THE EFFECT OF CORRELATED VARIABILITY
ON THE ACCURACY OF A POPULATION CODE

L. F. Abbott
Volen Center and Department of Biology
Brandeis University
Waltham MA 02254-9110

and

Peter Dayan
Department of Brain and Cognitive Sciences
Massachusetts Institute of Technology
Cambridge, MA 02139

Abstract

We study the impact of correlated neuronal firing rate variability on the accuracy with which an encoded quantity can be extracted from a population of neurons. Contrary to a widespread belief, correlations in the variabilities of neuronal firing rates do not, in general, limit the increase in coding accuracy provided by using large populations of encoding neurons. Furthermore, in some cases, but not all, correlations improve the accuracy of a population code.

Introduction

In population coding schemes, the activities of a number of neurons jointly encode the value of a quantity. A frequently touted advantage of population coding is that it suppresses the effects of neuronal variability. The observation of correlations in the trial-to-trial fluctuations of simultaneously recorded neurons (Gawne & Richmond, 1993; Zohary *et al*, 1994; Lee *et al*, 1998) has raised some doubt as to whether this advantage is actually realized in real nervous systems. The dramatic effects of correlated variability can be seen by examining its impact on the average of N neuronal firing rates. When the fluctuations of individual neurons about their mean rates are uncorrelated, the variance of the average decreases like $1/N$ for large N. In contrast, correlated fluctuations cause the variance of the average to approach a fixed limit as the number of neurons increases. While illustrative, this example is not conclusive because the value of an encoded quantity can be extracted from a population of neurons by methods that do not require averaging their firing rates. Statements in the literature suggest that correlated variability can either decrease or increase (Snippe & Koenderink, 1992; Shadlen & Newsome, 1994; Shadlen *et al*, 1996; Reike et al., 1996; Oram et al., 1997; Lee *et al*, 1998) the accuracy of a population code. The purpose of this paper is to clarify this situation by addressing two questions: 1) Does correlation necessarily increase or decrease the accuracy with which the value of an encoded quantity can be extracted from a population of N neurons? 2) Does this accuracy approach a fixed limit as N increases?

This issue of correlated variability was first addressed by Johnson (1980b), who discussed circumstances under which correlation is either helpful or harmful for discrimination. Snippe and Koenderink (1992) studied the effect of correlated variability on optimal linear discrimination and also found some cases in which correlation improved discrimination and others where discrimination was degraded by correlation. We will study the effects of correlation on population coding accuracy by computing the Fisher information (Cox & Hinckley, 1974; Paradiso, 1988; Seung & Sompolinsky, 1993). The inverse of the Fisher information is the minimum averaged squared error for any unbiased estimator of an encoded variable. It thus sets a limit on the accuracy with which a population code can be read out by an unbiased decoding method.

Two simple examples illustrate the subtleties involved in analyzing the effects of correlation. Consider a set of N neurons with firing rates r_i, for $i = 1, 2, \ldots N$, which have mean values f_i, identical variances σ^2, and correlated variabilities so that

$$\langle (r_i - f_i)(r_j - f_j) \rangle = \sigma^2 \left[\delta_{ij} + c(1 - \delta_{ij}) \right] \tag{1}$$

with correlation coefficient c satisfying $0 \leq c < 1$. The angle brackets on the left side of this equation denote an average over trials, and $\delta_{ii} = 1$ for all i while

$\delta_{ij} = 0$ if $i \neq j$. In this case, the variance of the average of the rates

$$\overline{R} = \frac{1}{N} \sum_{i=1}^{N} r_i \tag{2}$$

is

$$\sigma_{\overline{R}}^2 = \frac{\sigma^2}{N} [1 + c(N-1)] \,. \tag{3}$$

This illustrates two negative effects of correlation. For fixed N, the variance increases as a function of the degree of correlation c, and beyond $N \approx 1/c$ the variance approaches a fixed limit $\sigma_{\overline{R}}^2 \to c\sigma^2$, as discussed in the opening paragraph. Correlation among the activities of neurons in area MT of monkeys viewing moving random dot displays has been estimated at about $c = 0.1 - 0.2$ (Zohary *et al*, 1994; Shadlen *et al*, 1996). This has led to the suggestion that coding accuracy will not improve for populations of more than about 100 neurons (Shadlen & Newsome, 1994).

The second example may seem a bit contrived but is nevertheless illustrative. Consider the sign-alternating sum

$$\tilde{R} = \frac{1}{N} \sum_{i=1}^{N} (-1)^i r_i \,. \tag{4}$$

The variance of this quantity (for even N) is

$$\sigma_{\tilde{R}}^2 = \frac{\sigma^2}{N} (1 - c) \,. \tag{5}$$

For this variable, positive correlation always decreases the variance, and the variance is proportional to $1/N$ whether or not correlation is present.

One reason to think that correlation need not always be harmful is that it generally *reduces* the entropy of the variability in a neural population, suggesting that it should therefore *increase* the accuracy of a population code. Our results on population coding generally concur with this entropy analysis. For all of the cases we consider, the lower limit on the averaged squared decoding error provided by the Fisher information is proportional to $1/N$ for large N, similar to the behavior of equation (5) not (3). For additive or multiplicative noise with uniform correlations, the dependence on the degree of correlation c also resembles that of equation (5), and thus correlation improves population coding accuracy. We also consider correlations of limited range for which coding accuracy displays both increasing and decreasing behavior (Snippe & Koenderink, 1992).

2 The Model

We consider a population code in which N neurons respond to a stimulus with firing rates that depend on a variable x that parameterizes some stimulus attribute (Johnson, 1980a&b; Georgopoulos, Schwartz & Kettner, 1986; Paradiso, 1988; Baldi & Heiligenberg, 1988; Snippe & Koenderink, 1992; Seung & Sompolinsky, 1993; Salinas & Abbott, 1994; Snippe, 1996; Sanger, 1996). The activity of neuron i, averaged over trials that use the stimulus x, is $f_i(x)$, and its activity on any given trial is

$$r_i = f_i(x) + \eta_i \, . \tag{6}$$

We interpret this as the number of spikes fired by the neuron over a fixed time period. We do not discuss encoding that involves the fine-scale temporal structure of spike trains. The random terms η_i for $i=1, 2, \ldots, N$ are generated from a Gaussian probability distribution with zero mean and covariance matrix $\mathbf{Q}(x)$. We consider three different models of variability. In the additive noise model (Johnson, 1980b), the covariance matrix is identical to (1),

$$Q_{ij} = \sigma^2[\delta_{ij} + c(1 - \delta_{ij})] \, . \tag{7}$$

For multiplicative noise, the variability in the firing rate is still described by Equation (6), but the covariance matrix is scaled by the average firing rates,

$$Q_{ij}(x) = \sigma^2[\delta_{ij} + c(1 - \delta_{ij})] f_i(x) f_j(x) \, . \tag{8}$$

This produces variances that increase as a function of firing rate, and larger correlations for neurons with overlapping tuning curves, as is seen in the data (Lee *et al*, 1998). We also consider a model in which the correlations can have an even more dramatic dependence on the distance between tuning curves. This is the limited range correlation model (Snippe & Koenderink, 1992) with the correlation matrix written as

$$Q_{ij} = \sigma^2 \rho^{|i-j|} \tag{9}$$

where the parameter ρ (with $0 < \rho < 1$) determines the range of the correlations between different neurons in the population.

We use the notation $\mathbf{Q}$ to denote the matrix with elements Q_{ij}, and $\mathbf{r}$ and $\mathbf{f}(x)$ to denote the vectors of firing rates with elements r_i and $f_i(x)$ respectively. Note that in the additive and limited range cases, $\mathbf{Q}$ does not depend on x, while for multiplicative noise it does.

The average firing rates $\mathbf{f}(x)$ are the tuning curves of the neurons in the population. We imagine the the tuning curves are arranged to cover a range of x values, with different tuning curves localized to different ranges of x. We assume that the coverage is dense and roughly uniform (we define these terms below), but otherwise leave the exact nature of these tuning curves relatively unrestricted.

3 Fisher Information

The Fisher information provides a useful measure of the accuracy of a population code. Through the Cramér-Rao bound, the Fisher information limits the accuracy with which any unbiased decoding scheme can extract an estimate of an encoded quantity from the activity of a population of neurons. The average value of an unbiased estimate is equal to the true value, x, of the encoded quantity, but the estimate will typically differ from x on a trial-to-trial basis. For an unbiased estimate, the average squared decoding error is equal to the variance of these trial-to-trial deviations. The Cramér-Rao bound states that the average squared decoding error for an unbiased estimate is greater than or equal to $1/I_F(x)$, where $I_F(x)$ is the Fisher information. Although, in some cases, an biased decoding scheme may outperform a biased method, no biased estimate can do better than the Cramér-Rao lower bound. To compute the Fisher information, we need to know the conditional probability distribution $P[\mathbf{r}|x]$ which determines the probability that a given response $\mathbf{r}$ is evoked by the stimulus x. The Fisher information is given in terms of $P[\mathbf{r}|x]$ by

$$I_F(x) = -\int d\mathbf{r}\, P[\mathbf{r}|x] \frac{d^2 \log P[\mathbf{r}|x]}{dx^2} \,. \tag{10}$$

The maximum likelihood estimator, which chooses for an estimate the value of x that maximizes $P[\mathbf{r}|x]$, asymptotically saturates the Cramér-Rao bound as $N \to \infty$. Thus, the bound sets a realizable limit making it a good measure of the accuracy of a population code (see Paradiso, 1988; Seung & Sompolinsky, 1993; and Pouget *et al*, 1998 for discussions of the use of ML techniques and Fisher information for population codes in the absence of correlation). The psychophysical measure of discriminability d' that quantifies how accurately discriminations can be made between two slightly different values x and $x+\Delta x$ based on $\mathbf{r}$ is related to the Fisher information by the formula

$$d' = \Delta x \sqrt{I_F(x)} \,. \tag{11}$$

The larger the Fisher information, the better the discriminability and the smaller the minimum unbiased decoding error.

When the random η terms are generated from a Gaussian probability distribution as discussed above,

$$P[\mathbf{r}|x] = \frac{1}{\sqrt{(2\pi)^N \det \mathbf{Q}(x)}} \exp\left[-\frac{1}{2}[\mathbf{r} - \mathbf{f}(x)]^T \mathbf{Q}^{-1}(x)[\mathbf{r} - \mathbf{f}(x)]\right] \,. \tag{12}$$

This equation does not give zero probability for negative firing rates, but we assume that the means and variances are such that this has a small effect. Substituting (12) into Equation (10), we find (see, for example, Kay, 1993)

$$I_F(x) = \mathbf{f}'(x)^T \mathbf{Q}^{-1}(x)\mathbf{f}'(x) + \frac{1}{2}\mathrm{Tr}\left[\mathbf{Q}'(x)\mathbf{Q}^{-1}(x)\mathbf{Q}'(x)\mathbf{Q}^{-1}(x)\right] \tag{13}$$

where Tr stands for the trace operation,

$$\mathbf{Q}'(x) = \frac{d\mathbf{Q}(x)}{dx} \quad \text{and} \quad \mathbf{f}'(x) = \frac{d\mathbf{f}(x)}{dx}\,. \tag{14}$$

When $\mathbf{Q}$ is independent of x, as it is for additive noise and limited range correlations, this reduces to

$$I_F(x) = \mathbf{f}'(x)^T \mathbf{Q}^{-1} \mathbf{f}'(x)\,. \tag{15}$$

Equations (15) and (13) are the basis for all our results. To apply them we need the inverses of the covariance matrices which are, in the additive case

$$[\mathbf{Q}^{-1}]_{ij} = \frac{\delta_{ij}(Nc+1-c) - c}{\sigma^2(1-c)(Nc+1-c)} \tag{16}$$

in the multiplicative case,

$$[\mathbf{Q}^{-1}(x)]_{ij} = \frac{\delta_{ij}(Nc+1-c) - c}{f_i(x) f_j(x) \sigma^2(1-c)(Nc+1-c)} \tag{17}$$

and, in the case of correlations with limited range

$$[\mathbf{Q}^{-1}]_{ij} = \frac{1+\rho^2}{\sigma^2(1-\rho^2)} \left[\delta_{ij} - \frac{\rho}{1+\rho^2} \left(\delta_{i+1,j} + \delta_{i-1,j} \right) \right]\,. \tag{18}$$

4 Results

4.1 Additive Noise

The Fisher information in the additive case is computed by substituting the correlation matrix (7) into Equation (15) and doing the necessary algebra. The result depends on two sums,

$$F_1(x) = \frac{1}{N} \sum_i \left(f_i'(x) \right)^2 \quad \text{and} \quad F_2(x) = \left(\frac{1}{N} \sum_i f_i'(x) \right)^2\,. \tag{19}$$

These have been scaled to be of order one for the case of uniform tuning curve placement. In terms of these quantities,

$$I_F(x) = \frac{cN^2[F_1(x) - F_2(x)] + (1-c)NF_1(x)}{\sigma^2(1-c)(Nc+1-c)} \tag{20}$$

As N tends to infinity,

$$I_F(x) \to \frac{N[F_1(x) - F_2(x)]}{\sigma^2(1-c)} \tag{21}$$

which grows with N and c provided that $F_1(x) - F_2(x) > 0$. Note that, aside from the factor of $F_1(x) - F_2(x)$, the inverse of this expression matches the variance of Equation (5). For large N, a uniform array of tuning curves should generate functions $F_1(x)$ and $F_2(x)$ that are insensitive to the values of either x or N (indeed, this is our definition of uniform tuning curve placement). Tuning curve arrays that are symmetric with respect to the sign flip $x \to -x$ (that is, for every neuron with tuning curve $f(x)$ there is another neuron with tuning curve $f(-x)$) have $F_2(x) = 0$.

The inverse of the Fisher information for the case $F_2 = 0$ is plotted in Figure 1A. This shows that the Cramér-Rao bound decreases as a function of the degree of correlation in the additive noise case. Note that the Fisher information diverges, and the minimal error goes to zero as c approaches one. As $c \to 1$, any slight difference in the tuning curves can be exploited to calculate the noise exactly and remove it.

Zhang and Sejnowski (1998) have noted an interesting scaling property of the Fisher information that also appears in our results. They considered the simultaneous encoding of D variables by a population of neurons, and studied the effect of changing tuning curve width on encoding accuracy. If the widths of the tuning curves of the encoding population are scaled by a parameter w, we expect F_1 to scale like w^{D-2}. The factor of w^D reflects the number of responding neurons, while the factor w^{-2} is due to the squared derivative. For simplicity, we take $F_2 = 0$. Then, the Fisher information satisfies $I_F \propto N w^{D-2}/\sigma^2$ in agreement with the results of Zhang and Sejnowski (1998).

The Fisher information we have computed increases as a function of c and N unless $F_1(x) - F_2(x) = 0$ or $F_1(x) - F_2(x) \to 0$ for large N. By the Cauchy-Schwartz inequality $F_1(x) \geq F_2(x)$. For large N, $F_1(x) - F_2(x)$ could go to zero if both $F_1(x) \to 0$ and $F_2(x) \to 0$. We define the tuning curve coverage as being dense if $F_1(x) \not\to 0$ for any x, since this implies that as more neurons are added, a fixed fraction respond significantly to x. By the condition for equality in the Cauchy-Schwartz inequality, the other alternative, $F_1(x) = F_2(x)$, requires $f_i'(x)$ to be independent of i or, equivalently,

$$f_i(x) = p(x) + q_i \tag{22}$$

for any function p and numbers q_i. Thus, the Fisher information will fail to grow as a function of c and N only if there is an additive separation of dependency between the value x and the index i. This means that correlation is only harmful in the case where all the neurons share the same tuning dependence on x. This is not normally the case since neurons almost always have some variability in their stimulus preferences. Of course, we must assume that the mechanism that reads out the encoded quantity takes advantage of this variability and does not simply perform an averaging operation.

4.2 Multiplicative Noise

When $\mathbf{Q}(x)$ is given by Equation (8), the Fisher information defined by Equation (13) depends on the logarithmic derivatives of the average firing rate tuning curves

$$g_i'(x) = \frac{1}{f_i(x)}\frac{df_i(x)}{dx} = \frac{d\ln f_i(x)}{dx}\,. \tag{23}$$

In particular, it depends on the sums

$$G_1(x) = \frac{1}{N}\sum_i \left(g_i'(x)\right)^2 \quad \text{and} \quad G_2(x) = \left(\frac{1}{N}\sum_i g_i'(x)\right)^2 \tag{24}$$

and is given by

$$\begin{aligned} I_F(x) &= \frac{cN^2[[G_1(x) - G_2(x)] + (1-c)NG_1}{\sigma^2(1-c)(Nc+1-c)} \\ &+ \frac{[N^2c(2-c) + 2N(1-c)^2]G_1(x) - c^2N^2G_2(x)}{(1-c)(Nc+1-c)}\,. \end{aligned} \tag{25}$$

For large N, this approaches the limit

$$I_F(x) \to \frac{N[G_1(x) - G_2(x)]}{\sigma^2(1-c)} + \frac{N[(2-c)G_1(x) - cG_2(x)]}{(1-c)}\,. \tag{26}$$

The Fisher information for multiplicative noise contains one term that depends on the noise variance σ^2, and another term that, surprisingly, is independent of σ^2. This second term arises because, with multiplicative noise, the encoded variable can be estimated from second-order quantities, not merely from measurements of the firing rates themselves.

The Fisher information of Equation (26) is proportional to N and is an increasing function of c provided that $G_1(x) > G_2(x)$. Since $G_1(x) \geq G_2(x)$ by the Cauchy-Schwartz inequality, the only way to modify this behavior is if $G_1(x) = G_2(x)$. This condition only holds if $g_i'(x)$ is independent of i or, equivalently, if

$$f_i(x) = p(x)q_i + r(x) + s_i \tag{27}$$

for any functions p and r and numbers q_i and s_i. This is multiplicative separability rather than the additive separability of Equation (22). As in the case of additive noise, the Fisher information with multiplicative noise increases with correlation and grows linearly with N unless a contrived set of tuning curves is used. An example of the decrease in the minimal unbiased decoding error as a function of c for the case of multiplicative noise is plotted in Figure 1B.

4.3 Limited Range Correlations

In both of the cases we have studied thus far, the accuracy of the population code, quantified by the Fisher information, increase both as a function of N and c. While the linear dependence of the Fisher information on N appears quite general, there are cases in which introducing correlation decreases rather than increases I_F (Johnson, 1980b; Snippe & Koenderink, 1992). One example is provided by the limited range correlations described by the matrix of Equation (9). The Fisher information for this case is

$$I_F(x) = \frac{N}{\sigma^2}\frac{(1-\rho)^2}{1-\rho^2} + \frac{1}{N\sigma^2}\frac{\rho}{1-\rho^2}F_3(x) \tag{28}$$

where F_1 is as given above and (provided that x is away from the edges of the range covered by the population tuning curves)

$$F_3(x) = N\sum_{i=1}^{N}\left(f'_{i+1}(x) - f'_i(x)\right)^2 . \tag{29}$$

Note the difference between $F_1(x)$ and $F_3(x)$ in their scaling with respect to N. As N gets large, the tuning functions of neighboring cells get very close, and so the difference between their derivatives gets small.

The Fisher information for this case depends on the range of the correlations, which in turn is determined by the value of ρ. The range is best expressed in terms of a correlation length L. If Δ is the spacing between the peaks of adjacent tuning curves,

$$\rho = \exp(-\Delta/L) . \tag{30}$$

The first term in Equation (28) is a decreasing function of ρ and hence L, while the second term has the opposite dependence. For a fixed N, the first term dominates for small L, and the second dominates for large L. The Fisher information is thus a nonmonotonic function of the correlation length. Typical behavior is shown in Figure 1C. The second term scales like $1/N$, and so as N increases it becomes insignificant. For any fixed $\rho < 1$, the overall Fisher information scales with N, although, for instance, if ρ tends to 1 as $c^{1/N}$, for $0 < c < 1$, then the Fisher information tends to a constant with N and does not grow.

5 Discussion

We have studied how correlations between the activities of neurons within a coding population affect the accuracy with which an encoded quantity can be determined or discriminated (Johnson, 1980b). We find that, generically, correlations between units do not prevent the Fisher information from growing linearly with the number of encoding neurons, and correlations can either improve or degrade decoding accuracy depending on the form of the correlation matrix. Only in the limit as the correlations get very close to 1 can this behavior change in some circumstances. Since our results are based on the Fisher information, they apply only to unbiased estimates. However, biased estimates would presumably be used only to further improve accuracy, and thus the increase in accuracy with N would not be eliminated by using a biased estimate. There are nevertheless possible lacunæ in our analysis. We only considered relatively simple noise models. We also used noise with Gaussian statistics. Poisson noise would be an obvious alternative and would entail slightly different calculations. Finally, we did not consider the computational complexity or biological implementation of the optimal decoding algorithms, although a good point of departure would be Pouget *et al* (1998)'s work showing how to perform ML inference using a recurrent network in the case without correlations.

The most relevant requirement for retaining the improved accuracy provided by large populations of encoding neurons is that the neurons should have different selectivities to the quantity they are jointly encoding. In particular, their tuning curves must not be additively or multiplicatively separable, and therefore, they should not be combined by simple, unweighted, averaging. Tuning functions that are commonly adopted in modeling work and seen in biological systems do not appear to have these problematic features.

Acknowledgments

We are grateful to Peter Latham, Alex Pouget, Sebastian Seung, Michael Shadlen and Haim Sompolinsky for discussions and comments. Research supported by the National Science Foundation (DMS-9503261) and the WM Keck Foundation for LFA, and by the National Institute of Mental Health (MH-55541) and the National Science Foundation (IBM-9634339) for PD.

<u>References</u>

Baldi, P & Heiligenberg, W (1988) How sensory maps could enhance resolution through ordered arrangements of broadly tuned receivers. *Biological Cybernetics* **59**:313-318.

Cox, DR & Hinckley, DV (1974) *Theoretical Statistics.* London:Chapman & Hall.

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

Georgopoulos, AP, Schwartz, AB & Kettner, RE (1986) Neuronal population coding of movement direction. *Science* **243**:1416-1419.

Gawne, TJ & Richmond, BJ (1993) How independent are the messages carried by adjacent inferior temporal cortical neurons? *Journal of Neuroscience* **13**:2758-2771.

Johnson, KO (1980a) Sensory discrimination: decision process. *Journal of Neurophysiology* **43**:1771-1792.

Johnson, KO (1980b) Sensory discrimination: neural processes preceding discrimination decision. *Journal of Neurophysiology* **43**:1793-1815.

Kay, SM (1993) *Fundamentals of Statistical Signal Processing: Estimation Theory.* Englewood Cliffs, NJ:Prentice-Hall.

Lee, D, Port, NL, Kruse, W & Georgopoulos, AP (1998) Variability and correlated noise in the discharge of neurons in motor and parietal areas of the primate cortex. *Journal of Neuroscience* **18**:1161-1170.

Oram, MW, Foldiak, P, Perett, DI & Sengpiel, F (1997) The 'ideal homunculus': decoding neural population signals. Trends in Neuroscience (in press).

Paradiso, MA (1988) A theory for the use of visual orientation information which exploits the columnar structure of striate cortex. *Biological Cybernetics* **58**:35-49.

Pouget A, Zhang K, Deneve S & Latham P (1998) Statistically Efficient Estimation Using Population Coding. *Neural Computation* **10**:2-8.

Reike, F, Warland, D, de Ruyter van Steveninck, R & Bialek, W (1996) *Spikes: Exploring the Neural Code.* Cambridge MA:MIT Press.

Salinas, E & Abbott, LF (1994) Vector reconstruction from firing rates. *Journal of Computational Neuroscience* **1**:89-107.

Sanger, TD (1996) Probability density estimation for the interpretation of neural population codes. *Journal of Neurophysiology* **76**:2790-2793.

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

Shadlen, MN, Britten, KH, Newsome, WT & Movshon, JA (1996) A computational analysis of the relationship between neuronal and behavioral responses to visual motion. *Journal of Neuroscience* **16**:1486-1510.

Shadlen, MN & Newsome, WT (1994) Noise, neural codes and cortical organization. *Current Opinion in Neurobiology* **4**:569-579.

Snippe, HP (1996) Parameter extraction from population codes: a critical assessment. *Neural Computation* **8**:511-530.

Snippe, HP & Koenderink, JJ (1992) Information in channel-coded systems: correlated receivers. *Biological Cybernetics* **67**:183-190.

Zhang K & Sejnowski T (1998) Neural tuning: to sharpen of broaden? *Neural Computation* (in press).

Zohary E, Shadlen MN, & Newsome, WT (1994) Correlated neuronal discharge rate and its implications for psychophysical performance. *Nature* **370**:140-143.

Figure 1

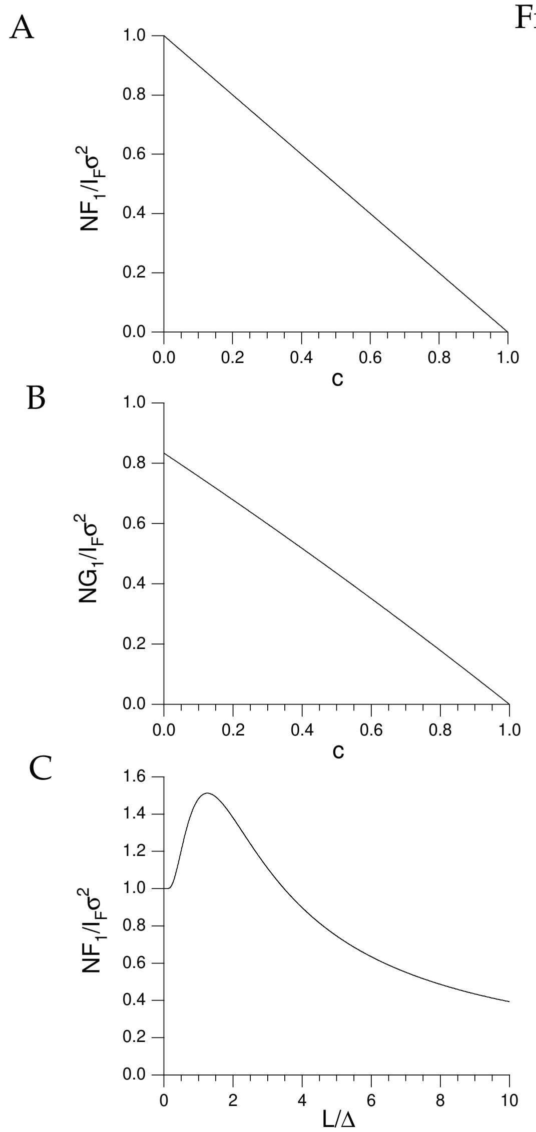

Figure 1: Dependence of the minimal unbiased decoding error on the degree and length of correlation. The inverse of the Fisher information, scaled by an appropriate factor, is plotted. A) The inverse of the Fisher information for the additive noise model with $F_2 = 0$ as a function of the degree of correlation c. The minimal unbiased decoding error decreases linearly with the degree of correlation. B) The inverse of the Fisher information for the multiplicative noise model with $G_2 = 0$ as a function of the degree of correlation c. The minimal unbiased decoding error decreases with c. For this example, we set $\sigma^2 = 0.1$. C) The inverse of the Fisher information as a function of the correlation length L in the limited range model. The minimal unbiased decoding error first increases and then decreases as L increases. L is expressed in units of the spacing Δ between successive tuning curves. For this example, we set $F_3 = N^2 F_1/2$.