DECODING NEURONAL FIRING
AND
MODELING NEURAL NETWORKS

L.F. Abbott

Center for Complex Systems
Brandeis University
Waltham, MA 02254

Published in *Quart. Rev. Biophys.* **27**:291-331 (1994).

1. Introduction

Biological neural networks are large systems of complex elements interacting through a complex array of connections. Individual neurons express a large number of active conductances (Connors *et al.*, 1982; Adams & Gavin, 1986; Llinás, 1988; McCormick, 1990; Hille, 1992) and exhibit a wide variety of dynamic behaviors on time scales ranging from milliseconds to many minutes (Llinás, 1988; Harris-Warrick & Marder, 1991; Churchland & Sejnowski, 1992; Turrigiano *et al.*, 1994). Neurons in cortical circuits are typically coupled to thousands of other neurons (Stevens, 1989) and very little is known about the strengths of these synapses (although see Rosenmund *et al.*, 1993; Hessler *et al.*, 1993; Smetters & Nelson, 1993). The complex firing patterns of large neuronal populations are difficult to describe let alone understand. There is little point in accurately modeling each membrane potential in a large neural circuit unless we have an effective method for interpreting (or even visualizing) the enormous numerical output of such a model. Thus, major issues in modeling biological circuits are: i) How do we describe and interpret the activity of a large population of neurons and how do we model neural circuits when ii) individual neurons are such complex elements and iii) our knowledge of the synaptic connections is so incomplete. This review covers some of the approaches and techniques that have been developed to try to deal with these problems. Although they cannot be considered solved, progress has been made and, by combining a number of different techniques, we can put together a fairly comprehensive description of at least some aspects of how neural systems function. In this review, I will concentrate on mathematical and computational methods with applications provided primarily to illustrate the methods. The reader can consult the references for applications to specific neural systems.

Four general techniques have proven particularly valuable for analyzing complex neural systems and dealing with the problems mention above:

1. Patterns of activity in neural networks only gain functional and behavioral significance when they are interpreted correctly. This is a two step process. First, the spike trains of individual neurons must be considered and then the activity of an entire population of neurons must be interpreted collectively. The second step is described in 2) below.

The spike train produced by a single neuron can be extremely complex, reflecting in part the complexity of the underlying neuronal dynamics, problem ii). A method for analyzing neuronal spike trains based on a linear filter (Bialek 1989; Bialek *et al.*, 1991; Reike, 1991) has been developed and applied to the problem of decoding the information contained in the spike trains of single neurons (Bialek 1989; Bialek *et al.*, 1991; Reike, 1991; Warland *et al.*, 1991, Theunnisen, 1993). This approach is reviewed in section 3.

In the vast majority of cases, neuronal spike trains are characterized by a firing rate. This characterization is often thought to eliminate all information arising from spike timing and to be quite distinct from the linear filter approach, but this is not the case. As shown in section 3, precisely defined firing rate variables are a special case of decoding schemes based on linear filters. Firing rates that can be generated by linear filters (see section 3) provide a concise description of the output of a single neuron that often contains enough information, either individually or in conjunction with population decoding methods discussed in sections 5 and 6, to describe and interpret neuronal activity.

2. In many cases, information is represented by the activity of an ensemble of neurons in a collective and distributed manner (Knudsen *et al.*, 1987; Konishi, 1987, Sejnowski, 1988; Konishi, 1991; Eichenbaum, 1993) as discussed in section 4. Various methods for decoding the information represented by neural ensembles have been developed and applied to a number of systems (Georgopolous *et al.*, 1986; Wilson & McNaughton, 1993; Salinas & Abbott, 1994). These are reviewed in sections 5 and 6. The ability to decode the activity of a neural network is a tremendous advantage both for experimentalists and modelers. In addition to providing a functional interpretation of neuronal output, decoding allows the complex activity of a neural circuit to be described much more concisely without loss of functionally relevant information. The impact that various neuronal and synaptic properties have on the activity of a large network is typically both complex and subtle. However, the effect of these same properties on the value of a coded quantity can sometimes be understood and interpreted quite easily and clearly (Abbott & Blum, 1996). Examples are given in section 9. In cases where it can be applied, decoding provides an effective solution to problem i). By combining the optimal linear filter with efficient methods for decoding populations of neurons, it is possible to extract the maximum amount of information from a neural circuit.

   It should be stressed that decoding is considered to be a tool for analyzing what a neural network is doing. The methods used by the nervous system to interpret and respond to patterns of neuronal firing do not necessarily have any relationship to the decoding methods used here. However, we do assume that the result of the decoding procedure provides an accurate measure of the information being represented by a neural network (for a cautionary note see Salzman and Newsome, 1994).

3. Problem iii), the lack of knowledge of synaptic strengths, is the most severe of the three problems being addressed. The difficulty of measuring the strength of even a single synapse and the enormous number of synapses severely limit the amount of information that can be obtained about connections in large neural networks. However, a great deal of information is available about how synaptic strengths can change due to activity through long-term potentiation (LTP) (Bliss & Collingridge, 1993; Malenka & Nicoll, 1993) and depression (LTD) (Artola & Singer; 1993). Thus, it makes more sense to

study the effect of synaptic changes than of synaptic values. Synaptic learning rules (see section 8) based on or inspired by LTP and LTD data can be used to compute the changes in synaptic strengths that arise from specific training experiences. The effect of these changes can be studied by looking at the resulting modifications in network activity and decoding them. This allows us to examine the impact of training experience and learning on network function, as shown in section 9. If these changes are small, we have the added advantage that their impact can be determined by a linear computation which resolves many of the problems associated with intrinsic neuronal complexity, problem ii). If the changes are larger, calculating their effects is more ambiguous but the linear approximation can still serve as a basic guideline. This approach is taken in sections 10 and 11.

4. Describing and decoding neuronal activity is only one aspect of the problem of modeling a neural network. Another challenge is to build a model that accounts for that activity and describes how it changes with time. In cases where firing rates are used to describe activity, it is tempting to try to build a dynamic model based purely on firing rates, as is frequently done (Wilson & Cowan, 1972 & 1973; Hopfield, 1984; Frolov & Medvedev, 1986; Abbott, 1991a; Amit & Tsodyks, 1991a & 1991b; Ermentrout, 1994). The firing of a neuron is affected by a large number of variables such as the state of activation and inactivation of numerous ion channels and the concentration of calcium and various other second messenger molecules inside the cell, that cannot be described purely in terms of the firing rate itself. However, a firing rate model does not have to compute rates from first principles on the basis of ion channels and other fundamental neuronal characteristics. Instead, mathematical fits of measured rates can be incorporated into these models phenomenologically and augmented with various approximations for unmeasured elements of the model. This is the approach used in sections 7-13. The dynamics describing how the firing rate changes over time is greatly simplified if the integration time used to define the firing rate is longer than any intrinsic neuronal time scale affecting firing, as discussed in section 13. In this case, measured and calculated static properties can be used to construct a dynamic model. Although firing-rate models are often highly simplified, this is not an essential feature. It is possible to calculate and include nonlinear effects, some of which are considered in section 12.

It is difficult to interpret results from modeling studies of neural networks unless the patterns of firing that they produce can be decoded in some behaviorally relevant way. For this reason, I will consider cases where the firing of a single neuron or a population of neurons can be correlated with some external variable x which may be a sensory input, the coded value of an input from or an output to another network, or a motor output. The quantity x can be a single number or it can be a vector array of numbers. I will mainly consider the case of a single coded quantity, leaving the discussion of coded vectors to section 6. A classic experimental procedure in neuroscience is to find a neuron with a firing pattern that is related to some sensory stimulus and then to study the response properties of that neuron while varying the stimulus. The same paradigm can be carried out replacing the sensory stimulus with a motor output. Experiments of this sort have built up a large catalog of neuronal responses and their external correlates, discussed in section 4. For clarity, I will refer to the quantity x being coded as the stimulus or stimulus value. However, it should be kept in mind that x could equally well represent a motor output.

Although this review will concentrate on methods, some results will be given to illustrate these methods. Section 9 considers the effect that training has on the coded value of a stimulus. This phenomenon has applications to memory and learning. Section 10 provides a simple example of the activity-dependent development of a neuronal response tuning curve. In sections 11 and 13, a resonance phenomenon that may be used as an amplification mechanism for recognition of an input is analyzed.

2. Spike Coding

The fundamental unit of information and signaling in the nervous system is the action potential. Because an action potential is a stereotyped response, its shape and amplitude do not convey information. Instead, all information is carried by action potential timing. The response of a single neuron can be specified by listing the times $t_1, t_2, t_3, \ldots$ when it fires action potentials. The spiking output of a neuron can be characterized as a series of delta functions at the times when action potentials occur,

$$s(t) = \sum_i \delta(t - t_i) \ . \tag{2.1}$$

This representation of neuronal firing is complete but awkward to work with because $s(t)$ is not a well-behaved function and because a complete listing of individual spike times may be more information than can be handled practically or computationally. These problems are resolved by spike decoding.

3. Spike Decoding

In the most general terms, spike decoding is a method for generating a time-dependent variable $r(t)$ that describes, in some useful way, the spike train being produced by a neuron. One way to define $r(t)$ is by using a linear filter (Bialek, 1989; Bialek *et al.*, 1991; Reike, 1991). This expresses the variable $r(t)$ characterizing the spike train $s(t)$ as a convolution of the spike train signal with a specific kernel,

$$r(t) = \int_0^\infty dt' \ K(t')s(t - t') \ . \tag{3.1}$$

The best kernel K to use in equation (3.1) depends on what information $r(t)$ is suppose to convey. Suppose that a neuron is firing in response to some stimulus $x(t)$. It is common for modeling studies to address the question: How does the stimulus produce the observed spike train? However, decoding allows us to ask the question: How is information about the stimulus encoded in the series of spikes being produced? To address this question we can decode the spike train by constructing an estimate $x_{est}(t)$ of the stimulus value. One method of decoding is to use $r(t)$ itself as a direct estimate of the value of the stimulus (Bialek, 1989; Bialek *et al.*, 1991; Reike, 1991),

$$x_{est}(t) = r(t) \ . \tag{3.2}$$

The key is to find the kernel that generates the $x_{est}(t)$ providing the best possible direct estimate of the value of the stimulus. To do this we demand that $x_{est}(t)$ be as close to the stimulus value $x(t)$ as possible by minimizing the square of the difference between them averaged over time,

$$\int dt \ [x_{est}(t) - x(t)]^2 = \text{minimum} \ . \tag{3.3}$$

Such an approach has been used with success to reconstruct coded stimuli from the spike trains of movement-sensitive neurons in the blowfly (Bialek *et al.*, 1991), as well as neurons in other systems (Bialek, 1989; Warland *et al.*, 1991; Reike, 1991; Theunnisen, 1993). The kernel that accomplishes this task can be computed by the method of Fourier transforms (Bialek, 1989; Bialek *et al.*, 1991) or it can be constructed by applying a modification rule that gradually changes the kernel to improve the match between the true stimulus and the predicted value (Sen & Abbott, unpublished).

To describe the kernel it is easiest to divide time up into discrete intervals $t = n\Delta t$ for integer n and to define the kernel in terms of a discrete array by

$$K(n\Delta t) = K_n \ . \tag{3.4}$$

For practical reasons, the kernel is defined to be nonzero over a finite time range so that $K_n = 0$ for $n > M$ with M some fixed value. The spike train is similarly defined by a variable

$$S_n = \int_{(n-1)\Delta t}^{n\Delta t} dt\ s(t) \tag{3.5}$$

that is equal to the number of action potentials occurring in the interval $(n-1)\Delta t < t < n\Delta t$. If Δt is small enough, S_n will be either zero or one. In terms of these variables, the response function which is to provide the estimate can be written as

$$x_{est}(n\Delta t) = \sum_{m=0}^{M} K_m S_{n-m} \ . \tag{3.6}$$

The kernel modification rule is based on the minimization condition (3.3). Using the gradient decent algorithm we find that the modification

$$K_m \to K_m + \epsilon\left[x_{est}(n\Delta(t)) - x(n\Delta t)\right] S_{n-m} \tag{3.7}$$

for $m = 0, \ldots, M$, with ϵ a small parameter, will reduce the squared difference between $x_{est}(t)$ and $x(t)$.

The procedure for generating a kernel is to start with random elements K_n, to compute the resulting estimate $x_{est}(t)$ corresponding to a stimulus $x(t)$ and then to apply the modification rule (3.7) repeatedly until the difference between $x_{est}(t)$ and $x(t)$ has decreased to a desired level of accuracy. Figure 1 shows the results of such a process. A quasi-periodic signal $x(t)$ was generated and used as a current input to an integrate-and-fire model neuron. The decoded signal $x_{est}(t)$ was then computed from the spike train produced by the model neuron using equation (3.6). The kernel used initially in figure 1a was random, as shown in figure 1b, and the resulting estimate does not match the input data. Figure 1c shows a comparison of $x_{est}(t)$ and $x(t)$ after the modification process constructed the kernel seen in figure 1d. This kernel provides a good estimate of the signal for large positive values that produce many spikes. For smaller or subthreshold inputs the reconstruction is not as accurate. There is also a small delay in the reconstructed value caused by the latency to the first spike in response to an input. Because the modification rule develops the kernel over an extended period, the final form of the kernel reflects both the firing properties of the neuron and the statistics of the input data stream.

Generating a direct estimate of the stimulus signal, $x_{est}(t) = r(t)$, is not the only thing that can be done with a linear filter. Other measures of the spike train given by $r(t)$ may also be useful. For example, the method can be used when the stimulus signal is unknown. In such cases, K can be chosen so that the information content or entropy of $r(t)$ is maximized (Abbott & Salinas, unpublished). This is done by making the distribution of r values over time as constant as possible. Alternately, the kernel can be chosen to generate an $r(t)$ that can be considered to be a measure of the firing rate of the neuron.

In many cases, the rate of action potential firing is the desired measure of neuronal response to a stimulus, although additional information may be encoded in other characteristics of the firing patterns (Optican & Richmond, 1987; Gozani & Miller, 1994; Middlebrooks *et al*;, 1994). Despite the fact that firing rate is used in this way in most experimental and theoretical papers, often little care is taken in defining what is meant by a firing rate. Since neurons do not fire with perfect periodicity indefinitely, some care must be taken in defining what we mean by the firing rate of a neuron. One method used to define the firing rate is to choose a time interval, count the number of spikes appearing during this period of time and then define the rate as the number of spikes divided by the length of the interval. This measure of the firing rate has the undesirable property that it is only defined at discrete values of the time. Some neuron models actually do use such a definition and update the firing rate discontinuously at discrete time intervals (McCullock & Pitts, 1943; Hopfield, 1982; Amit, 1989; Hertz *et al.*, 1990). However, it is often more useful to generate a firing rate variable that is defined at all times. The oxymoron 'instantaneous firing rate' is often used to denote such variables.

The linear filter can be used to define a firing rate variable if an appropriate constraint is place on the kernel. Suppose that the neuron is firing at a constant rate with period T. We require that, if $r(t)$ is to be considered a firing rate variable, its value averaged over one period T must be equal to the firing rate $1/T$. From definition (3.1), this means that we must have

$$\frac{1}{T}\int_0^T dt \int_0^\infty dt' \; K(t')s(t-t') = \frac{1}{T} \; . \tag{3.8}$$

By switching the order of integration and eliminating the common factor of $1/T$, this can be rewritten as the condition

$$\int_0^\infty dt' \; K(t') \int_0^T dt \; s(t-t') = 1 \; . \tag{3.9}$$

The second integral is always equal to one when the neuron is firing with period T so we find that $r(t)$ will qualify as a firing rate variable provided that

$$\int_0^\infty dt \; K(t) = 1 \; . \tag{3.10}$$

Such a firing-rate variable is just a special case of a linear filter with the kernel required to satisfy condition (3.10). Conversely, every decoded signal produced by a linear filter is proportional to a quantity that can be considered to be a firing rate. Condition (3.10) obviously leaves an infinite variety of kernels that can be used. Some possibilities are:

$$K(t) = \frac{1}{\tau}\exp(-t/\tau) \tag{3.11}$$

or

$$K(t) = \frac{t}{\tau^2}\exp(-t/\tau) \tag{3.12}$$

or

$$K(t) = \frac{1}{\tau_1 - \tau_2}\left[\exp(-t/\tau_1) - \exp(-t/\tau_2)\right] \tag{3.13}$$

with $\tau_1 > \tau_2$. All three of these have obvious interpretations in terms of down-stream processing of the spike train since forms like these are often used to characterize postsynaptic responses. The exponential kernel (3.11) defines the firing rate as a low-pass filtered version of the spike train. The resulting firing rate jumps up every time there is a spike and decays exponentially toward zero with time constant τ when no spikes are present. The discontinuous jump in the firing rate produced by this kernel is eliminated by the other two kernels, the alpha-function (3.12) and difference of exponentials (3.13), because they do not respond instantaneously. The second of these, (3.13), has the advantage that the rise time τ_2 and fall time τ_1 can be adjusted independently.

4. Population Coding

An extremely widespread form of neural coding involves large arrays neurons that represent information collectively as an ensemble (Knudsen *et al.*, 1987; Konishi, 1987; Sejnowski, 1988; Konishi, 1991; Eichenbaum, 1993). In such arrays, individual neurons are broadly-tuned to a stimulus value. Individual firing rates depend on and reflect the value of the coded information, but to be accurate the decoding should be based on action potentials coming from many neurons. Many examples of information coded by the firing rates of arrays of tuned neurons can be given. A partial list of neuronal arrays and the corresponding coded quantities includes: neurons in the primary visual cortex of a cat coding the orientation of a bar of light (for a review see Orban, 1984), the direction of arm movements of a monkey coded by neurons in its motor cortex (Georgopolous *et al.*, 1986, 1988 & 1993), neurons in the superior colliculus encoding saccade direction (Van Gisbergen *et al.*, 1987; Lee *et al.*, 1988), interneurons in the cercal system of the cricket coding the direction of a wind stimulus (Bacon & Murphey, 1984; Miller *et al.*, 1991; Theunissen & Miller, 1991; Theunissen, 1993) and a similar system in the cockroach (Camhi & Levy, 1989), the direction of a sound stimulus coded in neural arrays of the barn owl (Knudsen & Konishi, 1978; Konishi, 1991), the position of a rat in its environment coded by hippocampal place cells (O'Keefe & Dostovsky, 1971; O'Keefe & Nadel, 1978; O'Keefe, 1979; Wilson & McNaughton, 1993), MT neurons coding visual motion direction (Maunsell & Newsome, 1987) and neurons coding echo delay in the bat (O'Neill & Suga, 1982).

Population coding has been the subject of a number of theoretical analyses (Paradiso, 1988; Vogels, 1990; Foldiak, 1991 & 1993; Zohary, 1992; Snippe & Koenderink, 1992; Seung & Sompolinsky, 1993; Touretzky *et al.*, 1993; Salinas & Abbott, 1994), some related to the issue of hyperacuity (Baldi & Heiligenberg, 1988; Altes, 1989; Lehky & Sejnowski, 1990; Zhang & Miller, 1991) and others connected with network models (Lukashin, 1990; Burnod *et al.*, 1992; Van Opstal & Kappen, 1993).

In experiments involving coding by arrays of tuned neurons, the average responses or tuning curves of individual neurons are measured as functions of the value of the stimulus variable x. I will use an index that runs from one to N, the total number of neurons in the array, to label individual neurons. The average firing response of neuron i to a stimulus that takes the value x is denoted by $f_i(x)$. For the purposes of this review, I will consider an idealized array of tuned neurons with Gaussian tuning curves of identical shape but varied optimal stimulus value,

$$f_i(x) = r_{max} \exp\left(-(x - y_i)^2/2\sigma_f^2\right) \tag{4.1}$$

as shown in figure 2. The parameter y_i is the stimulus value at which neuron i responds with its maximum average firing rate r_{max}. The width of the tuning curves is $2\sigma_f$.

If the tuning curves cover the range of x densely and they overlap as in figure 2, we can approximate sums over neurons by integrals over the positions y where the neurons fire at their maximum rates, as follows

$$\sum_{i=1}^{N} \to \rho \int dy \ . \tag{4.2}$$

Here ρ is the neuronal tuning curve density which is the number of neurons covering a unit range of x. We will ignore the edges of the array where the coded value reaches the limits of its range by making this range infinite. Some useful results of these approximations are:

$$\sum_{i=1}^{N} f_i(x) f_i(y) = \sqrt{\pi}\rho\sigma_f r_{max}^2 \exp\left(-(x - y)^2/4\sigma_f^2\right) \tag{4.3}$$

from which we find

$$\sum_{i=1}^{N} f_i'(x) f_i(y) = \frac{\sqrt{\pi}\rho r_{max}^2}{2\sigma_f}(y - x) \exp\left(-(x - y)^2/4\sigma_f^2\right) \tag{4.4}$$

and

$$\sum_{i=1}^{N} \left[f_i'(x)\right]^2 = \frac{\sqrt{\pi}\rho r_{max}^2}{2\sigma_f} \tag{4.5}$$

where the prime denotes differentiation. Another useful formula for the derivations that follow is

$$\int dz \exp\left(-\frac{(x - z)^2}{2\sigma_1^2} - \frac{(y - z)^2}{2\sigma_2^2}\right) = \frac{\sqrt{2\pi}\sigma_1\sigma_2}{\sqrt{\sigma_1^2 + \sigma_2^2}} \exp\left(-\frac{(x - y)^2}{2(\sigma_1^2 + \sigma_2^2)}\right) \tag{4.6}$$

These equations will be used to derive the results given in the following sections.

5. Population Decoding

When information about a stimulus or motor output is coded by an entire ensemble of neurons, decoding the full population response requires procedures for combining the firing rates of many neurons into a population ensemble estimate. Georgopoulos and collaborators developed one such approach by defining the population vector relating activity in the motor cortex of the monkey to the direction of arm movements (Georgopoulos *et al.*, 1986, 1988 & 1993). The population vector method developed for this system has been used in similar arm movement studies of premotor cortex (Caminiti *et al.*, 1991), parietal area 5 (Kalaska *et al.*, 1983) and cerebellum (Fortier *et al.*, 1989). In addition, the vector method has been applied to other systems such as primary visual cortex (Gilbert & Wiesel, 1990), parietal visual neurons (Steinmetz *et al.*, 1987) and inferotemporal neurons in the monkey responding to human faces (Young & Yamane, 1992). This last application involved the use of multi-dimensional scaling to relate the responses of face cells to the encoding of a two-dimensional vector. Wilson and McNaughton (1993) have used a different method to decode the output of position selective place cells (O'Keefe & Dostovsky, 1971; O'Keefe & Nadel, 1978; O'Keefe, 1979) in the hippocampus of the rat. A number of different methods can be used to decode neural population activity in a variety of systems (for a review see Salinas & Abbott, 1994).

Decoding the activity of a neural ensemble means computing an estimate of the coded quantity x_{est} from the firing rates of neurons in the ensemble. This is similar in spirit to the procedure discussed in section 3 but this time the estimate is based on the firing of many neurons, not one. Typically, this calculation makes use of the tuning curves $f_i(x)$ of individual neurons $i = 1, 2, \ldots, N$ responding to the stimulus x. If the decoding schemes works well, x_{est} will typically be close to the true value of the stimulus x. However, this need not always be the case. We will use decoding in another way as well. If something in a neural network changes, such as the strengths of synapses between neurons, this may shift the value of x_{est} constructed by the decoding scheme. In this case, the shift in x_{est} may tell us something about the meaning and significance of the synaptic changes that have occurred. As discussed in section 9, if x_{est} is different from x this may have functional significance rather than reflecting inaccuracy of the decoding scheme.

Several methods can be used to construct a population-based estimate of a coded signal. We will denote the collection of all firing rates r_i for $i = 1, 2, \ldots N$ by the notation $\mathbf{r}$ and likewise the collection of tuning curves by $\mathbf{f}$. The most sophisticated methods require knowledge of the probability that the stimulus value is x given that the population firing response was $\mathbf{r}$, denoted $P(x|\mathbf{r})$. The commonly used maximum likelihood method consists of finding the value of x with the maximum probability given a set of firing rates and using this for the estimate. Thus, $P(x_{est}|\mathbf{r})$ is maximized. $P(\mathbf{r}|x)$, the probability of a given stimulus producing a set of firing rates, can be obtained from experimental measurements and then $P(x_{est}|\mathbf{r})$ can be determined from this by Bayes' theorem

$$P(x|\mathbf{r})P(\mathbf{r}) = P(\mathbf{r}|x)P(x) \ . \tag{5.1}$$

Because the maximum likelihood procedure maximizes the probability with respect to the value of x, the factor $P(\mathbf{r})$ is not relevant and we can express the maximum likelihood condition as

$$P(\mathbf{r}|x_{est})P(x_{est}) = \text{maximum}. \tag{5.2}$$

If the population of neurons is large and the necessary probabilities are know, the maximum likelihood method is probably the optimal decoding procedure to use. However, if the population of neurons being used in the decoding is not large, there is no rigorous theoretical justification for using the maximum likelihood method. Instead, it may be better to use a Bayesian decoding scheme that is guaranteed to minimize the average squared difference between the estimated value and the true stimulus value. This estimate is just the average value of x corresponding to the probability distribution $P(x|\mathbf{r})$,

$$x_{est} = \int dx\, x P(x|\mathbf{r}) \ . \tag{5.3}$$

It many cases, the probability $P(\mathbf{r}|x)$ may not be known. A decoding method that has been used in this case (Churchland & Sejnowski 1992; Wilson & McNaughton 1993) chooses the estimated value so that

$$\sum_{i=1}^{N} r_i f_i(x_{est}) \left[\sum_{i=1}^{N} r_i^2\right]^{-1/2} \left[\sum_{i=1}^{N} f_i^2(x_{est})\right]^{-1/2} = \text{maximum} \ . \tag{5.4}$$

Both $\mathbf{r}$ and $\mathbf{f}$ can be considered to be vectors in the N-dimensional space of neural responses. Equation (5.4) minimizes the angle between the true response vector $\mathbf{r}$ and the mean response vector $\mathbf{f}$. This method is particularly useful if there are population wide variations in the excitability of the neurons being decoded. Since this review covers theoretical applications of decoding, this type of variability will not be a problem and I will use the simpler least squares approach discussed next.

A good general method for population decoding is to choose x_{est} so that the mean firing rates $\mathbf{f}(x_{est})$ provide the best possible match to the observed rates $\mathbf{r}$. This can be done by performing a least squares fit of the average rates $\mathbf{f}(x_{est})$ to the observed rates $\mathbf{r}$

$$\sum_{i=1}^{N} [r_i - f_i(x_{est})]^2 = \text{minimum}. \tag{5.5}$$

The functions $\mathbf{f}(x)$ represent the average firing responses of the neurons in the array when the coded quantity takes the value x. However, due to noise and neuronal variability there will be variations around these mean firing rates. Because of this variability, the decoded value x_{est} will not always be identical to the true value x. On average the two quantities will be the same, but x_{est} will fluctuate around its average due to the variability of the firing rates. Suppose that these fluctuations are uncorrelated and have zero mean and standard deviation σ_r. The variance of the estimate x_{est} given by the least-squares method, σ_x^2, is then

$$\sigma_x^2 = \sigma_r^2 \left(\sum_{i=1}^{N} [f_i'(x)]^2 \right)^{-1} \tag{5.6}$$

where the prime denotes differentiation. For the case of Gaussian tuning curves, as in figure 2, this reduces to

$$\sigma_x^2 = \frac{2\sigma_r^2 \sigma_f}{\sqrt{\pi} \rho r_{max}^2} \,. \tag{5.7}$$

The higher the density of coverage, the lower the decoding error and the dependence of σ_x on ρ indicates that this is a typical square root of N effect. The error depends directly on the variability of the individual neuronal rates but only on the square root of the tuning curve width.

6. Decoding Direction

The coding and decoding of direction is an interesting special case of population coding. In this case the coded quantity is a vector $\vec{x}$ with multiple components x_μ for $\mu = 1, 2, \ldots, d$ when a d-dimensional vector is being coded. The dimension d is typically either 2 or 3. If only the direction and not the magnitude of $\vec{x}$ is being coded, it is convenient to make $\vec{x}$ a unit vector so that $|\vec{x}| = 1$. The Cartesian components x_μ of the vector $\vec{x}$ are the projections of $\vec{x}$ onto unit vectors $\vec{C}_\mu$ that form orthogonal coordinate axes,

$$x_\mu = \vec{x} \cdot \vec{C}_\mu = \cos(\theta_\mu) \tag{6.1}$$

where θ_u is the angle between $\vec{x}$ and $\vec{C}_\mu$. The inverse procedure, reconstruction of a vector from its components, is accomplished by summing the component values times the appropriate unit vectors

$$\vec{x} = \sum_{\mu=1}^{d} x_\mu \vec{C}_\mu \,. \tag{6.2}$$

Vectors are typically represented mathematically as in equation (6.1), by components projected along orthogonal, Cartesian coordinate axes. Remarkable, many neuronal circuits construct essentially the same representation (Touretzky *et al.*, 1993; Salinas & Abbott, 1994). The indication that this is happening is that the tuning curves $\mathbf{f}(\vec{x})$ are in the form of cosine functions just like equation (6.1) so that the firing rates are proportional to the projections of the coded vector along preferred axes defined by the neurons. Cosine tuning curves are found in many systems. There are interneurons in the cricket cercal system that code wind direction at low velocity by firing at rates that are proportional to the Cartesian coordinates of the wind direction vector (Theunissen & Miller, 1991; Theunissen, 1993). The systems where the vector method has been applied to monkey arm movements, in motor cortex (Schwartz *et al.*, 1988), premotor cortex (Caminiti *et al.*, 1991), parietal area 5 (Kalaska *et al.*, 1983) and cerebellum (Fortier *et al.*, 1989) all exhibit cosine tuning curves. Cosine responses are also found in neurons coding body (Suzuki *et al.*, 1985) and head position in cats (Shor *et al.*, 1984) and head direction in rats (Chen *et al.*, 1990; Taube *et al.*, 1990). Parietal visual neurons fire at rates proportional to the cosine of the angle of movement direction of a visual stimulus (Steinmetz *et al.*, 1987). Of course, cosine functions are both positive and negative while firing rates cannot be negative. In some systems, this means that only half of the cosine function can be represented by one neuron. For example, in the cercal system of the cricket, one neuron represents the positive part of the cosine function and a second neuron is used for the negative part (Theunissen & Miller, 1991; Theunissen, 1993). In other systems like the monkey motor cortex (Schwartz *et al.*, 1988), there is a background rate of firing. If this is subtracted out in the definition of the firing rate, then $r(t)$ can become negative. As a result, a larger portion or even all of the cosine curve can be represented by a single neuron.

The population vector decoding method (Georgopoulos *et al.*, 1986, 1988 & 1993) is based on the standard formula for reconstructing a vector from its components, equation (6.2). If neuron i responds to stimulus $\vec{x}$ by firing at a rate proportional to $\vec{x} \cdot \vec{C}_i$, the vector method estimate of $\vec{x}$ is

$$\vec{x}_{est} = \sum_{i=1}^{N} r_i \vec{C}_i \tag{6.3}$$

just like (6.2). It can be shown that for a uniform distribution of preferred direction vectors $\vec{\mathbf{C}}$, the direction of $\vec{x}_{est}$ will converge to that of $\vec{x}$ provided that enough neurons are included in the sum (Georgopoulos *et al.*, 1988).

There is a complication with neural coding of direction that is not present in the usual Cartesian system. The number of neurons typically exceeds the number of coordinate axes, $N \gg d$, and the axes $\vec{\mathbf{C}}$ defined by the neurons do not form an orthogonal system. We can correct for this by using a method similar in spirit to the Bayesian method of decoding but linear in the firing rates. We express the estimated vector as a linear sum similar to equations (6.2) and (6.3)

$$\vec{x}_{est} = \sum_{i=1}^{N} r_i \vec{D}_i \tag{6.4}$$

except that the basis vectors $\vec{\mathbf{D}}$ are not set equal to the projection vectors $\vec{\mathbf{C}}$ but are determined by requiring that this linear estimate minimizes the average squared error $(\vec{x}_{est} - \vec{x})^2$.

The resulting basis vectors are (Salinas & Abbott, 1994)

$$\vec{D}_i = \sum_{j=1}^{N} Q_{ij}^{-1} \vec{L}_j \tag{6.5}$$

with

$$\vec{L}_j = \int d\vec{x}\, \vec{x} f_j(\vec{x}) \tag{6.6}$$

and

$$Q_{ij} = \sigma_r^2 \delta_{ij} + \int d\vec{x}\, f_i(\vec{x}) f_j(\vec{x}) \,. \tag{6.7}$$

The vectors $\vec{\mathbf{L}}$ represent the 'center of mass' of the tuning curves. For symmetric tuning curves they point in the same directions as the preferred direction vectors $\vec{\mathbf{C}}$. $\mathbf{Q}$ is the correlation matrix of firing rates. This method provides the optimal linear estimate of the coded signal.

7. Network Models

As well as being a powerful experimental tool, decoding is a useful technique for studying model neural networks because it provides a concise description of population activity that has direct functional significance. We will consider a model network with the architecture shown in figure 3. Information about the stimulus value x is carried into the network by a set of input firing rates. The firing rate of input a when the stimulus takes the value x is $g_a(x)$ and the synaptic coupling from input a to neuron i has strength M_{ia}. As a result of these inputs, neuron i fires at a rate $f_i(x)$ in response to the stimulus. The firing of neuron i is also affected by recurrent synapses between network neurons. The strength of the synaptic connection from neuron j to neuron i is given by the weight W_{ij}.

Using this architecture we can address many issues concerning network function and plasticity. How do the response tuning curves $\mathbf{f}(x)$ arise from the afferent activity $\mathbf{g}(x)$ and the synaptic couplings $\mathbf{M}$ between the input lines and the coding neurons? How do modifications in the strengths of the input synapses affect the response properties of the network neurons and the ensemble-coded value of the stimulus? What is the role of the recurrent synapses within the network represented by the matrix $\mathbf{W}$? How do modifications in the strengths of recurrent synapses affect the response properties of the network neurons and ensemble-coded value of the stimulus? In the following sections, I will discuss general techniques for addressing these questions and will present some results obtained by applying them.

Some applications require us to know how the firing rates $\mathbf{r}$ of the network neurons depend on the input firing rates and on the synaptic weights connecting network neurons to their inputs and to each other. It is simplest to assume a linear formula,

$$r_i = \sum_a M_{ia} g_a(x) + \sum_j W_{ij} r_j \tag{7.1}$$

where the first term represents the effect of input firing and the second term recurrent network connections. I will discuss corrections to this linear approximation in section 12. Equation (7.1) refers to static or quasi-static rates. The case of dynamically varying rates will be considered in section 13.

8. Synaptic Modification and Learning Rules

It is widely believed that modification of synaptic strength is the basic phenomenon underlying learning and memory (Morris *et al.*, 1986; Byrne & Berry, 1989; Gluck & Rumelhart, 1990; Baudry & Davis, 1991; Hawkins *et al.*, 1993). In addition, activity-dependent synaptic plasticity plays an important role in the development of the nervous system (Shatz, 1990; Miller, 1990). A large body of experimental evidence supports the notion that neuronal activity can affect synaptic strength and specific features of both LTP (Bliss & Collingridge, 1993; Malenka & Nicoll, 1993) and LTD (Artola & Singer; 1993) of synapses have been characterized. In addition, evidence and ideas about the molecular basis for these phenomenon are accumulating (Lynch *et al.* 1983; Lisman & Goldring, 1988; Lisman, 1989; Gustafsson & Wigstrom, 1988; Jahr & Stevens, 1990; Madison *et al.*, 1991). To reveal the functional and behavioral significance of LTP and LTD we must understand how experience and training modify synapses and how the resulting modifications change patterns of neuronal firing and affect behavior. Work on the first of these two issues has led to a large number of neural network learning rules proposing how activity and training experience change synaptic weights (Hebb, 1949; Sejnowski, 1977; Bienenstock *et al.*, 1992; reviewed in Abbott, 1990). Proposals concerning the second of these questions include ideas about associative memory and pattern recognition (Marr, 1971; Hopfield, 1982 & 1984; Grossberg, 1988; Kohonen, 1984 & 1988; Rolls, 1989; Amit, 1989; Hertz *et al.*, 1990), storage of statistics (Sejnowski, 1977; Levy *et al.*, 1990) and storage and recall of paths or sequences of motor actions (Blum & Abbott, 1996; Abbott & Blum, 1996).

The form of synaptic modification that has attracted the most attention for models of learning and memory is Hebbian LTP (Hebb, 1949). For potentiation of a synapse to occur through this mechanism, both the presynaptic and postsynaptic neurons must be depolarized. This suggests a learning rule in which the strength of a synapse from neuron j to neuron i changes at a rate that is proportional to the product of the firing rates r_i and r_j,

$$\frac{dW_{ij}}{dt} = \epsilon_W r_i r_j \; . \tag{8.1}$$

Here ϵ_W is a parameter that determines the rate of the synaptic change. Similarly, the strength of a synapse from input a to neuron i will be modified at a rate

$$\frac{dM_{ia}}{dt} = \epsilon_M r_i g_a \tag{8.2}$$

with ϵ_M the corresponding rate constant for $\mathbf{M}$. Of course, these equations assume that a synapse is there to be modified. If no physical synapse exists, the corresponding element W_{ij} or M_{ia} must be kept at the value zero. In particular, synapses between a neuron and itself, $i = j$ in equation (8.1), should not be included. In practice, it is common to assume all-to-all coupling and ignore these constraints. This is a valid approximation for large, highly-coupled networks.

Equations (8.1) and (8.2) allow the synaptic strengths to grow without bound if the pre- and postsynaptic firing rates are nonzero for a long period of time. This is clearly unreasonable so some form of constraint must be imposed to prevent the synaptic weights from growing too large (Oja, 1982; Linsker, 1988 & 1993; Miller & MacKay, 1994). I will not discuss these constraints, assuming that synaptic potentiation stops before the constraint is reached. This requires a mechanism that is capable of turning synaptic potentiation on and off (Bortolotto *et al.*, 1994). We can also justify ignoring the constraints by assuming that the initial pattern of changes in synaptic strengths can be used to characterize the final outcome of those changes (MacKay & Miller, 1990).

We would like to relate the changes in synaptic strength characterized by the above equations to the experience of an animal during training or development. We will divide this problem into two parts. During development, we are interested in the large changes that give rise to the tuning curves $\mathbf{f}$. During training following development, the changes are apt to be more subtle and we can consider small deviations from the basic firing rates $\mathbf{f}$ that are induced by the training experience. Because of this, the computation of training induced changes requires fewer assumptions and we will begin with this case and discuss it further in section 9.

Equations (8.1) and (8.2) require a knowledge of the firing rates $\mathbf{r}$. If the tuning curves have been developed previously and if the changes in $\mathbf{W}$ and $\mathbf{M}$ are not too large, we can use our knowledge of the tuning curves to predict the firing rates. Thus, we use the approximation $\mathbf{r} = \mathbf{f}(x)$ in equations (8.1) and (8.2). Suppose that during the training period the stimulus is described by $x = X(t)$. Then, the firing rates of the network neurons are given by $\mathbf{f}(X(t))$ so, from (8.1) and (8.2),

$$\frac{dW_{ij}}{dt} = \epsilon_W f_i(X(t)) f_j(X(t)) \tag{8.3}$$

and

$$\frac{dM_{ia}}{dt} = \epsilon_M f_i(X(t)) g_a(X(t)) \ . \tag{8.4}$$

These equation can be integrated to determine the weight changes resulting from the entire training period (see section 9).

Equations (8.3) and (8.4) ignore a feature of LTP that may have very important consequences. The efficacy of LTP depends critically on the relative timing between the pre- and postsynaptic activity (Gustafsson *et al.*, 1987). If these are simultaneous, LTP occurs as discussed above. However strict simultaneity of activity is not required. Postsynaptic activity can follow presynaptic activity by 100 to 200 ms and LTP can still occur. On the other hand, if postsynaptic activity precedes presynaptic activity, LTP does not occur. Instead, LTD may result from this sequence of stimulation (Debanne *et al.*, 1994). If we let $H(t)$ represent the relative efficacy of LTP when postsynaptic activity follows presynaptic activity by a time t, the rate of change in synaptic strengths is given by

$$\frac{dW_{ij}}{dt} = \epsilon_W \int dt' \, H(t') f_i(X(t+t')) f_j(X(t)) \tag{8.5}$$

and

$$\frac{dM_{ia}}{dt} = \epsilon_M \int dt' \, H(t') f_i(X(t+t')) g_a(X(t)) \tag{8.6}$$

which replace equations (8.3) and (8.4). The temporal asymmetry of H has a dramatic effect on the synaptic weight changes produced by training because it causes them to depend on the temporal order of the activity generated during training. This allows synaptic weights to store information about temporal sequences and has been proposed as a mechanism for learning paths and sequential motor actions (Blum & Abbott, 1996; Abbott & Blum, 1996). However, in the cases considered here, temporal ordering will not play an important role and equations (8.3) and (8.4) will lead to the same results as equations (8.5) and (8.6).

In studying development of tuning curves, we cannot rely on the approximation $\mathbf{r} = \mathbf{f}(x)$ used above because this assumes that the functions $\mathbf{f}$ are already well established. Instead, we will use the linear approximation for the firing rates of equation (7.1). To simplify the discussion of tuning curve development I will ignore the effect of recurrent synapses by setting $\mathbf{W} = 0$. The basic synaptic modification rule is obtained by substituting the firing rates given by (7.1) with $\mathbf{W} = 0$ into (8.2),

$$\frac{dM_{ia}}{dt} = \epsilon_M \sum_b g_a(X(t))g_b(X(t))M_{ib} \, . \tag{8.7}$$

Here $X(t)$ is the set of stimulus values presented during the development period. This equation for the development of tuning curves will be analyzed in section 10.

9. The Effects of Synaptic Modification on Coding

The responses of neurons in a population coding the variable x are characterized by the set of firing rate tuning curves $\mathbf{f}(x)$. These tuning curves are shaped by both input connections and recurrent synapses between network neurons. Recurrent synaptic connections between network neurons cause the firing rates of the neurons to depend on each other as well as on the input firing rates. The effect of the firing rates on each other is considered in section 11. Here we analyze the effect of changes in the strengths of synapses. Suppose that we have mapped out the average firing responses $\mathbf{f}$ and then some modifications occur due to experience and learning that change the synaptic weights by amounts $\Delta\mathbf{W}$ and $\Delta\mathbf{M}$. These changes can be computed by integrating equations (8.3) and (8.4) (or 8.5 and 8.6) over time. If they are small, we can Taylor expand and write

$$r_i = f_i(x) + \sum_a \Delta M_{ia} g_a(x) + \sum_j \Delta W_{ij} f_j(x) \, . \tag{9.1}$$

This result is valid to first order in the synaptic changes assuming that the corresponding dependence on the firing rates is linear as in equation (7.1). Equation (9.1) is a useful result because it allows us to compute the effect of the training experience on the response properties of neurons.

Suppose that after the changes represented by $\Delta\mathbf{M}$ and $\Delta\mathbf{W}$ we decode using the original response functions $\mathbf{f}$ as in section 5. The result we will get is no longer $x_{est} = x$ as it was before because of the shift in neuronal responses implied by the second term on the right side of equation (9.1). Instead

$$x_{est} = x + \frac{\sum_i f_i'(x) \left[\sum_a \Delta M_{ia} g_a(x) + \sum_j \Delta W_{ij} f_j(x)\right]}{\sum_i [f_i'(x)]^2} \, . \tag{9.2}$$

This result is obtained by minimizing the squared difference (5.5) with respect to x_{est} using equation (9.1) for the firing rates $\mathbf{r}$ and computing to lowest order in the weight changes. Equation (9.2) allows us to evaluate the effect of synaptic modifications on the value of a coded quantity. This effect is a shift, given by the second term on the right side of equation (9.2), in the coded value away from the stimulus value.

To decode the firing rates given by (9.1) and derive the result (9.2), we used the original tuning curves $\mathbf{f}$ to match with the actual firing rates $\mathbf{r}$ despite the fact that these are now given by (9.1) not by $\mathbf{r} = \mathbf{f}$. This assumes that downstream neural networks have not themselves been changed by the training experience that induced the changes of synaptic weights represented by $\Delta\mathbf{M}$ and $\Delta\mathbf{W}$. If, at a later time, these networks are modified so that the appropriate tuning curves to use in decoding are those given by equation (9.1) rather than by the original $\mathbf{f}(x)$, the shift in equation (9.2) will be eliminated and we will find $x_{est} = x$ once again. This is an effective way of erasing the effects of learning without having to reset the synaptic weights that were modified during the training process.

The results of this and the previous section give us enough information to predict how exposure to a stimulus during a training period will affect synaptic strengths, how changes in synaptic strengths modify patterns of firing and how the modified firing patterns affect the value of a quantity encoded by the neural ensemble. After a training period during which the stimulus takes a sequence of values $X(t)$, if we subsequently expose the animal to a stimulus x, neuron i will fire at a rate given by equation (9.1) with $\Delta\mathbf{W}$ and $\Delta\mathbf{M}$ determined by integrating equations (8.3) and (8.4) (or 8.5 and 8.6). On the basis of these equations, we can determine how network firing and the value of x_{est} obtained from that firing are affected by exposure to a time-dependent stimulus $X(t)$ during training. This relates the value of a coded quantity that is of behavioral relevance to the stimulus experience of the animal during training and allows a direct computation of the effects of training on coding. The results of some such calculations are given below and in figures 4 and 5 (see also Abbott & Blum, 1996).

In the examples given here, I will only consider the effects of changes of recurrent synapses between network neurons and will set $\Delta\mathbf{M} = 0$. The first case considered is an $X(t)$ that varies randomly but uniformly over the entire x range. This corresponds to exposing the animal to a random sequence of stimuli during the training period. Using the results given at the end of section 4, we find that such exposure changes the recurrent synaptic strengths by an amount

$$\Delta W_{ij} = g_W \sqrt{\pi} r_{max}^2 \sigma_f \exp\left(-(y_i - y_j)^2/4\sigma_f^2\right) \tag{9.3}$$

where g_W is a constant related to ϵ_W and the duration of the training period. The resulting effect of this sort of training on the neuronal response tuning curves is determined by inserting this expression into equation (9.1) and again using the results at the end of section 4. We find that after this random training

$$r_i = f_i(x) + \frac{2\pi g_W \rho r_{max}^3 \sigma_f^2}{\sqrt{3}} \exp\left(-(x - y_i)^2/6\sigma_f^2\right) \,. \tag{9.4}$$

The second term is a wider Gaussian curve than the original tuning curve $f_i(x)$. Thus, exposure to random stimuli during training tends to broaden the tuning curves. This has no effect on the value of the coded quantity because the tuning curves are not shifted or asymmetrically distorted. As a result, decoding equation (9.4) produces the result $x_{est} = x$.

We now consider a case where the value of the coded quantity is changed by the experience during training. Here the training period consists of exposure to a single stimulus value $X(t) = 0$. Such training changes the recurrent synapses by

$$\Delta W_{ij} = g_W r_{max}^2 \exp\left(-(y_i^2 + y_j^2)/2\sigma_f^2\right) \,. \tag{9.5}$$

The constant g_W defined here is somewhat different than the one appearing in (9.3) but again depends on the rate of synaptic change and on the length of the training period. The rate of firing in response to an input x after this training, computed with the same techniques used for equation (9.4), is

$$r_i = f_i(x) + g_W \sqrt{\pi} r_{max}^3 \rho \sigma_f \exp\left(-(x^2 + 2y_i^2)/4\sigma_f^2\right) . \tag{9.6}$$

This causes an increase in the strength of the response for tuning curves centered at $x = 0$ and also an asymmetric distortion of the tuning curves near this point. The impact of these changes on the decoded value x_{est} as a function of the true stimulus value, computed using equation (9.2), is

$$x_{est} = x\left(1 - g_W \sqrt{\pi} \rho r_{max}^2 \sigma_f \exp\left(-x^2/2\sigma_f^2\right)\right) . \tag{9.7}$$

This is plotted in figure 4. Far from $x = 0$, the coded and actual stimulus values agree. However, stimuli near zero, produce a coded value that is different from the stimulus value. After exposure to the stimulus, the output of the network does not merely represent the stimulus but rather a combination of its value and the training value.

What is the significance of the shift that causes x_{est} to differ from x in the region around the training point $x = 0$, as see in figure 4? Information about the stimuli experienced during training affects the perception of new stimuli. The network interprets inputs near $x = 0$ as if they were located nearer to $x = 0$ than they actually are. This relates to ideas about pattern completion and associative memory in neural networks (Hopfield, 1982 & 1984; Grossberg, 1988; Kohonen, 1984 & 1988; Amit, 1989; Hertz *et al.*, 1990). The point $x = 0$ can be recovered from initial points near zero if we feed the value x_{est} coded by the output of the network back so that it acts as a new input x value. The result of iterating this feedback loop is shown in figure 5. Iteration leads to a coded output value of $x_{est} = 0$ for any initial input x near zero. Feedback of the network coded output to its input could be achieved internally at the network level, or it could arise if the response of an animal to the coded output leads to a new stimulus value equal to the coded value. In network associative memories, only the final fixed-point value, $x = 0$ in this example, is given any functional significance. However, at every value of x the deviation between the coded value and the true value contains information. This difference identifies the direction toward the training point $x = 0$ from any nearby x position.

Recently, it has been shown that networks of this type can store and recall information about paths or sequences of motor actions (Blum & Abbott, 1996; Abbott & Blum, 1996). Suppose, for example, that the input signal x corresponds to sensory information about the position of a limb. Initially, network activity codes for a value x_{est} identical to the true position x. Now suppose that, during a training session, the motion $X(t)$ is repeated while synaptic potentiation takes place. Temporal inputs like this lead to an asymmetric, sequence-dependent potentiation of synapses between coding neurons through the LTP processes described by equation (8.5). After this potentiation, the result of decoding network activity, x_{est}, is shifted away from the input value x by an amount given by equation (9.2). This shift causes the coded value x_{est} to lead the true value x and thus, to provided information about where the limb should go in order to execute the learned motion $X(t)$. If the coded value is fed to motor networks controlling limb motion, the learned sequence will automatically be recalled and generated.

This idea has also been applied to the recall of paths through a spatial environment by the hippocampus (Blum & Abbott, 1996). In the hippocampus of the rat, place cells collectively code information about the spatial position of the animal (O'Keefe & Dostovsky, 1971; O'Keefe & Nadel, 1978; O'Keefe, 1979; Wilson & McNaughton, 1993). If LTP occurs while the rat travels down a particular path, a path-dependent pattern of synaptic potentiation will be established. This LTP has been shown to shift the location coded by place cell ensemble activity away from the location of the rat in a direction toward and along the learned path. Thus, the position of the coded point can be used to guide navigation down the previously learned path.

10. Development of Tuning Curves

In the previous section, we studied how existing tuning curves are modified by synaptic changes and the impact that this has on coded quantities. Now we will consider how larger synaptic modifications can lead to the development of tuning curves. This issue has been analyzed in great depth in studies of the development of ocular dominance and directional selectivity in the primary visual cortex (Malsburg, 1971 & 1979; Swindale, 1980; Linsker, 1986; Miller, 1990, 1992 & 1994; Bienenstock *et al.*, 1992). I will discuss a simpler, generic example to illustrate the general mechanism.

We will examine development without recurrent collateral input so that $\mathbf{W} = 0$. The tuning curves in this case are just the firing rates given by equation (7.1) with $\mathbf{W} = 0$,

$$r_i = f_i(x) = \sum_a M_{ia} g_a(x) \ . \tag{10.1}$$

If the rate parameter ϵ_M controlling the speed of synaptic changes is small so that the synaptic changes are much slower than changes in the stimulus value, we can average over the stimulus history and rewrite (8.7) as

$$\frac{dM_{ia}}{dt} = \epsilon_M \sum_b < g_a g_b > M_{ib} \ . \tag{10.2}$$

Here the brackets $< >$ signify time averaging. The quantity $< g_a g_b >$ is the correlation matrix of input firing rates. Since this matrix has positive eigenvalues, the synaptic weights $\mathbf{M}$ will grow exponentially. The combination of synaptic weights that grows most rapidly will ultimately dominate so we can approximate the final weights by considering only this mode of equation (10.2) (Hertz *et al.*, 1990; MacKay & Miller, 1990). This means that ultimately $\mathbf{M}$ will be proportional to the eigenvector of the matrix $< g_a g_b >$ with the maximum eigenvalue. This result has several interesting interpretations. First, it means that the weights automatically select the combination of inputs with the maximum input correlation (or input variance if the mean of the input signal is zero). The resulting firing rate is proportional to the maximal principal component of the input data set (Oja, 1982; Hertz *et al.*, 1990). If certain conditions apply, it can be shown that this assures that the neuronal response carries the maximum amount of information about the input signal (Linsker, 1988 & 1993). It is remarkable that such a simple developmental rule leads to an optimal neuronal response.

Figure 6 shows the development of a neuronal response tuning curve for a single neuron on the basis of equation (10.2). The inputs $\mathbf{g}$ in this case, are a set of Gaussian firing rate curves like those shown in figure 2 and given by equation (4.1). Initially all the input weights $\mathbf{M}$ were set equal to each other and the resulting firing rate as a function of x, given by equation (10.1), is flat except at the edges (figure 6a). During development the stimulus consisted of the single value $X(t) = 0$ so that

$$< g_a g_b >= r_{max}^2 \exp\left(-(y_a^2 + y_b^2)/2\sigma_f^2\right) \ . \tag{10.3}$$

This leads to the development of a Gaussian firing response curve centered about the point $x = 0$ (figures 6b and 6c). In this case, the final tuning curve has the property that the neuron responds optimally to the value of the stimulus seen during development.

11. Effects of Recurrent Synapses on Firing Rates

In the last section, we saw that neuronal responses tuned to a specific stimulus value could be established using only the input synaptic weights and ignoring recurrent synapses between network neurons. In section 9, we noted the effect of training-induced changes in the strengths of recurrent synapses, but did not consider the impact of the recurrents themselves. What then is the role of recurrent network synapses in tuning and refining the neuronal and network responses to inputs? Suppose that the input synaptic weights have developed to produce tuning curves given by equation (10.1). When we add recurrent synapses, the firing rates of network neurons depend on each other as well as on the stimulus value,

$$r_i = f_i(x) + \sum_{j=1}^{N} W_{ij} r_j \ . \tag{11.1}$$

To determine the neuronal responses in the presence of recurrent synapses, we solve this equation for the firing rates,

$$r_i = \sum_{j=1}^{N} (\delta_{ij} - W_{ij})^{-1} f_j(x) \ . \tag{11.2}$$

We assume that $\mathbf{W}$ can be diagonalized and define a complete set of eigenvectors ξ_i^n and eigenvalues λ_n labeled by the index n and satisfying

$$\sum_{j=1}^{N} W_{ij} \xi_j^n = \lambda_n \xi_i^n \ . \tag{11.3}$$

In terms of these, we can write

$$f_j(x) = \sum_n A_n(x) \xi_j^n \tag{11.4}$$

and the solution for the firing rates becomes

$$r_i = \sum_n \frac{A_n(x) \xi_i^n}{1 - \lambda_n} \ . \tag{11.5}$$

Equation (11.5) displays the interesting phenomenon of resonance. Suppose that one or more of the eigenvalues is near, but slightly less than, one. Any stimulus that causes the value of $A_n(x)$ corresponding to one of these eigenvalues to be appreciably greater than zero will elicit a very strong firing response. The firing response is amplified by the factor $1 - \lambda_n$ in the denominator of (11.5) for eigenvalues near one. This strong response will be proportional to ξ^n. The network is exhibiting pattern (or in this simple case, value) recognition. Special values of the stimulus, those that excite modes with eigenvalues near one, elicit extremely strong responses from the system. In addition, there will be a strong bias of the coded value of a stimulus toward values corresponding to eigenvectors of $\mathbf{W}$ with eigenvalues near one.

If any of the eigenvalues is larger than one, the resonance phenomenon is even more dramatic. In this case, if we do not have inhibition or some mechanism to limit the firing rates, a dynamic model (section 13) shows that the network becomes unstable and the firing rates diverge. However, if these limiting mechanisms are incorporated into the model a strong resonance phenomenon is observed. This is shown in section 13 where the implications of resonant responses are also discussed.

12. Network Models - Nonlinear Effects

Many of the results discussed in previous sections have been based on an assumption about how the firing rate of an individual neuron depends on the firing rates of other neurons that synapse onto it. We have assumed in equation (7.1) that this dependence is linear, but this assumption should be compared with experimental results and corrected if it is not accurate. Unfortunately, this is not something that is easily done. In neurophysiology experiments, the firing rate is usually measured as a function of the amount of current injected through a microelectrode into the soma of a recorded neuron (Gustafsson & Wigstrom, 1981; Lanthorn *et al.*, 1984). However, in a natural setting neurons receive current from synapses not through electrodes. There are three parts to the computation that determines how neuronal firing rates depend on each other. First, we need to know how the firing rate of a neuron depends on the amount of current injected into its soma. Second, we must determine how presynaptic firing changes the conductance at the postsynaptic terminal leading to a synaptic current in the postsynaptic neuron. Finally, we must relate the amount of current entering the soma to these synaptically induced conductance changes.

There is evidence that in some circuits and situations neurons may operate in a fairly linear range (Jagadeesh *et al.*, 1993). However, there are many potential sources of nonlinearities along the pathway from presynaptic to postsynaptic firing. The firing rates of neurons do not have a strictly linear dependence on the amount of current injected into their somata. Below a threshold current, no firing occurs. At the threshold, the firing rate may suddenly jump to a nonzero value or it may rise continuously from zero (Rinzel & Ermentrout, 1989). Bifurcation theory indicates that in the latter case for a wide class of models, the firing rate should depend on the square root of the difference between the injected current and the threshold current near the transition to firing (Rinzel & Ermentrout, 1989). At the high current end, firing rates do not rise indefinitely but saturate because of the effects of the refractory period. Under constant current injection, the firing rate may not remain steady but may fall due to spike-rate adaptation. Even the simplest of model neurons, the integrate-and-fire model, has a firing rate that is a nonlinear (logarithmic) function of injected current (see for example Tuckwell, 1988).

Presynaptic firing results in the release of transmitter from the presynaptic neuron and this induces a change in the membrane conductance of the postsynaptic neuron at the site of the synapse. The amount of current per unit area that enters a neuron due to a synaptic conductance can be written as

$$I_{syn} = \overline{g}_{syn} m \left(E_{syn} - V\right) \tag{12.1}$$

where E_{syn} is the synaptic reversal potential. Because this current depends on the presynaptic voltage V, it produces potential nonlinear effects that are discussed below. The postsynaptic conductance per unit area has been written in the form $\overline{g}_{syn} m$ where $\overline{g}_{syn}$ is a constant and

m is a dynamic variable in the range $0 \leq m \leq 1$. The variable m is the probability that a synaptic receptor channel is in an open, conducting state.

The first step in analyzing the dependence of the synaptic conductance on the presynaptic firing rate is to determine how the amount of transmitter released from the presynaptic terminal depends on this rate. For simplicity, we might take this dependence to be linear, but it need not be. Suppose that the transmitter interacts with the receptor through a simple binding reaction

$$\text{closed receptor} + n \text{ transmitter molecules} \rightleftharpoons \text{open receptor} \tag{12.2}$$

with forward rate constant k_1 and the backward rate constant k_{-1}. In terms of the open probability for the receptor m, these reactions imply that

$$\frac{dm}{dt} = k_1 T^n (1 - m) - k_{-1} m \tag{12.3}$$

with T the transmitter concentration. The steady-state open probability given by this equation is

$$m = m_\infty = \frac{k_1 T^n}{k_{-1} + k_1 T^n} \ . \tag{12.4}$$

This is only linear in T if $n = 1$ and T is small.

Synaptic conductances induce currents over the surface of large, elaborate dendritic trees. The total current I_{soma} feeding into the soma from the trunk of the dendritic tree, shown in figure 7a, is the analog of the current injected through an electrode in an experimental setting. Currents entering through dendritic synapses must travel down dendritic cables before reaching the soma. They are thus subject to losses from the longitudinal resistance of the dendritic cable and due to leakage through the dendritic membrane. Since synaptic conductances along the dendritic cable can affect the amount of membrane leakage, there is a potential nonlinear impact on the amount of current entering the soma (Rall, 1964; Koch *et al.*, 1983; Segev & Parnas, 1989; Abbott, 1991b; Bernander *et al.*, 1991; Rapp *et al.*, 1992).

I will assume that the neuron being modeled has a passive dendritic tree. This is not a particularly valid assumption but it provides at least a lower limit on the nonlinear effects to be expected. Active dendritic trees can introduce additional sources of nonlinearity. In the passive case, the current I_{soma} entering the soma from the dendritic tree can be computed using methods of cable theory (Jack *et al.*, 1974; Rall, 1977; Tuckwell, 1988; Butz & Cowan, 1974; Koch & Poggio, 1985; Abbott, 1992). In the examples given here, I will compute the current which flows out of the trunk of a dendritic tree into the soma when synapses along the dendrite introduce a conductance change per unit area given by $\overline{g}_{syn} m$ and a synaptic current density I_{syn}. This computation is greatly simplified if we make use of a very convenient simplification due to Rall (1962 & 1977). The idea is to replace a complete dendritic tree with a single cylindrical cable having the same total electrotonic length as the original tree and the same total surface area. This simplification is shown in figures 7b and 7c.

I will give results for the total dendritic current entering the soma, I_{soma}, for an equivalent cable of radius a and length L with two different synaptic placements. First, suppose that the synapses are uniformly spread along the entire dendritic cable as in figure 7b. In this case (Abbott, 1991b),

$$I_{soma} = 2\pi a \lambda(m) I_{syn} \tanh(L/\lambda(m)) \tag{12.5}$$

with

$$\lambda(m) = \sqrt{\frac{aR_M}{2R_L(1+\overline{g}_{syn} m R_M)}} \tag{12.6}$$

where $1/R_M$ is the conductance per unit area of the membrane and R_L is the longitudinal resistivity of the intracellular fluid. For a short cable and small synaptic conductance, this formula is in fact linear in the synaptic current,

$$I_{soma} \approx 2\pi a L I_{syn} \ . \tag{12.7}$$

This is just the area of the cable times the synaptic current density. However, if the cable length is longer and the synaptic conductance is larger, the somatic current given by (12.5) grows only like the square root of the synaptic conductance.

The nonlinear dependence of I_{soma} on the firing rates is more dramatic if the synaptic input is located at the end of the equivalent cable, as in figure 7c. In this case (Abbott, 1991b),

$$I_{soma} = \frac{\pi a^2 I_{syn}}{\cosh(L/\lambda_0) + \overline{g}_{syn} m R_L \lambda_0 \sinh(L/\lambda_0)} \tag{12.8}$$

with

$$\lambda_0 = \sqrt{\frac{aR_M}{2R_L}} \ . \tag{12.9}$$

For a short cable, this becomes

$$I_{soma} \approx \frac{\pi a^2 I_{syn}}{1 + \overline{g}_{syn} m R_L L} \ . \tag{12.10}$$

Both (12.8) and (12.10) are linear in I_{syn} for small synaptic conductances but approach finite limiting values for large synaptic conductances.

In summary there are many possible sources of nonlinearity in the relation between pre- and postsynaptic firing rates. Transmitter release can be a nonlinear function of presynaptic firing rate, synaptic conductance a nonlinear function of transmitter concentration, somatic current a nonlinear function of synaptic conductance and postsynaptic firing rate a nonlinear function of somatic current. Perhaps, with luck, all of these combine to produce a fairly linear dependence and, of course, over a small enough range most nonlinear functions can be approximated by linear expressions. Nevertheless, there are situations in which nonlinearities may be important, as seen in the next section.

13. Network Dynamics

Perhaps the single feature of biological neural systems most responsible for making them difficult to model is the tremendous range of time scales for their dynamics (Churchland & Sejnowski, 1992). The fastest currents in a typical neuron, those responsible for initiating the action potential, turn on in less than a millisecond. The slowest currents and some synaptic currents have time constants of hundreds of milliseconds or more. Cultured neurons have shown activity-dependent changes over periods of about one hour (Turrigiano *et al.*, 1994). Thus, the dynamic range that must be modeled extends over three to seven orders of magnitude. The most drastic of the many approximations made to produce idealized neural network models is the collapsing of this huge range to a single time scale. This approximation has been made rigorous in cases where synaptic couplings are much slower than any other time constants affecting neuronal dynamics (Ermentrout, 1994). I will take another approach here and assume that the time scale used in the kernel of equation (3.1) to define the spike rate variable $r(t)$ is significantly longer than any time scale affecting neuronal or network dynamics.

If the time scale of the kernel K is slower than all other intrinsic and synaptic time scales, the problems associated with the complexity of neuronal dynamics disappear. We will consider at first a single neuron responding to a stimulus. Suppose that the stimulus is described by a slowly varying function $x(t)$ and the average, quasi-steady-state firing rate response to this stimulus is $f(x(t))$. If the neuronal dynamics is faster than the time scale for the kernel, the actual firing rate will settle down to something close to its steady-state value in a time shorter than this scale. In addition, for a slow kernel, many spike will be integrated before the kernel changes appreciably. As a result, we can replace the actual spike response function $s(t)$ in equation (3.1) by the average firing rate at a particular value of $x(t)$ and write

$$r(t) = \int_0^\infty dt'\; K(t') f(x(t-t')) \; . \tag{13.1}$$

This relates the dynamics of the firing rate variable $r(t)$ directly to the stimulus dynamics $x(t)$ through the steady-state response function. By using a slowly varying kernel to define the firing rate, we can avoid the complications of neuronal and network dynamics and build a dynamic model purely on the basis of steady-state properties.

If we take the exponential kernel of equation (3.11), r(t) can also be express in the form of a differential equation

$$\tau \frac{dr(t)}{dt} = -r(t) + f(x(t)) \; . \tag{13.2}$$

Equation (13.2) has been discussed and derived many times (Wilson & Cowan, 1972 & 1973; Hopfield, 1984; Frolov & Medvedev, 1986; Abbott, 1991a; Amit & Tsodyks, 1991a & 1991b; Ermentrout, 1994). It should be stress that, according to the viewpoint presented here, the time constant τ appearing in equation (13.2) is not directly related to any time constant characterizing the neuron, the membrane time constant in particular. Rather it is a parameter related to properties of the neuronal spike train and its characterization in terms of an integrated variable.

Differential equations can be written corresponding to the kernels (3.12) and (3.13) as well. For example, if we use the kernel (3.12) instead of the exponential kernel, $r(t)$ is given by a pair of differential equations,

$$\tau \frac{dr(t)}{dt} = -r(t) + a(t) \tag{13.3}$$

and

$$\tau \frac{da(t)}{dt} = -a(t) + f(x(t)) \; . \tag{13.4}$$

For the double exponential kernel, the firing rate is described by

$$r(t) = \frac{a(t) - b(t)}{\tau_1 - \tau_2} \tag{13.5}$$

where

$$\tau_1 \frac{da(t)}{dt} = -a(t) + \tau_1 f(x(t)) \tag{13.6}$$

and

$$\tau_2 \frac{db(t)}{dt} = -b(t) + \tau_2 f(x(t)) \ . \tag{13.7}$$

The effects of recurrent synapses between network neurons can also be included in this type of firing-rate model. For example, if we use equation (11.1) to describe how network firing rates depend on each other and we use the exponential kernel, firing rates in the network will be obey the equations

$$\tau \frac{dr_i(t)}{dt} = -r_i(t) + f_i(x) + \sum_{j=1}^{N} W_{ij} r_j(t) \ . \tag{13.8}$$

Nonlinear effects like those discussed in the last section can be incorporated by using more complex expressions for the last term in this equation (Abbott, 1991a & 1991b).

A dynamic model of the neuronal response allows us to study, among other things, the resonance phenomenon discussed in section 11. There it was pointed out that if there are eigenvalues of the synaptic weight matrix greater than one, the linear network is unstable to uncontrolled growth of the firing rates. This can be seen by performing a linear stability analysis of equation (13.8). Various ways of controlling this explosive, unstable rise in firing have been proposed. The basic mechanism is inhibition and involves either a nonlinear firing rate for inhibitory neurons (Buhmann, 1989), nonlinear effects of fast inhibition (Abbott, 1991a) or the relative timing of excitatory and inhibitory responses (Douglas *et al.*, 1989; Douglas & Martin, 1991). In the model shown in figures 8 and 9, the unbounded growth of firing rates in the linear networks during a resonance is controlled by nonlinear effects of inhibition computed as in the last section (Abbott, 1991a & 1991b).

The typical response of a population of cortical or hippocampal neurons to stimulation is induced firing followed by two kinds of inhibition (Avoli, 1986; Connors *et al.*, 1988). The two inhibitory responses are a fast component that can terminate the firing followed by a slower, long-lasting hyperpolarization. These are the result of the neurotransmitter GABA activating two types of receptors labeled A and B respectively. The fast inhibition is essential for the normal operation of excitatory neurons, if it is blocked these tend to fire in epileptic bursts (Dichter & Ayala, 1987; Miles & Wong, 1987). The existence of two forms of inhibition and the time scales associated with them suggest that, if resonances actually occur, they might be controlled by GABA_A inhibition and terminated by GABA_B inhibition.

Figure 8a shows the non-resonant response of network neurons in a model that includes excitatory synapses and both slow and fast inhibition (see Abbott, 1991a for details of the model). The stimulation induces firing of the principal neurons of the network and this in turn causes a rise in both the fast and slow inhibition. In this case, the stimulus does not produce a strong response and, in particular, the response does not outlast the duration of the stimulus. Figure 8b shows a resonant response achieved by increasing the strength of excitatory synapses until they cause an instability in the network firing following stimulation. This instability can be seen in the rapid rise of the firing rate following stimulation. The response is stabilized by the rapid activation of fast GABA_A inhibition that stops the exponential rise of the firing rate. Firing is temporarily self-sustaining but ultimately the slower GABA_B inhibition terminates the firing. The result is a much stronger response that lasts significantly longer than the stimulus.

Figure 9 shows how the resonance effect can be used to achieve pattern recognition. In this figure, a model network was exposed to random patterns of neuronal stimulation lasting for 10 ms and occurring every 100 ms. These produced the modest patterns of firing shown. However, at 500 ms a special pattern of stimulation was introduced, one that led to the activation of strong excitatory coupling between network neurons and resulted in a resonance. This is seen in the greatly enhanced activity corresponding to this pattern of stimulation. Interpreted using the concept of decoding, the resonance phenomenon means that stimulus values close to values experienced during a training period will elicit a much stronger response from the network than values that have not been previously experienced. The resonance phenomenon is an amplification process caused by the recurrent collaterals as discussed in section 11 and, in the model shown, it is controlled by some of the nonlinear process reviewed in section 12.

14. Conclusions

The possibility of decoding spike trains from single neurons as well as firing from an entire ensemble of neurons has important implications for both experiments and theoretical studies. One the theoretical side, it allows us to interpret the results of our models in a way that has both behavioral and experimental significance. Instead of reporting the results of modeling studied in terms of abstract patterns of neural firing, we can provided the values of coded stimuli or motor outputs.

Our knowledge of synaptic plasticity is growing continually and these results, when translated into synaptic weight learning rules, provide us with the possibility of avoiding the dilemma caused by our ignorance of synaptic strengths in real neuronal circuits. In some cases, detailed knowledge of how a neural network is affected by training experience coupled with information about the initial state of the system before training (provided for example by neuronal tuning curves) may be sufficient to allow us to predict and analyze how training changes behavior.

Although most neural network models are extremely simplified, this is not an essential feature of such modeling. Nonlinear and dynamic effects can be included in these models if they are needed to account for a particular phenomenon being studied. An example of this is the nonlinear resonance effect that may be related to recognition of previously learned stimuli. However, it should be emphasized that simplicity is one of the great virtues of these network models since it allows phenomena to be studied in enough detail to provide new insights. People eager for more complexity can always study the biological system itself.

I am extremely grateful to my collaborators Emilio Salinas, Kenneth Blum, Kamal Sen and Carl van Vreeswijk who helped develop many of these ideas. Research supported by NSF grant DMS-9208206.

References

Abbott, L.F. (1990) Learning in Neural Network Memories. Network: Comp. Neural Sys. 1:105-122.

Abbott, L.F. (1991a) Firing-Rate Models for Neural Populations. In Benhar, O., Bosio, C., Del Giudice, P. and Tabet, E., eds., *Neural Networks: From Biology to High- Energy Physics.* ETS Editrice, Pisa. pp. 179-196.

Abbott, L.F. (1991b) Realistic Synaptic Inputs for Network Models. Network: Comp. Neural Sys. 2:245-258.

Abbott, L.F. (1992) Simple Diagrammatic Rules for Solving Dendritic Cable Problems. Physica A185:343-356.

Abbott, L.F. and Blum, K.I. (1996) Functional Significance of Long-Term Potentiation for Sequence Learning and Prediction. Cerebral Cortex 6:406-416.

Adams, P.R. and Gavin, M. (1986) Voltage-dependent currents of vertebrate neurons and their role in membrane excitability. Adv. Neurol. 44:37-170.

Altes, R.A. (1989) Ubiquity of hyperacuity. J. Acoust. Soc. Am. 85:943-952.

Amit, D.J. (1989) *Modelling Brain Function.* Cambridge University Press, NY.

Amit, D.J. and Tsodyks, M.V. (1991a) Quantitative study of attractor neural network retrieving at low spike rates I. and II. Network 2:259-294 and Effective neurons and attractor neural networks in cortical environment. Network 3:121-138.

Amit, D.J. and Tsodyks, M.V. (1991b) Effective neurons and attractor neural networks in cortical environment. Network 3:121-138.

Artola, A. and Singer, W. (1993) Long-term depression of excitatory synaptic transmission and its relationship to long-term potentiation. Trends Neurosci. 16:480-487.

Avoli, M. (1986) Inhibitory potentials in neurons of the deep layers of the *in vitro* neocortical slice. Brain Res. 370:165-170.

Bacon, J.P. and Murphey, R.K. (1984) Receptive fields of cricket (*Acheta domesticus*) interneurons are related to their dendritic structure. J. Physiol. (Lond.) 272:779-797.

Baldi, P. and Heiligenberg, W. (1988) How sensory maps could enhance resolution through ordered arrangements of broadly tuned receivers. Biol. Cybern. 59:313-318.

Baudry, M. and Davis, J.L., eds. (1991) *Long-Term Potentiation.* MIT Press, Cambridge MA.

Bienenstock, E.L., Cooper, L.N. and Munro, P.W. (1982) Theory for the development of neuron selectivity: orientation specificity and binocular interaction in visual cortex. J. Neurosci. 2:32-48.

Bernander, O., Douglas, R.J., Martin, K.A.C. and Koch, C. (1991) Synaptic background activity determines spatio-temporal integration in single pyramidal cells. Proc. Natl. Acad. Sci. USA 88:11569-11573.

Bialek, W. (1989) Theoretical physics meets experimental neurobiology. In Jen, E. ed., *Lectures in Complex Systems, SFI Studies in the Science of Complexity, Vol. 2.* Addison-Wesley, Redwood City CA. pp. 413-595.

Bialek, W., Rieke, F., de Ruyter van Steveninck, R.R. and Warland, D. (1991) Reading a neural code. Science 252:1854-1857.

Bliss, T.V.P. and Collingridge, G.L. (1993) A synaptic model of memory: long-term potentiation in the hippocampus. Nature 361:31-39.

Blum, K.I. and Abbott, L.F. (1996) A Model of Spatial Map Formation in the Hippocampus of the Rat. Neural Comp. 8:85-93.

Bortolotto, Z.A., Bahir, Z.I., Davies, C.H. and Collingridge, G.L. (1994) A molecular switch activated by metabotropic glutamte receptors regulates induction of long-term potentiation. Nature 368:740-743.

Buhmann, J. (1989) Oscillations and low firing rates in associative memory neural networks. Phys. Rev. A40:4145-4151.

Butz, E.G. and Cowan, J.D. (1974) Transient potentials in dendritic systems of arbitrary geometry. Biophys. J. 14:661-689.

Burnod, Y., Grandguillaume, P., Otto, I., Ferraina, S. and Johnson, P.B. and Camaniti R. (1992) Visuomotor transformation underlying arm movements toward visual targets: A neural network model of cerebral cortical operations. J. Neurosci. 12:1435-1452.

Byrne, J.H. and Berry, W.O. (1989) *Neural Models of Plasticity.* Academic Press, San Diego.

Camhi, J.M. and Levy, A. (1989) The code for stimulus direction in a cell assembly in the cockroach. J. Comp. Physiol. A165:83-97.

Caminiti, R., Johnson, P.B., Galli, C., Ferraina, S. and Burnod, Y. (1991) Making arm movements within different parts of space: The premotor and motor cortical representations of a coordinate system for reaching to visual targets. J. Neurosci. 11:1182-1197.

Chen, L.L., McNaughton, B.L., Barnes, C.A. and Ortiz, E.R. (1990) Head-directional and behavioral correlates of posterior cingulate and medial prestriate cortex neurons in freely-moving rats. Soc. Neurosci. Abst. 16:441.

Churchland, P.S. and Sejnowski, T.J. (1992) *The Computational Brain.* MIT Press, Cambridge MA.

Connors, B.W., Gutnick, J.M. and Prince, D.A. (1982) Electrophysiological properties of neocortical neurons *in vitro.* J. Neurophysiol. 48:1302-1320.

Connors, B.W., Malenka, R.C. and Silva, L.R. (1988) Two inhibitory postsynaptic potentials and GABA_A and GABA_B receptor-mediated responses in neocortex of rat and cat. J. Physiol. 406:443-468.

Debanne, D., Gahwiler, B.H. and Thompson, S.M. (1994) Asynchronous pre- and postsynaptic activity induces associative long-term depression in area CA1 of the rat hippocampus *in vitro.* Proc. Natl. Acad. Sci. USA 91:1148-1152.

Dichter, M.A. and Ayala, G.F. (1987) Cellular mechanisms of epilepsy: a status report. Science 237:157-164.

Douglas, R.J., Martin, K.A.C. and Whitteridge, D. (1989) A canonical microcircuit for neocortex. Neural Comp. 1:480-488.

Douglas, R.J. and Martin, K.A.C. (1991) A funtional microcircuit for cat visual cortex. J. Physiol. 440:735-769.

Eichenbaum, H. (1993) Thinking about brain cell assemblies. Science 261:993-994.

Ermentrout, G.B. (1994) Reduction of Conductance Based Models with Slow Synapses to Neural Nets. Neural Comp. 6:679-695.

Foldiak, P. (1991) *Models of Sensory Coding.* Cambridge Univ. PhD Thesis.

Foldiak, P. (1993) The 'ideal homunculus': statistical inference from neural population responses. In Eeckman, F.H. and Bower, J., eds., *Computation and Neural Systems.* Kluwer Academic Publishers, Norwell MA. pp. 55-60

Fortier, P.A., Kalaska, J.F. and Smith, A.M. (1989) Cerebellar neuronal activity related to whole-arm reaching movements in the monkey. J. Neurophysiol. 62:198-211.

Frolov, A.A. and Medvedev, A.V. 1986. Substantiation of the 'point approximation' for describing the total electrical activity of the brain with the use of a simulation model. Biophysics 31:332-335.

Georgopoulos, A.P., Kettner, R.E. and Schwartz, A. (1988) Primate motor cortex and free arm movements to visual targets in three-dimensional space. II. Coding of the direction of movement by a neuronal population. Neurosci. 8:2928-2937.

Georgopoulos, A.P., Schwartz, A. and Kettner, R.E. (1986) Neuronal population coding of movement direction. Science 233:1416-1419.

Georgopoulos, A.P., Taira, M. and Lukashin, A. (1993) Cognitive neurophysiology of the motor cortex. Science 260:47-52.

Gilbert, C.D. and Wiesel, T.N. (1990) The influence of contextual stimuli on the orientation selectivity of cells in primary visual cortex of the cat. Vision Res. 30:1689-1701.

Gluck, M.A. and Rumelhart, D.E. (1990) *Neuroscience and Connectionist Theory.* Lawrence Erlbaum, Hillsdale NJ.

Grossberg, S. (1988) Nonlinear neural networks: principles, mechanisms and architectures. Neural Networks 1:17-61.

Gozani, S.N. and Miller, J.P. (1994) Ensemble coding of information by eight primary sensory interneurons in the cricket cercal system I. and II. (submitted).

Gustafsson, B. and Wigstrom, H. (1981) Shape of frequency-current curves in CA1 pyramidal cells in the hippocampus. Brain Res. 223:417-421.

Gustafsson, B. and Wigstrom, H. (1988) Physiological mechanisms underlying long-term potentiation. Trends Neurosci. 11:156-162.

Gustafsson, B., Wigstrom, H., Abraham, W.C. and Huang, Y.-Y. (1987) Long-term potentiation in the hippocampus using depolarizing current pulses as the conditioning stimulus to single volley synaptic potentials. J. Neurosci. 7:774-780.

Harris-Warrick, R.M. and Marder, E. (1991) Modulation of neural networks for behavior. Annu. Rev. Neurosci. 14:39-57.

Hawkins, R.D., Kandel, E.R. and Siegelbaum, S.A. (1993) Learning to modulate transmitter release: Themes and variations in synaptic plasticity. Annu. Rev. Neurosci. 16:625-665.

Hebb, D.O. (1949) *The Organization of Behavior: A Neuropsychological Theory.* J. Wiley, NY.

Hertz, J., Krough, A. and Palmer, R.G. (1990) *Introduction to the Theory of Neural Computation.* Addison-Wesley, NY.

Hessler, N.A., Shirke, A.M. and Malinow, R. (1993) The probability of transmitter release at a mammalian central synapse. Nature 366:569-572.

Hille, B. (1992) *Ionic Channels of Excitable Membranes.* Sinauer Assoc., Sunderland, MA.

Hopfield, J.J. (1982) Neural networks and systems with emergent selective computational abilities. Proc. Natl. Acad. Sci. USA 79:2554-2558.

Hopfield, J.J. (1984) Neurons with graded response have collective computational properties like those of two-state neurons. Proc. Natl. Acad. Sci. USA 81:3088-3092.

Jack, J.J.B., Noble, D. and Tsien, R.W. (1974) *Electrical Current Flow in Excitable Cells.* Oxford Univ. Press, Oxford.

Jagadeesh, B., Wheat, H.S. and Ferster, D. (1993) Linearity of summation of synaptic potentials underlying direction selectivity in simple cells of the cat visual cortex. Science 262:1901-1904.

Jahr, C.E. and Stevens, C.F. (1990) A quantitative description of NMDA receptor channel kinetic behavior. J. Neurosci. 10:1830-1837.

Kalaska, J.F., Caminiti, R. and Georgopoulos, A.P. (1983) Cortical mechanisms related to the direction of two-dimensional arm movements: relations in parietal area 5 and comparison with motor cortex. Exp. Brain. Res. 51:247-260.

Knudsen, E. and Konishi, M. (1978) A neural map of auditory space in the owl. Science 200:795-797.

Knudsen, E., DuLac, S. and Esterly, S.D. (1987) Computational maps in the brain. Annu. Rev. Neurosci. 10:41-65.

Koch, C. and Poggio, T. (1985) A simple algorithm for solving the cable equation in dendritic trees of arbitrary geometry. J. Neurosci. Meth. 12:303-315.

Koch, C., Poggio, T. and Torre, V. (1983) Nonlinear interactions in a dendritic tree: localization, timing and role in information processing. Proc. Natl. Acad. Sci. USA 80:2799-2802.

Kohonen, T. (1984) *Self Organization and Associative Memory.* Springer Verlag, Berlin.

Kohonen, T. (1988) An introduction to neural computing. Neural Networks 1:3-16.

Konishi, M. (1987) Centrally synthesized maps of sensory space. Trends Neurosci. 9:163-168.

Konishi, M. (1991) Deciphering the brain's codes. Neural Comp. 3:1-18.

Lanthorn, T., Storm, J. and Andersen, P. (1984) Current-to-frequency transduction in CA1 hippocampal pyramidal cells: Slow prepotentials dominate the primary firing range. Exp. Brain Res. 53:431-443.

Lee, C., Rohrer, W.H. and Sparks, D.L. (1988) Population coding of saccadic eye movements by neurons in the superior colliculus. Nature 332:357-360.

Lehky, S.R. and Sejnowski, T.J. (1990) Neural model of stereoacuity and depth interpolation based on a distributed representation of stereo disparity. J. Neurosci. 10:2281-2299.

Levy, W.B., Colbert, C.M. and Desmond, N.L. (1990) Elemental adaptive processes of neurons and synapses: a statistical/computational perspective. In Gluck, M.A. and Rumelhart, D.E. eds., *Neuroscience and Connectionist Theory.* Lawrence Erlbaum, Hillsboro, N.Y. pp. 187-236.

Lisman, J.E. (1989) A mechanism for the Hebb and the anti-Hebb processes underlying learning and memory. Proc. Natl. Acad. Sci. USA 86:9574-9578.

Lisman, J.E. and Goldring, M.A. (1988) Feasibility of long-term storage of graded information by the Ca^{2+}/calmodulin-dependent protein kinase molecules of the postsynaptic density. Proc. Natl. Acad. Sci. USA 85:5320-5324.

Linkser, R. (1986) From basic network principles to neural architecture. Proc. Natl. Acad. Sci. USA 83:7508-7512, 8390-8394 & 8779-8783.

Linsker, R. (1988) Self-organization in a perceptual network. Computer 21:105-117.

Linsker, R. (1993) Local synaptic learning rules suffice to maximize mutual information in a linear network. Neural Comp. 4:691-702

Llinás, R.R. (1988) Intrinsic electrophysiological properties of mammalian neurons: insights into central nervous system function. Science 242:654 -1664.

Lukashin, A.V. (1990) A learned neural network that simulates properties of the neural population vector. Biol. Cybern. 63:377-382.

Lynch, G., Larson, J., Kelso, S., Barrionuevo, G. and Schottler, F. (1983) Intracellular injections of EGTA block induction of hippocampal long-term potentiation. Nature 305:719-721.

Madison, D.V., Malenka, R.C. and Nicoll, R.A. (1991) Mechanisms underlying long-term potentiation of synaptic transmission. Annu. Rev. Neurosci. 124:379-397.

Malenka, R.C. and Nicoll, R.A. (1993) MBDA-receptor-dependent synaptic plasticity: Multiple forms and mechanisms. Trends Neurosci. 16:521-527.

Malsburg, C.v.d. (1971) Self-organization of orientation sensitive cells in the striate cortex. Kybernetik 14:85-100.

Malsburg, C.v.d. (1979) Development of ocularity domains and growth behavior of axon terminals. Biol. Cybern. 32:49-62.

Marr, D. (1971) Simple memory: A theory for archicortex. Phil. Trans. Royal Soc. London B262:21-81.

Maunsell, J.H.R. and Newsome, W.T. (1987) Visual processing in monkey extrastriate cortex. Annu. Rev. Neurosci. 10:363.

McCormick, D.A. (1990) Membrane properties and neurotransmitter actions. In Shepherd, G.M. ed., *The Synaptic Organization of the Brain.* Oxford Univ. Press, NY. pp. 32-66.

McCulloch, W.S. and Pitts, W. (1943) A logical calculus of the ideas immanent in nervous activity. Bull. Math. Biophys. 5:115-133.

MacKay, D.J.C. and Miller, K.D. (1990) Analysis of Linsker's application of Hebian rules to linear networks. Network 1:257-299.

Middlebrooks, J.C., Clock, A.E., Xu, L. and Green, D.M. (1994) A panoramic code for sound location by coritcal neurons. Science 264:842-844.

Miles, R. and Wong, R.K.S. (1987) Inhibitory control of local excitatory circuits in the guinea-pig hippocampus. J. Physiol. 388:611-629.

Miller, J.P., Jacobs, G.A. and Theunissen, F. (1991) Representation of sensory information in the cricket cercal sensory system. I. Response properties of the primary interneurons. J. Neurophysiol. 66:1680-1689.

Miller, K.D. (1990) Correlation-based models of neural development. In Gluck, M.A. and Rumelhart, D.E., eds., *Neuroscience and Connectionist Theory.* Lawrence Erlbaum, Hillsdale NJ. pp. 267-354.

Miller, K.D. (1992) Models of activity-dependent neural development. Seminars in Neurosci. 4:61-73.

Miller, K.D. (1994) A model for the development of simple cell receptive fields and the ordered arrangement of orientation columns through activity-dependent competition between on- and off-center inputs. J. Neurosci. 14:409-441.

Miller, K.D. and MacKay, D.J.C. (1994) The role of constraints in Hebbian learning. Neural Comp. 6:100-126.

Morris, R.G.M., Anderson, E., Lynch, G.S. and Baudry, M. (1986) Selective impairment of learning and blockade of long-term potentiation by an N-methyl-D-aspartate receptor antagonist, AP5. Nature 319: 774-776.

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

O'Keefe, J. (1979) A review of the hippocampal place cells. Prog. Neurobiol. 13:419-439.

O'Keefe J. and Dostovsky J. (1971) The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely moving rat. Brain Res. 34:171-175.

O'Keefe J. and Nadel L. (1978) *The Hippocampus as a Cognitive Map.* Clarendon, Oxford.

O'Neill, W.E. and Suga, N. (1982) Encoding of target range information and its representation in the auditory cortex of the mustache bat. J. Neurosci. 2:17-31.

Optican, L.M. and Richmond, B.J. (1987) Temporal encoding of two-dimensional patterns by single units in primate inferior temporal cortex III. Information theoretic analysis. J. Neurophysiol. 57:163-178.

Orban, G.A. (1984) *Neuronal Operations in the Visual Cortex.* Springer, Berlin.

Paradiso, M.A (1988) A theory for the use of visual orientation information which exploits the columnar structure of striate cortex. Biol. Cybern. 58:35-49.

Rall, W. (1962) Theory of physiological properties of dendrites. Ann. N.Y. Acad. Sci. 96:1071-1092.

Rall, W. (1964) Theoretical significance of dendritic trees for neuronal input-output relations. In Reiss, R., ed., *Neural Theory and Modeling.* Stanford Univ. Press, Stanford CA. pp. 73-97.

Rall, W. (1977) Core conductor theory and cable properties of neurons. In Kandel, E.R., ed., *Handbook of Physiology Vol. 1.* Amer. Physiol. Soc., Bethesda. pp. 39-97.

Rapp, M., Yarom, Y. and Segev, I. (1992) The impact of parallel background activity on the cable properties of cerebellar Purkinje cells. Neural. Comp. 4:518-533

Reike, F.M. (1991) *Physical Principles Underlying Sensory Processing and Computation.* Univ. of California Berkeley PhD Thesis.

Rinzel, J. and Ermentrout, G.B., (1989) Analysis of neural excitability and oscillations In Koch, C. and Segev, I., eds., *Methods in Neuronal Modeling.* MIT Press, Cambridge MA.

Rolls, E.T. (1989) Parallel distributed processing in the brain: Implications of the functional architecture of neuronal networks in the hippocampus. In Morris, R.G.M., ed., *Parallel Distributed Processing: Implications for Pschology and Neuroscience.* Oxford Univ. Press, Oxford. pp. 286-308.

Rosenmund, C., Clements, J.D. and Westbrook, G.L. (1993) Nonuniform probability of glutamate release at a hippocampal synapse. Science 262:754-758.

Salinas, E. and Abbott, L.F. (1994) Vector reconstruction from firing rates. J. Comp. Neurosci. 1:89-107.

Salzman, C.D. and Newsome, W.T. (1994) Neural mechanisms for forming a perceptual decision. Science 264:231-237.

Schwartz, A., Kettner, R.E. and Georgopoulos, A.P. (1988) Primate motor cortex and free arm movements to visual targets in three-dimensional space. I. Relations between single cell discharge and direction of movement. Neurosci. 8:2913-2927.

Segev, I. and Parnas, I. (1989) Synaptic integration mechanisms: a theoretical and experimental investigation of temporal postsynaptic interaction between excitatory and inhibitory input. Biophys. J. 41:41-50.

Sejnowski, T.J. (1977) Storing covariance with nonlinearly interacting neurons. J. Math. Biol. 4:303-321.

Sejnowski, T.J. (1988) Neural populations revealed. Nature 332:308.

Seung, H.S. and Sompolinsky, H. (1993) Simple models for reading neuronal population codes. Proc. Natl. Acad. Sci. USA 90:10749-10753.

Shatz, J.C.(1990) Impulse activity and the patterning of connections during CNS development. Neuron 5:745-756.

Shor, R.H., Miller, A.D. and Tomko, D.L. (1984) Responses to head tilt in cat central vestibular neurons. I. Direction of maximum sensitivity. J. Neurophysiol. 51:136-146.

Smetters, D.K. and Nelson, S.B. (1993) Estimates of functional synaptic convergence in rat and cat visual cortical neurons. Neurosci. Abstr. 16:263.6.

Snippe, H.P. and Koenderink, J.J. (1992) Discrimination thresholds for channel-coded systems. Biol. Cybern. 66:543-551.

Steinmetz, M.A., Motter, B.C., Duffy, C.J. and Mountcastle, V.B. (1987) Functional properties of parietal visual neurons: Radial organization of directionalities with the visual field. J. Neurosci. 7:177-191.

Stevens, C.F. (1989) How cortical interconnectedness varies with network size. Neural Comp. 1:473-479.

Suzuki, I., Timerick, J.B. and Wilson, V.J. (1985) Body position with respect to the head or body position in space is coded in lumbar interneurons. J. Neurophysiol. 54:123-133.

Swindale, N.V. (1980) A model for the formation of ocular dominance stripes. Proc. Roy. Soc. London B208:243-264.

Taube, J.S., Muller, R.I. and Ranck, J.B.J. (1990) Head direction cells recorded from the postsubicullum in freely moving rats. I. Description and quantitative analysis. J. Neurosci. 10:420-435.

Theunissen, F. (1993) *An Investigation of Sensory Coding Principles Using Advanced Statistical Techniques.* Univ. of California Berkeley PhD Thesis.

Theunissen, F. and Miller, J.P. (1991) Representation of sensory information in the cricket cercal sensory system. II. Information theoretic calculation of system accuracy and optimal tuning-curve widths of four primary interneurons. J. Neurophysiol. 66:1690-1703.

Touretzky, D.S., Redish, A.D. and Wan, H.S. (1993) Neural representation of space using sinusoidal arrays. Neural Comp. 5:869-884.

Tuckwell, H.C. (1988) *Introduction to Theoretical Neurobiology.* Cambridge Univ. Press, Cambridge.

Turrigiano, G., Abbott, L.F. and Marder, E. (1994) Activity-dependent changes in the intrinsic properties of cultured neurons. Science 264:974-977.

Van Gisbergen, J.A.M., Van Opstal, A.J. and Tax, A.M.M. (1987) Collicular ensemble coding of saccades based on vector summation. Neuroscience 21:541-555.

Van Opstal, A.J. and Kappen, H. (1993) A two-dimensional ensemble coding model for spatial-temporal transformation of saccades in monkey superior colliculus. Network 4:19-38.

Vogels, R. (1990) Population coding of stimulus orientation by cortical cells. J. Neurosci 10:3543-3558.

Warland, D., Landolfa, M.A., Miller, J.P. and Bialek, W. (1991) Reading between the spikes in the cercal filiform hair receptors of the cricket. In Eeckman. F. and Bower. J., eds., *Analysis and Modeling of Neural Systems.* Kluwer Academic Publishers, Norwell, MA.

Wilson, H.R. and Cowan, J.D. (1972) Excitatory and inhibitory interactions in localized populations of model neurons. Biophys. J. 12:1-24.

Wilson, H.R. and Cowan, J.D. (1973) A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. Kybern. 13:55-80.

Wilson, M.A. and McNaughton, B. (1993) Dynamics of the hippocampal ensemble code for space. Science 261:1055-1058.

Young, M.P. and Yamane, S. (1992) Sparce population coding of faces in the inferotemporal cortex. Science 256:1327-1331.

Zhang, J. and Miller, J.P. (1991) A mathematical model for resolution enhancement in layered sensory systems. Biol. Cybern. 64:357-364.

Zohary, E. (1992) Population coding of visual stimuli by cortical neurons tuned to more than one dimension. Biol. Cybern. 66:265-272.

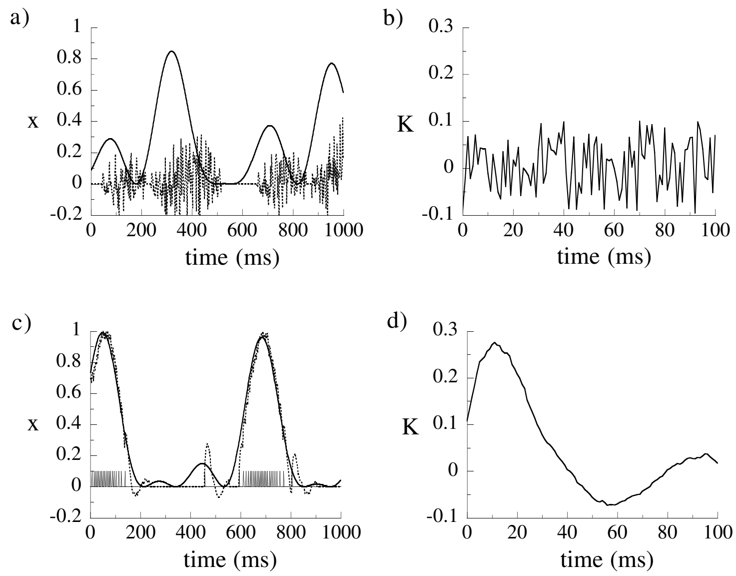

Fig. 1: Spike decoding of a stimulus by a linear filter. The stimulus, shown by the solid lines in a) and c), is quasiperiodic. Spike trains were generated by an integrate-and-fire model neuron with the stimulus acting as an input current. These were decoded by a linear filter to estimate the stimulus. The estimate is indicated by the dashed lines in figures a) and c). The spike train is also shown in c). In a) the random kernel plotted in b) is used and the estimate does not match the signal. The good match between the estimate and the stimulus value seen in c) is produced by the kernel shown in d) that resulted from repeated application of the modification rule discussed in section 3.

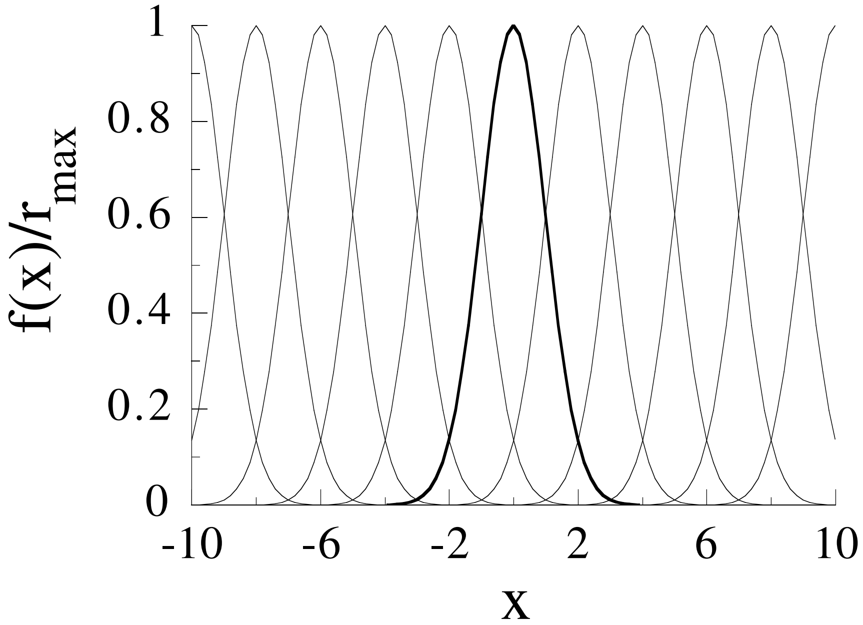

Fig. 2: An array of Gaussian tuning curves. Tuning curves are all identical in shape and are overlapping. The tuning curve centered at $x = 0$ is darkened for clarity.

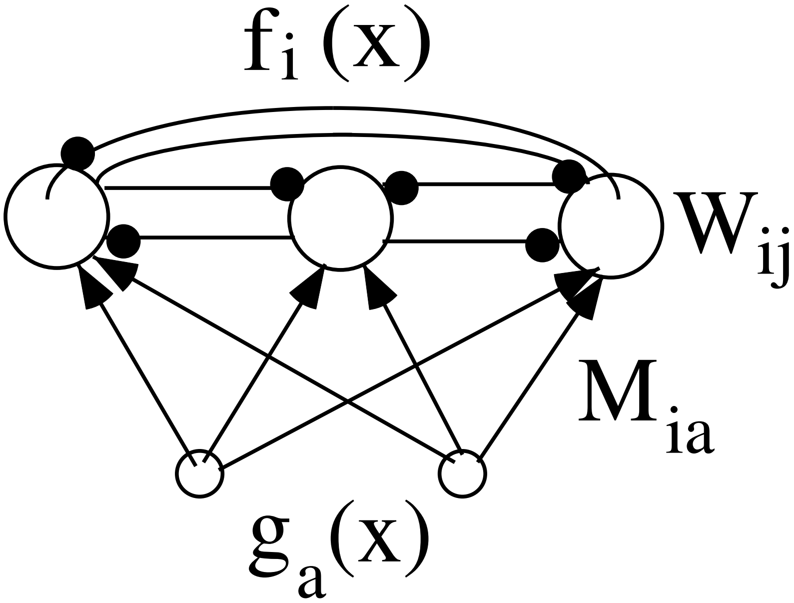

Fig. 3: The network model. Neurons, label by the index i and shown as stippled circles, respond to a stimulus value x by firing at a rate $f_i(x)$. Information about x is carried into the network from inputs shown as small striped circles through synapses depicted by arrow heads. Firing rates along the input lines, label by the index a, are given by $g_a(x)$. The strength of the synapse from input a to neuron i is M_{ai}. Synapses between network neurons are denoted by filled circles. The strength of the synapse from neuron j to neuron i is given by W_{ij}.

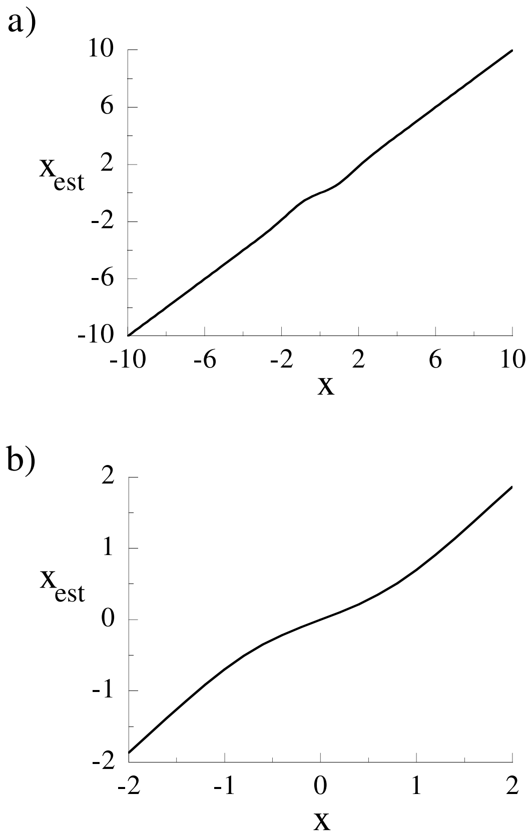

Fig. 4: Effect of training with a single stimulus value on the coded variable. The coded value x_{est} is plotted as a function of the true stimulus value x. The training period consisted of exposure to the stimulus $X = 0$. This results in the distortion of the coded value near $x = 0$ seen in a) and magnified in b).

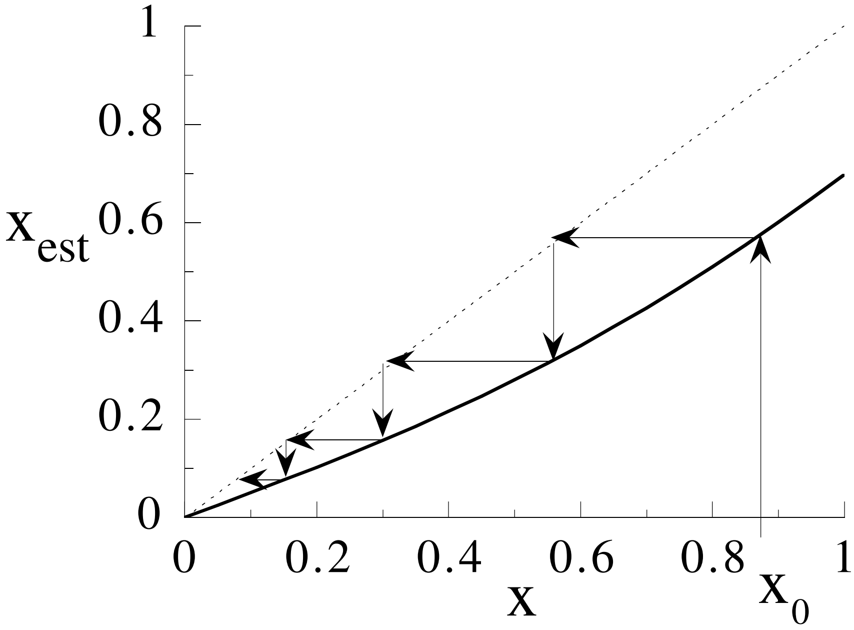

Fig. 5: Iteration of the trained network. It the coded value x_{est} is fed back so that it becomes a new stimulus value, this results in a steady progression from the initial stimulus value x_0 to a final coded value $x_{est} = 0$ which is equal to the training stimulus.

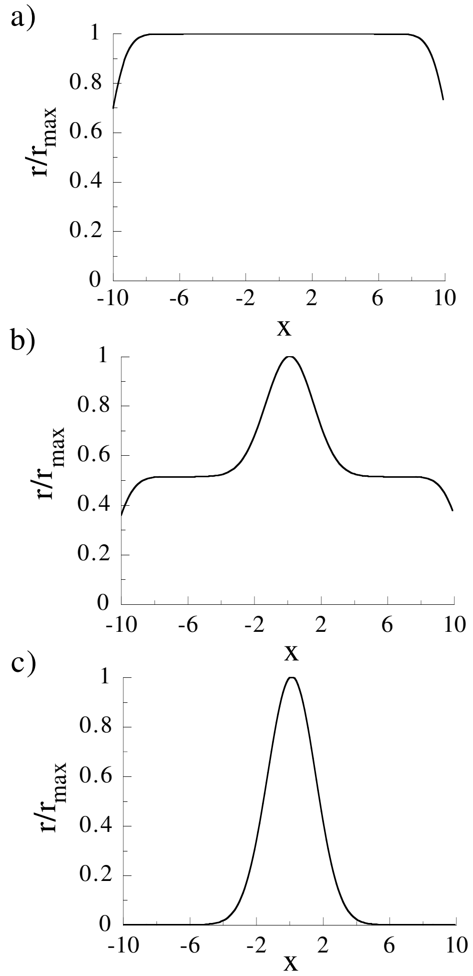

Fig. 6: Development of a Gaussian tuning curve. Correlation-based synaptic modification of input synapses during a developmental period when the stimulus value was zero results in the appearance of a Gaussian tuning curve centered at this point. a) The initial tuning curve is flat, except near the edges. b) The Gaussian curve begins to appear. c) The final Gaussian tuning curve. In all three plots, the firing rates are divided by the maximum rates for each case. In absolute terms, the tuning curve in c) represents a much larger firing rate that the curve shown in a).

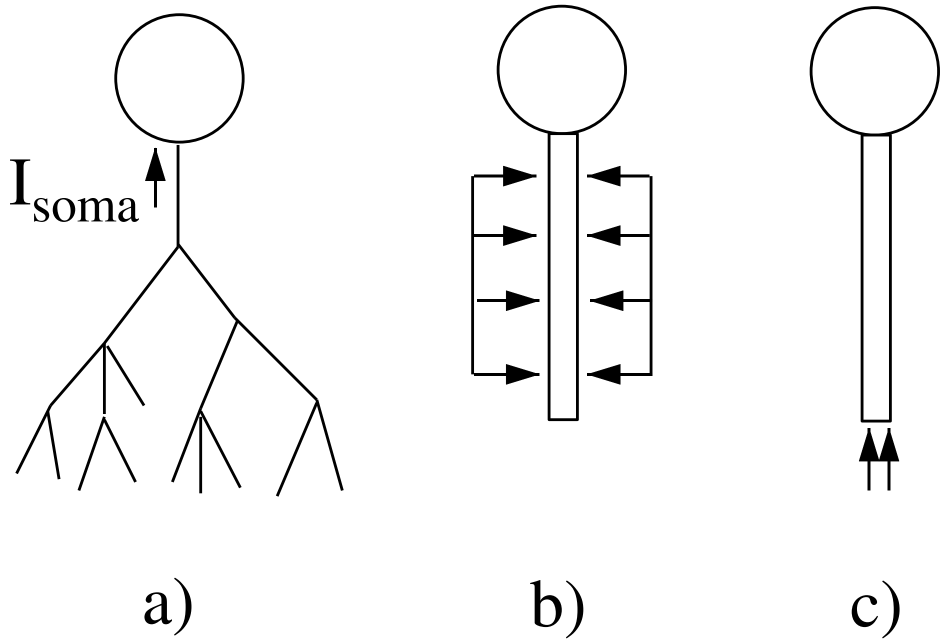

Fig. 7: The model neuron and synaptic placements considered in section 12. a) A soma and typical passive dendritic tree. The current I_{soma} being computed is the current flowing from the dendritic tree into the soma. b) and c) show the soma with a single passive dendritic cable equivalent to the full tree in a). For b) synapses are uniformly distributed over the equivalent cable while in c) synapses are located at the end of the cable.

Fig. 8: Resonance in a firing-rate model with two forms of inhibition. For network (excitatory) neurons, both spikes and a firing rate are plotted. The levels of fast and slow inhibition as a function of time are also indicated. A brief 10 ms stimulation current excites the system initially. a) Stimulation results in a transient, weak firing response with modest increases in inhibition. b) Stronger excitatory couplings between network neurons cause the firing rate to climb sharply but fast inhibition controls this rate resulting in sustained firing. Ultimately, slow inhibition increases to the point where it terminates the firing. This is a typical resonance response to a stimulus. (Not available on Web.)

Fig. 9: A network resonance and pattern recognition. Excitatory couplings between network neurons were chosen so that one particular pattern of activity resulted in particularly strong network excitation. Each trace represents the firing of one unit and tick marks indicate action potentials. Every 100 ms a brief current is injected into a random collection of neurons. This elicits a small number of spikes from the stimulated neurons. However, at time 500 ms the specific pattern leading to strong excitatory network coupling is stimulated. As a result, the system responds with a firing resonance. (Not available on Web.)