input neurons given a desired level of postsynaptic activity is roughly normal in its central region and (2) the estimate of the number of input neurons needed to achieve a desired level of average postsynaptic activity (12) over a wide range of desired postsynaptic activity and input statistics is acceptably accurate.

Figure 3 is the histogram of the number of input neurons per postsynaptic neuron when an average postsynaptic activity level of 6.40 is desired. (This level of postsynaptic activity was chosen as it corresponds to approximately 100 input neurons, the same number chosen for illustration in Fig. 1 of the distribution of average postsynaptic activity.) The average postsynaptic activity is calculated using (5) (c_w and c_y equal to 1) with λ_j^1 determined using the iterative algorithm of Sect. 4. The data are obtained from 2000 output neurons. The input environment consists of 32 neuronal firing patterns (the average entry of the input correlation matrix is 6.27×10^{-2} with variance 1.17×10^{-3}). An appropriately scaled normal distribution with identical mean and variance is superimposed upon the histogram of m_j . It is clear from Fig. 3 that the normal distribution is an acceptable approximation of the distribution of m_j , especially in the central region. The actual distribution will only be approximately normal due to the discrete nature of m_j and because its distribution is bound to be asymmetric, for small numbers of inputs will never fire a postsynaptic neuron enough. In fact, we can give a recursive algorithm for the distribution of m_j where the probability of each bin is calculated by differencing the successive cumulative normal distributions obtained by integrating up to the appropriate m_j . Such differencing will reproduce an approximately normal distribution

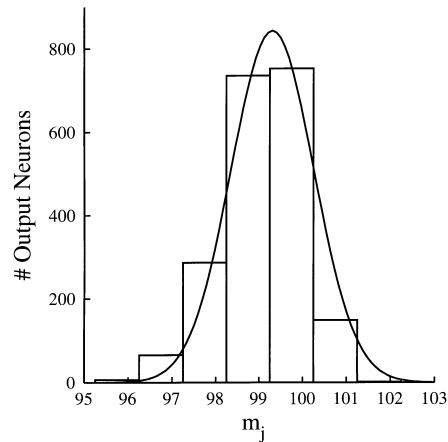


Fig. 3. Histogram of the number of input neurons (m_j) needed to achieve $E[\mathbf{Y}_j] = 6.40$. The data are from 2000 output neurons. The input environment consists of 32 neuronal firing patterns. The average entry of the input correlation matrix is 6.27×10^{-2} with variance 1.17×10^{-3} . Superimposed upon the distribution is an appropriately scaled normal distribution with a mean and variance identical to the histogram data set

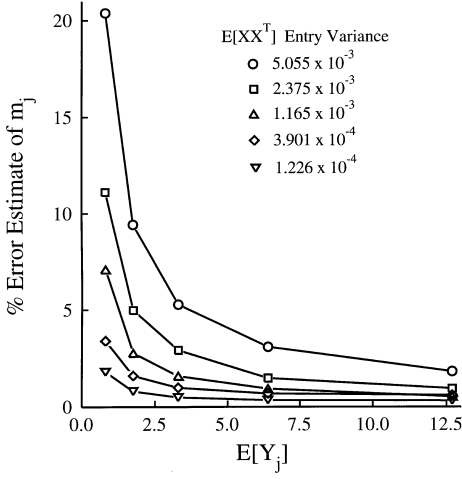


Fig. 4. Percentage error estimate of m_j as a function of $E[\mathbf{Y}_j]$ and the variance of the elements of $E[\mathbf{X}\mathbf{X}^T]$. Each plotted value is the average of 500 output neurons. The input environments consist of 8 (\circ), 16 (\square), 32 (\triangle), 100 (\diamond), or 400 (∇) neuronal firing patterns. The percentage error estimate of m_j is calculated as $|m_j - \hat{m}_j|/m_j$

when incrementing m_j by 1 changes the integral by only a small amount.

Figure 4 depicts the sensitivity of the estimate of m_j (\hat{m}_j) as a function of the average postsynaptic activity and the variance of the entries of the correlation matrix $E[\mathbf{X}\mathbf{X}^T]$. The desired levels of average postsynaptic activity values chosen, specifically 0.80, 1.75, 3.30, 6.40, and 12.70, correspond approximately to 10, 25, 50, 100, and 200 input neurons, respectively (based on the first set of simulations). For each desired $E[\mathbf{Y}_j]$, input neurons are accrued one at a time until it is determined that the average postsynaptic activity of the output is greater than or equal to the desired $E[\mathbf{Y}_j]$ (using the iterative method of Sect. 4 to determine λ_j and with c_w and c_y equal to 1). Then, given the actual $E[\mathbf{Y}_j]$ produced by the selected input neurons, (12) is used to calculate the approximation of the number of input neurons needed to achieve the particular $E[\mathbf{Y}_j]$. The percent error of the estimate of m_j is then calculated as $|m_j - \hat{m}_j|/m_j$. The averages are obtained over 500 output neurons and plotted against $E[\mathbf{Y}_j]$. Figure 4 illustrates that as $E[\mathbf{Y}_j]$ increases, the error in \hat{m}_j decreases. This decrease is not surprising because of the assumption that leads to (8) and (9). That is, as the number of synapses on each postsynaptic neuron grows larger so as to increase $E[\mathbf{Y}_j]$, the assumption of asymptotically large connectivity is better satisfied. When the number of input neurons innervating each output neuron averages at least 50, there is less than a 5% difference between the true and estimated number of input neurons for each of the five input environments tested.

6 Extension to recurrent nonlinear networks

Since most biological networks in the brain are recurrent and use nonlinear activation rules, it is desirable to extend our results to such networks. Gibson and Robinson (1992) have approached this problem from

an associative memory perspective, including a simple Hebbian learning rule. Several strategies for setting the threshold in a fractionally connected nonlinear network were proposed in Buckingham and Willshaw (1993). A generalization to sparsely connected recurrent nonlinear networks can be made by employing the following variant of the shunting inhibition neuron model in Minai and Levy (1993a):

$$\mathbf{Y}_j(t) = \frac{\sum_{i=1}^{m_j} w_{ij} c_{ij} \mathbf{Z}_j(t-1)}{\sum_{i=1}^{m_j} w_{ij} c_{ij} \mathbf{Z}_j(t-1) + K a(t-1)},$$

$$1 \leq a(t-1) \leq n \quad (21)$$

$$\mathbf{Z}_j(t) = \begin{cases} 1, & \text{if } \mathbf{Y}_j(t) \geq \theta \\ 0, & \text{otherwise} \end{cases}$$

where c_{ij} is an indicator function indicating the presence or absence of a synaptic connection from i to j , $a(t)$ is the number of active neurons in the network at time t , K is the global inhibitory weight, and θ is the firing threshold. \mathbf{Y}_j and \mathbf{Z}_j denote the excitation and firing status of the postsynaptic neuron j at time t , respectively. If $a(t-1) = 0$, then $\mathbf{Y}_j(t) = 0$, for all j .

Let $a(t-1) = A$, where A is very large, and assume that the neurons fire independently, which is generally the case in large and sparsely connected networks (Minai and Levy 1993a,b). Then, using the corollary in Appendix A, and employing a Gaussian approximation to the binomial distribution and a hyperbolic tangent approximation to the error function (Minai and Levy 1993a, 1994), we conclude that the following equation holds for the average firing probability ρ_j :

$$\rho_j = \frac{1}{2} \left(1 - \tanh \frac{\sqrt{A}}{T} \right) \quad (22)$$

where

$$T = \frac{1}{\alpha - p} \sqrt{\frac{\pi p - (1-p)}{2}}$$

where $\alpha = \theta K / (1 - \theta) E[x_i]$, assuming $E[x_i]$ is approximately identical for all i , and p is the probability that $c_{ij} = 1$.

Equation (22) provides us with a recipe for setting the parameter values to obtain a desirable activity level for a recurrent nonlinear network with fixed weights. Naturally, the tighter the distribution of the synaptic weights, the better an approximation (22) will be of the average firing probability of a recurrent nonlinear network with time-varying weights.

7 Discussion

When working with a large system driven by stochastic inputs, it is natural to describe system performance in terms of statistics instead of a necessarily incomplete listing of particular states that correspond to a particular environment. In this report, we considered the relationships between input and output neural activity and

network connectivity. We restricted our analysis to fractionally connected networks, where the synaptic connections are feed-forward and excitatory, as we were motivated by our research into biologically plausible, adaptive mechanisms for the construction of such networks. We found that the empirical average of input correlations alone produces a relatively accurate number of the input neurons needed to innervate an output neuron with a prespecified level of average activity.

We showed that the synapses of such a neuron will take on values proportional to the dominant eigenvector of the correlation matrix of the inputs after experiencing a stationary, ergodic input for a long enough time; the average activity of the neuron will be proportional to the dominant eigenvalue of the same matrix.

Such statistical statements can be used by neuroscientists to gain a better understanding of natural networks. For instance, nature may dynamically regulate the number of synapses innervating an output neuron so that the neuron maintains some predefined average activity level. (Different neuron types in the different areas of the brain may have varying predetermined activity levels.) The neuroscientist can test this hypothesis by altering input neuron activities and measuring the subsequent changes in network connectivity. To quantify the effect of this experimental manipulation accurately however, a statistical description of the network connectivity is needed.

The statistical relationships between network connectivity and input/output activities are also of concern from the engineering perspective. For example, design limitations may limit the number of synaptic connections per neural element or the average number of active elements. Engineers need to know the interaction between these design limitations and the networks' computational characteristics. Limits on input activity or synapse number may impose reduced activity levels in the output elements. This may result in a loss of representational information as measured by entropy or a loss in computational capacity.

Ultimately, the research presented here provides the scientist with a prescription for estimating various network parameters through either analytical means or simulations. The only piece of information that is needed in the calculations is the summary statistics of empirical correlations.

Acknowledgments. This work was supported by NIH MH48161 and MH00622, EPRI RP8030-08, and Pittsburgh Supercomputing Center Grant BNS950001 to W.B.L, and by the Department of Neurological Surgery, University of Virginia Health Sciences Center, Dr. John A. Jane, Chairman.

Appendix A

In this section, we formally state the theorem of (4). A more detailed version of the proof is found in Levy and Geman (1982). A similar result for a slightly different modification rule was reported in Oja (1982).

Theorem: Suppose that the following conditions are met:

- [1] $\mathbf{W}_j(0) > 0$ where the notation implies that the inequality holds for every element of \mathbf{W}_j .

- [2] $E_j[\mathbf{X}]$ and $E_j[\mathbf{X}\mathbf{X}^T]$ are constant (that is, we assume that the statistics of the environment are stationary over the time period of interest).

- [3] $E_j[\mathbf{X}\mathbf{X}^T] > 0$.

- [4] $E_j[\mathbf{X}\mathbf{X}^T]$ is positive definite.

Then,

$$\lim_{t \rightarrow \infty} \mathbf{W}_j(t) = \frac{c_w \mathbf{e}_j^1 \lambda_j^1}{E_j[\mathbf{X}] \mathbf{e}_j^1} \quad (23)$$

where λ_j^1 and \mathbf{e}_j^1 are the dominant eigenvalue and its associated eigenvector of $E_j[\mathbf{X}\mathbf{X}^T]$, respectively. Hence the weight vector \mathbf{W}_j is aligned with the dominant eigenvector as $t \rightarrow \infty$.

Sketch of the proof: We begin our proof by first establishing some preliminary bounds on $\|\mathbf{W}_j\|$, $\mathbf{W}_j \cdot E[\mathbf{X}]$, and $\mathbf{W}_j \cdot \mathbf{e}_j^1$. Specifically, there exists a $\delta > 0$ such that

$$\|\mathbf{W}_j\| > \delta, \mathbf{W}_j \cdot E[\mathbf{X}] > \delta \|\mathbf{W}_j\| \text{ and } \mathbf{W}_j \cdot \mathbf{e}_j^1 > \delta \|\mathbf{W}_j\|, \forall t \geq 0 \quad (24)$$

and

$$\overline{\lim}_{t \rightarrow \infty} \|\mathbf{W}_j\| < \infty \quad (25)$$

To prove (24), first note that $\mathbf{W}_j(0) > 0 \Rightarrow \mathbf{W}_j(t) \geq 0, \forall t \geq 0$ (apply condition 3 of the theorem). Furthermore,

$$\begin{aligned} \frac{d}{dt} \|\mathbf{W}_j\|^2 &= 2\mathbf{W}_j \cdot \left(\frac{d}{dt} \mathbf{W}_j \right) \geq 2 \|\mathbf{W}_j\|^2 (c_w \lambda_j^1 \\ &\quad - \|\mathbf{W}_j\| \|E_j[\mathbf{X}]\|) > 0 \end{aligned}$$

so that

$$\underline{\lim}_{t \rightarrow \infty} \|\mathbf{W}_j\| \geq \frac{c_w \lambda_j^1}{\|E_j[\mathbf{X}]\|} > 0$$

This implies the first part of (24). Also note that condition 3 implies $E_j[\mathbf{X}] > 0$. Since $\mathbf{W}_j > 0, \forall t \geq 0$,

$$\mathbf{W}_j \cdot E_j[\mathbf{X}] \geq \left(\min_{1 \leq k \leq n} E_j[\mathbf{X}_k] \right) \|\mathbf{W}_j\|$$

This establishes the second part of (24). The third part of (24) follows similarly using the observation that $\mathbf{e}_j^1 > 0$.

Equation (25) is proven by first expanding $\frac{d}{dt} \|\mathbf{W}_j\|^2$, then using the second part of (24) to prove that $\overline{\lim}_{t \rightarrow \infty} \|\mathbf{W}_j\| \leq c_w \lambda_j^1 / \delta$.

Now our theorem is simply a consequence of two lemmata about the asymptotic behavior of the angle and magnitude of the weight vector \mathbf{W}_j . The first lemma

$$\lim_{t \rightarrow \infty} \frac{\mathbf{W}_j \cdot \mathbf{e}_j^1}{\|\mathbf{W}_j\|} = 1$$

demonstrates the alignment of the vector \mathbf{W}_j with \mathbf{e}_j^1 .

The second lemma

$$\lim_{t \rightarrow \infty} \|\mathbf{W}_j\| = \frac{c_w \lambda_j^1}{E_j[\mathbf{X}] \cdot \mathbf{e}_j^1}$$

provides the asymptotic length of the vector of synaptic weights. The proofs of lemmata 1 and 2 are provided in the 1982 technical report by Levy and Geman.

With lemmata 1 and 2 holding, $\lim_{t \rightarrow \infty} (\mathbf{W}_j \cdot \mathbf{e}_j^k / \|\mathbf{W}_j\|) = 0, \forall k \neq 1$, since the set of eigenvectors $\{\mathbf{e}_j^k\}$ forms an orthonormal basis. Consequently,

$$\begin{aligned} \lim_{t \rightarrow \infty} \mathbf{W}_j &= \lim_{t \rightarrow \infty} \|\mathbf{W}_j\| \sum_k \frac{\mathbf{W}_j \cdot \mathbf{e}_j^k}{\|\mathbf{W}_j\|} \mathbf{e}_j^k \\ &= \lim_{t \rightarrow \infty} \|\mathbf{W}_j\| \frac{\mathbf{W}_j \cdot \mathbf{e}_j^1}{\|\mathbf{W}_j\|} \mathbf{e}_j^1 = \lim_{t \rightarrow \infty} \|\mathbf{W}_j\| \mathbf{e}_j^1 = \frac{c_w \lambda_j^1 \mathbf{e}_j^1}{E_j[\mathbf{X}] \cdot \mathbf{e}_j^1} \end{aligned}$$

proving the theorem. \square

Corollary: For a recurrent nonlinear network, as modeled in Sect. 6, (21) implies

$$\lim_{t \rightarrow \infty} \mathbf{W}_j(t) = E_j[\mathbf{X}]. \quad \square$$

Proof of the corollary: Multiplying both sides of (23) by $E_j[\mathbf{X}\mathbf{X}^T]$ and using the eigenequation in Sect. 2 yields

$$\lim_{t \rightarrow \infty} \mathbf{W}_j(t) = \frac{E_j[\mathbf{X}\mathbf{Y}_j]}{E[\mathbf{Y}_j]}. \quad \square$$

Appendix B

Proof of (12): Let the number of synapses m_j be asymptotically large, and assume that $E_j[x_i^2] \approx \zeta, \forall i$. Then the matrix $E_j[\mathbf{X}\mathbf{X}^T]$ can be partitioned as follows:

$$E_j[\mathbf{X}\mathbf{X}^T] = (m_j - 1)\xi\mathbf{A} + \mathbf{B} \quad (26)$$

where

$$\mathbf{A} = \begin{bmatrix} 0 & \frac{E_j[x_1x_2]}{(m_j-1)\xi} & \frac{E_j[x_1x_3]}{(m_j-1)\xi} & \cdots & \frac{E_j[x_1x_{m_j}]}{(m_j-1)\xi} \\ \frac{E_j[x_2x_1]}{(m_j-1)\xi} & 0 & \frac{E_j[x_2x_3]}{(m_j-1)\xi} & \cdots & \frac{E_j[x_2x_{m_j}]}{(m_j-1)\xi} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \frac{E_j[x_{m_j}x_1]}{(m_j-1)\xi} & \frac{E_j[x_{m_j}x_2]}{(m_j-1)\xi} & \cdots & \frac{E_j[x_{m_j}x_{m_j-1}]}{(m_j-1)\xi} & 0 \end{bmatrix}$$

and

$$\mathbf{B} = \begin{bmatrix} E_j[x_1^2] & 0 & 0 & \cdots & 0 \\ 0 & E_j[x_2^2] & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & E_j[x_{m_j}^2] \end{bmatrix}$$

Since $\sum_{k=1}^{m_j} E_j[x_i x_k] = (m_j - 1)\xi, \forall i$ and $\sum_{i=1}^{m_j} E_j[x_i x_k] = (m_j - 1)\xi,$

$\forall k$ by (9), \mathbf{A} is a stochastic matrix with a dominant eigenvalue of 1. Recognizing that the diagonal matrix \mathbf{B} has a repeated eigenvalue of ζ , (26) implies that the dominant eigenvalue λ_j^1 of $E_j[\mathbf{X}\mathbf{X}^T]$ becomes

$$\lambda_j^1 = (m_j - 1)\xi + \zeta \quad (27)$$

Combining (5) and (27) finally gives the desired equation. \square

References

Adelsberger-Mangan DM, Levy WB (1993) Adaptive synaptogenesis can construct networks which maintain information and reduce statistical dependence. *Biol Cybern* 70:81–87

Adelsberger-Mangan DM, Levy WB (1994) The influence of limited presynaptic growth and synapse removal in adaptive synaptogenesis. *Biol Cybern* 71:461–468

Aldous D (1989) Probability approximations via the poisson clumping heuristic. Springer Berlin Heidelberg, New York

Amaral DG, Ishizuka N, Clairborne B (1990) Neurons, numbers and the hippocampal network. In: Storm-Mathisen J, Zimmer J, Ottersen OP (eds) *Progress in brain research*, Vol 83. Elsevier, Amsterdam, pp 1–11

Barlow HB (1959) Sensory mechanisms, the reduction of redundancy, and intelligence. In: *The mechanization of thought processes*. (National Laboratory Symposium no 10) HMS Stationery Office, London, pp 537–559

Buckingham J, Willshaw D (1993) On setting unit thresholds in an incompletely connected associative net. *Network Comput Neural Syst* 4:441–459

Field DJ (1994) What is the goal of sensory coding? *Neural Comp* 6:559–601

Földiák P (1990) Forming sparse representations by local anti-hebbian learning. *Biol Cybern* 64:165–170

Gibson WG, Robinson J (1992) Statistical analysis of the dynamics of a sparse associative memory. *Neural Networks* 5:645–661

Grossberg S (1976) Adaptive pattern classification and universal recoding. I. Parallel development and coding of neural feature detectors. *Biol Cybern* 23:121–134

Hillman DE, Chen S (1984) Reciprocal relationship between size of postsynaptic densities and their number: constancy in contact area. *Brain Res* 295:325–343

Isaacson E, Keller HB (1966) *Analysis of numerical methods*. Wiley, New York

Levy WB (1992) A calculation of neural network statistics: the probability distribution of average cell firing in a feedforward neural network. *Proc Int Joint Conf Neural Networks III*:341–346

Levy WB, Baxter RA (1996) Energy efficient neural codes. *Neural Comp* 8:531–543

Levy WB, Colbert CM (1991) Adaptive synaptogenesis can complement associative potentiation/depression. In: Commons ML (ed) *Neural network models of conditioning: quantitative analyses of behavior*, Vol 13. Lawrence Erlbaum, Hillsdale, NJ, pp 53–68

Levy WB, Desmond NL (1985) Associative potentiation/depression in the hippocampal dentate gyrus. In: Buzsáki G, Vanderwolf CH (eds) *Electrical activity of the archicortex*. Akademiai Kiado, Budapest, Hungary, pp 359–373

Levy WB, Geman S (1982) Limit behavior of experimentally derived synaptic modification rules. *Reports in pattern analysis*, no. 121. Division of Applied Mathematics, Brown University, Providence, RI

Levy WB, Steward O (1983) Temporal contiguity requirements for long term associative potentiation/depression in the hippocampus. *Neuroscience* 8:791–797

Levy WB, Colbert CM, Desmond NL (1990) Elemental adaptive processes of neurons and synapses: a statistical/computational perspective. In: Gluck M, Rumelhart D (eds) *Neuroscience and connectionist theory*. Lawrence Erlbaum, Hillsdale, pp 187–235

Mead C (1989) *Analog VLSI implementation of neural systems*. Kluwer, Boston, Mass

Minai AA, Levy WB (1993a) Predicting complex behavior in sparse asymmetric networks. In: Giles CL, Hanson SJ, Cowan JD (eds) *Advances in neural information processing systems* 5. Morgan Kaufmann, San Mateo, CA, pp 556–563

Minai AA, Levy WB (1993b) The dynamics of sparse random networks. *Biol Cybern* 70:177–187

Minai AA, Levy WB (1994) Setting the activity level in sparse random networks. *Neural Comput* 6:85–99

Mitchison GJ (1991) Neuronal branching patterns and the economy of cortical wiring. *Proc R Soc [Biol]* 245:151–158

Oja E (1982) A simplified neuron model as a principal component analyzer. *J Math Biol* 15:267–273

Rauschecker JP, Singer W (1979) Changes in the circuitry of the kitten visual cortex are gated by postsynaptic activity. *Nature* 280:58–60

Rauschecker JP, Singer W (1981) The effects of early visual experience on the cat's visual cortex and their possible explanation by hebb synapses. *J Physiol* 310:215–239

Rolls ET (1989) Functions of neuronal networks in the hippocampus and cerebral cortex in memory. In: Cotterill RMJ (ed) *Models of brain function*. Cambridge University Press, Cambridge, UK, pp 15–33