

## Vector Reconstruction from Firing Rates

EMILIO SALINAS

SALINAS@BINAH.CC.BRANDEIS.EDU.

L.F. ABBOTT

ABBOTT@BINAH.CC.BRANDEIS.EDU.

*Biology Department and Center for Complex Systems, Brandeis University, Waltham, MA 02254*

*Received December 29, 1993; Revised March 16, 1994; Accepted (in revised form) March 18, 1994*

*Action Editor: A. Georgopoulos*

**Abstract.** In a number of systems including wind detection in the cricket, visual motion perception and coding of arm movement direction in the monkey and place cell response to position in the rat hippocampus, firing rates in a population of tuned neurons are correlated with a vector quantity. We examine and compare several methods that allow the coded vector to be reconstructed from measured firing rates. In cases where the neuronal tuning curves resemble cosines, linear reconstruction methods work as well as more complex statistical methods requiring more detailed information about the responses of the coding neurons. We present a new linear method, the optimal linear estimator (OLE), that on average provides the best possible linear reconstruction. This method is compared with the more familiar vector method and shown to produce more accurate reconstructions using far fewer recorded neurons.

### Introduction

To determine how information is represented by nervous systems, we need to understand how neural activity is correlated with various external cues and actions such as sensory input or motor output. This is a crucial first step before more complex issues like information processing and storage can be addressed adequately. Neural encoding of information can be studied by measuring and cataloging neuronal responses as a function of external correlates. Our understanding of the neural code can be tested by solving the inverse problem, inferring sensory input or motor output from a given set of neuronal activities. To decode information we must extract an estimate of the external correlate from measured neuronal spike trains and determine its accuracy.

Decoding a neural population response requires procedures for extracting information from individual spike trains and for combining these results into a population ensemble estimate. Optimal methods for reconstructing information from the spike train of a single neuron have been developed (Bialek, 1989; Bialek et al.,

1991; Reike, 1992) and applied to movement-sensitive neurons in the blowfly (Bialek et al., 1991), as well as to other systems (Bialek, 1989; Warland et al., 1991). In this work, a linear filter is used to extract the maximum possible information from each spike. In the cases we consider, information is encoded by a population of neurons with responses that are 'tuned' to various features of the coded information (Knudsen, DuLac and Esterly, 1987; Konishi, 1987) and the decoding will be based on multiple action potentials from each recorded neuron. In situations such as this, the bulk of the information is typically carried by the firing rates, although additional information may be encoded in other characteristics of the firing patterns (Optican and Richmond, 1987; Gozani and Miller, 1993). We will assume that the firing rates alone can be used and focus on how to combine a number of single neuron firing rates into a population-based estimate of the coded quantity.

Pioneering work on population decoding was done by Georgopoulos and collaborators who developed a population vector relating activity in the motor cortex of the monkey to the direc-

tion of arm movements (Georgopoulos et al., 1983; Georgopoulos, Schwartz and Kettner, 1986; Georgopoulos, Kettner and Schwartz, 1988; Georgopoulos, Taira and Lukashin, 1993). The approach developed for this system, which we will call the vector method, was used in similar studies of premotor cortex (Caminiti et al., 1991), parietal area 5 (Kalaska, Caminiti and Georgopoulos, 1983) and cerebellum (Fortier, Kalaska and Smith, 1989). In addition, the vector method has been used in studies of primary visual cortex (Gilbert and Wiesel, 1990) and applied to parietal visual neurons coding for the direction of motion of an object in the visual field (Steinmetz et al., 1987), neurons in the superior colliculus encoding saccade direction (Van Gisbergen, Van Opstal and Tax, 1987; Lee, Rohrer and Sparks, 1988) and inferotemporal neurons in the monkey responding to human faces (Young and Yamane, 1992). A different approach was used recently by Wilson and McNaughton (1993) to determine **the position of a rat in an enclosure using recorded activities of place cells in the CA1 region of the hippocampus.**

There have been a number of theoretical analyses of population decoding in a variety of contexts (Paradiso, 1988; Vogels, 1990; Foldiak, 1991 and 1993; Zohary, 1992; Snippe and Koenderink, 1992; Seung and Sompolinsky, 1993; Touretzky, Redish and Wan, 1993), **addressing the issue of acuity** (Baldi and Heiligenberg, 1988; Altes, 1989; Lehky and Sejnowski, 1990; Zhang and Miller, 1991) and in network models (Lukashin, 1990; Burnod et al., 1992; Van Opstal and Kappen, 1993). For further discussions see (Sejnowski, 1988; Churchland and Sejnowski, 1992; Eichenbaum, 1993). The theoretical papers often use methods that are optimal in some statistical sense usually based on the probability distributions of the neuronal firing rates. Experimental papers typically employ simpler methods that require less detailed knowledge of the system, such as the vector method. As a result, a gap has developed between methods that have been analyzed theoretically and methods that are actually employed in population reconstructions. One purpose of this paper is to fill this gap by comparing more powerful statistical methods

with simpler and more practical methods on a variety of realistic examples.

In contrast to the statistical methods, **the estimate produced by the vector method is a linear function of the neuronal responses.** In addition to being simple to construct and evaluate, a linear estimator is much more likely to be related to computations actually carried out by the nervous system than nonlinear statistical calculations which seem far removed from anything that neural circuits could actually do. As we will show, linear decoding schemes are ideally suited for situations where **coding is accomplished by an array of neurons with tuning curves (average firing responses) that resemble cosine functions** (Touretzky, Redish and Wan, 1993). Cosine neuronal response curves are surprisingly widespread. They occur, for example, in cricket cercal interneurons coding wind direction (Bacon and Murphey, 1984; Miller, Jacobs and Theunissen, 1991; Theunissen and Miller, 1991). The systems where the vector method has been applied to monkey arm movements, in motor cortex (Schwartz, Kettner and Georgopoulos, 1988), premotor cortex (Caminiti et al., 1991), parietal area 5 (Kalaska, Caminiti and Georgopoulos, 1983) and cerebellum (Fortier, Kalaska and Smith, 1989), all show cosine tuning curves. **Cosine tuning curves are also found in neurons coding body** (Suzuki, Timerick and Wilson, 1985) **and head position in cats** (Shor, Miller and Tomko, 1984) **and head direction in rats** (Chen et al., 1990; Taube, Muller and Ranck, 1990). Parietal visual neurons exhibit cosine responses to movement direction (Steinmetz et al., 1987). Often these response functions do not trace out a full cosine curve, even when shifted by a background firing rate, because firing rates cannot be negative, but we will see that even in these cases a linear decoding scheme works very well.

In view of the widespread use of a particular linear decoding method, the vector method, and of the appearance of cosine tuning curves in so many different neural systems, it is useful to consider other linear reconstruction techniques. We present here a new linear decoding method that, by construction, produces the smallest average error of any linear method. In many cases, this

performs considerably better than the standard vector method. By means of computer simulations, we will apply this method, which we call the optimal linear estimator or OLE to a variety of systems. We will also present a recent application of this new method to the reconstruction of arm movements from recordings of cortical motor neurons in the monkey, done and kindly provided by G. Pellizzer and A. Georgopoulos, which shows that the OLE produces more accurate predictions of arm movement direction using far fewer recorded neurons than the vector method.

In most of the systems we will discuss, the quantity being encoded is the direction of a vector in space. The experiments done on monkey arm movements (Georgopoulos et al., 1983; Georgopoulos, Schwartz and Kettner, 1986; Georgopoulos, Kettner and Schwartz, 1988; Georgopoulos, Taira and Lukashin, 1993; Caminiti et al., 1991; Kalaska, Caminiti and Georgopoulos, 1983; Fortier, Kalaska and Smith, 1989) on saccades (Van Gisbergen, Van Opstal and Tax, 1987; Lee, Rohrer and Sparks, 1988) and on primary (Gilbert and Wiesel, 1990) and parietal visual neurons (Steinmetz et al., 1987) all involved the encoding of the direction of either two-dimensional or three-dimensional vectors. The cercal system of the cricket (Bacon and Murphey, 1984; Miller, Jacobs and Theunissen, 1991; Theunissen and Miller, 1991), and a similar system in the cockroach (Camhi and Tom, 1978; Camhi and Levy, 1989), code the direction of a two-dimensional vector, the horizontal direction of the wind. However, in some systems the magnitude of the encoded vector is as important as its direction. For example, hippocampal place cells (O'Keefe and Nadel, 1978; Wilson and McNaughton, 1993) code for both the direction and magnitude of a two-dimensional position vector. The vector method is not well suited to situations where the magnitude of a vector as well as its direction must be decoded, but the OLE and other methods we discuss can be used in this case as well.

We will restrict our attention to situations where **the quantity being encoded is a spatial vector**. In some cases, it may be possible to rep-

resent more general and complex information as a vector in some abstract space of attributes. For example, by using multi-dimensional scaling Young and Yamane (1992) were able to relate the responses of face cells to the encoding of a two-dimensional vector. The methods we discuss can be applied to the reconstruction of any encoded quantity that can be characterized as some type of vector.

We will begin by analyzing the decoding problem in the cricket cercal system which has been studied in great detail (Bacon and Murphey, 1984; Miller, Jacobs and Theunissen, 1991). The elegant work on this system by Theunissen and Miller (1991) (see also Theunissen, 1993) provides an excellent starting point for our analysis of population decoding. We then consider a more general situation where **the direction of a vector is represented by an arbitrary number of neurons with partial or complete cosine tuning curves**. This will lead to the development of the new optimal linear method (OLE) and its application to the monkey motor cortex and arm movements. Finally, we analyze an array of neurons with non-cosine tuning curves coding the magnitude of a scalar (or one-dimensional vector) quantity. This situation is similar to one encountered, for example, in the neurons coding echo delay in the bat (O'Neill and Suga, 1982; Suga and Horikawa, 1986) and is the one-dimensional analog of position decoding from rat hippocampal neurons (Wilson and McNaughton, 1993).

## Methods

### Notation

We use the following notation:

$\vec{V}$  = **the vector being encoded.**

$\vec{V}_{\text{est}}$  = the estimate of the encoded vector.

$N$  = the number of recorded neurons upon which the estimate is based.

$i$  = an index to denote a particular neuron,  $i = 1, 2, \dots, N$ .

$R_i$  = the measured firing rate of neuron  $i$ .

$r_i$  = firing response of neuron  $i$  which is related

to  $R_i$  (see below).

$\mathbf{r}$  = the collection of all firing responses  $r_1, r_2, r_3, \dots, r_N$  for the  $N$  neurons.

$\vec{C}_i$  = the preferred direction vector for neuron  $i$  which is the value of  $\vec{V}$  that produces the maximum firing response in this neuron.

$\vec{V} \cdot \vec{C}_i$  = the dot product of  $\vec{V}$  with  $\vec{C}_i$  which is equal to the length of  $\vec{V}$  times the length of  $\vec{C}_i$  times the cosine of the angle between these two vectors. If  $\vec{V}$  and  $\vec{C}_i$  are unit vectors (vectors with length equal to one) then the dot product is just the cosine of the angle between them.

$P(\mathbf{r} | \vec{V})$  = the probability of obtaining the firing responses  $\mathbf{r}$  given that the encoded vector takes the value  $\vec{V}$ .

$f_i(\vec{V})$  = average firing response or tuning curve of cell  $i$  when the encoded vector takes the value  $\vec{V}$ . Equivalently,

$$f_i(\vec{V}) = \int d\mathbf{r} r_i P(\mathbf{r} | \vec{V}). \quad (2.1)$$

$L_i$  = the 'center of mass' vector for the tuning curve function  $f_i$  defined by

$$\vec{L}_i = \int d\vec{V} \vec{V} f_i(\vec{V}). \quad (2.2)$$

$Q_{ij}$  = the correlation of the firing rates of neurons  $i$  and  $j$  defined by

$$Q_{ij} = \int d\mathbf{r} d\vec{V} r_i r_j P(\mathbf{r} | \vec{V}). \quad (2.3)$$

$Q_{ij}^{-1}$  = the matrix inverse of  $Q_{ij}$ .  
 $\delta_{ij}$  = Kronecker delta defined by

$$\delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases} \quad (2.4)$$

$[x]_+$  = a notation defined by

$$[x]_+ = \begin{cases} x & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases} \quad (2.5)$$

for any quantity  $x$ .

### Statement of the Problem

Suppose that the firing rates of  $N$  recorded

neurons are correlated with some vector  $\vec{V}$ .  $\vec{V}$  is coded by the set of firing rates  $R_i$  with  $i = 1, 2, \dots, N$  and we wish to compute an estimate of  $\vec{V}$ , which we will call  $\vec{V}_{\text{est}}$ , on the basis of these firing rates. This is the decoding problem.

### Neuronal Responses

A preliminary issue concerns the measure used to quantify the firing rate of each neuron. We denote the raw firing rate of neuron  $i$  by  $R_i$ . However, it may be better to use some other variable  $r_i$  to represent the firing of neuron  $i$ . Numerous possibilities exist. Of course, we could just use the raw firing rate itself

$$r_i = R_i \quad (2.6)$$

or we may want to subtract the background firing rate of neuron  $i$ ,  $B_i$ , and define

$$r_i = R_i - B_i. \quad (2.7)$$

This neuronal response can be negative. It may be convenient to normalize  $r_i$  so that it is never greater than one by defining

$$r_i = \frac{R_i - B_i}{R_i^{\max} - B_i} \quad (2.8)$$

where  $R_i^{\max}$  is the maximum firing rate for neuron  $i$ . Finally, in some cases the average firing rate of the entire population of neurons may be variable. In this case, it may be useful to normalize the firing rate for neuron  $i$  relative to the firing rates for the entire population by defining, for example,

$$r_i = R_i \left( \sum_{j=1}^N R_j^2 \right)^{-1/2} \quad (2.9)$$

In our analysis we will primarily use definitions (2.6) and (2.8) although the methods and discussion apply to all of the above definitions and any others that might prove useful. To distinguish  $r$  from the firing rate  $R$ , we will refer to it as the neuronal or firing response.

Once a choice is made on which neuronal response to use, we must devise a method for combining the individual neuronal responses into



a population-based estimate. We will compare several different approaches applied to a number of different systems. We will assume throughout that the **vectors  $\vec{V}$  being encoded are chosen randomly and uniformly over some allowed range.** If this were not the case, the probability of a particular  $\vec{V}$  occurring,  $P(\vec{V})$ , would appear in some of our formulas. We begin by introducing each of the methods and discussing them briefly.

### Methods That Use Response Probabilities

The maximum likelihood method and the Bayesian estimator are standard statistical approaches to problems like decoding. They are likely to perform as well or better than any other method. A disadvantage of both of these methods is that **they require knowledge of the conditional probability function  $P(\mathbf{r} | \vec{V})$ .**

**Maximum Likelihood Method.** The maximum likelihood method is a standard statistical approach for obtaining parameter estimates from experimental data. Under certain circumstances, it can be shown to be optimal in the sense that **it minimizes the square of the difference between the estimate  $\vec{V}_{\text{est}}$  and the true vector  $\vec{V}$  averaged over all responses for each value of  $\vec{V}$ .** Some care must be exercised in applying theorems about the maximum likelihood method. For example, the Cramer-Rao theorem (Kiefer, 1987) is sometimes quoted to justify its use, but this theorem only establishes the superiority of the method relative to unbiased estimators, and the other estimators we consider are not always unbiased. For large numbers of neurons, however, it is likely that this method will be optimal and theoretical papers (Paradiso, 1988; Foldiak, 1993; Seung and Sompolinsky, 1993) tend to use it. **The estimated vector is simply the vector that maximizes the probability that the neuronal responses take the observed values  $\mathbf{r} = r_1, r_2, \dots, r_N$**

$$P(\mathbf{r} | \vec{V}_{\text{est}}) = \text{maximum.} \quad (2.10)$$

(Note that the quantity that we really want to maximize is  $P(\vec{V}_{\text{est}} | \mathbf{r})$  but this is related to the  $P(\mathbf{r} | \vec{V}_{\text{est}})$  by Bayes theorem and when  $P(\vec{V})$

is a constant, maximizing one is equivalent to maximizing the other. If  $P(\vec{V})$  is not constant,  $P(\vec{V}_{\text{est}} | \mathbf{r})$  should be maximized.)

**Bayesian Estimator.** Unless the population of neurons being recorded is large, there is no rigorous theoretical justification for using the maximum likelihood method. However, it is possible to write down an estimator that is guaranteed to minimize the squared difference  $(\vec{V}_{\text{est}} - \vec{V})^2$  averaged over all responses and over all values of  $\vec{V}$ ,

$$\int d\mathbf{r} d\vec{V} (\vec{V} - \vec{V}_{\text{est}})^2 P(\mathbf{r} | \vec{V}). \quad (2.11)$$

This estimator is obtained by minimizing this expression with respect to  $\vec{V}_{\text{est}}$  and is **just the average value of  $\vec{V}$**  for the probability distribution  $P(\mathbf{r} | \vec{V})$ ,

$$\vec{V}_{\text{est}} = \int d\vec{V} \vec{V} P(\mathbf{r} | \vec{V}) / P(\mathbf{r}). \quad (2.12)$$

(If  $P(\vec{V})$  is not constant, it should appear multiplying the integrand in this formula.) The Bayesian estimator always minimizes  $(\vec{V} - \vec{V}_{\text{est}})^2$  averaged over all responses and all vectors  $\vec{V}$  while the maximum likelihood estimator sometimes minimizes the same quantity averaged over responses for each value of  $\vec{V}$ .

### Methods That Use Response Tuning Curves

**Least Squares Estimator.** If we do not know the probability distribution  $P(\mathbf{r} | \vec{V})$ , we cannot use the maximum likelihood or Bayesian approaches as presented above. However, we can make a guess about the form of  $P(\mathbf{r} | \vec{V})$  and the most reasonable assumption is Gaussian statistics. If  $P(\mathbf{r} | \vec{V})$  is assumed to be **a product over all the neurons of Gaussian distributions with means  $f_i(\vec{V})$  and standard deviations  $\sigma_i$ ,** the maximum likelihood estimator can be determined by the simpler condition,

$$\sum_{i=1}^N \left( \frac{r_i - f_i(\vec{V}_{\text{est}})}{\sigma_i} \right)^2 = \text{minimum.} \quad (2.13)$$

This is just a least squares fit of the expected rates  $f_i(\vec{V}_{\text{est}})$  to the observed rates  $r_i$ .

**Optimal Linear Estimator (OLE).** During the course of this work, we constructed an estimator much like the Bayesian estimator except that it is required to **be linear** in the responses  $\mathbf{r}$  (see appendix A). It is the linear estimator that minimizes  $(\vec{V}_{\text{est}} - \vec{V})^2$  averaged over all trials and all vectors  $\vec{V}$ , formula (2.11), and is given by

$$\vec{V}_{\text{est}} = \sum_i r_i \vec{D}_i \quad (2.14)$$

where the vectors  $\vec{D}$  are determined by

$$\vec{D}_i = \sum_j Q_{ij}^{-1} \vec{L}_j. \quad (2.15)$$

with

$$\vec{L}_j = \int d\vec{V} \vec{V} f_j(\vec{V}) \quad (2.16)$$

and

$$Q_{ij} = \sigma_i^2 \delta_{ij} + \int d\vec{V} f_i(\vec{V}) f_j(\vec{V}) \quad (2.17)$$

(appendix A). The vectors  $\vec{L}$  represent the ‘center of mass’ of the tuning curves. For symmetric tuning curves they point in the same directions as the preferred direction vectors  $\vec{C}$ .  **$Q$  is the correlation matrix of firing rates defined,** in general, by equation (2.3). However, if the variances of the firing rates of the responding neurons about **their mean values  $f_i$**  are uncorrelated, the above formula for  $Q$  can be used (appendix A).

**Projection Method.** Another decoding method (Churchland and Sejnowski, 1992; Wilson and McNaughton, 1993) uses the neural response definition (2.6) and then chooses the estimate vector so that

$$\sum_{i=1}^N r_i f_i(\vec{V}_{\text{est}}) \left( \sum_{i=1}^N r_i^2 \right)^{-1/2} \left( \sum_{i=1}^N f_i^2(\vec{V}_{\text{est}}) \right)^{-1/2} = \text{maximum}. \quad (2.18)$$

This method minimizes the angle between the response vector  $\mathbf{r}$  and the mean response vector  $f_i$  for  $i = 1, 2, \dots, N$  in the  $N$ -dimensional space of neural responses. However if we use instead the neural response definition (2.9) and if all variances are equal for all cells, this method is

essentially (although not exactly) equivalent to the least squares method. We prefer to use the least squares method because it has a firmer theoretical basis.

### Method That Uses Preferred Direction Vectors

**Vector Method.** The method most commonly used in experiments involving population decoding is the vector method (Georgopoulos et al., 1983; Georgopoulos, Schwartz and Kettner, 1986). Suppose that neuron  $i$  responds with its maximal firing response when  $\vec{V} = \vec{C}_i$ . The vector method estimator is constructed from these preferred direction vectors by

$$\vec{V}_{\text{est}} = \sum_{i=1}^N r_i \vec{C}_i. \quad (2.19)$$

Clearly the normalization of the estimated vector depends on the specific definition of  $r_i$  being used. As a result, the method is mainly used in situations where **only the direction** and not the magnitude of  $\vec{V}_{\text{est}}$  is needed. It can be shown that for a uniform distribution of preferred direction vectors, the direction of  $\vec{V}_{\text{est}}$  will converge to that of  $\vec{V}$  provided that enough neurons are included in the sum (Georgopoulos, Kettner and Schwartz, 1988). Advantages of this method are that the estimator is a linear function of the responses  $r_i$  and the method only requires knowledge of the preferred direction vectors. As we will see below, the primary disadvantages of this method are that **it requires a large number of neurons** and that it is prone to errors produced by a non-uniform distribution of the  $\vec{C}$  vectors.

### Simulation Procedure

Many of the figures in this paper **show the average error for different decoding schemes.** To generate these results we simulated a series of experiments. The vector being coded,  $\vec{V}$ , **was chosen randomly and a set of neuronal responses was generated using mathematical fits of the response tuning curves for the system being studied.** We included variability in the

firing response by using a random number generator. The decoding method being evaluated was then applied to these simulated responses to obtain the estimate  $\vec{V}_{\text{est}}$ . An error was computed by comparing the estimated vector  $\vec{V}_{\text{est}}$  with the original vector  $\vec{V}$ . This procedure was repeated (typically a few thousand times) until an average error had been produced. In cases where we are only interested in the direction of the vector, the error is the average angle (in degrees) between the estimated vector  $\vec{V}_{\text{est}}$  and the 'true' vector  $\vec{V}$ . When we are interested in the magnitude as well as the direction of  $\vec{V}$ , the error reported is the distance between the  $\vec{V}_{\text{est}}$  and  $\vec{V}$  divided by the maximum length of  $\vec{V}$ .

## Results

### Coding Direction—The Cricket Cercal System

In the cricket cercal system, a wind stimulus is detected by a large number of filiform hair receptors located on the cerci. These synapse onto a small number of interneurons (Bacon and Murphey, 1984). At low wind velocities, the wind direction is primarily encoded by four interneurons, the left and right 10-2 and 10-3 neurons, and the responses of these neurons to wind stimuli with a fixed velocity profile but variable direction have been measured (Miller, Jacobs and Theunissen, 1991). The wind angle  $\theta$  in the horizontal plane is measured relative to the forward direction. We will define the response following definition (2.8) (with  $B = 0$ ) so that the maximal response is one. The average response as a function of position for neuron  $i$  is denoted by  $f_i(\theta)$ . A pulse of wind at angle  $\theta$  produced an average firing rate (Theunissen and Miller, 1991)

$$f_i(\theta) = \left[ \frac{\cos(\theta - \theta_i) - a}{1 - a} \right]_+ \quad (3.1)$$

where  $i = 1, 2, 3, 4$  represents the four interneurons,  $\theta_1 = 45^\circ$  and  $\theta_2 = 135^\circ$ ,  $\theta_3 = -135^\circ$ ,  $\theta_4 = -45^\circ$  and  $a = -0.14$ . The  $[\ ]_+$  notation, defined in equation (2.5), prevents the firing rate from being negative. The variability of the

firing rate about this mean was also determined (Theunissen and Miller, 1991). (In our simulations,  $f_i$  is actually the position of the peak of a Gaussian distribution of firing rates. Since rates cannot be negative, there is a slight difference between the average value and the peak of the Gaussian distribution at wind angles producing low firing rates.)

Because the firing statistics as well as the mean firing rate are known, the cercal system is a rare case where we have all the information we need to compare all of the different reconstruction methods. Figure 1 shows the results of such a comparison. All four methods give average errors around 5 or 6 degrees. Theunissen and Miller (1991) determined the intrinsic accuracy of this system using information theory and got results in the same range. This means that the errors shown for all four methods are caused primarily by the intrinsic error due to the variability of the neuronal responses and not by the decoding method. That is, the reconstructions are essentially as good as they can be.

The maximum likelihood and least squares reconstructions shown in figure 1 are virtually identical and both are very close to the Bayesian reconstruction. This follows from the fact that the probability distribution  $P(\mathbf{r} | \theta)$  is a product of four Gaussians except that the Gaussians are clipped to exclude negative firing rates (Theunissen and Miller, 1991). If they were exactly Gaussian distributions, the maximum likelihood and least squares methods would give identical results. For reasons discussed below and in appendix B, the vector and OLE decoding methods are identical in this case. These much simpler linear methods give surprisingly accurate results. When the wind is aligned with the preferred direction  $\theta_i$  for one of the interneurons, the other methods give a better estimate than the vector method. However, when the wind falls half way between these preferred directions ( $\theta = 0^\circ, \pm 90^\circ, 180^\circ$ ) the vector method is more accurate than the more complicated methods. In fact, the linear vector method is best in the minimax sense, that is, it has the smallest maximum error.

It is easy to understand why the vector method is so accurate for decoding the cer-

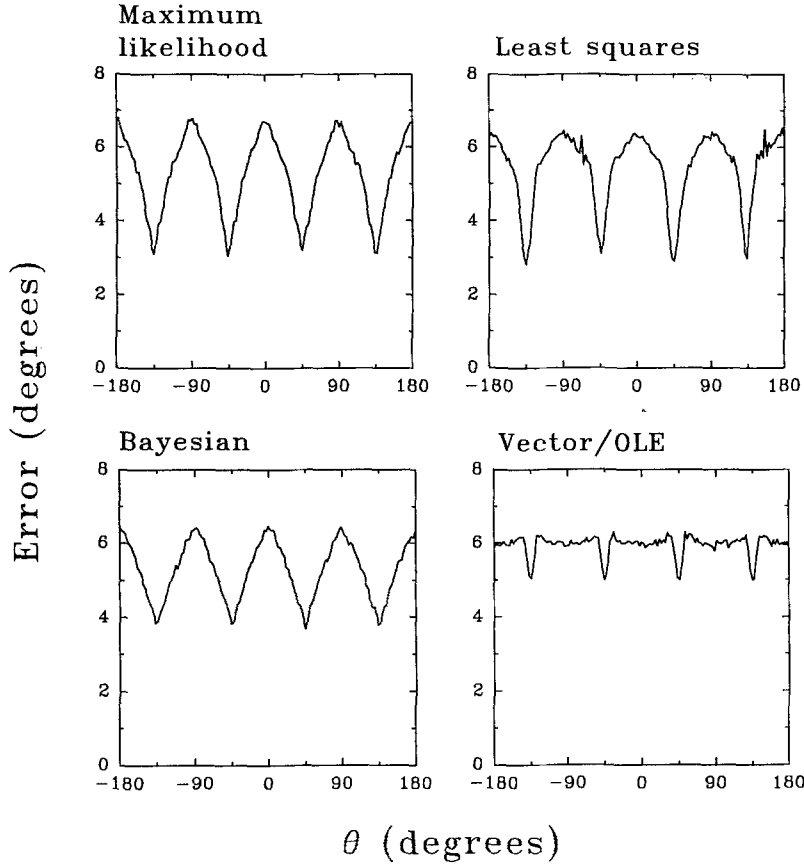


Fig. 1. Four decoding methods applied to the cricket cercal system. The average difference (in degrees) between the direction of the estimated vector and the true wind direction is plotted as a function of wind direction angle. Each point is the average of between 3000 and 6000 trials. Firing rates were generated by computer using the fits of Theunissen and Miller (1991) and reconstructions were made using the methods described in the text.

cal interneurons. We define the preferred direction vectors for the four interneurons,  $\vec{C}_i$  with  $i = 1, 2, 3, 4$ , as the wind directions producing maximal responses in the corresponding interneurons.  $\vec{C}_1, \vec{C}_2, \vec{C}_3, \vec{C}_4$  are vectors of unit length pointing in the directions given by the angles  $\theta_1, \theta_2, \theta_3, \theta_4$  respectively. Likewise, we define the wind direction vector  $\vec{V}$  as a vector of unit length pointing in the direction of the wind  $\theta$ . In terms of these vectors, the average neuronal response functions of equation (3.1) can be written as

$$f_i(\vec{V}) = \left[ \frac{\vec{V} \cdot \vec{C}_i - a}{1 - a} \right]_+ \quad (3.2)$$

If  $a$  were equal to zero, the average response

of neuron  $i$  would be the component of the vector  $\vec{V}$  along the direction  $\vec{C}_i$ . Since the preferred direction vectors  $\vec{C}_i$  are separated by right angles, the set of responses  $r_i$  for  $i = 1, 2, 3, 4$  would then be the positive  $x$ , positive  $y$ , negative  $x$  and negative  $y$  components of the vector  $\vec{V}$  in a coordinate system tipped  $45^\circ$  relative to the body axes of the cricket. In other words, the cricket cercal system is using ordinary Cartesian coordinates to represent the wind direction vector. The only difference between this representation and the one we would normally use, is that the cricket must use four coordinates instead of two because the interneuron firing rates cannot be negative.

If the Cartesian coordinates  $r_i$  of a vector



along orthogonal coordinate axes with directions  $\vec{C}_i$  are known, the vector can be reconstructed from the coordinates just by multiplying the components by the coordinate direction vectors exactly as in equation (2.19) defining the estimate for the vector method. The only thing unfamiliar about this formula is the use of four rather than two coordinates caused by the lack of negative numbers. Thus, the vector method in the case of the cricket cercal system is just the standard Cartesian reconstruction of a vector from its components.

The above discussion applies to the case  $a = 0$  but in reality  $a = -0.14$  (Theunissen and Miller, 1991). However, as shown in figure 2B, the difference between the tuning curves with  $a = 0$  and with  $a = -0.14$  is quite small. In fact, the nonzero value of  $a$  actually improves the performance of the system (Theunissen and Miller, 1991) by compensating for noise effects at low firing rates. In figure 2A, we have plotted both the average error and the maximum error (the error for the worst case wind direction) for the vector reconstruction with different values of  $a$ . The average error reaches a minimum very close to  $a = 0$ , but the maximum error is minimized by setting  $a$  equal to the value  $-0.12 \pm 0.01$  which is very close to the measured value of  $-0.14$  (Theunissen and Miller, 1991). The results of figures 1 and 2 indicate that, with the observed value of  $a$ , the vector method minimizes the worst case error in the reconstruction of wind direction.

### Linear Reconstruction – General Case

The success of the linear reconstruction in the case of the cricket cercal system led us to consider linear reconstructions in other cases. Under what conditions is a linear reconstruction method appropriate? In appendix C we answer a related question, what type of neuronal response tuning curve allows for the best possible reconstruction by a linear estimator? The answer is a cosine tuning curve, that is, an average response proportional to the projection of the coded vector  $\vec{V}$  along the preferred direction vector for neuron  $i$ ,

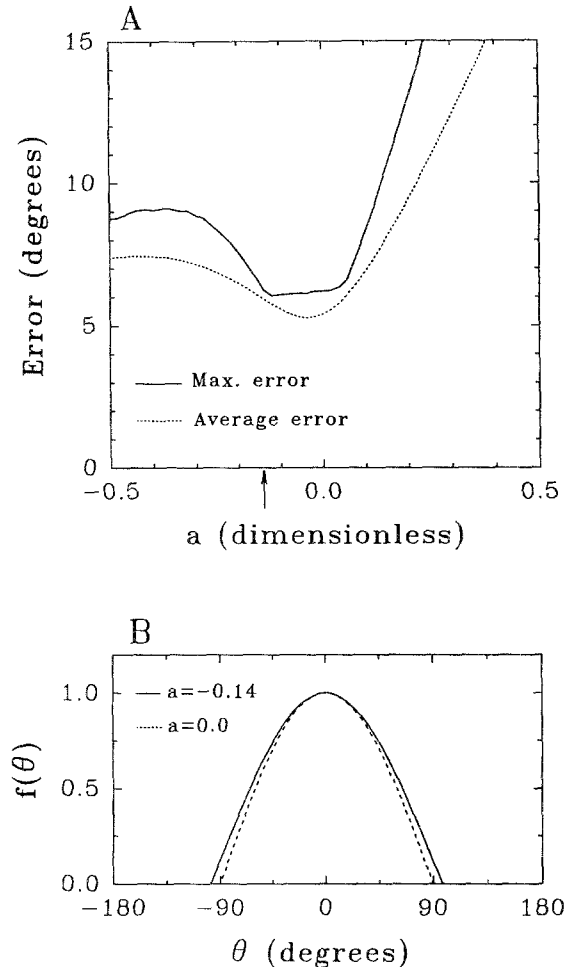


Fig. 2. Comparison of decoding for different values of the parameter  $a$  appearing in the cercal interneuron tuning curves. A) The average error and maximum error of the vector method of reconstruction as a function of  $a$ . The minimum value of the average error occurs near  $a = 0$  where the tuning curve is a perfect half-cosine function. The maximum error is the magnitude of the reconstruction error for the worst case wind direction. The maximum error is minimized when  $a = -0.12$  very close to the measured value (arrow)  $a = -0.14$  (Theunissen and Miller, 1991). B) A comparison of the interneuron average response tuning curves with  $a = 0$  (dashed) and with  $a = -0.14$  (solid). When  $a > 0$  the tuning curve is thinner than a half cosine and when  $a < 0$  it is broader than a half cosine.

$$f_i(\vec{V}) = \vec{V} \cdot \vec{C}_i \quad (3.3)$$

( $r_i$  can always be defined so that the proportionality constant in this equation is one). Many examples of cosine or approximately cosine tuning curves exist as discussed in the introduction. In these cases, the raw firing rates are fit by the form

$$R_i = [B_i + K_i \vec{V} \cdot \vec{C}_i]_+ \quad (3.4)$$

so defining

$$r_i = \frac{R_i - B_i}{K_i} \quad (3.5)$$

we obtain a response  $r_i$  with average tuning curve

$$f_i(\vec{V}) = \begin{cases} \vec{V} \cdot \vec{C}_i & \text{if } \vec{V} \cdot \vec{C}_i \geq -B_i/K_i \\ -B_i/K_i & \text{if } \vec{V} \cdot \vec{C}_i < -B_i/K_i \end{cases} \quad (3.6)$$

In such a situation, a linear reconstruction is strongly suggested.

If  $B/K < 1$ , the tuning curve of equation (3.6) does not describe a complete cosine function but **only a portion of it**. We will nevertheless refer to this as **a cosine tuning curve**. We will mainly consider two types of cosine tuning curves in our analysis: those with  $B_i/K_i > 1$  for all  $i$  so that the average firing rate describes **a full cosine function**,

$$f_i(\vec{V}) = \vec{V} \cdot \vec{C}_i, \quad (3.7)$$

and those with  $B_i = 0$  for all  $i$  so that the tuning curves **are half cosines**

$$f_i(\vec{V}) = [\vec{V} \cdot \vec{C}_i]_+. \quad (3.8)$$

This second case is fairly close to what is found in the cricket cercal system. When we consider the monkey motor cortex, neurons are found with **a wide distribution of  $B$  and  $K$  values** (Schwartz, Kettner and Georgopoulos, 1988) and thus we must include all values of  $B/K$  in our analysis.

In the cricket cercal system at low wind velocities, the wind direction is coded by four interneurons accurately tuned to directions located symmetrically, at  $90^\circ$  intervals around the circle. In figure 3, we examine what happens

when **a larger number of nonsymmetrically tuned neurons is used**. Here, we have randomly chosen five preferred directions for five neurons responding much like the cricket interneurons with **half-cosine tuning curves**. Despite the fact that the variability in the firing rates of these neurons is similar to that of the cricket interneurons, the error in the vector method reconstruction, shown in the upper panel of figure 3, is very large. As in the previous case, **the responses of the neurons give the components of  $\vec{V}$  along the directions defined by the preferred direction vectors  $\vec{C}_i$** . However, in this case the vectors  $\vec{C}_i$  do not define an orthogonal coordinate system and they are not uniformly distributed in angle. As a result, the reconstruction formula (2.19) is not correct.

The OLE in this generalized case is not equivalent to the vector method. The vectors  $\vec{D}$  used in the OLE reconstruction of (2.14), shown in the lower panel of figure 3, do not point in the same directions or have lengths equal to the  $\vec{C}_i$  vectors shown in the upper panel. The correlation matrix  $Q$  used in the OLE reconstruction corrects for the bias and non-orthogonality of the preferred direction vectors and produces an estimate of high accuracy as seen in the lower panel of figure 3. The lesson from this example is that linear reconstruction procedures can produce accurate estimates in cases where neurons have cosine tuning curves, but the vector method is subject to errors due to nonuniform distributions of neuronal preferred directions that can be corrected by the OLE approach.

In appendix B, we show that when the vectors  $\vec{C}_i$  are uniformly distributed, the  $\vec{D}$  vectors are proportional to the  $\vec{C}_i$  vectors and  $\vec{V}_{\text{est}}$  for the OLE method will point in the same direction as it does for the vector method. However, if the preferred directions of the neurons are not uniformly placed around the circle (or sphere in three dimensions) the methods are not equivalent. It is unlikely that in a random sampling of  $N$  cells, a uniform distribution of  $\vec{C}_i$  vectors will be found even if the full population of neurons has a uniform distribution, unless  $N$  is very large. If the full population is not uniform, then of course, even large samples are unlikely to be uniform.

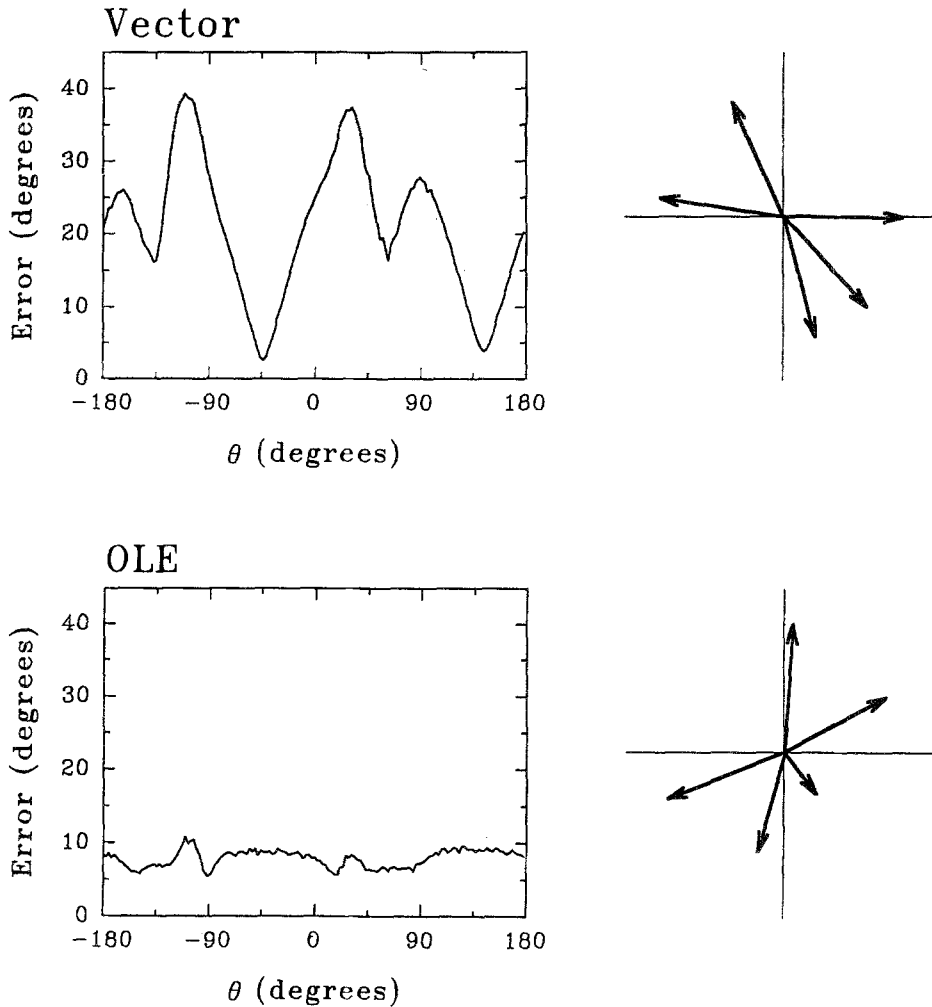


Fig. 3. The vector and OLE reconstructions of the direction of a two-dimensional vector using five neurons with random preferred direction vectors and half-cosine tuning curves. In these runs, all cells had the standard deviation of their firing rates set to  $\sigma = 0.1$ . The randomly chosen preferred direction vectors are shown in the plot at the upper right. The average error for the vector method is plotted as a function of wind angle at the upper left. The vector method gives a poor reconstruction because the vectors are not uniformly placed around the circle. The  $\vec{D}$  vectors of the OLE method corresponding to these preferred direction vectors are shown at the bottom right. The OLE error shown at the bottom left is quite small and relatively constant over wind directions. Results are based on 500 trials per direction.

#### Arrays of Direction Sensitive Neurons—The Monkey Motor Cortex

We now consider the reconstruction of a vector  $\vec{V}$  encoded by a large population of neurons with cosine or half-cosine responses. The preferred direction vectors  $\vec{C}_i$  are chosen randomly and, in our initial example, we allow a variability in the neuronal response that is uniformly 10%

of the peak response. The statistics of this variability is Gaussian. In this case, as in the case of the cricket cercal system, the maximum likelihood, Bayesian and least squares methods give practically identical results. Therefore, we will only show the results of the simpler **least squares method** in our figures.

Figure 4 compares the accuracy of the least squares, vector and OLE reconstructions for ar-

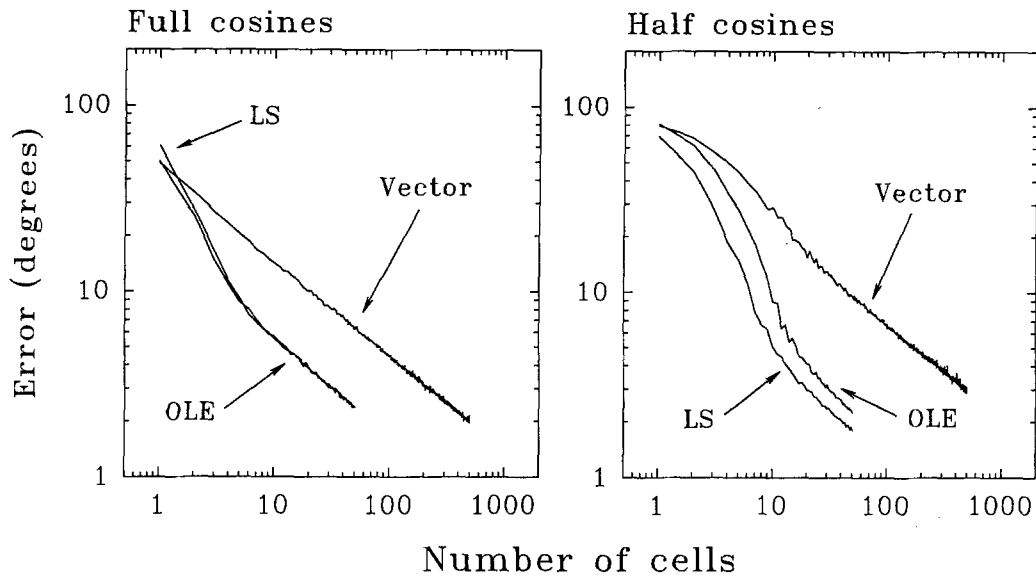


Fig. 4. Comparison of least squares, OLE and vector methods for arrays of cells with full- and half-cosine tuning curves coding the direction of a two-dimensional vector. All cells had  $\sigma = 0.1$  and the preferred direction vectors were chosen randomly. The average angle between the direction of the coded vector and its reconstructed estimate is plotted as a function of the number of cells used in the decoding on a log-log plot. Results are based on between 1500 and 3750 trials per point. The accuracy of the full- and half-cosine cases should not be compared directly with each other because the  $\sigma = 0.1$  variability has a different impact in the two cases.

rays of full cosine and half cosine tuning curves as a function of the number of neurons used in the reconstruction,  $N$ . For full cosine tuning curves, the least squares and OLE reconstructions are equally accurate and are significantly better than the vector method of decoding. With the vector method, it takes roughly ten times as many neurons to decode with an accuracy equivalent to that of the other two methods. The OLE is not quite as accurate as the least squares methods for half cosine tuning curves. This is because the restriction to half cosines introduces some nonlinearity into the coding. However, the OLE still does an excellent job of decoding and is again significantly better than the vector method. For large  $N$ , all three curves in both plots of figure 4 are linear on these log-log plots with slopes of  $-1/2$ . This means that the accuracy is decreasing like  $N^{-1/2}$  in this range, so for these values of  $N$  statistical fluctuations are the primary factor limiting the accuracy of the reconstructions.

Figure 5 shows the limitations of the vector method when there is a bias in the preferred

direction vectors of the neurons. To produce this figure, we used an array of neurons with full cosine tuning curves and randomly chosen preferred directions, except that the preferred directions were excluded from the range  $0 < \theta < 57.3^\circ$  (one radian). This introduces a bias that does not go away even for large  $N$ . The bias has a large effect on the accuracy of the vector method. Instead of decreasing like  $N^{-1/2}$  for large  $N$ , the accuracy levels off and approaches a constant. The accuracy of the other two methods is essentially unaffected by the bias.

We have studied the effects of different amounts of firing response variability on the accuracy of the reconstructions. As the variability increases the accuracy of all the methods decreases as does the difference between the various methods. With high degrees of variability, all the methods we have considered give essentially the same result and require a large number of recorded neurons to produce an accurate reconstruction.

The accuracy of the OLE for full and half cosine tuning curves suggests that it would be an

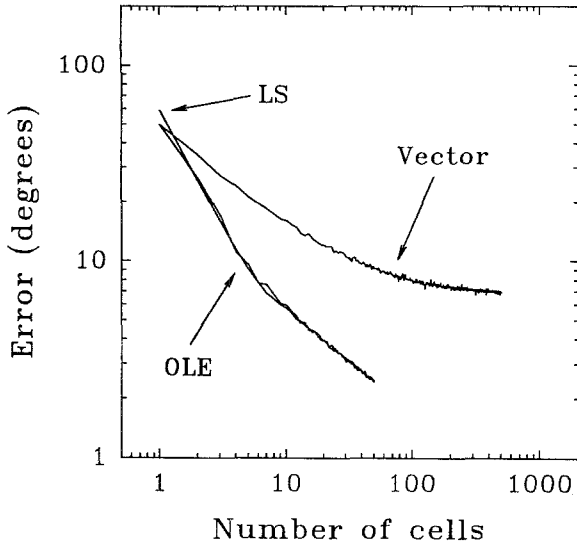


Fig. 5. Comparison of least squares, OLE and vector methods for arrays of cells coding the direction of a two-dimensional vector with **full-cosine tuning curves** having a bias in the choice of preferred direction vectors. Preferred direction vectors were chosen randomly except that they were excluded from the region between zero and one radian. All  $\sigma = 0.1$ . The bias has a large impact on the vector method but little on the other two reconstructions.

ideal method for decoding the arm movement vector from the population of neurons in the monkey motor cortex (Georgopoulos, Schwartz and Kettner, 1986) and other similar systems. Figure 6 shows a comparison of the vector and OLE methods as applied to recordings of monkey motor cortex during a task involving three-dimensional reaching movements to targets. This analysis was done by G. Pellizzer and A. Georgopoulos who kindly allowed us to use this figure. The results are based on recordings from 189 tuned cells made during reaching movements to the eight corners of a cube. The OLE method is clearly a dramatic improvement over the vector method. In particular, the OLE allows a reconstruction of the movement vector using far fewer recorded cells. The accuracy of the OLE using 5 cells is better than that of the vector method using all 189 recorded cells. This raises the exciting possibility of decoding in real time without having to sum over repeated trials. Hopefully, this will allow further studies of the dynamics of the population vector like those reported in Georgopoulos et al. (1989).

### Coding Values and Positions

Up to now, we have only considered examples in which the direction of the vector  $\vec{V}$  is being coded but not its length. Because the normalization of the estimate vector  $\vec{V}_{\text{est}}$  is arbitrary for the vector method, it is not clear how to apply it to cases where the length of  $\vec{V}$  needs to be estimated as well. However, all of the other methods we have discussed can be used to decode vector lengths in addition to directions. When both the length and direction of a vector are encoded by an array of neurons with cosine tuning curves, the OLE should be an excellent method of reconstruction. Indeed, as derived in appendix A, it is optimized for just this case.

However, there are important examples where the length of a vector, or value of a scalar quantity, is coded by arrays of neurons with tuning curves that do not resemble cosines. In the rat hippocampus, place cells code for the position of the animal within its enclosure (O'Keefe and Nadel 1978; Wilson and McNaughton 1993) but they do not act like cosine-tuned neurons. Instead their average response is a function of the distance between the position of the rat  $\vec{V}$  and a preferred position (as opposed to preferred direction) for a particular place cell. In the bat (O'Neill and Suga 1982; Suga and Horikawa 1986), echo delay is coded by an array of tuned neurons but again the tuning curves are not cosines. We will approximate such a situation by considering a population of  $N$  neurons with average responses given by Gaussian tuning curves

$$f_i(\vec{V}) = \exp\left(-\frac{1}{2}((\vec{V} - \vec{C}_i)/\Delta)^2\right) \quad (3.9)$$

where  $\Delta$  is the width of the tuning curve and  $\vec{C}_i$  is the preferred position vector for neuron  $i$ . We set the width of all the tuning curves to the same value and choose the preferred positions  $\vec{C}_i$  randomly. For simplicity, we will study this problem in only one dimension, so the 'vectors'  $\vec{V}$  and  $\vec{C}_i$  will actually represent scalar quantities.

The results of applying the least squares and OLE methods to the scalar coding problem are shown in figure 7. Since the tuning curves are not cosines, the OLE reconstructions are



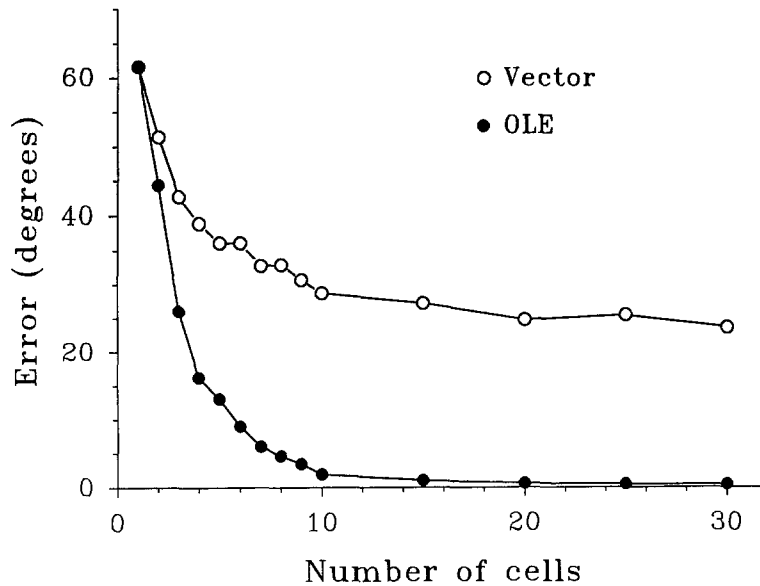


Fig. 6. Comparison of the OLE and vector methods for reconstruction of arm movement directions from recordings of direction-sensitive neurons in the monkey motor cortex. The plot was made by G. Pellizzer and A. Georgopoulos based on data from the Georgopoulos lab. Each plotted point represents the average angular reconstruction error as a function of the number of cells used. For each data point shown, 20 random sets of the given number of cells were selected from the 189 neurons that were recorded during reaching movements to targets located at the eight corners of a cube. For each set, the error was averaged over the eight different movement directions and then was averaged over the 20 random sets.

significantly worse than those of the least squares method. In particular, the error for the OLE method approaches a limit at large  $N$  rather than decreasing like  $N^{-1/2}$ . This reflects the limit of a linear estimator in a situation where the coding is not linear and serves as a reminder that linear decoding schemes are not always the best strategy. Nevertheless, the OLE method works fairly well even in this situation since the limiting error is fairly small. The least squares method again works extremely well.

## Discussion

The different methods we have discussed appear to be capable of extracting a fairly accurate estimate of a coded vector from a set of spike rates in a wide variety of systems. Approaches like the maximum likelihood, Bayesian and least squares methods can provide accurate reconstructions for a wide variety of systems. When the coding neurons have complete or

partial cosine tuning curves, linear reconstructions are appropriate. The new OLE method removes the sensitivity to nonuniform distributions of preferred directions that is a weakness of the vector method. In cases where the vector method of decoding shows a systematic discrepancy between the reconstructed and true vectors, it is likely the result of biases in the preferred directions of neurons being recorded. Use of the OLE is likely to eliminate the discrepancy.

In situations where neurons have broadly tuned responses, it is often said that individual neurons carry little information about the external correlate and that accurate reconstruction requires summing over a large population of neurons. In the cases we have studied, this is a misrepresentation of what is going on. A neuron with a cosine response is certainly broadly tuned. However, knowledge of its firing response gives the projection of the coded vector along a particular preferred direction which is as much information as a single number can carry about a vector. In the absence of noise, a  $d$ -dimensional vector

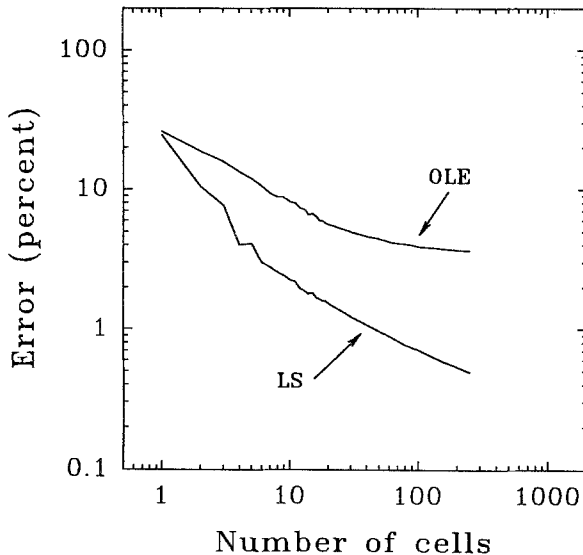


Fig. 7. Comparison of the least squares and OLE reconstructions for a scalar quantity coded by an array of neurons with Gaussian tuning curves. The coded quantity took values between zero and one and the Gaussians were centered about random positions within this range and had widths  $\Delta = 0.25$ . All  $\sigma = 0.1$ . Error reported is the root mean square difference between the estimated and coded quantities expressed in percent.

could be reconstructed from just  $d$  such firing rates (or  $2d$  for the case of half-cosine tuning curves). This is the coding strategy used by the cricket cercal interneurons. If more neurons are needed for the reconstruction, this may be due to a high degree of variability in the neuronal responses or to the use of an inefficient reconstruction method which is prone to errors caused by a nonuniform and non-orthogonal distribution in the preferred direction vectors of the sampled neurons. In the absence of large firing-rate fluctuations, an efficient algorithm like the OLE or least squares method should be able to provide accurate decoding using a relatively small number of neurons.

If it is possible to code vectors with a small number of neurons (as the cricket does), why do other systems employ large neural populations? Our work suggests that when a neural network tries to encode a vector using more than the minimum number of neurons (perhaps to reduce statistical fluctuations) two problems arise. First, it may be difficult (or even im-

possible in three dimensions, see appendix B) to achieve a uniform distribution of preferred direction vectors. If a uniform distribution cannot be realized, down-stream decoding networks cannot rely on a simple linear sum such as the vector method, but must face the task of inverting the correlation matrix  $Q$  if they attempt something like the OLE reconstruction. Learning rules developed for artificial neural networks (Linsker, 1993) can invert a matrix and it is possible to devise a learning algorithm that develops an OLE reconstruction. However, a simpler solution to this problem is to use a large population of neurons with an approximately uniform distribution to encode the vector so that the vector method can be used. Nevertheless, the experimenter recording only a small fraction of these neurons is better off using a different reconstruction technique like the OLE.

### Acknowledgments

We are extremely grateful to G. Pellizzer and A. Georgopoulos for helpful comments and for providing figure 6. We also thank F. Theunissen and J. Miller for discussions about their work on the cercal system, S. Seung and M. Wilson for telling us about their approaches to decoding and S. Guttman for advice on statistics. This research was supported by the National Science Foundation (DMS-9208206) and the Conacyt-Fulbright-IIE program.

### Appendix A—Derivation of the OLE

We express the population vector  $\vec{V}_{\text{est}}$  as in equation (2.14). The vectors  $\vec{D}$  are determined by finding the set of vectors that makes this linear estimate as accurate as possible. This is done by minimizing the square of the difference between the estimated vector and the true vector averaged over responses and over  $\vec{V}$  values. Thus, we minimize

$$\int d\mathbf{r} d\vec{V} P(\mathbf{r} | \vec{V}) \left( \vec{V} - \sum_i r_i \vec{D}_i \right)^2. \quad (8.1)$$

We find the minimum by taking a derivative with respect to  $\vec{D}_i$  and setting the result to zero. This gives the condition

$$\int d\mathbf{r} d\vec{V} P(\mathbf{r} | \vec{V}) r_i \left( \vec{V} - \sum_j r_j \vec{D}_j \right) = 0. \quad (8.2)$$

Using equation (2.1) and the definitions of the vectors  $\vec{L}$  and the matrix  $Q$  in equations (2.2) and (2.3), this becomes

$$\vec{L}_i = \sum_j Q_{ij} \vec{D}_j. \quad (8.3)$$

To solve this equation we invert the matrix to obtain the solution (2.15).

It is important to realize that the variances  $\sigma_i^2$  may play an important role in defining the inverse of the correlation matrix  $Q$ . In some situations, if we set all the variances to zero, (8.3) may have more than one set of solutions so that  $Q$  cannot be inverted. For example, suppose that two neurons have preferred direction vectors pointing in virtually the same direction. Since these two neurons are providing the same information there is an ambiguity about how to weight their firing rates in the population vector sum. This ambiguity is resolved by non-zero variances because we then weight the neurons proportional to the inverse of their variances. The bottom line is that it may be essential to include the variances to assure that  $Q$  can be inverted.

If the variations in the firing rates of the neurons are independent of each other, we can express the firing rates in response to a vector  $\vec{V}$  as

$$r_i = f_i(\vec{V}) + z_i \quad (8.4)$$

where  $z_i$  is an independent random variable with zero mean and standard deviation  $\sigma_i$ . Substituting this into equation (2.3) gives the result (2.17).

## Appendix B—Equivalence of OLE and Vector Methods for Uniform Distributions

Suppose we have a set of neurons with identical firing rate variances,  $\sigma_i^2 = \sigma^2$  for all  $i$ , and tuning curves that depend only on the dot product

between the encoded vector  $\vec{V}$  and the preferred direction vector  $\vec{C}_i$  for neuron  $i$ . We take the **preferred direction vectors to be of unit length**. Then by rotational invariance, the correlation matrix element  $Q_{ij}$ , and the corresponding inverse element  $Q_{ij}^{-1}$  can only depend on  $\vec{C}_i \cdot \vec{C}_j$ , so that

$$Q_{ij}^{-1} = F(\vec{C}_i \cdot \vec{C}_j) \quad (9.1)$$

where  $F$  is an arbitrary function. Assume also that the preferred direction vectors  $\vec{C}$  are described by or chosen from a distribution  $P(\{\vec{C}\})$  that is uniform over the circle (in two-dimensions) or sphere (in three dimensions) (The notation  $\{\vec{C}\}$  stands for the complete set of vectors  $\vec{C}_i$  for all  $i$ .) In this case, the  $\vec{D}$  vectors of the OLE method are proportional to the preferred direction vectors,  $\vec{D}_i = k\vec{C}_i$ , so the OLE and vector methods will give identical predictions for the direction of the estimated vector  $\vec{V}_{\text{est}}$ . To prove that  $\vec{D}_i = k\vec{C}_i$  for all  $i$ , we will show that, when averaged over the distribution of preferred direction vectors,  $\vec{D}_i \times \vec{C}_i = 0$  and  $\vec{D}_i \cdot \vec{C}_i = k$ . The first identity shows the  $\vec{D}_i$  is parallel to  $\vec{C}_i$  and the second that the constant of proportionality between these two vectors does not depend on  $i$ . This is sufficient to prove the identity of the two methods as far as the direction of the estimated vector is concerned.

Both proofs are trivial. From the definition of  $\vec{D}$ , the statement that the average of the cross product of  $\vec{D}$  and  $\vec{C}$  is zero amounts to

$$\int d\{\vec{C}\} P(\{\vec{C}\}) \sum_{j=1}^N F(\vec{C}_i \cdot \vec{C}_j) \vec{C}_j \times \vec{C}_i = 0. \quad (9.2)$$

The integral above is a vector. However, because  $P(\{\vec{C}\})$  is uniform its definition involves no fixed vectors. The only vectors appearing in the integrand, the  $\vec{C}_j$ , are being integrated over. Thus, there is no vector to which the integral could be equal and it must be zero.

The average of the dot product of  $\vec{D}_i$  with  $\vec{C}_i$  is

$$\int d\{\vec{C}\} P(\{\vec{C}\}) \sum_{j=1}^N F(\vec{C}_i \cdot \vec{C}_j) \vec{C}_j \cdot \vec{C}_i = k. \quad (9.3)$$

Again this cannot depend on  $i$  because the vector  $\vec{C}_i$  is being integrated over and the assumed uniformity of the probability distribution  $P(\{\vec{C}\})$  does not permit the integral to distinguish between different values of the index  $i$ .

The above proof is couched in statistical terms, that is, the proof requires integrating over a distribution  $P(\{\vec{C}\})$ . In two dimensions this is irrelevant, we could prove the same results in a more deterministic form. However, in three dimensions, distributions of vectors that actually satisfy the uniformity criterion are extremely rare, the  $\vec{C}$  vectors must point to the vertices of a Platonic solid. Thus, in general it is impossible to realize the uniformity criterion with a finite number of vectors in three dimensions. However, the  $\vec{C}$  vectors can be drawn from a uniform distribution and, for large  $N$ , the OLE and vector methods should converge. The difference between the two and three dimensional case could be quite important and lead to different coding strategies for vectors of different dimension.

### Appendix C – Optimal Tuning Curve for Linear Decoding

Suppose that we use a linear decoding method so that the estimated vector is given by equation (2.14) with arbitrary vectors  $\vec{D}$ . What form of the tuning curve  $f$  will assure that this estimator is as accurate as possible? To answer this question we set  $r_i$  equal to its expected value  $f_i$  (ignoring fluctuations around this mean) and find the function  $f_i$  that minimizes the error

$$\int d\vec{V} \left( \vec{V} - \sum_{i=1}^N f_i(\vec{V}) \vec{D}_i \right)^2. \quad (10.1)$$

Taking the derivative with respect to  $f_i$  gives the condition

$$\sum_{j=1}^N \vec{D}_i \cdot \vec{D}_j f_j(\vec{V}) = \vec{D}_i \cdot \vec{V} \quad (10.2)$$

A solution of this equation is

$$f_i(\vec{V}) = \vec{V} \cdot \vec{C}_i \quad (10.3)$$

where, using  $\alpha = 1, 2$  in two dimensions and  $\alpha = 1, 2, 3$  in three dimensions to denote the components of the vectors,

$$C_i^\alpha = \sum_{\beta} P_{\alpha\beta}^{-1} D_i^\beta \quad (10.4)$$

with

$$P_{\alpha\beta} = \sum_{i=1}^N D_i^\alpha D_i^\beta. \quad (10.5)$$

Thus, cosine tuning curves are the optimal case for linear decoding methods.

### References

- Altes R.A. (1989) Ubiquity of hyperacuity. *J. Acoust. Soc. Am.* 85:943-952.
- Bacon J.P. and Murphey R.K. (1984) Receptive fields of cricket (*Acheta domestica*) interneurons are related to their dendritic structure. *J. Physiol. (Lond.)* 272:779-797.
- Baldi P, Heiligenberg W. (1988) How sensory maps could enhance resolution through ordered arrangements of broadly tuned receivers. *Biol. Cybern.* 59:313-318.
- 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., Warland, D. (1991) Reading a neural code. *Science* 252:1854-1857.
- Burnod Y., Grandguillaume P., Otto, I., Ferraina S., Johnson P.B., Caminiti R. (1992) Visuomotor transformation underlying arm movements toward visual targets: A neural network model of cerebral cortical operations. *J. Neurosci.* 12:1435-1452.
- Camhi J.M., Levy A. (1989) The code for stimulus direction in a cell assembly in the cockroach. *J. Comp. Physiol.* A165:83-97.
- Camhi, J.M. and Tom, W. (1978) The escape behavior of the cockroach *Periplaneta americana* I. Tuning response to wind puffs. *J. Comp. Physiol.* A128:193-201.
- Caminiti R., Johnson P.B., Galli C., Ferraina S., 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., Ortiz E.R. (1990) Head-directional and behavioral correlates of posterior cingulate and medial prefrontal cortex neurons in freely-moving rats. *Soc. Neurosci. Abst.* 16:441.
- Churchland P.S., Sejnowski T.J. (1992) *The Computational Brain*. MIT Press, Cambridge MA.
- Eichenbaum H. (1993) Thinking about brain cell assemblies. *Science* 261:993-994.
- 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.
- Foldiak P. (1991) Models of Sensory Coding. Cambridge Univ. Ph.D. Thesis.
- Fortier P.A., Kalaska J.F., Smith A.M. (1989) Cerebellar neuronal activity related to whole-arm reaching movements in the monkey. *J. Neurophysiol.* 62:198–211.
- Georgopoulos A.P., Kettner R.E., 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., Lurito J.T., Petrides M., Schwartz A., Massey J.T. (1989) Mental rotation of the neuronal population vector. *Science* 243:234–236.
- 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., Lukashin A. (1993) Cognitive neurophysiology of the motor cortex. *Science* 260:47–52.
- Gilbert C.D., 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.
- Gozani S.N., Miller J.P. (1993) Ensemble coding of information by eight primary sensory interneurons in the cricket cercal system (in preparation).
- Kalaska J.F., Caminiti R., 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.
- Kiefer J.C. (1987) *Introduction to Statistical Inference*. Springer-Verlag, NY.
- Knudsen E.I., DuLac S., Esterly S.D. (1987) Computational maps in the brain. *Annu. Rev. Neurosci.* 10:41–65.
- Konishi M. (1987) Centrally synthesized maps of sensory space. *Trends Neurosci.* 9:163–168.
- Lee C., Rohrer W.H., Sparks D.L. (1988) Population coding of saccadic eye movements by neurons in the superior colliculus. *Nature* 332:357–360.
- Lehky S.R., Sejnowski T.J. (1990) Neural model of stereoacuity and depth interpolation based on a distributed representation of stereo disparity. *J. Neurosci.* 10:2281–2299.
- Linsker R. (1993) Local synaptic learning rules suffice to maximize mutual information in a linear network. *Neural Comp.* 4:691–702.
- Lukashin, A.V. (1990) A learned neural network that simulates properties of the neural population vector. *Biol. Cybern.* 63:377–382.
- Miller J.P., Jacobs G.A., 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.
- O'Keefe J., Nadel L. (1978) *The Hippocampus as a Cognitive Map*. Clarendon, Oxford.
- O'Neill W.E., 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., 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.
- 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.
- Rieke F.M. (1991) *Physical Principles Underlying Sensory Processing and Computation*. Univ. of California Berkeley PhD Thesis.
- Schwartz A., Kettner R.E., 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. *J. Neurosci.* 8:2913–2927.
- Sejnowski T.J. (1988) Neural populations revealed. *Nature* 332:308.
- Seung H.S., Sompolinsky H. (1993) Simple neural network models of psychophysical tasks. *Proc. Natl. Acad. Sci. (USA)* (in press).
- Shor R.H., Miller A.D., Tomko D.L. (1984) Responses to head tilt in cat central vestibular neurons. I. Direction of maximum sensitivity. *J. Neurophysiol.* 51:136–146.
- Snippe H.P., Koenderink J.J. (1992) Discrimination thresholds for channel-coded systems. *Biol. Cybern.* 66:543–551.
- Steinmetz M.A., Motter B.C., Duffy C.J., Mountcastle V.B. (1987) Functional properties of parietal visual neurons: Radial organization of directionalities with the visual field. *J. Neurosci.* 7:177–191.
- Suga N., Horikawa J. (1986) Multiple time axes for representation of echo delays in the auditory cortex of the mustached bat. *J. Neurophysiol.* 55:776–805.
- Suzuki I., Timerick J.B., 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.
- Taube J.S., Muller R.I., Ranck J.B.J. (1990) Head direction cells recorded from the postsubiculum 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., 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., Wan H.S. (1993) Neural representation of space using sinusoidal arrays. *Neural Comp.* 5:869–884.
- Van Gisbergen J.A.M., Van Opstal A.J., Tax A.M.M. (1987) Collicular ensemble coding of saccades based on vector summation. *Neuroscience* 21:541–555.
- Van Opstal A.J., 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., Landolfi M.A., Miller J.P., 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 M.A., McNaughton B. (1993) Dynamics of the hippocampal ensemble code for space. *Science* 261:1055–1058.
- Young M.P., Yamane S. (1992) Sparse population coding of faces in the inferotemporal cortex. *Science* 256:1327–1331.
- Zhang J., 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.