

# An Information Theoretic Approach to the Contributions of the Firing Rates and the Correlations Between the Firing of Neurons

Edmund T. Rolls, Leonardo Franco, Nicholas C. Aggelopoulos, and Steven Reece

*Department of Experimental Psychology, Oxford University, Oxford OX1 3UD, United Kingdom*

Submitted 27 November 2002; accepted in final form 10 January 2003

**Rolls, Edmund T., Leonardo Franco, Nicholas C. Aggelopoulos, and Steven Reece.** An information theoretic approach to the contributions of the firing rates and the correlations between the firing of neurons. *J Neurophysiol* 89: 2810–2822, 2003. First published January 15, 2003; 10.1152/jn.01070.2002. To analyze the extent to which populations of neurons encode information in the numbers of spikes each neuron emits or in the relative time of firing of the different neurons that might reflect synchronization, we developed and analyzed the performance of an information theoretic approach. The formula quantifies the corrections to the instantaneous information rate that result from correlations in spike emission between pairs of neurons. We showed how these cross-cell terms can be separated from the correlations that occur between the spikes emitted by each neuron, the auto-cell terms in the information rate expansion. We also described a method to test whether the estimate of the amount of information contributed by stimulus-dependent synchronization is significant. With simulated data, we show that the approach can separate information arising from the number of spikes emitted by each neuron from the redundancy that can arise if neurons have common inputs and from the synergy that can arise if cells have stimulus-dependent synchronization. The usefulness of the approach is also demonstrated by showing how it helps to interpret the encoding shown by neurons in the primate inferior temporal visual cortex. When applied to a sample dataset of simultaneously recorded inferior temporal cortex neurons, the algorithm showed that most of the information is available in the number of spikes emitted by each cell; that there is typically just a small degree (approximately 12%) of redundancy between simultaneously recorded inferior temporal cortex (IT) neurons; and that there is very little gain of information that arises from stimulus-dependent synchronization effects in these neurons.

## INTRODUCTION

To analyze how neurons encode information about stimuli or other events, it is useful to apply information theory, because this allows the contributions of different possible factors (such as the number of spikes vs. the relative timing of spikes from different cells) to be measured quantitatively and with the same metric (Rolls and Deco 2002). Simultaneously recorded neurons sometimes show cross-correlations in their firing, i.e., the firing of one cell is systematically related to the firing of the other cell. One example of this is neuronal response synchronization. The cross-correlation, to be defined below, shows the time difference between the cells at which the systematic relation appears. A significant peak or trough in the cross-correlation function could reveal a synaptic connection from one cell to the other, a common input to each of the cells, or

any of a considerable number of other possibilities. If the synchronization occurred for only some of the stimuli, the presence of the significant cross-correlation for only those stimuli could provide additional evidence separate from any information in the firing rate of the neurons about which stimulus had been shown. Information theory in principle provides a way of quantitatively assessing the relative contributions from these two types of encoding by expressing what can be learned from each type of encoding in the same units—bits of information. An information theory-based approach to this has been developed by Panzeri et al. (1999).

When applying information theory to the responses of two or more simultaneously recorded neurons, the number of possible combinations of the relative times of the spikes of the different cells becomes very large. That is, the dimensionality of the space that must be filled adequately with real neurophysiological data to obtain reliable estimates of the information becomes so large that the information estimates become unreliable, and in fact, are biased upward. Even bias correction measures (Panzeri and Treves 1996; Treves and Panzeri 1995) cannot completely correct for this amount of undersampling. In this situation, the dimensionality of the space in which the neuronal responses are measured must be reduced. A recent approach to this issue has been to simply count the number of spikes in a single short time window from the simultaneously recorded cells and to use these spike counts to estimate the information that is contributed by different factors, including factors such as synchronization of the spikes of different cells. This is the approach taken by Panzeri et al. (1999). The new contributions of the present paper are as follows. First, we extended the previous approach by describing a method for calculating and separating the effects arising from cross-correlations between the spikes of the simultaneously recorded cells from the effects produced by autocorrelations arising from the spikes produced by each cell interacting with its own spikes. Second, we introduced a method that allows the statistical significance of the synchronization-related information that is measured to be quantified, which has not been available previously (Panzeri et al. 1999). Third, having shown how these cross- and auto-cell terms can be separated and calculated, we tested the whole algorithm with simulated neuronal data and showed how the different terms can be interpreted. We note that the cross- or between-cell terms were not separated from the auto- or within-cell terms in the simulations

Address reprint requests to E. T. Rolls (E-mail: Edmund.Rolls@psy.ox.ac.uk); web: www.cns.ox.ac.uk

The costs of publication of this article were defrayed in part by the payment of page charges. The article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

shown by Panzeri et al. (1999), and that separating these terms greatly enhances the interpretation of what can be measured with this approach to neuronal encoding. Fourth, we showed results of the application of the algorithm to real neuronal data from the primate inferior temporal visual cortex and showed how particular effects evident with real neuronal data can arise.

The information theoretic approach used and developed in this paper measured the different contributions to the total information that arise from stimulus-dependent and stimulus-independent cross-cell correlations in contrast to earlier methods (Gawne and Richmond 1993). Stimulus-dependent cross-cell information can potentially provide evidence about which stimulus was presented by reflecting information that can be extracted from the co-modulation of the spikes of two neurons. A case of this might be if on a trial by trial basis, for one stimulus two cells might both give many (or few) spikes, whereas for another stimulus the spike firings might be uncorrelated. Stimulus-independent cross-cell information might reflect for example a correlation between the firing rates of two cells independently of which stimulus was shown, and would in this case typically introduce redundancy into the encoding. In addition, there are within-cell information measures provided by the approach. One measure, the “stimulus-independent auto-term,” reflects the neuronal response variability. If this term is high, the implication is that each cell encodes little information about the stimuli.

METHODS

*Measuring the information available from simultaneously recorded cells*

It is first necessary to describe the approach taken by Panzeri et al. (1999), which limits the dimensionality problem by taking short time epochs for the information analysis, in which low numbers of spikes, typically 0–2, are likely to occur from each neuron. In this case, in which at most two spikes are emitted from the population, the response probabilities can be calculated in terms of pairwise correlations. These response probabilities are inserted into the Shannon information formula shown in Eq.1 to obtain expressions quantifying the impact of the pairwise correlations on the information  $I(t)$  transmitted in a short time  $t$  by groups of spiking neurons

$$I(t) = \sum_{s \in S} \sum_{\mathbf{r}} P(s, \mathbf{r}) \log_2 \frac{P(s, \mathbf{r})}{P(s)P(\mathbf{r})} \quad (1)$$

where  $\mathbf{r}$  is the firing rate response vector comprised of the number of spikes emitted by each of the simultaneously recorded cells in the population in the short time  $t$ , and  $P(s, \mathbf{r})$  refers to the joint probability distribution of stimuli with their respective neuronal response vectors. The firing rate response vector  $\mathbf{r}$  for a single trial consists of the number of spikes  $n_i$  emitted by each cell  $i$  in a short time  $t$ .

The approach consists then, in the short timescale limit, of using the first ( $I_1$ ) and second ( $I_{11}$ ) information derivatives to describe the information  $I(t)$  available in the short time  $t$

$$I(t) = tI_1 + \frac{t^2}{2} I_{11} \quad (2)$$

(The 0th order, time-independent term is 0, because no information can be transmitted by the neurons in a time window of 0 length.

Higher order terms are also excluded because they become negligible in a short time window.) We develop below the expansion of Eq.2 in terms of two types of correlation, which we introduce first. These two correlations are as follows.

CORRELATIONS IN THE NEURONAL RESPONSE VARIABILITY FROM THE AVERAGE TO EACH STIMULUS (SOMETIMES CALLED “NOISE” CORRELATIONS)  $\gamma$ . This type of correlation is high if on one trial for a given stimulus  $S$  the spike rates of two neurons being considered are higher than the average to that stimulus, whereas on another trial to the same stimulus the rates of both neurons are lower than the average to that stimulus. Neurophysiologically, this type of effect could be produced by cross-coupling between the neurons. It is called a “noise” correlation (Gawne and Richmond 1993; Shadlen and Newsome 1994, 1998) because it reflects the trial by trial co-variation in the responses of the neurons, but it is also called the “scaled cross-correlation density” (Aertsen et al. 1989; Panzeri et al. 1999). More formally, where the two cells are indexed by  $i$  and  $j$ ,  $\gamma_{ij}(s)$  (for  $i \neq j$ ) is the fraction of coincidences above (or below) that expected from uncorrelated responses, relative to the number of coincidences in the uncorrelated case [which is  $\bar{n}_i(s)\bar{n}_j(s)$ , the bar denoting the average across-trials belonging to stimulus  $S$ , where  $n_i(s)$  is the number of spikes emitted by cell  $i$  to stimulus  $s$  on a given trial]

$$\gamma_{ij}(s) = \frac{\overline{n_i(s)n_j(s)}}{(\bar{n}_i(s)\bar{n}_j(s))} - 1 \quad (3)$$

It can vary from  $-1$  to  $\infty$ ; negative values of  $\gamma_{ij}(s)$  indicate anticorrelation, whereas positive values of  $\gamma_{ij}(s)$  indicate correlation.<sup>1</sup>  $\gamma_{ij}(s)$  can be thought of as the amount of trial by trial concurrent firing of the cells  $i$  and  $j$  compared with that expected in the uncorrelated case.

We will be interested to measure the effects of synchronization between cells in contributing to the information available from simultaneously recorded cells in the case where  $\gamma_{ij}(s)$  is different for different stimuli. Cells that are synchronized will tend to have a positive value of  $\gamma_{ij}(s)$ , as shown in Fig. 1. Although the  $\gamma_{ij}(s)$  measure utilizes the numbers of spikes from the different neurons and thus reflects rate co-modulation, this will almost always with real neurons (as contrasted with possible artificial scenarios) capture any synchronization that is present. This is because in a sufficiently short time window (and the information measures are for this reason plotted in the figures in different time windows in the range 0–100 ms), in the unlikely event that cell  $i$  fires if cell  $j$  also fires, this is likely to reflect synchronization. (We note that in the general case, cells with synchronized spikes will show co-modulation of the number of spikes on single trials in short time windows. However, if a scenario ever occurred in which there was stimulus-dependent synchronization but no co-modulation of the number of spikes obtained in a short time window, then  $\gamma_{ij}(s)$ , and the algorithm described here would not detect it.) An advantage of the measure  $\gamma_{ij}(s)$  of the covariation between the number of spikes of different cells in a short time window is that it can be positive independently of any particular time lag in the cross-corrologram and will be negative if the two cells anti-covary (e.g., reflecting inhibition of 1 cell by the other).

<sup>1</sup>  $\gamma_{ij}(s)$  is an alternative, which produces a more compact information analysis, to the neuronal cross-correlation based on the Pearson correlation coefficient  $\rho_{ij}(s)$ , which normalizes the number of coincidences above independence to the standard deviation of the number of coincidences expected if the cells were independent. The normalization used by the Pearson correlation coefficient has the advantage that it quantifies the strength of correlations between neurons in a rate-independent way. For the information analysis, it is more convenient to use the scaled correlation density  $\gamma_{ij}(s)$  than the Pearson correlation coefficient, because of the compactness of the resulting formulation, and because of its scaling properties for small  $t$ .  $\gamma_{ij}(s)$  remains finite as  $t \rightarrow 0$ , thus by using this measure we can keep the  $t$  expansion of the information explicit. Keeping the time-dependence of the resulting information components explicit greatly increases the amount of insight obtained from the series expansion (see Panzeri et al, 1999).

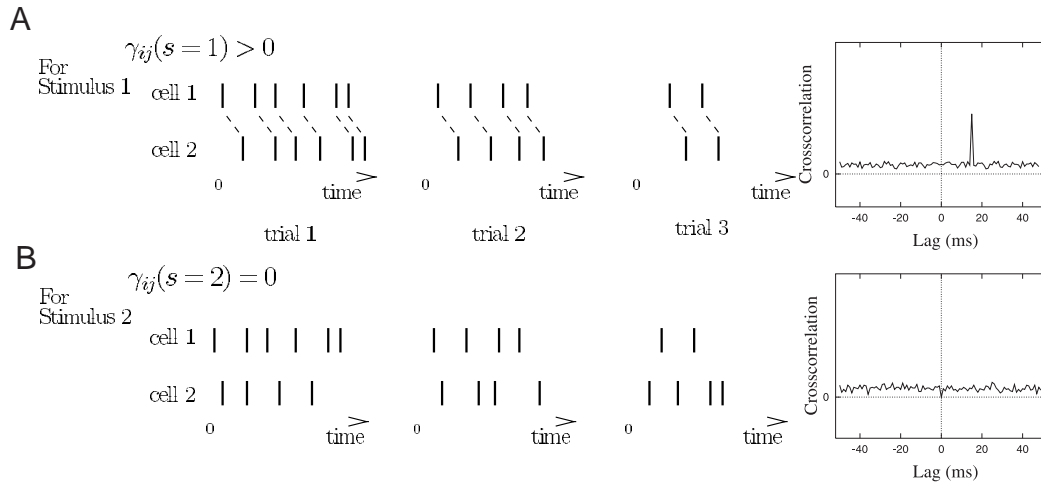


FIG. 1. Illustration to show that  $\gamma_{ij}(s)$  reflects synchronization between cells. Cells have mean rates of 4 spikes per trial to stimulus 1 (A) and to stimulus 2 (B). For stimulus 1, cell 1 fires “noisily” above its average rate to stimulus  $s = 1$  on trial 1, at its average rate on trial 2, and below its average rate on trial 3. Because with stimulus 1 there is a significant synchronization effect between cell 1 and cell 2, cell 2 also has an above average rate on trial 1, and a below average rate on trial 3. This results in a high  $\gamma_{ij}(s)$  for  $s = 1$ . Synchronization is reflected by a peak in the cross-correlation function (cross-correlogram) between the 2 cells, as shown on the right. In contrast, for stimulus 2 shown on the bottom, there is no covariation between the spikes of the 2 cells on a trial by trial basis,  $\gamma_{ij}(s = 2)$  is 0, and the cross-correlogram shows that there is no synchronization (reflected by an absence of peaks in the cross-correlogram).

There is also an autocorrelation term  $\gamma_{ii}(s)$ , which reflects the variability of the number of spikes emitted by a cell to a given stimulus  $s$  from trial to trial. We measure this by

$$\gamma_{ii}(s) = \frac{\overline{n_i(s)n_i(s)} - \overline{n_i(s)}^2}{(\overline{n_i(s)}\overline{n_i(s)})} - 1 \quad (4)$$

The reasons for defining  $\gamma_{ii}$  in this way [including the subtraction of  $\overline{n_i(s)}$  performed to express the result relative to the variance of the spike count that is expected in the independent spikes case] are to quantify correctly the number of spikes from the same cell, as set out by Panzeri et al. (1999).  $\gamma_{ii}(s)$  takes negative values if there is no or small variability from trial to trial for a particular stimulus, and zero for a random variation from trial to trial produced by a Poisson process (in which the variability equals the mean; Panzeri et al. 1999).  $\gamma_{ii}(s)$  can be positive if there is more variability than in an independent spike generation process, and this can be produced in real neurophysiological data if for example there are some trials with exceptionally few spikes.

CORRELATIONS IN THE MEAN RESPONSES OF THE NEURONS ACROSS THE SET OF STIMULI (SOMETIMES CALLED “SIGNAL” CORRELATIONS)  $\nu_{ij}$ .  $\nu_{ij}$  can be thought of as the degree of similarity in the mean response profiles (averaged across-trials) of the cells  $i$  and  $j$  to different stimuli.  $\nu_{ij}$  is sometimes called the “signal” correlation (Gawne and Richmond 1993; Shadlen and Newsome 1994, 1998). It is defined by

$$\nu_{ij} = \frac{\langle \overline{n_i(s)}\overline{n_j(s)} \rangle_s}{\langle \overline{n_i(s)} \rangle_s \langle \overline{n_j(s)} \rangle_s} - 1 = \frac{\langle \overline{r_i(s)}\overline{r_j(s)} \rangle_s}{\langle \overline{r_i(s)} \rangle_s \langle \overline{r_j(s)} \rangle_s} - 1 \quad (5)$$

where  $\overline{r_i(s)}$  is the mean rate of response of cell  $i$  [ $\overline{n_i(s)}$  is the mean number of spikes of cell  $i$  in the interval considered; among  $C$  cells in total] to stimulus  $s$  over all the trials in which that stimulus was present. It can vary from  $-1$  to  $\infty$ . ( $\langle \dots \rangle_s$  indicates the ensemble average over the  $s$  stimuli.)

If  $\nu_{ij}$  is zero, the cells have uncorrelated response profiles to the stimuli, and there is no redundancy. If  $\nu_{ij}$  is either positive or negative, it always reflects redundancy between the cells, as both cases mean that the two cells  $i$  and  $j$  are conveying the same information about the stimuli. For example, if the responses of cell 1 to four stimuli  $a, b, c,$

and  $d$  are 100, 50, 25, and 1 spikes/s, and of cell 2 are 1, 25, 50, and 100 spikes/s, then  $\nu_{ij}$  is negative, and the two cells have redundancy.

The autoterms,  $\nu_{ii}$ , reflect the degree to which a single cell  $i$  responds differently to the different stimuli. If the cell responds equally to all stimuli,  $\nu_{ii}$  is 0 (and the cell will contribute nothing by its firing rate differences to different stimuli to the information about the stimulus). If the cells respond with very different rates to the different stimuli, then  $\nu_{ii}$  is positive. It is simply the reciprocal (minus 1) of the sparseness  $a$  of the representation of the stimuli by a neuron as defined by Treves (1993), Rolls and Treves (1998), and Rolls and Deco (2002).<sup>2</sup>

MEASURING THE INFORMATION IN SHORT TIME PERIODS. As noted above, in the short timescale limit, the first ( $I_t$ ) and second ( $I_{tt}$ ) information derivatives describe the information  $I(t)$  available in the short time  $t$

$$I(t) = tI_t + \frac{t^2}{2} I_{tt} \quad (7)$$

The instantaneous information rate  $I_t$  is<sup>3</sup>

$$I_t = \sum_{i=1}^C \left\langle \overline{r_i(s)} \log_2 \frac{\overline{r_i(s)}}{\langle \overline{r_i(s')} \rangle_{s'}} \right\rangle_s \quad (8)$$

This term is just a simple sum across-the  $C$  cells in the population of the instantaneous information rate of each single cell (Bialek et al. 1991; Skaggs et al. 1993), and thus this term does not take into account any interactions (arising from any of the correlations) between the neurons. Nor does this term reflect the trial by trial vari-

2

$$a = \frac{\left( \sum_{s=1}^S r(s)/S \right)^2}{\sum_{s=1}^S r(s)^2/S} \quad (6)$$

where  $r(s)$  is the firing rate of the neuron to stimulus  $s$  in the set of  $S$  stimuli.

<sup>3</sup> Note that  $s'$  is used in Equations 8 and 9 just as a dummy variable to stand for  $s$ , as there are two summations performed over  $s$ .



ability in the responses of each cell taken individually (which is reflected in  $\gamma_{ii}$ ).

The effect of (pairwise) correlations between the cells begins to be expressed in the second time derivative of the information. The expression for the instantaneous information “acceleration”  $I_{tt}$  (the second time derivative of the information) breaks up into three terms as described by Panzeri et al. (1999)

$$I_{tt} = \frac{1}{\ln 2} \sum_{i=1}^c \sum_{j=1}^c \langle \bar{r}_i(s) \bar{r}_j(s) \rangle_s \left[ v_{ij} + (1 + v_{ij}) \ln \left( \frac{1}{1 + v_{ij}} \right) \right] + \sum_{i=1}^c \sum_{j=1}^c [\langle \bar{r}_i(s) \bar{r}_j(s) \gamma_{ij}(s) \rangle_s] \log_2 \left( \frac{1}{1 + v_{ij}} \right) + \sum_{i=1}^c \sum_{j=1}^c \left\langle \bar{r}_i(s) \bar{r}_j(s) (1 + \gamma_{ij}(s)) \log_2 \left[ \frac{(1 + \gamma_{ij}(s)) \langle \bar{r}_i(s') \bar{r}_j(s') \rangle_{s'}}{\langle \bar{r}_i(s') \bar{r}_j(s') \rangle_{s'} (1 + \gamma_{ij}(s))} \right] \right\rangle_s \quad (9)$$

*Cross- or between-cells terms*

We consider here the case when  $i \neq j$ , which is the interesting case in terms of interactions between the neurons and could result in synergy (perhaps reflecting synchronization) and redundancy.

The first of these terms (which we will call  $I_{tta}$ ) is all that survives of the cross-cell terms if there is no noise correlation at all [i.e., if  $\gamma_{ij}(s) = 0$ ]. Thus the rate component of the between-cell information is given by the sum of  $I_t$  (which is always greater than or equal to 0) and of  $I_{tta}$  (which is always less than or equal to 0).  $I_{tta}$  is thus less than zero whenever  $v_{ij}$  is not equal to 0.  $I_{tta}$  thus reflects the redundancy between the cells introduced by the similarity (or anticorrelation) of their response profiles to different stimuli. The “rate” component of the information we note does still not reflect the trial by trial variability of the responses of each cell to a stimulus (which is reflected by the  $\gamma_{ii}$  to be described below). Nor does it reflect stimulus-dependent cross-correlations (reflected in nonzero  $\gamma_{ij}$ ), which are taken into account in the next two terms in the expansion.

The second term (which we will call  $I_{ttb}$ ) is nonzero if there is some “noise” cross-correlation between the cells independently of which stimulus is present (i.e., if  $\langle \gamma_{ij}(s) \rangle_s$  weighted by the average spike counts, which we denote by  $\langle \gamma_{ij} \rangle_w$  and define as  $\langle \bar{n}_i(s) \bar{n}_j(s) \langle \gamma_{ij}(s) \rangle_s \rangle_s$ , is not equal to 0. For the case  $i \neq j$ , we call the contribution of this term the “stimulus-independent cross-term.” One way to think about  $I_{ttb}$  is in terms of synergy versus redundancy with respect to the case in which the spikes of the two cells vary independently from trial to trial, which is what is expressed in  $I_{tta}$ .  $I_{ttb}$  is a term that corrects relative to the independent spike case for the average stimulus-independent cross-correlation between two cells. If we consider the case when  $v_{ij} > 0$  (i.e., when the response profiles of the cells to the stimuli have some positive correlation), then if  $\langle \gamma_{ij} \rangle_w$  is  $> 0$  (meaning that on a trial by trial basis the numbers of spikes from the two cells tend to be correlated), then this correlation reduces the extent to which the cell can discriminate between the stimuli. Conversely, if the response profiles are anticorrelated ( $v_{ij} < 0$ ), then  $\langle \gamma_{ij} \rangle_w > 0$  makes the responses to different stimuli more separate, and synergy between the cells arises. For the case when  $\langle \gamma_{ij} \rangle_w$  is  $< 0$ , then this trial by trial anticorrelation between the spike numbers synergistically increases the information (reflected in positive  $I_{ttb}$ ) when the response profiles are correlated ( $v_{ij} > 0$ ), and decreases the information when the response profiles are anti-correlated ( $v_{ij} < 0$ ).

Equation 9 shows that the sign of  $I_{ttb}$  is the opposite of the product of the signs of  $\langle \gamma_{ij} \rangle_w$  and  $v_{ij}$ .  $I_{ttb}$  is negative, indicating redundancy, if  $v_{ij}$  and  $\langle \gamma_{ij} \rangle_w$  have the same sign. Conversely,  $I_{ttb}$  is positive, indicating synergy, if  $v_{ij}$  and  $\langle \gamma_{ij} \rangle_w$  have opposite signs. Although when this term is positive this has been thought of as synergy (Panzeri et al. 1999), for the case of negative  $\langle \gamma_{ij} \rangle_w$ , and positive  $v_{ij}$ , the effect might better be thought of as less redundancy than the Poisson case. ( $\langle \gamma_{ij} \rangle_w$

is negative when there is less variability than Poisson.) We thus think of synergy related to  $I_{ttb}$  as arising in the case when  $\langle \gamma_{ij} \rangle_w$  is positive and  $v_{ij}$  is negative.

Furthermore, if  $v_{ij}$  is zero (reflecting uncorrelated response profiles of the cells to the set of different stimuli), the “stimulus-independent noise cross-term” will be zero (independently of the value of  $\langle \gamma_{ij} \rangle_w$ ).

The third component of  $I_{tt}$  (which we will call  $I_{ttc}$ ) represents the contribution of stimulus-dependent correlation, because it becomes nonzero only for stimulus-dependent noise correlations, i.e., where  $\gamma_{ij}(s)$  is different for different stimuli.  $I_{ttc}$  is the term (and the only term in this expansion) in which stimulus-dependent synchronization of neuronal firing would be apparent, if present. In simulated data, this term is positive if two neurons are selected to receive strong synaptic inputs for only some stimuli. In a new procedure introduced in this paper, we provide a method to estimate the statistical significance of the value of this term. The method involves repeated estimates of the value of this stimulus-dependent cross-cell term after different shufflings of the trials randomly within a stimulus in a number of Monte Carlo iterations. Each shuffle removes any stimulus-dependent cross-correlation information provided that there are more than a few trials. The mean and SD of the information estimates calculated from many different shufflings allow a significance test of the actual information value collected without shuffling from the neuron. The statistical test implemented is to provide confidence limits for the values that would be obtained by chance with no possibility for real cross-correlation effects because the trials have been shuffled within a stimulus. The confidence limit was set to 2 SD from the mean value, which, as this is a one-sided test, corresponds to a 97.8% confidence limit. We checked that the information values from the shuffled data were sufficiently normally distributed for this to be a good estimate. We also showed that 30–50 Monte Carlo shufflings and information estimates were sufficient to provide a good and stable estimate of the mean and SD of the information value obtained after shuffling. In the figures in this paper, we show the stimulus-dependent cross-cell information term with the mean value obtained from 50 Monte Carlo shufflings subtracted and 2 SD from this mean value (see Figs. 4, 6, and 7). If the stimulus-dependent information measured without shuffling is greater than the 2 SD limit shown on the figures obtained with the Monte Carlo procedure, the information measured without shuffling is significant at  $P < 0.022$ .

*Auto- or within-cell terms*

For the first term  $I_{tta}$  of the second derivative, if we consider the within-cell part of this, when  $v_{ii}$  is positive, this indicates that the cell discriminates well between the different stimuli. The first term evaluates to a negative contribution to the information in the first derivative, and this reduces the single cell information to what would be expected with independent trial-by-trial variability [i.e., with  $\langle \gamma_{ii}(s) \rangle_s = 0$ , as in a Poisson case]. If the trial by trial variability is different from the Poisson case, the second term corrects for this.

For the second term, ( $I_{ttb}$ ), we considered above the case where  $i \neq j$ , which we can call the contribution of the “stimulus-independent cross-term” to the information. For the case  $i = j$ , the contribution can be called the “stimulus-independent auto-term.” The neuronal response variability is captured in part by this stimulus-independent auto-term. In particular, if the cell has high trial to trial variability, this is reflected in a high  $\langle \gamma_{ii}(s) \rangle_s$  value and becomes a negative contribution to the total information because it is weighted by the log factor of the second term, which in fact weights the negative contribution according to how much the neuron has different mean firing rates to the different stimuli. (This is because  $v_{ii}$  is large and positive if the neuron has different mean firing rates to the different stimuli.) This is thus the main way in which “noise,” i.e., trial to trial variability of the neuronal response decreases the total information available from the cell.

If the data are made more variable by zeroing all the spikes on

TABLE 1. Contributions to the information encoding

Exper.	Number of Cells	Max Rate	Total Information	Rate Information	Stim. Dep. (Cross Correl.)	Stim. Indep (Cross Correl.)	Stim. Indep (Auto Correl.)
bj006	2	28.75	0.35	0.16	0.00	0.00	0.16
bj008	2	33.25	0.12	0.08	0.00	0.00	0.06
bj017a	4	15.00	0.55	0.28	0.00	-0.04	-0.03
bj017b	2	12.75	0.34	0.31	0.00	0.00	0.03
bj019	3	12.25	0.32	0.26	0.00	-0.04	0.09
bj022a	3	22.00	0.65	0.40	0.07	0.00	0.05
bj024b	3	30.08	0.55	0.44	0.00	-0.11	0.22
Mean	2.7	22.01	0.41	0.28	0.01	-0.03	0.08

The average contributions (in bits) of different components of Eqs. 8 and 9 to the information available in a 100-ms time window from seven sets of simultaneously recorded inferior temporal cortex neurons when shown five stimuli effective for the cells. The column labelled "total information" is all of the information from the Taylor expansion and includes in addition to the other information terms in the Table also the stimulus-dependent auto term. The average firing rate to the most effective stimulus of the most responsive cell in each experiment is shown in the column labeled "Max Rate."

occasional trials, this auto-part of  $I_{\text{tb}}$  becomes negative, thus making a negative contribution to the total information. For Poisson simulated spike trains, the "stimulus-independent auto-term" may be positive if just a few trials are utilized. However, for real neuronal data, in some recordings, this term can be negative, as shown in Table 1. The effect is due to the fact that on some trials, especially in short time windows, real neuronal data has no spikes (in cases where this would not be predicted if the real neuronal firing were Poisson based), as shown by Treves et al. (1999). Indeed, the "stimulus-independent auto-term" becomes negative in simulated data if the spikes on occasional trials are zeroed. We note that  $v_{ii}$  is generally positive definite for real data, so that if the "stimulus-independent auto-term" is negative as with real data, then  $\langle \gamma_{ii}(s) \rangle_s$  must be positive, reflecting more trial by trial variability in the neuronal response than occurs with Poisson data.

$I_{\text{tc}}$  has a "stimulus-dependent auto-term" for the case when  $i = j$  in  $\gamma_{ij}(s)$ . This term is subtracted out by the Monte Carlo procedure, but it can be estimated by subtracting the sum of the other components of the information (rate, stimulus-independent, and stimulus-dependent-cross) from the total information. This term is normally close to zero, both in simulations and in real data. There is a special case in which the stimulus-dependent auto-term could be positive. This would arise for example if the trial by trial variability to a given stimulus,  $\gamma_{ii}(s)$ , was different for different stimuli. If the brain could measure this trial by trial variability and found that it was large, this might give information about which stimulus was shown. However, it is not clear how such a measurement could be implemented in the brain.

We note that the "total information" shown on the graphs is the total information from the full expansion, that is  $I_{\text{t}} + (r^2/2)I_{\text{tt}}$  as shown in Eq. 2, and that in practice the stimulus-dependent auto-term generally makes very little contribution to the total information. We also note that the bias correction procedure described by Panzeri et al. (1999) was applied as part of the information analysis used throughout this paper.

### Neurophysiological methods

To examine how this approach operates with examples of real neuronal data, we made recordings from the macaque inferior temporal visual cortex while a set of five visual images (or a set of 20 images in 2 experiments not included in Table 1) was being shown in a visual fixation task. The images included objects and faces, which are known to activate some inferior temporal cortex neurons. Twenty trials of data for each stimulus, presented for 500 ms, were acquired. The methods and the types of neuron analyzed were similar to those described by Rolls et al. (1997), except that simultaneous recordings from small sets of single neurons were made with two to four separate independently movable microelectrodes (using an Alpha-Omega recording system). Because the images were of objects and faces, the visual system had to bind together the component features of the

stimuli to produce the normal neuronal response to the stimuli. It has been shown that spatially rearranging the features in such stimuli leads the neurons that normally respond to these stimuli to fail to respond (Rolls et al. 1994). The recording system (Neuralynx) filtered and amplified the signal and stored spike waveforms that were later sorted and cluster cut off-line using the Datawave Discovery software. The neurophysiological methods used here, and the recording region (in the anterior part of the ventral lip of the superior temporal sulcus in areas TEa and TE<sub>m</sub>, in area TE, or in the cortex deep in the superior temporal sulcus) have been described in detail by Booth and Rolls (1998). All procedures, including preparative and subsequent ones, were carried out in accordance with the National Institutes of Health Guide for the Care and Use of Laboratory Animals, the guidelines of The Society for Neuroscience, and were licensed under the UK Animals (Scientific Procedures) Act 1986.

## RESULTS

### Validation and analysis of the approach using simulated neuronal data

To validate and further analyze the approach, we measured the values of the different terms with simulated data.<sup>4</sup> This is the first description of the performance of the algorithm with the cross-cell and auto-cell terms separated out on simulated data.

First, we considered the case of independent spike trains in two simulated neurons tested with 20 stimuli with 20 trials. This low number of trials per stimulus was investigated because this is in the range of what can be obtained when recording from real neurons in primates (Rolls et al. 2002). Information in the firing rates was implemented by setting the average firing rates of the Poisson-based spike generation process to those obtained for a set of 20 stimuli in a real neurophysiological experiment when the recordings were from

<sup>4</sup> The program used to perform the information theoretic analysis, corrinfo3, implemented the algorithm described by Panzeri et al. (1999). The original program used by Panzeri et al. (1999) was written in C, but was rewritten in Matlab by Drs. S. Panzeri (University of Newcastle, UK) and R. S. Petersen (SISSA, Trieste, Italy), used for research on the rat somatosensory cortex (Panzeri, Petersen, Schultz, Lebedev & Diamond 2001, Petersen, Panzeri & Diamond 2001), and kindly made available to us. We, in the research described in this paper, developed this Matlab code to separate the auto- and cross-cell terms (in a different way to that used by Panzeri et al. (2001) and Petersen et al. (2001)); incorporated the Monte Carlo procedure which allows the statistical significance of the cross-cell stimulus-dependent term to be evaluated; and evaluated the performance of the algorithm using simulated data, as described in this paper.

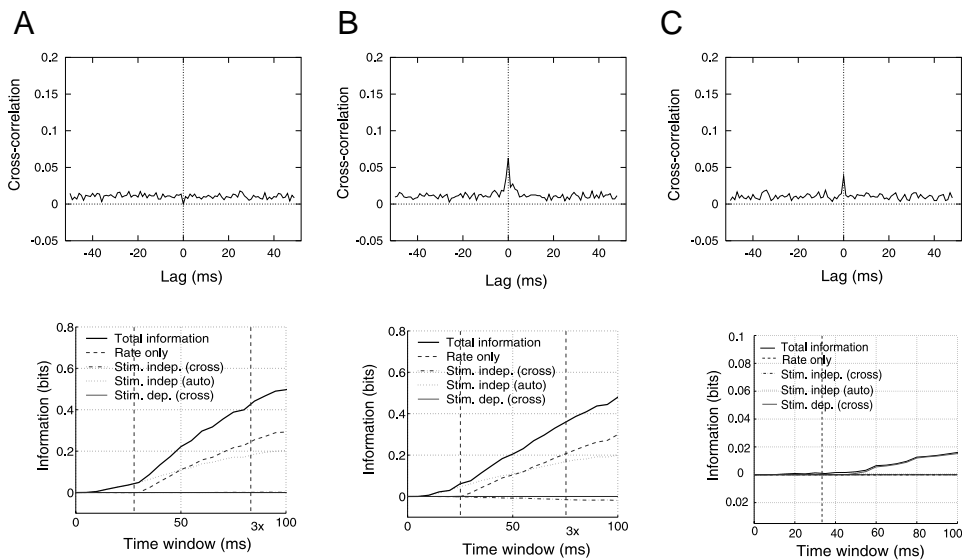


FIG. 2. Analysis of the performance of corrinfo3 using simulated test data with 20 trials per stimulus. *A*: independent Poisson-generated spike trains were used for each of 2 neurons. The cross-correlogram (shown above) is flat. (Abcissa is in milliseconds) *Bottom graph*: output of corrinfo3 as a function of the width of the time window between 0 and 100 ms. Separate plots show the total information from the sum of the 1st and 2nd derivatives shown in Eq. 2 and the contributions of the cross-cell and auto-(within-cell) stimulus-independent and cross-cell stimulus-dependent terms. Dashed vertical lines at  $\times$  and  $3\times$  show 1 and 3 times the mean interspike interval of the fastest firing neuron to its most effective stimulus. It is expected that the Taylor expansion shown in Eq. 2 will work at least as far as  $\times$  and may start to become inaccurate by  $3\times$ . *B*: spike trains generated by an integrate-and-fire simulation with 66% of common input to the 2 neurons. Rate information is the same as in *A*, but there is a negative contribution of the stimulus-independent cross-term. *C*: spike trains generated by an integrate-and-fire simulation to produce stimulus-dependent correlational information. We simulated a correlational assembly with a constant firing rate of 20 spikes/s to all stimuli and a percentage of shared connections of either 0% or 90% for different stimuli. There were 5 cells in the assembly, and for each of the 5 stimuli, 1 pair of cells had common connections. Results are plotted for 1 pair of cells (which had common connections for 1 of the stimuli). The only contributor to the total information was the stimulus-dependent cross-term. (The stimulus-dependent plot is almost hidden, because it is essentially identical to the total information plot in this case.)

an inferior temporal cortex cell (with the data obtained as described later). The results are shown in Fig. 2A. The cross-correlogram confirms that there is no relation between the firing of the two cells. The algorithm shows that there is information in the “Rate” term [i.e., the first derivative of the information (Eq. 8) and the first term of Eq. 9, i.e.,  $I_{tta}$ ], and correctly, no information contribution is indicated in the cross-cell parts of  $I_{ttb}$  (the “stimulus-independent cross-term” in Fig. 2, or  $I_{ttc}$ , the “stimulus-dependent cross-term” in Fig. 2A). There is some information in the auto-part of the stimulus-independent term, which although zero with large data sets, can be positive with small data sets.

Second, we consider the case when common input is added to the case shown in Fig. 2A. (The spike trains were generated with integrate-and-fire neurons using a procedure similar to that of Shadlen and Newsome (1998) and as described by Panzeri et al. (1999). In brief, each cell received 300 excitatory and 300 inhibitory inputs, each a Poisson process in itself, whose (possibly stimulus-dependent) mean rate was constant across the set of inputs for any specific stimulus condition and contributed a fixed quantity to the membrane potential. When the membrane potential exceeded a threshold, it was reset to a baseline value. The common input was added by connecting 66% of the inputs of both cells to the same input source. This common input is expected to produce redundancy. The results are shown in Fig. 2B. The algorithm shows that there is again information in the “Rate” term and that there is a negative contribution to the total information that is made explicit in the stimulus-independent cross-term. This latter term thus reflects

the common input to the neurons that is provided in a stimulus-independent way. The common input is reflected in the cross-correlogram shown in Fig. 2B. The stimulus-independent auto-term is similar to that in the Poisson case, because it reflects the within-cell distribution of responses to the set of stimuli.

Third, we show in Fig. 2C the operation for the algorithm when the firing rates to the different stimuli are identical, but when different pairs of cells become correlated (or synchronized) for only some of the stimuli. The spike trains were generated by an integrate-and-fire simulation in which we simulated a correlational assembly with a constant firing rate of 20 spikes/s to all stimuli, and a percentage of shared connections of either 0 or 90% for different stimuli. There were five cells in the assembly, and for each of the five stimuli, one pair of cells had common connections. An example of the cross-correlogram is shown in Fig. 2C, and the information plot below correctly shows “stimulus-dependent cross” information (i.e., a positive value of  $I_{ttc}$  for  $i \neq j$ ), and that this is the only contributor to the total information.

To assess the power efficiency of the method, i.e., how many trials of neuronal data are needed for the different aspects of the information to be estimated accurately, we performed simulations similar to those shown in Fig. 2 with different numbers of trials. The numbers of trials required were obtained as follows. For the stimulus-dependent cross-correlation term, we used the integrate-and-fire simulation described above to generate multiple datasets of spike trains each with for example 20 trials for each stimulus. Then we calculated the value of the information that was available in each of 20 independent



datasets. (Each dataset consisted of the equivalent of a neurophysiological experiment with a number of simultaneously recorded neurons.) This showed how, if short spike trains of neuronal firing are available from neurons, the statistics of those spike trains reflect any underlying cross-correlation that may be generated between the firing of neurons because they have for example stimulus-dependent common input. The details of these integrate-and-fire simulations were that two neurons with two stimuli shared 90% of their connections for one of the stimuli and none for the other stimulus. Each neuron had 300 excitatory inputs and 300 inhibitory inputs as described earlier, and the spike trains of each of the neurons were approximately Poisson-distributed, reflecting the Poisson inputs received through the excitatory inputs and the underlying spike generation process. With 20 trials of data for each stimulus, stimulus-dependent information was detected in most of the cases (75%). (This compares with only 8% of false positives obtained from a different simulation where there was no stimulus-dependent common input.) Furthermore, the mean stimulus-dependent cross-cell information across the 20 repetitions of the experiment was  $0.040 \pm 0.015$  (SE) bits. Thus in 20 neurophysiological experiments performed with different neurons and 20 trials for each stimulus, the stimulus-dependent cross-cell information could be detected very reliably ( $P = 0.001$  by  $t$ -test compared with the datasets with no stimulus-dependent common input). We also showed that in 10 neurophysiological experiments performed with different neurons and 20 trials for each stimulus, the stimulus-dependent cross-cell information could be detected reliably ( $P < 0.022$  by  $t$ -test compared with the datasets with no stimulus-dependent common input). For comparison, we analyzed the case with 200 trials per stimulus, which provides approximately the asymptotic values, and found that the information was detected in 100% of the experiments (with no false positives), and the mean stimulus-dependent cross-cell information was  $0.041 \pm 0.003$  bits. To assess the power efficiency of the rate information measurements, we performed simulations with Poisson spike trains for a case in which two cells fired at 45 and 35 Hz to one stimulus and at 37 and 25 Hz to a second stimulus. From 20 repetitions of the experiment and 20 trials for each stimulus, we obtained  $0.07 \pm 0.01$  bits ( $0.074 \pm 0.002$  bits with 200 trials per stimulus). Thus the rate information can also be detected with approximately 20 trials of data for each stimulus in 20 experiments. In summary, in the equivalent of 20 neurophysiological experiments, the rate information, the stimulus-independent cross-cell contribution to the information, and the stimulus-dependent cross-cell contribution to the information can all be highly reliably detected by this approach to information measurement. Furthermore, the most difficult term to detect, the stimulus-dependent cross-cell term, can be reliably detected in as few as 10 experiments with just 20 trials for each stimulus.

We emphasize that the point being made here is about whether spike data generated with biologically relevant parameters is likely to reflect information that could be produced by stimulus-dependent inputs to neurons, and the point is not about the sensitivity of the information measurement algorithm itself. The sensitivity of the algorithm, with respect to whether stimulus-dependent cross-cell information is present in the particular spike trains provided to the algorithm, was assessed as described above using the trial shuffling Monte Carlo pro-

cedure. What we are addressing here is whether the actual spike trains generated probabilistically by the neurons would reflect stimulus-dependent inputs if only a limited number of trials was selected to test. The conclusion is that, provided that there are 20 trials per stimulus, stimulus-dependent common input to cells is likely to lead to spike trains that show stimulus-dependent cross-cell information, although there are strong benefits to having more trials of data for each stimulus. The underlying reason for this is that with short trains of Poisson-like spike firing, there is so much variability with just a few spikes available on any one trial for each neuron that the randomness prevents, on some trials, the effects of the cross-correlation being present in the actual spike trains. We note that if only limited numbers of trials of data for each stimulus are available, using an algorithm that can utilize more than a few spikes per trial will be potentially advantageous. One such approach to this is being developed, in which rather than a Taylor expansion approach, a decoding approach is used (Franco et al. 2003).

#### *Application of the approach to real neuronal data*

In Fig. 4, the information values obtained for three simultaneously recorded cells in *experiment bj288* are shown. Figure 4A shows the total information and the different contributors to it. The development of the components in a time window starting at 100-ms poststimulus and increasing up to a value that is 100 ms long is shown. The cross-correlogram for one pair of the cells shows (Fig. 3) that there is a significant cross-correlation between the firing of the two cells (recorded on 2 separate microelectrodes) with a lag centered close to 0 ms. The algorithm shows that the contribution of the "Rate" term [i.e., the first derivative of the information (Eq. 8) and the first term of Eq. 9, i.e.,  $I_{ta}$ ] is approximately (after 100 ms) 0.38 bits. There is a small negative contribution of the cross-cell part of  $I_{tb}$  (the "stimulus-independent cross" term in Fig. 4), although this appears only toward the end on the time window (beyond approximately  $2\times$  as defined in Fig. 2) and is  $-0.04$  bits. This term reflects a small amount of redundancy that is produced because the cross-cell signal correlations  $v_{ij}$  and noise correlations  $\langle \gamma_{ij}(s) \rangle_s$  were both positive. The stimulus-independent auto-term is close to zero, indicating that the trial by trial variability of each cell (averaged across stimuli) is close to the variability expected for a Poisson process (i.e., the variance is close to the mean).

A feature of the data shown in Fig. 4 is that the stimulus-dependent cross-term contribution is positive and appears quite large (at 0.13 bits). To analyze the extent to which this is a statistically significant contribution to the total information, we performed a Monte Carlo test in which the degree of variability of this term was measured with repeated different random re-pairings (i.e., re-assortment or scrambling) of the trials within each stimulus. For each rearrangement of the data from individual cells between different trials (which also breaks any synchronization between the spikes of the simultaneously recorded neurons), the algorithm calculates the apparent cross-cell stimulus-dependent information. From this, the SD of this contributor to the information can be obtained. We show twice this SD in the plot in Fig. 4B. Any value of the cross-cell stimulus-dependent information from the simultaneously recorded neuronal data that lies outside this confidence interval

Cross Correlations for Cells 1 and 31, num spikes=4549 3933

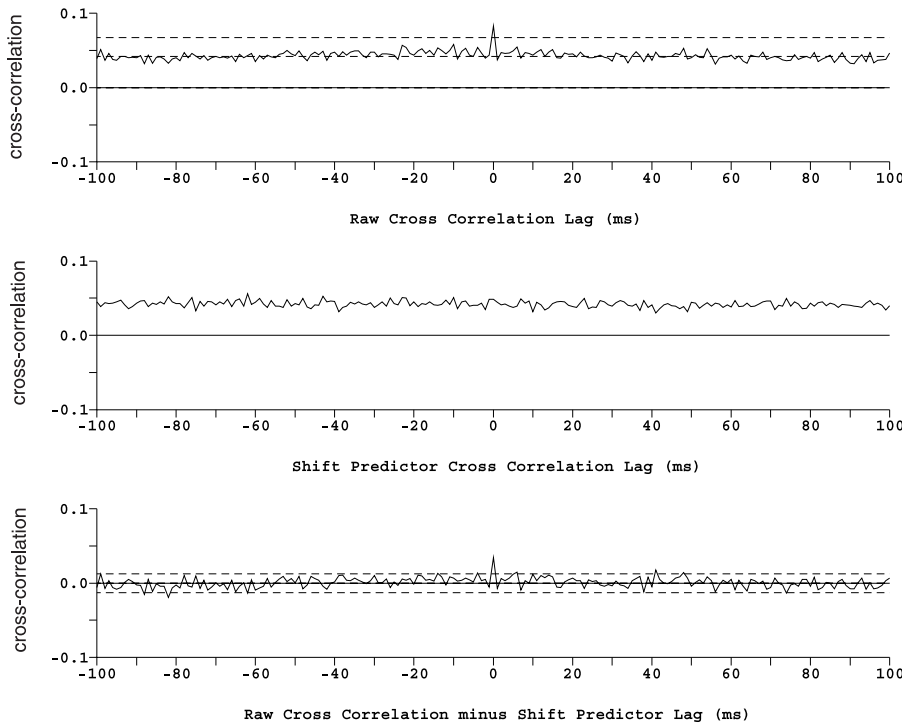


FIG. 3. Cross-correlogram between cells 1 and 31 for the set of simultaneously recorded cells recorded in *experiment bj288*. The raw cross-correlogram is shown at the *top*, the “shift predictor” cross-correlogram in the *middle* (which is produced by random re-pairings of the trials), and the corrected cross-correlogram (obtained by subtracting the shift predictor from the raw cross-correlogram) is at the *bottom*. The cross-correlogram was calculated by, for every spike that occurred in 1 neuron, adding to a histogram the relative times of occurrence, or lag, of all the spikes that occurred for the other neuron. Dashed lines show the 1% confidence limits, assuming that the counts in the bins of the cross-correlogram are Poisson-distributed. The cross-correlogram is for the period when the cells were responding to the 500-ms stimulus, namely in the period 100–600 ms after stimulus onset, given the response latencies of 100 ms.

would be significant. We see from Fig. 4B that the apparent information contribution from the stimulus-dependent cross-cell term is not generally more than 2 SD from the mean. This shows that there is considerable variability in what can be extracted in this stimulus-dependent cross-cell term, and in-

deed the estimated contribution is hardly statistically significant. This is a very useful feature of the details of the information analysis approach described here, because it enables a test to be performed of whether the stimulus-dependent cross-cell term, which could reflect stimulus-dependent synchroni-

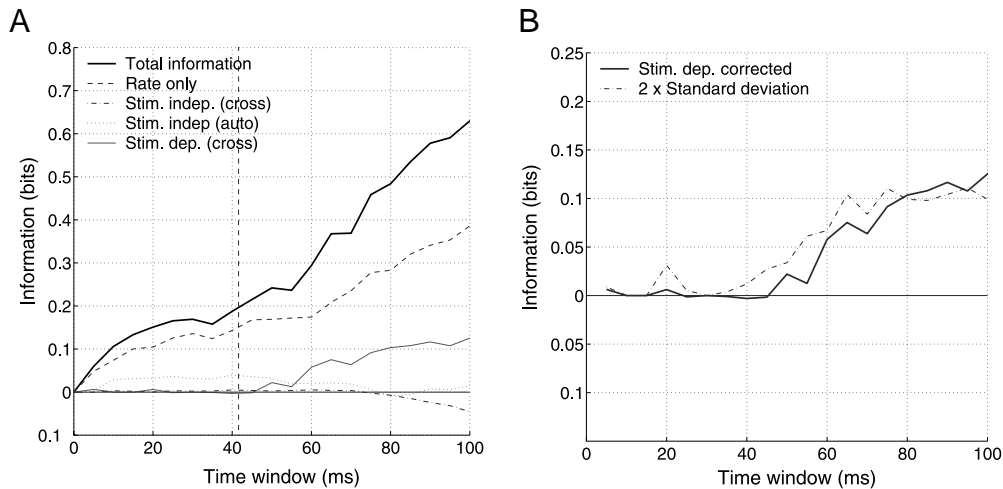


FIG. 4. The information analysis applied to real simultaneous recording data from 3 neurons in *experiment bj288*, in which 20 complex stimuli effective for inferior temporal cortex neurons (objects, faces, and scenes) were shown. *A*: graphs show the contributions to the information from the different terms in *Eqs. 8 and 9*, as a function of the length of the time window, which started 100 ms after stimulus onset when inferior temporal cortex (IT) neurons start to respond. Rate information is the sum of the term in *Eq. 8* and the 1st term of *Eq. 9*. Contribution of the stimulus-independent noise correlation to the information is the 2nd term of *Eq. 9* and is separated into components arising from the correlations between cells (the cross-component, for  $i \neq j$ ) and from the autocorrelation within a cell (the auto-component, for  $i = j$ ). This term is not 0 if there is some correlation in the variance to a given stimulus, even if it is independent of which stimulus is present. Contribution of the stimulus-dependent noise correlation to the information is the 3rd term of *Eq. 9*, and only the cross-cell term is shown (for  $i \neq j$ ), because this is the term of interest. *B*: value of the cross-cell term of the stimulus-dependent information and 2 SD of information obtained from 30 repetitions of the shuffling of the trials using the Monte Carlo procedure described in the text.



Cross Correlations for Cells 3 and 22, num spikes=1000 2150

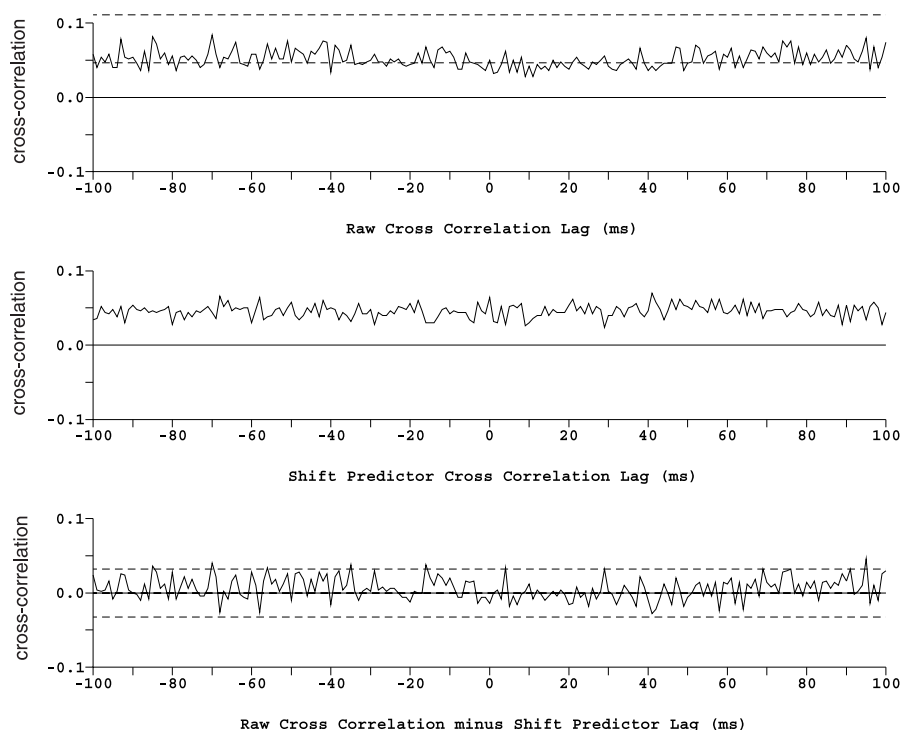


FIG. 5. Cross-correlogram between the responses of simultaneously recorded inferior temporal cortex neurons from an experiment in which 20 stimuli of the type that activate inferior temporal cortex neurons were shown in random sequence from *experiment 229*. *Top*: raw cross-correlogram. *Middle*: shift predictor data, showing values obtained for the cross-correlation coefficients when data for each cell were shuffled between trials. *Bottom*: cross-correlogram corrected by subtracting the shift predictor, with the  $P < 0.01$  confidence limits indicated.

zation of the simultaneously recorded cells and is reflected in whether  $\gamma_{ij}(s)$  is different for different stimuli, is significant.

The overall conclusion from the data analysis performed on the set of cells in *experiment bj288* is that most of the information is contained in the “Rate” term, that there is little redundancy between the cells in that the stimulus-independent cross-term is low, and that the stimulus-dependent cross-cell correlation contributes in a way that is barely statistically significant. It is also notable in Fig. 4 that the terms apart from the rate tend to start to contribute relatively far on in the time window (in this case after 50 ms). The vertical line in Fig. 4 shows the mean value of the interspike interval of the fastest firing neuron to its most effective stimulus.<sup>5</sup> At about this value, the number of spikes in the time window from each neuron may on some trials be more than 1, and beyond this value, the information expansion described in Eq. 2 may tend to break down. In practice, we find with simulated data (where the amount of information can be calculated) that the expansion often works reasonably  $\leq 2$ –3 times this period. A useful practical procedure in many cases is to note further that the region where the Taylor expansion will start to break down is when  $I_{tt}$  becomes as large as  $I_t$  (see Eq. 2). A related check is for whether there is an inflection in any of the information curve as time window increases in duration. We applied such checks to all data presented in the 100-ms time window in this paper. To show how many spikes were likely to be present for the most effective stimuli in the 100-ms time window for the set of neurons tested with five moving stimuli and hence included in Table 1, we showed in Table 1 the average firing

rate to the most effective stimulus for the most responsive neuron in each of these experiments.

Another type of effect of correlation that leads to stimulus-independent synergy contributions is illustrated in Fig. 6 from *experiment bj229*. (The cross-correlogram is shown in Fig. 5. Figure 6A shows that a considerable proportion of the information available in a 100-ms time period was available in the rates. In addition, there was a small positive value for the cross-cell stimulus-independent term, which reflected some anti-correlation between the response profiles of the cells to the set of stimuli. This is produced by a small negative value for  $v_{ij}$  (evident after approximately 50 ms into the time window, and indicating anticorrelated neuronal response profiles to the set of stimuli; shown in Fig. 6C) together with a positive value for  $\langle \gamma_{ij} \rangle_w$  [which is defined as  $\langle \bar{n}_i(s) \bar{n}_j(s) \gamma_{ij}(s) \rangle_s$ ; shown in Fig. 6D]. ( $\langle \gamma_{ij} \rangle_w$  is just  $\gamma_{ij}(s)$  for each stimulus weighted by the number of spikes and averaged across-stimuli.) Figure 6A shows that there is a small positive contribution to the information from the stimulus-dependent cross-cell term, and Fig. 6B shows that this is less than 2 SD from what is produced by random reassignment of the trials within those available for each stimulus, and therefore is not significant. We note that the cells were recorded on two different electrodes so that cells 1–3 mm apart can show these effects. The conclusion from the application of the information theoretic approach utilized in this paper is that the total information available from the cell is greater than that from the “Rate” term and that this is due to a stimulus-independent cross-cell term.

Another type of effect of correlation, which leads to stimulus-independent redundancy contributions, is illustrated in Fig. 7 from *experiment bj024b*. Figure 7A shows that a considerable

<sup>5</sup> 3x on some of the Figures indicates 3 times this value.

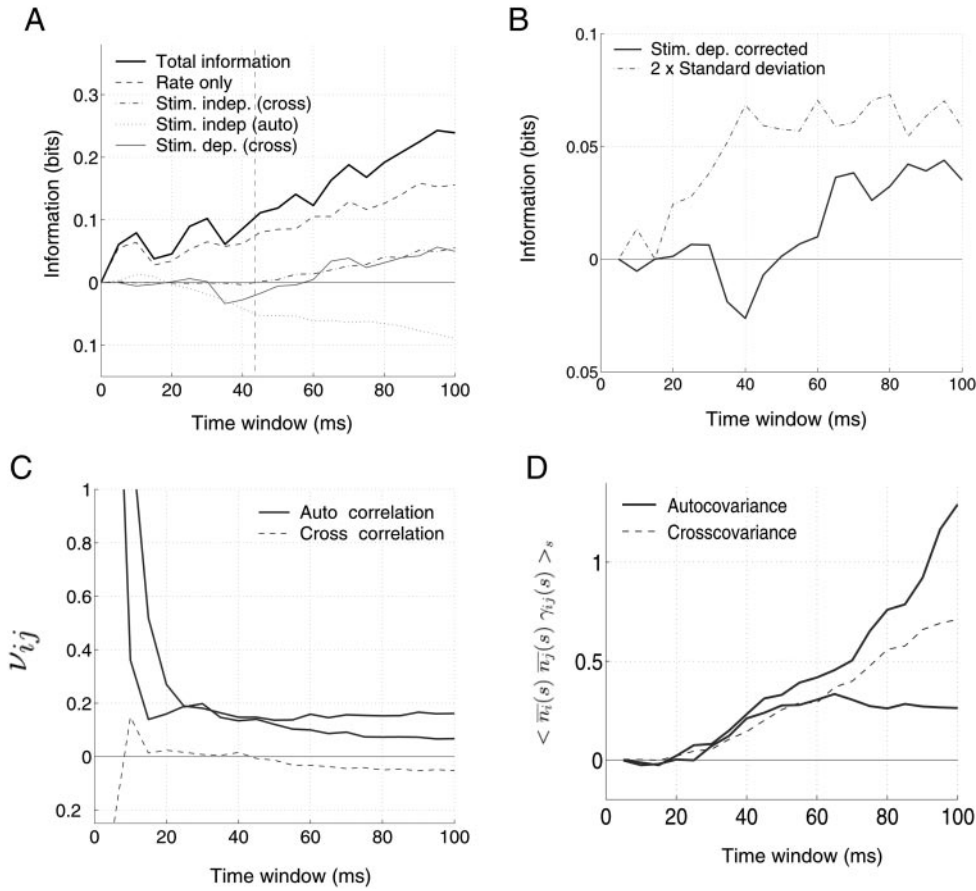


FIG. 6. Results of the information analysis on a set of 2 simultaneously recorded inferior temporal cortex neurons in *experiment* 229 in which 20 complex stimuli effective for IT neurons (objects, faces, and scenes) were shown. *A*: graphs show the contributions to the information from the different terms in *Eqs.* 8 and 9, as a function of the length of the time window, which started 100 ms after stimulus onset, when IT neurons start to respond. Rate information is the sum of the term in *Eq.* 8 and the 1st term of *Eq.* 9. Contribution of the stimulus-independent noise correlation to the information is the 2nd term of *Eq.* 9 and is separated into components arising from the correlations between cells (the cross-component, for  $i \neq j$ ) and from the autocorrelation within a cell (the auto-component, for  $i = j$ ). This term is not 0 if there is some correlation in the variance to a given stimulus, even if it is independent of which stimulus is present. Contribution of the stimulus-dependent noise correlation to the information is the 3rd term of *Eq.* 9, and only the cross-term is shown (for  $i \neq j$ ), because this is the term of interest. *B*: value of the cross-cell term of the stimulus-dependent information and 2 SD of its variation as estimated by the Monte Carlo method described in the text. *C*: value of  $v_{ij}$ , the signal correlations, measured both across-cell pairs (cross-, dashed lines) and within cells (auto-, i.e.,  $i = j$ , shown by a solid line for each cell). *D*: time course of the  $\gamma_{ij}$  term weighted by the firing rates of the neurons (in particular,  $\langle \bar{n}_i(s) \bar{n}_j(s) \gamma_{ij}(s) \rangle_s$ , where  $n_i$  is the number of spikes in time  $t$  from cell  $i$ , which corresponds to  $\langle \bar{r}_i(s) \bar{r}_j(s) \gamma_{ij}(s) t^2 \rangle_s$ , where  $r_i$  is the firing rate of cell  $i$  in time  $t$ . This term is defined as  $\langle \gamma_{ij} \rangle_w$  in the text.

proportion of the information available in a 100-ms time period was available in the rates. In addition, there was a small negative value for the cross-cell stimulus-independent term, which reflected some correlation between the response profiles of the cells to the set of stimuli. This is produced by a small positive value for  $v_{ij}$  (as shown in Fig. 7C), together with a positive value for  $\langle \gamma_{ij} \rangle_w$  (as shown in Fig. 7D). Although there was some redundancy related to the negative cross-cell stimulus-independent contribution to the information, the total information was in fact higher than the rate information because the total information included a positive stimulus-independent auto-cell contribution. The positive stimulus-independent auto-contribution reflects less variability in the neuronal responses than would be the case for independent spikes, generated by for example a Poisson process.

We now summarize a dataset of seven experiments the type of result that was obtained in a set of recordings from the macaque inferior temporal visual cortex when a set of five

moving visual stimuli were being shown in a visual fixation task. The images included moving faces and objects, which are known to activate some inferior temporal cortex neurons (Rolls 2000). Part of the interest of these data are that when stimulus-dependent synchronization effects have been described in early cortical visual areas, moving stimuli have often been used (Singer 1999, 2000). A total of 10–25 trials of data for each stimulus, presented for 500 ms, were acquired. The methods were similar to those described by Rolls et al. (1997), except that simultaneous recordings from small populations of single neurons were made with two to four separate movable micro-electrodes, and the stimuli were moving against complex backgrounds so that segmentation and motion, which potentially enhance the need for binding and thus perhaps for synchrony (Singer 1999, 2000), were involved.

The results for the seven experiments that were completed with groups of two to four simultaneously recorded inferior temporal cortex neurons are shown in Table 1. The total

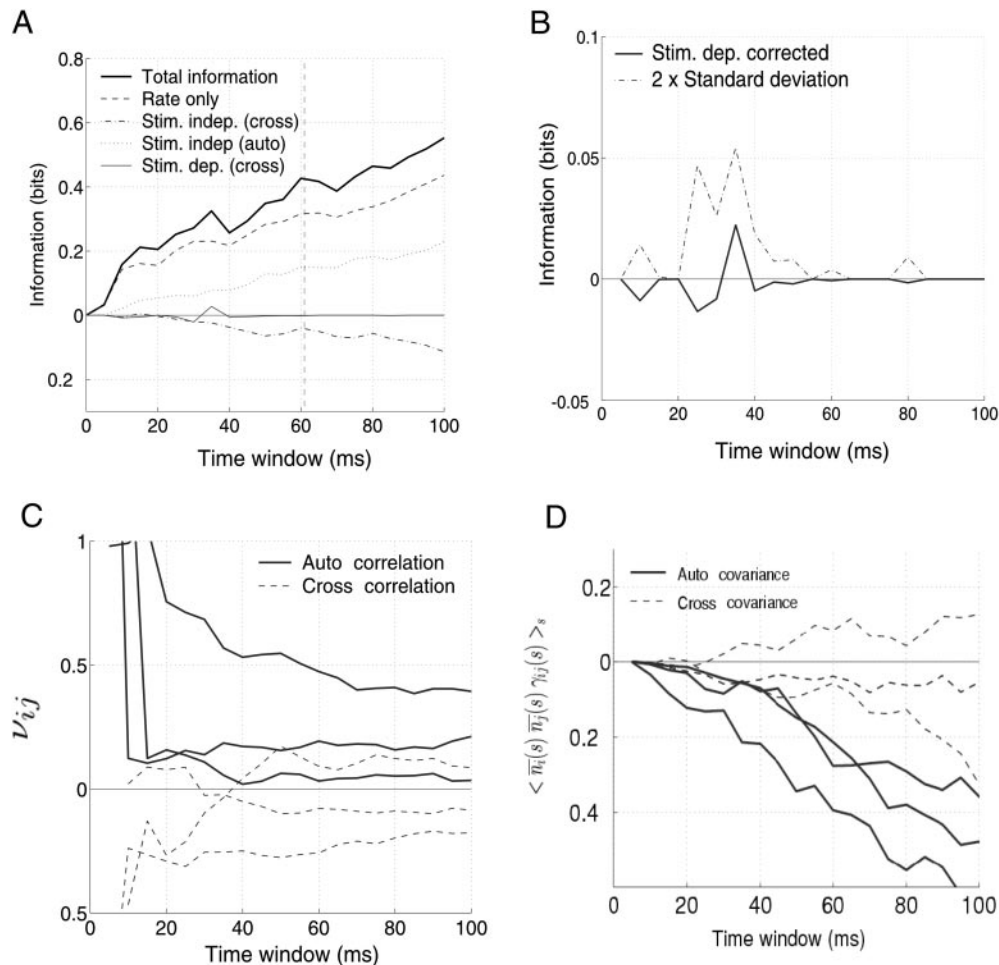


FIG. 7. Results of the information analysis on a set of 3 simultaneously recorded inferior temporal cortex neurons in *experiment bj024b* in which 5 moving stimuli effective for IT neurons (objects and faces) were shown against a complex natural background. Conventions as in Fig. 6.

information is the total from *Eqs. 8 and 9* in a 100-ms time window and is not expected to be the sum of the contributions shown in Table 1 because the stimulus-dependent auto-term is not shown in Table 1. (The latter term includes terms produced in the algorithm we used from the Monte Carlo correction procedure.) The results show that the greatest contribution to the information is that from the rates (0.28 bits on average in each experiment), i.e., from the numbers of spikes from each neuron in the time window of 100 ms. The average value of  $-0.03$  bits for the cross-term of the stimulus-independent “noise” correlation-related contribution is consistent with a small amount of common input to neurons in the inferior temporal visual cortex, producing redundancy. The cross-term of the stimulus-dependent contribution was on average very close to zero, and in no case was statistically significant. Thus stimulus-dependent synchronization made no contribution to the information reflected in the responses of this set of neurons. Table 1 also shows some positive contribution averaged across the datasets of the stimulus-independent auto-term. This reflects the fact that for this set of cells, the spike trains of each neuron taken individually were less variable than if they were generated by an independent process such as a Poisson process.

## DISCUSSION

We have shown in this paper how different aspects of the firing of simultaneously recorded neurons might contribute to the total information that could be extracted from the firing. We now consider what use different decoding procedures might make of these different sources of contribution and to what extent neurons might implement particular decoding procedures. The “rate” information plotted in the graphs does not include the trial-by-trial variability of the neuronal responses. However, this is reflected in the stimulus-independent cross-cell term and auto-term (of  $I_{\text{tb}}$ ). Let us consider the cross-cell term. If, for example, positive values for  $v_{ij}$ , reflecting some correlation in the firing rate response profiles of the cells to the set of stimuli, are combined multiplicatively with a contribution from the  $\gamma_{ij}$  terms, which indicates that the cells tend to respond in the same way on any given trial (reflecting stimulus-independent synchronization produced by common input), then redundancy arises. Now consider the auto-cell term of the stimulus-independent contribution. If this term is large, it reflects, calculated separately for each stimulus, the variability of the neuronal responses to that stimulus [measured by  $\gamma_{ij}(s)$ ] weighted by the firing rate to that stimulus and then averaged across-stimuli (see *Eq. 9*, second term). This term is simply added for different cells in a simultaneously recorded set. Let

us now consider a decoding procedure that operates on the average responses of each cell to each stimulus, as illustrated in Fig. 10.17 of Rolls and Treves (1998), in Fig. 5.9 of Rolls and Deco (2002), and in Fig. 9 of Franco et al. (2003). If we take a single test trial of the number of spikes available from each cell, as shown at the bottom of these figures, we can ask how much information can be obtained about which stimulus was present. If the decoding procedure takes just the dot product of the cell firing rate response vector on the single trial with the average response vectors to each stimulus, then this decoding will reflect what is described by the “Rate” term and by the stimulus-independent term ( $I_{\text{tb}}$ ), including both the cross- and auto-cell parts. This is one of the decoding procedures described by Rolls et al. (1997) and is used to calculate the information available from populations of neurons. We might call the combined contribution of the “Rate” and stimulus-independent contributions as described here the “average spike count information,” because it is available simply by counting the number of spikes from each of the cells on a given trial and comparing it to the average of the spike counts to each stimulus. What this does not take into account is the stimulus-dependent correlations between the spikes that occur. Such stimulus-dependent synchronization effects are not reflected in the dot product decoding procedure, because to detect the effects of synchronization would entail storing the correlation between the firing of every pair of cells for every stimulus, and this is reflected in the stimulus-dependent cross-cell term described in this paper.

We now discuss what this approach and the analyses of simulated data sets and of a small number of simultaneously recorded neuronal datasets reveal about redundancy. The baseline condition considered is independent spike generation of the simultaneously recorded neurons such as would be generated by Poisson processes operating in each neuron but where there are significant differences in the average number of spikes to each stimulus. If the signal correlations ( $v_{ij}$ ) are zero, there is information in the “Rate” terms as labeled in the diagrams in this paper of each cell [i.e., the first derivative of the information (Eq. 8) and the first term of Eq. 9, i.e.,  $I_{\text{ta}}$ ], and the “Rate” terms of simultaneously recorded cells add linearly. If the signal correlations ( $v_{ij}$ ) are different from zero, there is still information in the “Rate” terms of each single cell, but the cross-cell part of  $I_{\text{ta}}$  is negative, and the “Rate” terms of simultaneously recorded cells add sub-linearly, that is there is redundancy. Shuffling the trials makes no difference to this redundancy, because trial by trial correlations (reflected in  $\gamma$  terms) do not contribute to  $I_{\text{ta}}$ . This factor can only produce redundancy and not synergy. We note also that this “Rate” term does not include any reduction in the information that arises from the variability in the neuronal responses from trial to trial for a given stimulus.

We note that the redundancy arising from nonzero signal correlations can in principle be compensated for by the “noise” correlations averaged across stimuli. As described earlier, this compensation occurs when the signs of the signal and noise correlations are opposite and is reflected in  $I_{\text{tb}}$ , the stimulus-independent contribution to the total information that depends on the noise correlations. In practice, with most cell populations we have analyzed in the inferior temporal visual cortex, the signal correlations  $v_{ij}$  tend to be weakly positive or zero, the noise correlations averaged across stimuli,  $\langle \gamma_{ij}(s) \rangle_s$ , tend to be

zero or weakly positive, and the result is a small negative value for the cross-cell stimulus-independent term, reflecting a small amount of redundancy (see examples above, the data in Table 1, and the data in Rolls et al. (2003)).

The variability in the neuronal responses from trial to trial for a given stimulus, then averaged across-stimuli, is reflected in the auto-part of  $I_{\text{tb}}$ . This term is negative if the variability is greater than would be generated by a Poisson-like spike generation process.

Adding together the “Rate” contribution with both the cross-cell and auto-parts of  $I_{\text{tb}}$  gives what is generally thought of as the information that is present in the firing rates of cells (Rolls and Treves 1998; Rolls et al. 1997). As noted above, this information is available in a simple neuronally plausible dot product decoding.

The third way in which this approach identifies contributions to redundancy versus synergy is in the stimulus-dependent cross-cell term that is made explicit in  $I_{\text{tc}}$ . This term measures the effects of stimulus-dependent “noise” cross-correlations in the firing of different cells. This is not included in what is generally thought of as the information available from the firing rates of the cells. This term shows considerable variability in real data in realistic time epochs in the order of 100 ms that contain zero to several spikes from each cell. This variability means that, at least in our datasets, cross-correlations that may be evident when thousands of spikes are used to make the cross-correlogram (see e.g., Fig. 3) actually may in a short temporal epoch on a single trial make a relatively little, and only rarely a statistically significant, contribution to the information available from the spikes. One might argue that if there was an enormous pool of neurons showing these weak stimulus-dependent noise correlation effects, then overall this might become more important (Salinas and Sejnowski 2000). However, we note that by making the neuronal pool much larger, the information from the firing rates would also become much larger, leaving the relative contribution of spike timing between neurons small. (It is also possible that cross-cell temporal structure might contribute to encoding, and if present with different lags on different trials, might not be apparent in a cross-correlogram based on thousands of spikes. Such temporal encoding could in principle be detected by the methods described here, but in practice in the real neuronal dataset analyzed, no evidence for this type of encoding was found.)

The overall effect of correlations on the information available from simultaneously recorded cells can be estimated by comparing the “total information” for the cells recorded simultaneously with the sum across the cells for the “total information” measured separately for each cell. For the dataset shown in Table 1, this shows that any synergy from stimulus-dependent cross-correlations is smaller than the redundancy arising from stimulus-independent cross-correlations, resulting in a mean redundancy of 12% for these groups of two to four simultaneously recorded cells. This is smaller than the value of 20% for pairs of cells recorded from the same electrode in macaque inferior temporal cortex by Gawne and Richmond (1993), probably reflecting the fact that we recorded with several electrodes and thus were able to measure the information from neurons that were not adjacent in the cortex. As noted in the Introduction, the method described in this paper is able to show how this overall redundancy is contributed to by the stimulus-dependent and stimulus-independent correlations.



In conclusion, in this paper we have shown how to apply and statistically evaluate the Taylor expansion approach of the Shannon mutual information (Panzeri et al. 1999) to datasets from simultaneously recorded neurons. We have shown how the different terms in the expansion reflect stimulus-dependent and stimulus-independent covariation in the spike trains of the simultaneously recorded neurons, have shown how each of the terms can be statistically evaluated, and have performed a power efficiency analysis of the approach. The conclusion is that the approach can provide with 20 experiments statistically reliable estimates of the different contributions with as few as 20 trials of data for each stimulus in each experiment. We note that, even when steps are taken in the simulations to provide large amounts of stimulus-dependent input to pairs of neurons, that the resulting stimulus-dependent gain of information is typically low relative to the rate information term across the set of stimuli. We also found that, at least in the real neuronal datasets we examined from the inferior temporal visual cortex, the amount of information in the correlations is typically low relative to that in the rates. Moreover, the approach taken in this paper shows that the encoding of evidence available in neuronal firing about which stimulus is present must be assessed quantitatively using information theoretic measures, such as those described here.

This research was supported by Medical Research Council (MRC) Grant PG9826105, the Human Frontier Science Program, the MRC Interdisciplinary Research Centre for Cognitive Neuroscience, and the Wellcome Trust.

#### REFERENCES

- Aertsen AMHJ, Gerstein GL, Habib MK, and Palm G. Dynamics of neuronal firing correlation: modulation of "effective connectivity." *J Neurophysiol* 61: 900–917, 1989.
- Bialek W, Rieke F, de Ruyter van Steveninck RR, and Warland D. Reading a neural code. *Science* 252: 1854–1857, 1991.
- Booth MCA and Rolls ET. View-invariant representations of familiar objects by neurons in the inferior temporal visual cortex. *Cereb Cortex* 8: 510–523, 1998.
- Franco L, Rolls ET, and Treves A. The use of decoding to analyze the contribution to the information of the correlations between the firing of simultaneously recorded neurons. 2003 In press.
- Gawne TJ and Richmond BJ. How independent are the messages carried by adjacent inferior temporal cortical neurons? *J Neurosci* 13: 2758–2771, 1993.
- Panzeri S, Petersen RS, Schultz SR, Lebedev M, and Diamond ME. The role of spike timing in the coding of stimulus location in rat somatosensory cortex. *Neuron* 29: 769–777, 2001.
- Panzeri S, Schultz SR, Treves A, and Rolls ET. Correlations and the encoding of information in the nervous system. *Proc Roy Soc B* 266: 1001–1012, 1999.
- Panzeri S and Treves A. Analytical estimates of limited sampling biases in different information measures. *Network* 7: 87–107, 1996.
- Petersen RS, Panzeri S, and Diamond ME. Population coding of stimulus location in rat somatosensory cortex. *Neuron* 32: 503–514, 2001.
- Rolls ET. Functions of the primate temporal lobe cortical visual areas in invariant visual object and face recognition. *Neuron* 27: 205–218, 2000.
- Rolls ET, Aggelopoulos NC, Franco L, and Treves A. Information encoding in the inferior temporal visual cortex: contributions of the firing rates and the correlations between the firing of neurons. 2003 In press.
- Rolls ET and Deco G. *Computational Neuroscience of Vision*. Oxford, UK: Oxford, 2002.
- Rolls ET, Tovee MJ, Purcell DG, Stewart AL, and Azzopardi P. The responses of neurons in the temporal cortex of primates, and face identification and detection. *Exp Brain Res* 101: 474–484, 1994.
- Rolls ET and Treves A. *Neural Networks and Brain Function*. Oxford, UK: Oxford, 1998.
- Rolls ET, Treves A, and Tovee MJ. The representational capacity of the distributed encoding of information provided by populations of neurons in the primate temporal visual cortex. *Exp Brain Res* 114: 149–162, 1997.
- Salinas E and Sejnowski T. Impact of correlated synaptic input on output firing rate and variability in simple neuronal data. *J Neurosci* 20: 6193–6209, 2000.
- Shadlen M and Newsome W. Is there a signal in the noise? *Curr Opin Neurobiol* 5: 248–250, 1994.
- Shadlen M and Newsome W. The variable discharge of cortical neurons: implications for connectivity, computation and coding. *J Neurosci* 18: 3870–3896, 1998.
- Singer W. Neuronal synchrony: a versatile code for the definition of relations? *Neuron* 24: 49–65, 1999.
- Singer W. Response synchronisation: a universal coding strategy for the definition of relations. In: *The New Cognitive Neurosciences*, edited by M. Gazzaniga. Cambridge, MA: MIT Press, 2000, p. 325–338.
- Skaggs WE, McNaughton BL, Gothard K, and Markus E. An information theoretic approach to deciphering the hippocampal code. In: *Advances in Neural Information Processing Systems*, edited by Hanson S, Cowan JD, and Giles CL. San Mateo, CA: Morgan Kaufmann, vol. 5, 1993, p. 1030–1037.
- Treves A. Mean-field analysis of neuronal spike dynamics quantitative estimate of the information relayed by the Schaffer collaterals. *Network* 4: 259–284, 1993.
- Treves A and Panzeri S. The upward bias in measures of information derived from limited data samples. *Neural Comput* 7: 399–407, 1995.
- Treves A, Panzeri S, Rolls ET, Booth M, and Waksman EA. Firing rate distributions and efficiency of information transmission of inferior temporal cortex neurons to natural visual stimuli. *Neural Comput* 11: 611–641, 1999.