

Bruno B. Averbeck and Daeyeol Lee

J Neurophysiol 95:3633-3644, 2006. First published Mar 22, 2006; doi:10.1152/jn.00919.2005

You might find this additional information useful...

This article cites 58 articles, 34 of which you can access free at:

<http://jn.physiology.org/cgi/content/full/95/6/3633#BIBL>

This article has been cited by 4 other HighWire hosted articles:

Evaluating Information Transfer Between Auditory Cortical Neurons

B. Gourevitch and J. J. Eggermont

J Neurophysiol, March 1, 2007; 97 (3): 2533-2543.

[\[Abstract\]](#) [\[Full Text\]](#) [\[PDF\]](#)

Prefrontal Neural Correlates of Memory for Sequences

B. B. Averbeck and D. Lee

J. Neurosci., February 28, 2007; 27 (9): 2204-2211.

[\[Abstract\]](#) [\[Full Text\]](#) [\[PDF\]](#)

Sensory integration across space and in time for decision making in the somatosensory system of rodents

T. Celikel and B. Sakmann

PNAS, January 23, 2007; 104 (4): 1395-1400.

[\[Abstract\]](#) [\[Full Text\]](#) [\[PDF\]](#)

Probabilistic Encoding of Vocalizations in Macaque Ventral Lateral Prefrontal Cortex

B. B. Averbeck and L. M. Romanski

J. Neurosci., October 25, 2006; 26 (43): 11023-11033.

[\[Abstract\]](#) [\[Full Text\]](#) [\[PDF\]](#)

Updated information and services including high-resolution figures, can be found at:

<http://jn.physiology.org/cgi/content/full/95/6/3633>

Additional material and information about *Journal of Neurophysiology* can be found at:

<http://www.the-aps.org/publications/jn>

This information is current as of March 18, 2008 .

Effects of Noise Correlations on Information Encoding and Decoding

Bruno B. Averbeck and Daeyeol Lee

Department of Brain and Cognitive Sciences, Center for Visual Science, University of Rochester, Rochester, New York

Submitted 1 September 2005; accepted in final form 13 March 2006

Averbeck, Bruno B. and Daeyeol Lee. Effects of noise correlations on information encoding and decoding. *J Neurophysiol* 95: 3633–3644, 2006. First published March 22, 2006; doi:10.1152/jn.00919.2005. Response variability is often correlated across populations of neurons, and these noise correlations may play a role in information coding. In previous studies, this possibility has been examined from the encoding and decoding perspectives. Here we used *d prime* and related information measures to examine how studies of noise correlations from these two perspectives are related. We found that for a pair of neurons, the effect of noise correlations on information decoding can be zero when the effect of noise correlations on the information encoded obtains its largest positive or negative values. Furthermore, there can be no effect of noise correlations on the information encoded when it has an effect on information decoding. We also measured the effect of noise correlations on information encoding and decoding in simultaneously recorded neurons in the supplementary motor area to see how well *d prime* accounted for the information actually present in the neural responses and to see how noise correlations affected encoding and decoding in real data. These analyses showed that *d prime* provides an accurate measure of information encoding and decoding in our population of neurons. We also found that the effect of noise correlations on information encoding was somewhat larger than the effect of noise correlations on information decoding, but both were relatively small. Finally, as predicted theoretically, the effects of correlations were slightly greater for larger ensembles (3–8 neurons) than for pairs of neurons.

INTRODUCTION

The possibility that patterns of activity across neurons are important features of the neural code has led to their study from a number of perspectives. Many of these studies have focused on noise correlation, which is between neuron correlation in the variability of the neural response to a fixed stimulus, with the response often measured as a spike count (Gawne and Richmond 1993; Lee et al. 1998). Noise correlations should be distinguished from signal correlations, which are correlations in average spike count, and from coherent oscillations, which decompose noise correlations into multiple temporal or frequency points (Averbeck and Lee 2004). Here we examine the role of noise correlations in information encoding and decoding. Studying information encoding involves studying the mapping between the stimulus and the population neural response, and estimating the total amount of information present in the neural responses. To evaluate the effect of noise correlations on information encoding, we can determine whether neurons with correlated noise encode more or less information relative to those without correlated noise. Studying information decoding involves studying the mapping from the population neural response to a prediction of the stimulus. When we assess the

effects of noise correlations on information decoding, we calculate the amount of information lost when a decoding algorithm derived by ignoring correlations is applied to the neural responses with noise correlation.

Theoretical studies using Fisher Information have examined the effect of correlations on the information encoded by populations of neurons. Fisher Information bounds the variance with which a parameter encoded by a population of neurons can be estimated (Casella and Berger 1990). These studies have found that noise correlations can increase or decrease the information encoded with respect to an uncorrelated population, depending on their relationship with signal correlation (Johnson 1980; Snippe and Koenderink 1992). They have also found that information either grows with the number of neurons in a population (Abbott and Dayan 1999; Shamir and Sompolinsky 2004; Wilke and Eurich 2002) or saturates as the number of neurons goes to infinity (Sompolinsky et al. 2001; Zohary et al. 1994), depending on the structure of the correlations in the population. Theoretical work has also been done on the effect of noise correlations on information decoding (Shamir and Sompolinsky 2004; Wu et al. 2001). However, how the effects of noise correlation on information encoding and decoding are related has not been systematically investigated.

Empirical studies have focused on whether more or less information can be extracted from neural responses when trials are shuffled, destroying correlations, (Averbeck et al. 2003; Gawne and Richmond 1993; Gawne et al. 1996; Golledge et al. 2003; Panzeri and Schultz 2001; Panzeri et al. 1999, 2002; Petersen et al. 2001, 2002; Pola et al. 2003; Rolls et al. 2003; Romo et al. 2003) or whether correlations could be ignored by decoding algorithms without a loss of information (Averbeck and Lee 2003; Dan et al. 1998; Maynard et al. 1999; Nirenberg et al. 2001; Oram et al. 2001). These studies have generally found that noise correlations have little impact on information coding (Averbeck and Lee 2004). However, they have only analyzed interactions at the level of pairs of neurons. Finally, a number of other studies have considered neural coding at the ensemble level, but they have not directly addressed the effects of noise correlations on encoding or decoding (Brown et al. 1998, 2004; Nicolelis et al. 1997; Truccolo et al. 2005).

The effects of noise correlations on neural coding have been assessed using information measures and decoding algorithms. In general, there are many different ways to quantify information (Arndt 2001). In this study, we measured information using the square of *d prime* (d^2) and used d^2 to estimate the fraction or percent correct we obtained in corresponding decoding analyses. This allowed us to link directly the results of

Address for reprint requests and other correspondence: B. B. Averbeck, Dept. of Brain and Cognitive Sciences, Center for Visual Science, University of Rochester, Rochester, NY 14627 (E-mail: baverbeck@cvs.rochester.edu).

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.

using an information measure (d^2) and using a decoding analysis to study the effects of noise correlations. Although measures of the fraction correct are often closely correlated with Shannon information (Averbeck et al. 2003), it is theoretically possible to dissociate them (Thomson and Kristan 2005). The fraction correct is more directly related to behavior in experiments in which stimuli must be discriminated or movements must be produced. Because d^2 is the discrete analog of Fisher Information, the use of d^2 on our experimental data provides an assessment of the effect of correlations similar to that used in the theoretical studies cited above. Finally, d^2 is a simple measure, and therefore its interpretation is straightforward. We exploit this simplicity to examine how the effects of noise correlations on information encoding and decoding are related because both encoding and decoding have been studied in the literature, but they have not been linked.

In our results, we show extensively that d^2 provides an accurate measure of the amount of information in the neural activity recorded from the supplementary motor area of monkeys. In this study, information refers to the discriminability of the targets toward which the monkey reached. The analyses also showed that the decoding performance predicted by d^2 agreed closely with the results from actually carrying out the corresponding linear decoding analyses and that nonlinear decoding algorithms do not extract more information than a linear decoding algorithm. Thus both d^2 and the corresponding linear decoding algorithms provide an accurate measure of the information in the neural activity.

METHODS

General

The data set analyzed in this study has been described previously (Lee and Quesy 2003). All the procedures used in this study were approved by the University of Rochester Committee on Animal Research and conformed to the principles outlined in the Guide for the Care and Use of Laboratory Animals (National Institutes of Health Publications No. 85-23, revised 1996). Neurons analyzed in the present study were recorded from the left caudal supplementary motor area (SMA-proper or F3) in two rhesus macaques producing sequences of visually guided reaching movements.

Behavioral task

Two animals were trained on the serial reaction time task shown in Fig. 1. They sat facing a computer monitor on which a series of targets was presented in a 4×4 grid. The animals acquired each target by reaching toward the corresponding location on a touch screen placed horizontally in front of the animal. After acquisition of a target, the subsequent target was presented after a 250-ms delay. A correct trial consisted of a sequence of 10 target acquisitions after which a juice reward was given. All data analyzed in this study were obtained from the task condition in which the monkey repeated a deterministic

sequence of three movements three times (i.e., a single trial was 3 repeats of a 3 target sequence, for example, ABCABCABCA as shown in Fig. 1), with the first target of the sequence repeated at the end of the sequence. A new target sequence was selected randomly for each recording session. The minimum number of trials analyzed was 152, and the average was 267. Because each movement was repeated three times in each trial, ≥ 456 repetitions of each movement were available for analysis, with an average of 801.

Data preprocessing

We analyzed the responses of 193 pairs and 19 ensembles of simultaneously recorded neurons. The data for each trial were split into epochs corresponding to each of the ten movements of the sequence. Data from the first movement were not considered because they followed the inter-trial interval and varied from trial to trial. Neural activity in the period from 0 to 200 ms after target onset was used to predict the target toward which the animal was about to reach. Previously, we found that the optimal classification accuracy was obtained when we split the 200-ms epoch into 3 bins of 66-ms duration (Averbeck and Lee 2003). Thus the same three bins were used for the analyses shown in Figs. 4 and 5. For the analyses in which a single 66-ms window was considered (Figs. 7 and 8), the final 66 ms of the 200-ms epoch (i.e., from 134 to 200 ms) was used because this tended to contain the most information.

In our previous work (Averbeck and Lee 2003), the neural responses were used to predict one of the three possible movement targets (e.g., A, B, or C in Fig. 1). In this paper, we decoded targets in pairs because this is the case described by *d prime*, as discussed in the following text. Pairwise analyses of the three targets resulted in three separate analyses for each set of neural responses considered. For most of the analyses in the present study, we averaged the results from the three different pairs of targets. For the decoding analysis carried out on ensembles of neurons (Fig. 5), we show the results separately for each target pair because a relatively small number of ensembles were available for that analysis.

Analysis of *d prime*, encoding, and decoding

In the present study, we used d^2 to measure information. Figure 2A shows the components of d^2 , which is defined as

$$d^2 = \frac{(\mu_2 - \mu_1)^2}{\sigma^2} = \frac{\Delta\mu^2}{\sigma^2} \quad (1)$$

where μ_i indicates the mean spike count of a neuron to target i and σ^2 is the variance of spike count. This plot shows that d^2 is a measure of the discriminability of samples from two Gaussian response distributions, characterized by the same variance and different means. In our case, these response distributions correspond to the spike counts of a single neuron for two different movement directions. In RESULTS, because we always considered the responses of more than one neuron, we used a multivariate generalization (Poor 1994), given by

$$d^2 = \Delta\mu^T Q^{-1} \Delta\mu \quad (2)$$

where $\Delta\mu$ is the vector difference in mean responses to the pair of targets and Q is the pooled or average covariance matrix with the

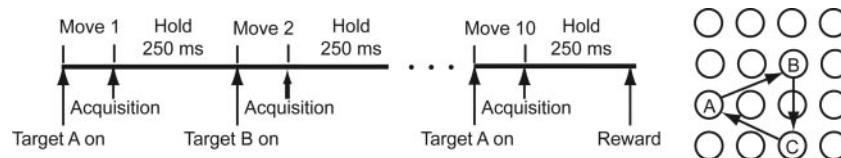


FIG. 1. Task. *Left*: time sequence of events in a trial. *Right*: movements for an example sequence. The trial starts when 1 of the targets is illuminated, for example, A. After acquisition of a target (i.e., when the monkey touches the corresponding location on the touch-screen), there is a 250-ms delay, and then the next target, in this case B, is illuminated, and the monkey has to move to B. This pattern continues until the monkey has completed the sequence of 3 movements 3 times.

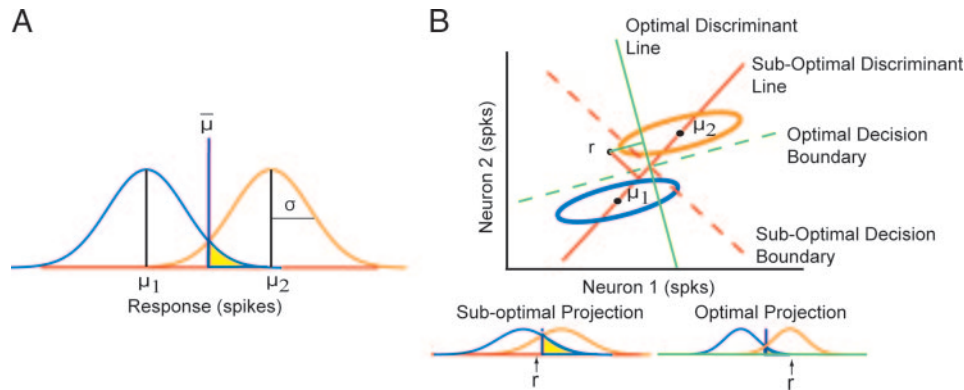


FIG. 2. Illustration of d^2 . *A*: d^2 measures the separability of 2 Gaussian distributions. The portion of the left distribution that is shaded in yellow would be incorrectly classified using the optimal decision boundary indicated at the midpoint between the distributions. Therefore an estimate of the amount of the left distribution lying to the right of the midpoint provides an estimate of the percentage of movements which would be misclassified. *B*: projection of multivariate response distributions onto a linear discriminant line (solid line). The dotted lines represent corresponding decision boundaries. The green lines correspond to the optimal decision boundary and the red lines to the sub-optimal decision boundary obtained by ignoring correlations. Classification is carried out by projecting responses (example r) onto the discriminant line perpendicular to the decision boundary.

average taken across targets. The dimensionality of $\Delta\mu$ and Q depends on the specific analysis. For example, when pairs of neurons and three time bins are analyzed, the dimensionality is 6. Multiplying by the inverse covariance matrix is analogous to dividing by the variance in Eq. 1. We also used two other measures that, when combined with d^2 , provide an estimate of the effect of noise correlations on information encoding and decoding. The first is a measure of the information that would be contained in the neural responses if they were uncorrelated. We call this d^2_{shuffled} because shuffling trials is often used to destroy correlations in neurophysiological data. It is defined as

$$d^2_{\text{shuffled}} = \Delta\mu^T Q_d^{-1} \Delta\mu \quad (3)$$

where Q_d is the diagonal covariance matrix obtained by setting the off-diagonal elements corresponding to correlations between neurons to 0. Finally, we defined d^2_{diag} as

$$d^2_{\text{diag}} = \frac{(\Delta\mu^T Q_d^{-1} \Delta\mu)^2}{\Delta\mu^T Q_d^{-1} Q_d^{-1} \Delta\mu} \quad (4)$$

This measures the amount of information that would be extracted by using a decoding algorithm that ignored correlations on the original *unshuffled* dataset. We refer to this as d^2_{diag} because it amounts to assuming a diagonal covariance matrix for the neural responses when deriving the decoding algorithm. In this case, the decoding algorithm is suboptimal. This quantity can be derived by computing the variance of the linear decoder obtained by ignoring correlations with respect to the real response distribution. It was derived for Fisher Information by Wu et al. (2001) as a local linear approximation. In our case, the formula is exact because the difference in the mean responses is necessarily linear.

To understand what is being measured by d^2_{diag} , we can consider the process of linear decoding. When carrying out linear decoding analyses, one projects the neural response vector onto a discriminant line that is perpendicular to the linear decision boundary (see Fig. 2*B*). This results in a scalar decision variable that is compared with a threshold to make the classification decision. In Fig. 2*B*, we show an example of a response r , projected onto both an optimal and a suboptimal discriminant line. The projection onto both lines is shown at the bottom of the plot. If this decision variable is compared with the decision boundary that bisects the distributions, the response will be properly classified in the optimal case because it was actually generated by target 2 but misclassified in the suboptimal case. If we project the entire distribution of responses onto the corresponding discriminant lines, we get the decision variable distributions shown at the bottom of Fig. 2*B*. The overlap in these distributions can be related to d^2 . Because the optimal decision boundary is defined as the one that

maximally separates the distributions, projection onto a sub-optimal decision boundary, for example one derived by ignoring correlations, always results in worse classification performance, in theory. Correspondingly, d^2_{diag} as it would be measured in the distribution on the left, cannot be larger than d^2 , as it would be measured in the distribution on the right, and decoding algorithms that ignore correlations cannot lead to fewer misclassifications.

Given d^2 , d^2_{shuffled} , and d^2_{diag} , we can estimate the effects of noise correlations on information encoding and decoding. The effect of noise correlations on decoding, which we will refer to as Δd^2_{diag} , is given by

$$\Delta d^2_{\text{diag}} = d^2 - d^2_{\text{diag}} \quad (5)$$

This quantity estimates the difference between the total amount of information that could be extracted from the neural responses using an optimal decoder, and the amount of information that would be extracted by a decoding algorithm which ignored correlations. Because information can only be lost by a decoding algorithm that ignores correlations, Δd^2_{diag} is always positive. Similarly, the effect of noise correlations on the information encoded, $\Delta d^2_{\text{shuffled}}$, is given by

$$\Delta d^2_{\text{shuffled}} = d^2 - d^2_{\text{shuffled}} \quad (6)$$

This quantity measures the difference in the information between the correlated neural responses, and the information that would be in a fictitious dataset of uncorrelated neural responses. $\Delta d^2_{\text{shuffled}}$ can be positive or negative.

The measures d^2 , d^2_{shuffled} , and d^2_{diag} , can each be converted to percent correct classification performance for the corresponding case (Poor 1994). In Fig. 2, the portion of the distribution for target 1 to the right of the classification boundary would be misclassified as having come from the distribution for target 2. Therefore we can write the probability of misclassification using the error function as

$$p(\hat{t} = 2 | t = 1) = (2\pi\sigma^2)^{-1/2} \int_{\bar{\mu}}^{\infty} \exp\left(-\frac{(x - \mu_1)^2}{2\sigma^2}\right) dx \quad (7)$$

where $\bar{\mu}$ refers to the decision boundary, \hat{t} is the predicted target, and t is the actual target. If we make the change of variables, $z = (x - \mu_1)/\sigma$, we get

$$p(\hat{t} = 2 | t = 1) = (2\pi)^{-1/2} \int_{\frac{\bar{\mu} - \mu_1}{\sigma}}^{\infty} \exp\left(-\frac{z^2}{2}\right) dz \quad (8)$$

Finally, noting that $\bar{\mu} - \mu_1 = (\mu_2 - \mu_1)/2$, we can rewrite Eq. 8 as the normalized error function, with a lower integration limit of $d/2$

$$p(\hat{t} = 2|t = 1) = (2\pi)^{-1/2} \int_{d/2}^{\infty} \exp\left(-\frac{z^2}{2}\right) dz = H(d/2) \quad (9)$$

Which shows that the fraction correct is only a function of d prime, where H is the complementary error function. From Eq. 9 we can calculate the percent correct performance or accuracy as

$$A = 1 - P(\hat{t} = 2|t = 1) = 1 - H(d/2) \quad (10)$$

Equation 10 shows that the percent correct is only a function of d^2 . Although Eq. 9 was derived for the univariate case, given by Eq. 1, it is the same for Eq. 2 as discussed in the preceding text. The predicted classification performance shown in the results was obtained by first calculating d^2 , d_{shuffled}^2 , and d_{diag}^2 per Eqs. 2–4 for each pair or ensemble of neurons. Then Eq. 10 was used to convert each of these to the corresponding decoding accuracy, denoted as A , A_{shuffled} , and A_{diag} . We then refer to the changes in accuracy related to the effect of noise correlations on encoding and decoding as $\Delta A_{\text{shuffled}}$ and ΔA_{diag} , respectively, and these were calculated as follows

$$\Delta A_{\text{shuffled}} = A - A_{\text{shuffled}} \quad (11)$$

$$\Delta A_{\text{diag}} = A - A_{\text{diag}} \quad (12)$$

Bhattacharyya distance

We also used the Bhattacharyya distance (BD) as an information measure (Basseville 1989) because it does not make the assumption of equal covariance matrices implicit in d^2 and thus provides a measure of the information in the differential variance and covariance of the neural response to different targets. The BD is a special case of the Chernoff distance, which was recently used in the analysis of V1 responses (Kang et al. 2004). The BD is given by

$$\text{BD}[p(r|t = 1), p(r|t = 2)] = -\log \int_r \sqrt{p(r|t = 1)p(r|t = 2)} dr \quad (13)$$

where the integral is over all possible responses, r . If we assume additive Gaussian noise, the response distributions are given by

$$p(r|t = i) = |2\pi Q_i|^{-\frac{1}{2}} \exp\left[-\frac{1}{2}(r - \mu_i)^T Q_i^{-1}(r - \mu_i)\right] \quad (14)$$

where r is a vector of spike counts for a given movement, μ_i is the vector of mean spike counts for target i , the superscript T indicates transpose, Q_i is the noise covariance matrix for target i , and $||$ indicates the determinant of the matrix. Substituting Eq. 14 into Eq. 13 leads to the BD for Gaussian distributions (Basseville 1989)

$$\text{BD}_G = \frac{1}{4} \Delta\mu^T(Q_1 + Q_2)^{-1}\Delta\mu + \frac{1}{2} \log \frac{|Q_1 + Q_2|}{2\sqrt{|Q_1 Q_2|}} \quad (15)$$

The first term of Eq. 15 is equal to $d^2/8$. Thus the second term indicates contributions due to the difference in covariance for different targets.

If covariance matrices vary between targets, the maximum likelihood estimator for the target is a quadratic Gaussian classifier (Johnson and Wichern 1998). This is referred to as a quadratic classifier because it contains terms that are products of firing rates between pairs of neurons. If the covariance matrices are the same across targets, the maximum likelihood estimator is a linear classifier, which does not contain interaction or product terms of the responses of individual neurons.

Decoding analyses

We compared the predicted percent correct classification performance based on d^2 (Eq. 10) to the results of carrying out decoding analyses and classifying the data movement by movement. Although it may seem counter-intuitive to use decoding analyses to estimate the effects of noise correlations on information encoding, an optimal decoder will extract all of the information available in the neural responses, and in our encoding analyses, we are trying to determine how much information is encoded. This was the approach adopted originally by Bialek and his colleagues (Bialek et al. 1991; Rieke et al. 1997). The effects of correlations on information decoding were examined using a sub-optimal decoder, specifically one that assumed that there were no noise correlations. In this case, the question is how much information is lost when the suboptimal decoder is used.

The Gaussian decoding analyses have been described in detail previously (Averbeck and Lee 2003). Two-fold cross validation was used whenever different decoding algorithms were being compared. In general, the target was predicted by selecting the target with the maximum probability from the conditional distribution of targets given the neural activity. This can be formalized as

$$\hat{t} = \arg \max_t p(t|r) \quad (16)$$

where \hat{t} is the estimated target for the subsequent movement and $p(t|r)$ is the conditional probability distribution of a target, t , given the response vector, r , that represents the response of one or more neurons across a given number of bins. The conditional probability of t is given by Bayes' rule

$$p(t|r) = \frac{p(r|t)p(t)}{p(r)} \quad (17)$$

where $p(t)$ is the prior probability of a given target, and $p(r)$ is a normalizing constant calculated as

$$p(r) = \sum_t p(r|t)p(t) \quad (18)$$

The likelihood for the Gaussian model is given by Eq. 14. We fit linear models by estimating a single, pooled covariance matrix for both targets, and we fit quadratic models by estimating separate covariance matrices for each target. These covariance matrices were the same as those used to calculate the predicted accuracy described in the preceding text, using d^2 .

The decoding analysis based on the linear Gaussian model was carried out under three conditions to derive values for the measured accuracy shown in Figs. 4 and 5. These conditions were in correspondence with the predicted accuracy, derived from d^2 , d_{shuffled}^2 , and d_{diag}^2 , given by Eq. 10. In this study, we only analyzed the effect of noise correlations, i.e., correlations between neurons, not autocorrelations which were analyzed in a previous study (Averbeck and Lee 2003). In the first condition, which was used as an estimate of A , the decoding analysis was carried out on the original dataset using a covariance matrix that used the measured values for the noise correlations. In the second condition, which was used as an estimate of A_{shuffled} , the analysis was carried out on a trial-shuffled dataset. The shuffling effectively destroyed the correlations between neurons. The shuffling analysis was carried out five times, and the average of these five analyses was used as the estimate. In the final analysis, a decoding model was used in which all off diagonal elements of the covariance matrix that correspond to inter-neuronal correlations were set to zero. This model was then applied to the original unshuffled dataset. This was used as an estimate of A_{diag} and corresponds to the independent model from our previous work (Averbeck and Lee 2003). Thus the decoding model was essentially the same for A_{shuffled} and A_{diag} because in both cases the off-diagonal elements of the covariance matrix were zero, but the decoders were applied to the shuffled trials

and the original datasets respectively. From these analyses, $\Delta A_{\text{shuffled}}$ and ΔA_{diag} were also calculated, per Eqs. 11 and 12.

We also used a multinomial decoding algorithm to determine how the results from the above analyses are influenced by the assumption of a Gaussian distribution for the neural responses. Because spike counts are discrete quantities, it was possible to tabulate the different responses for a pair of neurons, and generate a probability mass function for each target direction. For example, for target direction 1, how often did neuron 1 fire two spikes and neuron 2 fire one spike? A complete table of these probabilities specified the probability mass function for each target. In this case, the likelihood is estimated by

$$p(r_i|t=j) = \frac{n_{ij}}{N_j} \quad (19)$$

where n_{ij} is the number of times response r_i occurred when target j was presented and N_j is the number of times target j was presented. This is the same characterization of neural responses used in the direct method of estimating mutual information between neural responses and stimuli (Strong et al. 1998). These models provide the most detailed description of responses, at a given bin size, that is possible. It is important to note, however, that each different response type that occurs is a parameter of the model, and estimating these models for several response bins or multiple neurons is not possible without an extremely large dataset because the number of parameters necessary to estimate the models grows quickly. Therefore we estimated this model using a single 66-ms response bin for a pair of neurons. When we compared this model to the linear and quadratic models, those models were also fit on the same bin.

RESULTS

Noise correlations and information coding

The impact of noise correlations on information encoding and decoding can be compared by looking at $\Delta d^2_{\text{shuffled}}$ and Δd^2_{diag} , which measure the effect of noise correlations on information encoding and decoding respectively (see METHODS). $\Delta d^2_{\text{shuffled}}$ is defined as

$$\Delta d^2_{\text{shuffled}} = d^2 - d^2_{\text{shuffled}}$$

where d^2 is the information in the correlated neural responses, and d^2_{shuffled} is the information in the uncorrelated (shuffled) neural responses. The effects of noise correlations on information decoding have been studied by examining how much information is lost by a decoding algorithm that ignores correlations. Therefore Δd^2_{diag} is defined as

$$\Delta d^2_{\text{diag}} = d^2 - d^2_{\text{diag}}$$

where d^2_{diag} is the information that would be extracted by a decoding algorithm that ignores correlations, from the original correlated dataset. Thus Δd^2_{diag} measures the amount of information lost by a decoding algorithm that ignores correlations.

The relation between $\Delta d^2_{\text{shuffled}}$ and Δd^2_{diag} is shown in Fig. 3 (d^2 , d^2_{shuffled} , and d^2_{diag} are also shown for reference). In this figure, we used relatively large values for the correlation coefficient between neurons (0.6), so the effects of noise correlations can be clearly visualized. When noise correlations are smaller, the effects have the same periodicity, and are simply scaled down. Figure 3, B–D, shows three illustrative examples of the responses of two fictitious neurons to two different movement directions. The mean response vector for the corresponding movement direction is indicated as μ_i , and the ellipse represents the response distribution, or the variabil-

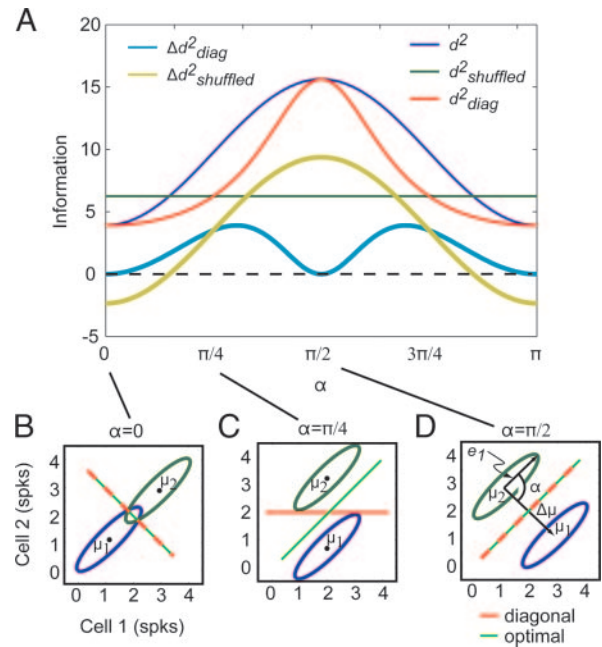


FIG. 3. Effects of noise correlations on encoding and decoding. A: components of d^2 vs. the angle α between e_1 , the eigenvector associated with the largest eigenvalue of the covariance matrix and $\Delta\mu$, the vector of the difference in mean responses. This shows how the elements of the information breakdown change as a function of α . For all calculations, the variance of both neurons was 1, and the correlation coefficient was 0.6. B–D: examples of covariance/signal structures. Horizontal and vertical axes represent the responses of 2 hypothetical neurons. Green lines represent optimal decision boundaries, red lines represent decision boundaries that would be obtained if correlations were ignored, referred to as diagonal. When the decision boundaries are the same (B and D), a dashed line is shown.

ity of the responses, for each movement direction. The ellipse also indicates the covariance of the pair of neurons because the orientation of the ellipse indicates whether the covariance is positive or negative. If the covariance was zero, i.e., no noise correlation, the ellipses would be circles because these neurons have the same variance. The primary axis of each ellipse is given by the eigenvector, e_1 (shown in Fig. 3D), which corresponds to the largest eigenvalue of the covariance matrix. The covariance in each panel of Fig. 3, B–D, is the same, as is the length of $\Delta\mu$, which defines the difference in the mean responses. However, α , the angle between e_1 and $\Delta\mu$, is different. As is shown in Fig. 3A, this is the key parameter for relating $\Delta d^2_{\text{shuffled}}$ and Δd^2_{diag} .

The effects of noise correlations on information encoding and decoding are different, and depend on the value of α . For example, for values of α near 0 or π , $\Delta d^2_{\text{shuffled}}$ is negative, whereas Δd^2_{diag} is zero. Thus when α is near zero, assessing the effect of correlations on the information encoded would suggest that they have a negative effect, whereas assessing the effect of correlations on information decoding would suggest no effect. Between approximately $\pi/7$ (a value that depends on the size of the correlation) and $\pi/2$, both $\Delta d^2_{\text{shuffled}}$ and Δd^2_{diag} are positive, whereas for $\alpha = \pi/2$, $\Delta d^2_{\text{shuffled}}$ takes on its largest value and Δd^2_{diag} is again zero. The effect of noise correlations on $\Delta d^2_{\text{shuffled}}$ can be seen by looking at the overlap of the response distributions. The more the distributions overlap, the less information encoded. To gain an intuition for the effect of noise correlations on Δd^2_{diag} , we have plotted the optimal decision boundaries in green, and the decision boundaries

derived by a decoding algorithm that ignored correlations (a diagonal decoding algorithm) in red. In Fig. 3, *B* and *D*, the optimal and the suboptimal (i.e., derived under the assumption of no correlation) decision boundaries are the same. Because the decision boundaries are the same in these cases, no information is lost by ignoring correlations, and Δd_{diag}^2 is zero. In Fig. 3*C*, however, the optimal and suboptimal decision boundaries are not the same, and so if the suboptimal decision boundary is used for classification, information will be lost, as indicated by the positive value of Δd_{diag}^2 at $\pi/4$.

Comparison of predicted and measured information

Implicit in the use of d^2 as an information measure is the assumption that the variance and covariance of the neural responses is the same for both targets. In this case, all of the information available in the neural responses can be extracted by computing the dot-product between the vector of spike counts of individual neurons, in time bins of the appropriate size (Averbeck and Lee 2003), and an appropriate weight vector. As shown in Fig. 3*C*, the weight vector can be affected by the correlations, i.e., it is not the same for correlated and uncorrelated neural responses. However, explicitly taking into account correlations between neurons by computing products or interactions between neural responses will not extract more information. Thus all of the information is available in the spike counts of neurons considered individually. This is a strong assumption about how information is encoded in neural responses. In the DISCUSSION, we consider the biophysical implications of this code.

As a first step toward examining how accurately d^2 predicts the information in neural responses, we compared the decoding

accuracy predicted by d^2 to the results of actually carrying out decoding analyses and classifying individual movements. Figure 4 shows the predicted and the measured values of the decoding accuracies for pairs of simultaneously recorded neurons in the supplementary motor area for each of the information measures shown in Fig. 3*A*. The predicted values of the accuracy, A , A_{shuffled} , and A_{diag} , were calculated by estimating d^2 , d_{shuffled}^2 , and d_{diag}^2 , based on the covariance, Q , and mean response vectors, $\Delta\mu_i$, estimated for each pair of neurons, and then converting these to estimates of percent correct classification performance or accuracy (see METHODS, Eqs. 9 and 10). The values plotted along the axis labeled *measured* accuracy were derived by explicitly classifying every movement using a linear decoding algorithm applied to the spike counts of pairs of neurons. Although d^2 is only a function of the covariance matrix and the difference in the mean response vectors, these measures predicted the outcome of actually carrying out the decoding analyses accurately (Fig. 4). The quantities $\Delta A_{\text{shuffled}}$ and ΔA_{diag} , which measure the effects of correlations on the classification performance, were also well approximated by the predicted values.

The histograms at the top of the scatter plots (Fig. 4, *bottom*) show the distribution of the corresponding *measured* accuracies. These plots show that, on average, correlations had almost no effect on the information encoded and only a small effect on decoding performance when correlations were ignored. The only distribution with a mean that was significantly deviated from zero was ΔA_{diag} (*t*-test, $P < 0.01$). Although the distribution was centered near zero, correlations did affect the information encoded in some cases, increasing it or decreasing it by up to a few percent. The negative values of *measured* accuracy for ΔA_{diag} are due to finite sampling and mismatches

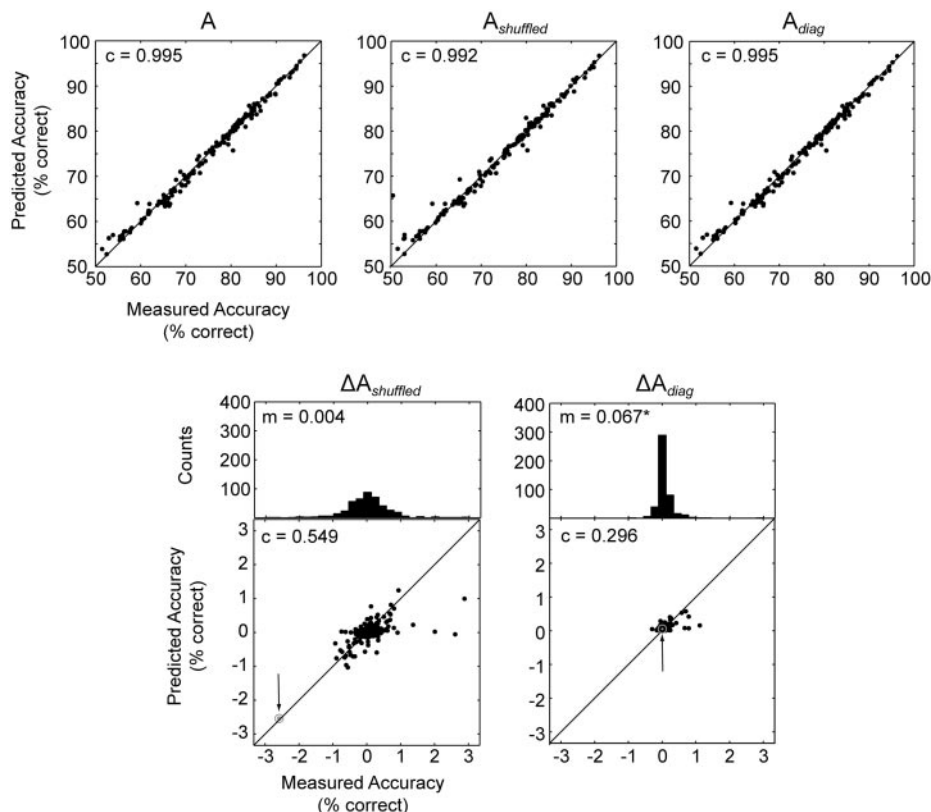


FIG. 4. Measured vs. predicted accuracy for pairs of neurons. Values in the top left corner of scatter plots are Spearman rank order correlation coefficients (c), and those in histograms are means (m) of distributions. The asterisk indicates that the mean was significantly different from 0 ($P < 0.05$). *Top*: measured (abscissa) and predicted (ordinate) accuracy values for A , A_{shuffled} , and A_{diag} respectively. *Bottom*: measured and predicted accuracies for $\Delta A_{\text{shuffled}}$ and ΔA_{diag} . Histograms of marginal distribution of decoded (measured) values are shown at the top of each plot. Data points associated with covariance matrices with negative eigenvalues, caused by numerical errors, were removed from the plots as these tended to be outliers. The outlined data points indicated by the arrows show a single pair of neurons for which the effect of noise correlations on information encoding (*left*) was relatively large, but the effect on decoding (*right*) was essentially 0.

between the linear model and the actual distribution of the data. As shown in Fig. 3, there could be a large effect of noise correlations on information encoding, when the effect on decoding is minimal. Similarly, the outlined data points indicated by the arrows in the plots of $\Delta A_{\text{shuffled}}$ and ΔA_{diag} in Fig. 4 show a pair of neurons for which the effect of noise correlations on information encoding was relatively large and the effect on information decoding was essentially zero.

To examine the role of noise correlations in ensembles of more than two neurons, the analyses were applied to groups of three to eight simultaneously recorded neurons (Fig. 5). As with the pair-wise analyses, the predicted and measured decoding performances again agreed closely. However, the largest effects of correlations at the ensemble level were larger than the effects for pairs of neurons (see histograms in Fig. 5, bottom). To ensure that this increase in the size of the effects of correlations was not due to the fact that we were fitting more complex models, we also re-ran the analysis using only half the data. The results were essentially the same (data not shown). Overall, correlations reduced the information encoded slightly, but the effect was small and the mean of the distribution was not significantly different from zero. Again the only distribution with a mean that was significantly deviated from zero was ΔA_{diag} (*t*-test, $P < 0.01$). These analyses show that noise correlations have a relatively small effect on either information encoding or decoding. The size of the effects will depend on the size of the noise correlations in the population. On average, noise correlations in our data are almost zero, although a few pairs do have correlations >0.2 (Fig. 6A). The signal correlations are somewhat larger, but close to zero on average (Fig. 6B).

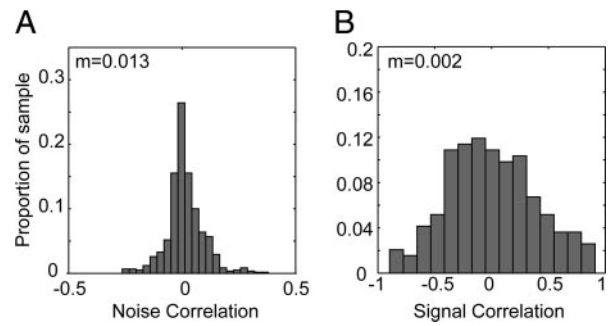


FIG. 6. Distributions of noise correlation (A) and signal correlation (B). The mean of the distribution is shown in the top left of each plot.

Contribution of target-related changes in covariance to encoded information

As described in the preceding text, even correlations that are the same for both targets can affect information encoding and decoding. In this case, the correlations themselves cannot be used directly to predict which target was presented. The correlations can, however, still affect information decoding by changing the optimal decision boundary used to estimate the target. For all the analyses described so far, it was assumed that the covariance of neural activity was identical for both targets. However, it is often observed that the variance of spike counts scales with the mean, and the covariances may change as well (Averbeck and Lee 2003; Tolhurst et al. 1981). This raises the possibility that additional information can be carried in the variances and the covariances of the neural responses. A classifier that explicitly computes interactions, a nonlinear operation, must be used to extract this information (Shamir and Sompolinsky 2004).

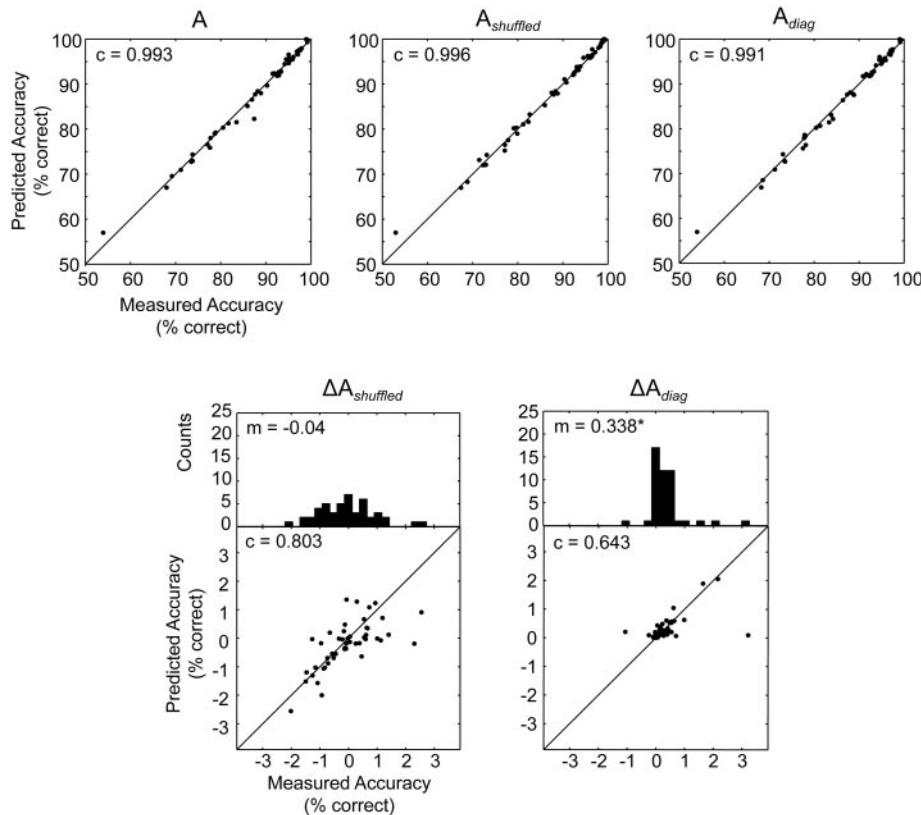


FIG. 5. Measured vs. predicted accuracy for ensembles of neurons. Conventions are as in Fig. 4.

We examined the possibility that changes in the covariances for different targets might carry additional information by computing the BD, and comparing it to d^2 . The BD has one term that is proportional to d^2 , and a second term that measures the amount of information in the covariances (see METHODS). Thus by comparing the BD to d^2 , we obtained an estimate of the additional information available in the covariances that cannot be extracted by a linear decoding algorithm. The results of this analysis suggested that there was additional information available in the covariances for many pairs of neurons (Fig. 7A). To test this further, we compared the classification performance of a nonlinear, quadratic classifier, which is a decoding model that can extract information from the target-dependent variances and covariances of the responses, to that of the linear classifier that uses only information available in the individual spike counts (Fig. 7B). In this analysis, only the accuracy in the unshuffled neural responses (A) was considered, and the analysis was applied to pairs of neurons to control model complexity. In contrast to the increase in information predicted by the BD, the performance of linear and nonlinear classifiers was essentially equivalent in most cases with the performance of the quadratic classifier being slightly worse on average than that of the linear classifier. We investigated possible reasons for the discrepancy between the information predicted by the BD and the actual performance of the decoding algorithms by looking at the decision boundaries produced by both classifiers (e.g., Fig. 7C). As expected, the classifica-

tion boundaries of both classifiers separated the data equally well despite the fact that the BD predicted suboptimal performance for the linear classifier. This is due to at least two factors. The first is that neural responses are discrete quantities (i.e., spike counts), so the exact position of the decision boundary does not affect the classification performance, i.e., whether the decision boundary is at 1.2 or 1.4 spikes does not matter because the responses never take any values between 1 and 2. Second, negative firing rates may not be properly classified by the linear model (see Fig. 7C), but they do not occur. Thus the BD does not predict the actual decoding performance accurately, as does d^2 .

Non-Gaussian, multinomial decoding algorithm

To further validate the results we obtained using the linear Gaussian decoding algorithm, we used a multinomial decoding algorithm. The multinomial decoding algorithm provides a general, assumption-free description of the neural responses, and as such allowed us to re-examine the effects of noise correlations on information encoding and decoding, without making the Gaussian assumption. We fit the multinomial model to the joint spike count distributions for pairs of neurons and a single 66-ms bin of neural responses. Comparison of the performance of the multinomial model to that of the linear model showed that the performance was similar, with a slight but not statistically significant advantage for the multinomial model (mean = 0.013%, paired t -test, $P = 0.4$; Fig. 8A). However, because the performance of these models was assessed with cross validation, the cases in which the multinomial performs better than the linear model are individually relevant. Interestingly, when the quadratic model outperformed the linear model for a given neuron pair, the multinomial model was also more likely to outperform the linear model (Fig. 8B). Thus in a small number of cases, extra information does appear to be available, beyond the spike counts of individual neurons. We also compared estimates of $\Delta A_{\text{shuffled}}$ and ΔA_{diag} obtained with the linear and multinomial decoding algorithms. As can be seen by the histograms in Fig. 8, C and E, the size of the effects of noise correlations on encoding and decoding assessed with the multinomial algorithm are similar to those assessed with the linear decoding algorithm (histograms in Fig. 4). Additionally, Fig. 8D shows that there is a fairly strong correspondence, on a pair by pair basis, between the multinomial and linear decoding algorithms for $\Delta A_{\text{shuffled}}$. However, the correspondence is poor for ΔA_{diag} . Thus the size of the effect of noise correlations on encoding and decoding is quite similar, independent of the algorithm with which it is assessed, but on a pair-by-pair basis, the estimates of the two algorithms differed somewhat for the effects on decoding.

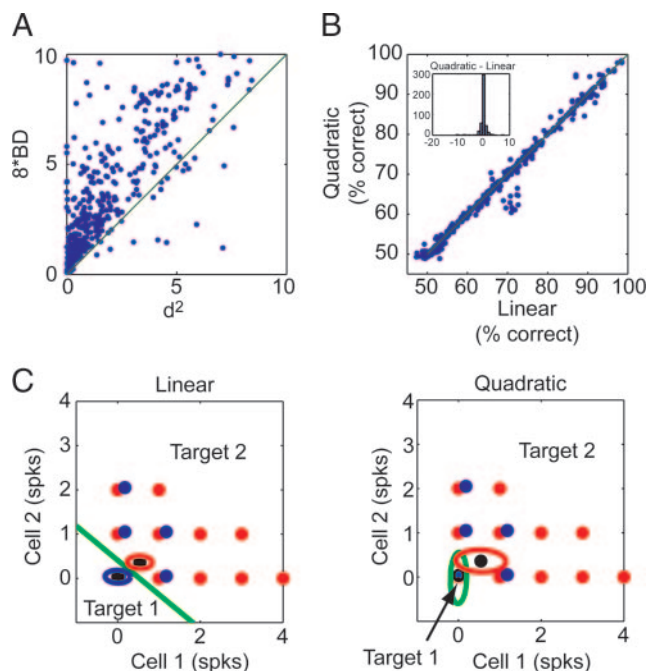


FIG. 7. A: Bhattacharyya distance (BD) vs. d^2 . We have plotted $8 \times \text{BD}$ because the 1st term of the BD is $d^2/8$. B: comparison of linear and quadratic classifiers. These comparisons are for the corresponding accuracy. Inset: histogram of differences between linear and quadratic classifiers. C: linear and quadratic decision boundaries (green) for a case in which BD predicted a large benefit of covariances, but the actual classification performance was the same for linear and quadratic decoders. The covariance ellipse for target 1 is indicated in blue and for target 2 in red; the mean of each distribution is indicated in black. The blue and red dots indicate individual responses for 2 targets with the blue dots shifted slightly rightward. In the linear case, the covariances are forced to be identical. In the quadratic case, the ellipse for target 1 is smaller than the response marker, and so is not visible.

Possibility of learning-related changes in noise correlation

The dataset used in this study was generated using a serial reaction-time task (Lee and Quessy 2003). This raises the possibility that learning-related changes in noise correlations could affect our decoding analyses. To examine this possibility, we divided each of our datasets into four parts, and compared the noise correlations in the first and last quarter of the data (Fig. 9). We found that there was a strong correlation between the noise correlations in the first and last blocks ($r =$

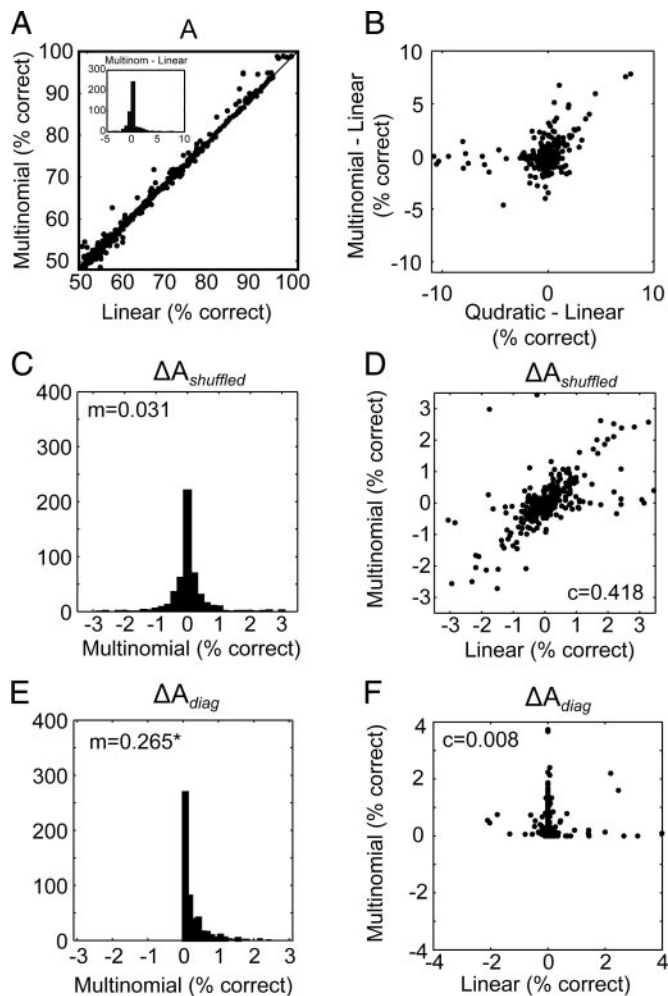


FIG. 8. Performance of multinomial classifier. *A*: decoding accuracy for linear vs. multinomial classifier. *Inset*: histogram of the difference between linear and multinomial classifiers. *B*: comparison of differences in decoding accuracy for linear vs. quadratic and for linear vs. multinomial classifiers. *C*: effect of noise correlations in information encoded, $\Delta A_{shuffled}$ assessed with the multinomial classifier. *D*: comparison of $\Delta A_{shuffled}$ assessed with the multinomial and linear classifiers. *E*: effect of noise correlations on information decoding, ΔA_{diag} assessed with the multinomial classifier. *F*: comparison of ΔA_{diag} for multinomial and linear classifiers.

0.933) but that the slope of the best fit line was not unity (95% confidence interval: 0.908–0.967). Thus there is a very small but significant difference in the noise correlations between the first and the last quarter of the data. To test how the effect of noise correlation on information coding is influenced by these small changes, we also computed $\Delta d_{shuffled}^2$ and Δd_{diag}^2 in the first and second half of the dataset for pairs of neurons. For this analysis, we used halves of the dataset to ensure a sufficient number of trials. The mean \pm SE ($n = 132$) for $\Delta d_{shuffled}^2$ were 0.336 ± 0.092 and $0.414 \pm 0.086\%$ for the first and second halves of the data, respectively, and the corresponding values for Δd_{diag}^2 were 0.074 ± 0.021 and $0.094 \pm 0.023\%$. The means for $\Delta d_{shuffled}^2$ were significantly different from zero, but the overall size of the effect was still quite small, with 95 and 93% of the distributions confined within $\pm 2\%$ for the first and second halves, respectively. Although this is slightly broader than the distribution shown in Fig. 4, reducing the sample size by dividing the dataset in half would be expected to broaden

the distribution. To examine this quantitatively, we generated a dataset with half as many trials by sampling randomly with replacement from the original dataset and calculated $\Delta d_{shuffled}^2$ in this bootstrapped dataset. The resulting distribution was somewhat broader with only 89% of the data between -2 and $+2\%$. Thus there is a slight shift in the mean of the distribution for $\Delta d_{shuffled}^2$ due to learning-related changes in neural activity or to nonstationarities in the neural responses not related to learning, but the overall width of the distribution, which is a measure of the largest positive and negative effects, appears to be a relatively stable feature of our dataset.

DISCUSSION

Noise correlations have been studied both empirically and theoretically using a variety of methods. One approach to the empirical study of noise correlations is to consider how much information is lost when neural responses are decoded using algorithms that assume that the response of individual neurons are uncorrelated. Theoretically, decoding algorithms can only do worse when correlations are ignored. Most of these studies have found that little information is lost when correlations are ignored (Averbeck and Lee 2003; Nirenberg et al. 2001; Oram et al. 2001), although some studies have shown that the effect of ignoring correlations can be larger (Dan et al. 1998; Maynard et al. 1999). Empirical analyses have also assessed the impact of noise correlations on information encoding. Romo and his colleagues (2003) have shown that for a subpopulation of their neurons correlations increase the amount of information encoded. As shown in this study (Fig. 3A), this occurs when the difference in the response vectors for different movements is nearly orthogonal to the longest axis of the noise covariance matrix, which was the case with the subset of the data considered by Romo et al. (2003). Another series of studies have examined the encoding effects of correlations using a decomposition of the Shannon information (Panzeri et al. 1999), which can be related to the decomposition used in this study. Their approach also considered the total effect of correlations, as well as splitting the correlation into two terms, one of which is related to Δd_{diag}^2 . Similar to our finding, which is based on neural activity recorded in the supplementary motor area, these studies have shown that correlations in pairs of neurons carry relatively little information in V1 (Golledge et

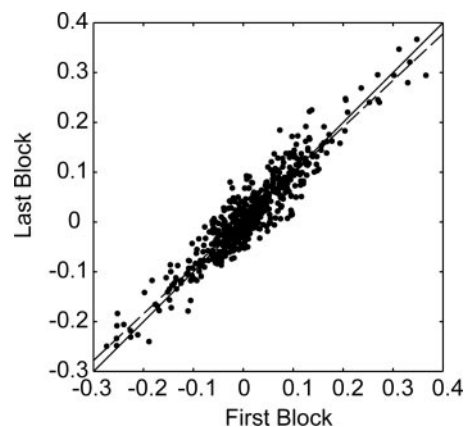


FIG. 9. Comparison of noise correlation in 1st and last block of data. The best fit line is shown (---; slope = 0.94) as well as the line with unit slope (—).

al. 2003), rat barrel cortex (Petersen et al. 2002), and inferior-temporal cortex (Rolls et al. 2003). Thus studies based on the analyses of pair of neurons have consistently demonstrated that the role of correlations in information coding is limited (Averbeck and Lee 2004).

Studying correlations from both the encoding and the decoding perspectives are useful. Assessing the effects of noise correlation on information encoding is valuable for at least two reasons. The first is to check predictions for the amount of information contained in the responses of populations of neurons, based on the recording of single neurons. If neurons are indeed independent, then extrapolation from single-cell recording studies, which by definition cannot estimate the effects of noise correlations, are valid. If, however, neurons are not independent, then these extrapolations are not valid. Second, information maximization models of information coding in the cortex often ignore correlations (Bell and Sejnowski 1995; Hyvarinen et al. 2001; Olshausen and Field 1996). However, maximizing the information contained in the responses of a population of neurons is not simply a matter of optimizing the mean responses of a population of neurons. It also requires the optimization of the distribution of the noise in the neural responses. Assessing the effects of noise on information encoding in small ensembles of neurons is a first step toward assessing the effects of noise in larger populations.

Consideration of the effect of noise correlations on decoding, in addition to the effects on encoding, is also valuable for several reasons. First, understanding the impact of noise correlations on information decoding is important in the design of algorithms for driving neural prosthetic devices (Musallam et al. 2004; Taylor et al. 2002). Second, insights into the biophysical or network mechanisms that would be necessary to extract all of the information from spike trains of upstream neurons can be gained from studying Δd_{diag}^2 . To carry out computation, the brain has to solve the same computational problem faced by our decoding algorithm. If Δd_{diag}^2 is small, the neural responses can be decoded reliably by assuming the upstream neurons are conditionally independent. This simplifies the computational task of defining the optimal decision boundaries, which presumably simplifies the problem to be solved by the biological system. When the decoding problem is considered from a more general, probabilistic perspective, estimation of the full joint distribution of neural responses is considerably simplified if the distribution can be factorized. This implies that the neurons can be considered conditionally independent for purposes of decoding. These simplifications are the basis for the recent success of graphical models (Jordan and Sejnowski 2001). Furthermore, if all the information in the neural responses resides in the spike counts of individual neurons, i.e., if d^2 describes all of the information available in the responses, they can be decoded linearly. This might obviate the need for computational machineries at the single neuron or network level that combine inputs nonlinearly. For example, if dendritic arbors combine their inputs linearly, they would not be able to extract information from differential covariances. Some results have suggested that dendritic arbors process their inputs relatively linearly (Cash and Yuste 1998, 1999), whereas others have shown that some level of nonlinearity can be found (Koch 1998; Margulis and Tang 1998; Nettleton and Spain 2000). In general, however, it is unlikely that the brain is limited to linear computations. Understanding the relation between the features

of neural responses that carry information and the processing capabilities of dendritic arbors and networks will provide important converging perspectives for understanding the neural code.

When considering whether or not correlations have an effect, studying information encoding and information decoding can lead to different answers (Fig. 3). For some pairs of simultaneously recorded neurons in the supplementary motor area, we found that noise correlations affected the information encoded. However, the effects were relatively small, and averaged across the population, the mean effect was not significantly different from zero. The effect of noise correlations on information decoding was similar in magnitude to the effect on information encoding. Although the mean of the distribution of effects on decoding was significantly different from zero, this term is in principle nonnegative, so this result is not surprising. At the ensemble level, the effects of noise correlations were somewhat larger, but the average effect for information encoding was again not significantly different from zero. It is important to point out that, as we and others have shown before (Averbeck and Lee 2003; Constantinidis and Goldman-Rakic 2002; Reich et al. 2001), measured correlations, and correspondingly the size of the effect of correlations, depend on the bin size used for their estimation. This is due to the fact that cross-covariances between neurons are much stronger at low frequencies and thus larger bins show stronger correlations between neurons (Averbeck and Lee 2004). We have chosen 66 ms in this study because our previous work (Averbeck and Lee 2003) showed that this was the optimum bin size for information extraction. The correlations we observed were similar in size to those that have been observed in other studies (Constantinidis and Goldman-Rakic 2002; Reich et al. 2001). Therefore a relatively small effect of noise correlation on information coding found in the present study may generalize to other brain areas and task conditions, although this remains to be investigated in future studies.

Implicit in the use of d^2 is the assumption that the conditional response distributions of the neurons are Gaussian, and have the same variance for different targets. Theoretically, this is a strong limitation of using d^2 as an information measure because the variance of neural responses tends to scale with the mean response (Averbeck and Lee 2003; Tolhurst et al. 1981; Werner and Mountcastle 1963), and response distributions are at best truncated Gaussians unless spike rates are high (Wiener and Richmond 1999). We have shown that the predicted decoding performance derived from d^2 closely matched the actual decoding performance of a linear decoding model. We have also shown, through a series of analyses, that linear decoding models can generally extract almost all of the information available in the neural responses. More general decoding models that assumed that variances can change for different targets, as well as a very general multinomial decoding model, were only able to do marginally better than the linear decoding model. The major discrepancy we found was between ΔA_{diag} measured with the linear and the multinomial decoding algorithms. Although the relative magnitude of the effects was similar across our population, the two decoding algorithms did not agree on a pair by pair basis. Continued investigation of the limitations of the Gaussian assumption will be important because most current theoretical models of information coding in the cortex make this assumption (Abbott and Dayan 1999;

Shamir and Sompolinsky 2004; Sompolinsky et al. 2001; Wilke and Eurich 2002; Wu et al. 2001).

Another information measure often computed on the responses of pairs of neurons is synergy/redundancy (Averbeck et al. 2003; Gawne and Richmond 1993; Latham and Nirenberg 2005; Narayanan et al. 2005; Puchalla et al. 2005; Schneidman et al. 2003). There are two important differences between this measure and our measures of the effects of noise correlations on encoding and decoding. The first is that synergy/redundancy is a function of both noise correlations and signal correlations. Specifically, even if noise correlations are zero, there can and likely will be redundancy in neural responses. Given that noise correlations play a relatively small role in information coding for pairs of neurons, and measured redundancy is normally large, signal correlations are presumably responsible for the reported redundancy effects. It has been shown that $\Delta d_{\text{shuffled}}^2$ can be small when there are large redundancies (Averbeck et al. 2003). Furthermore, the redundancy is largely a function of the finite entropy of discrete Shannon information, since discrete Shannon information saturates. We would get a similar effect if we calculated a synergy/redundancy statistic on the percent correct classification, which of course saturates at 100%. For example if we computed the statistic

$$\Delta A_{\text{synergy}} = A_{1,2} - (A_1 + A_2)$$

where A_1 is the percent correct classification for neuron 1, A_2 is the percent correct classification for neuron 2, and $A_{1,2}$ is the joint percent correct classification. If A_1 and A_2 individually perform at a 90% classification rate, but $A_{1,2}$ is only at 98%, the responses would be considered redundant. This could be true even if there were no noise correlations. However, if there are no noise correlations, a statistic based on d^2 , for example

$$\Delta d_{\text{synergy}}^2 = d_{1,2}^2 - (d_1^2 + d_2^2)$$

will be zero. Thus it seems to us that if synergy/redundancy is being calculated, the separate effects of signal and noise correlations should be examined, as has been done in some studies (Gawne and Richmond 1993; Gawne et al. 1996; Panzeri et al. 1999; Petersen et al. 2001; Pola et al. 2003).

In conclusion, noise correlations can differentially affect information encoding and decoding. Both perspectives are useful, and they address different questions about the nature of the neural code. We found that in general, the effects of noise correlations were relatively small in our population of SMA neurons. However, we have only considered analyses in small ensemble of neurons, and theoretical work (Shamir and Sompolinsky 2004) suggests that small effects of noise correlations in pairs of neurons can become substantial in large populations. Furthermore, because the output of a system cannot contain more information than the input, correlations must ultimately limit information, if the number of neurons becomes sufficiently large (Narayanan et al. 2005; Seriès et al. 2004). Consistent with this, the effects of correlations in our study were slightly larger in ensembles than in pairs of neurons. Perhaps this is also an explanation for the saturation effects seen in studies related to neural prosthetics that have attempted to use relatively small ensembles to decode hand kinematics (Averbeck et al. 2005; Paninski et al. 2004; Wessberg et al. 2000). Future studies with larger populations of neurons will help to answer these questions empirically.

ACKNOWLEDGMENTS

We thank S. Quessy for help with the experiment and A. Pouget for extensive conversations that led to many of the results in the paper.

GRANTS

This work was supported by National Institutes of Health Grants R01-MH-59216, T32-MH-19942, and P30-EY-01319.

REFERENCES

- Abbott LF and Dayan P.** The effect of correlated variability on the accuracy of a population code. *Neural Comput* 11: 91–101, 1999.
- Arndt C.** *Information Measures*. Berlin: Springer-Verlag, 2001.
- Averbeck BB, Chafee MV, Crowe DA, and Georgopoulos AP.** Parietal representation of hand velocity in a copy task. *J Neurophysiol* 93: 508–518, 2005.
- Averbeck BB, Crowe DA, Chafee MV, and Georgopoulos AP.** Neural activity in prefrontal cortex during copying geometrical shapes. II. Decoding shape segments from neural ensembles. *Exp Brain Res* 150: 142–153, 2003.
- Averbeck BB and Lee D.** Neural noise and movement-related codes in the macaque supplementary motor area. *J Neurosci* 23: 7630–7641, 2003.
- Averbeck BB and Lee D.** Coding and transmission of information by neural ensembles. *Trends Neurosci* 27: 225–230, 2004.
- Basseville M.** Distance measures for signal processing and pattern recognition. *Signal Process* 18: 349–369, 1989.
- Bell AJ and Sejnowski TJ.** An information-maximization approach to blind separation and blind deconvolution. *Neural Comput* 7: 1129–1159, 1995.
- Bialek W, Rieke F, de Ruyter van Steveninck RR, and Warland D.** Reading a neural code. *Science* 252: 1854–1857, 1991.
- Brown EN, Frank LM, Tang D, Quirk MC, and Wilson MA.** A statistical paradigm for neural spike train decoding applied to position prediction from ensemble firing patterns of rat hippocampal place cells. *J Neurosci* 18: 7411–7425, 1998.
- Brown EN, Kass RE, and Mitra PP.** Multiple neural spike train data analysis: state-of-the-art and future challenges. *Nat Neurosci* 7: 456–461, 2004.
- Casella G and Berger RL.** *Statistical Inference*. Belmont, CA: Duxbury, 1990.
- Cash S and Yuste R.** Input summation by cultured pyramidal neurons is linear and position-independent. *J Neurosci* 18: 10–15, 1998.
- Cash S and Yuste R.** Linear summation of excitatory inputs by CA1 pyramidal neurons. *Neuron* 22: 383–394, 1999.
- Constantinidis C and Goldman-Rakic PS.** Correlated discharges among putative pyramidal neurons and interneurons in the primate prefrontal cortex. *J Neurophysiol* 88: 3487–3497, 2002.
- Dan Y, Alonso JM, Usrey WM, and Reid RC.** Coding of visual information by precisely correlated spikes in the lateral geniculate nucleus. *Nat Neurosci* 1: 501–507, 1998.
- Gawne TJ, Kjaer TW, Hertz JA, and Richmond BJ.** Adjacent visual cortical complex cells share about 20% of their stimulus-related information. *Cereb Cortex* 6: 482–489, 1996.
- Gawne TJ and Richmond BJ.** How independent are the messages carried by adjacent inferior temporal cortical neurons? *J Neurosci* 13: 2758–2771, 1993.
- Golledge HD, Panzeri S, Zheng F, Pola G, Scannell JW, Giannikopoulos DV, Mason RJ, Tovee MJ, and Young MP.** Correlations, feature-binding and population coding in primary visual cortex. *Neuroreport* 14: 1045–1050, 2003.
- Hyvarinen A, Karhunen J, and Oja E.** *Independent Component Analysis*. New York: Wiley, 2001.
- Johnson KO.** Sensory discrimination: decision process. *J Neurophysiol* 43: 1771–1792, 1980.
- Johnson RA and Wichern DW.** *Applied Multivariate Statistical Analysis*. Saddle River, NJ: Prentice Hall, 1998.
- Jordan MI and Sejnowski TJ.** Graphical models. In: *Foundations of Neural Computation* (1st ed.). Cambridge, MA: MIT Press, 2001, p. 421.
- Kang K, Shapley RM, and Sompolinsky H.** Information tuning of populations of neurons in primary visual cortex. *J Neurosci* 24: 3726–3735, 2004.
- Koch C.** *Biophysics of Computation*. Oxford: Oxford, 1998.
- Latham PE and Nirenberg S.** Synergy, redundancy, and independence in population codes, revisited. *J Neurosci* 25: 5195–5206, 2005.
- Lee D, Port NL, Kruse W, and Georgopoulos AP.** Variability and correlated noise in the discharge of neurons in motor and parietal areas of the primate cortex. *J Neurosci* 18: 1161–1170, 1998.

- Lee D and Quesy S.** Activity in the supplementary motor area related to learning and performance during a sequential visuomotor task. *J Neurophysiol* 89: 1039–1056, 2003.
- Margulis M and Tang CM.** Temporal integration can readily switch between sublinear and supralinear summation. *J Neurophysiol* 79: 2809–2813, 1998.
- Maynard EM, Hatsopoulos NG, Ojakangas CL, Acuna BD, Sanes JN, Normann RA, and Donoghue JP.** Neuronal interactions improve cortical population coding of movement direction. *J Neurosci* 19: 8083–8093, 1999.
- Musallam S, Corneil BD, Greger B, Scherberger H, and Andersen RA.** Cognitive control signals for neural prosthetics. *Science* 305: 258–262, 2004.
- Narayanan NS, Kimchi EY, and Laubach M.** Redundancy and synergy of neuronal ensembles in motor cortex. *J Neurosci* 25: 4207–4216, 2005.
- Nettleton JS and Spain WJ.** Linear to supralinear summation of AMPA-mediated EPSPs in neocortical pyramidal neurons. *J Neurophysiol* 83: 3310–3322, 2000.
- Nicolelis MA, Fanselow EE, and Ghazanfar AA.** Hebb's dream: the resurgence of cell assemblies. *Neuron* 19: 219–221, 1997.
- Nirenberg S, Carcieri SM, Jacobs AL, and Latham PE.** Retinal ganglion cells act largely as independent encoders. *Nature* 411: 698–701, 2001.
- Olshausen BA and Field DJ.** Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature* 381: 607–609, 1996.
- Oram MW, Hatsopoulos NG, Richmond BJ, and Donoghue JP.** Excess synchrony in motor cortical neurons provides redundant direction information with that from coarse temporal measures. *J Neurophysiol* 86: 1700–1716, 2001.
- Paninski L, Fellows MR, Hatsopoulos NG, and Donoghue JP.** Spatiotemporal tuning of motor cortical neurons for hand position and velocity. *J Neurophysiol* 91: 515–532, 2004.
- Panzeri S, Pola G, Petroni F, Young MP, and Petersen RS.** A critical assessment of different measures of the information carried by correlated neuronal firing. *Biosystems* 67: 177–185, 2002.
- Panzeri S and Schultz SR.** A unified approach to the study of temporal, correlational, and rate coding. *Neural Comput* 13: 1311–1349, 2001.
- Panzeri S, Schultz SR, Treves A, and Rolls ET.** Correlations and the encoding of information in the nervous system. *Proc R Soc Lond B Biol Sci* 266: 1001–1012, 1999.
- Petersen RS, Panzeri S, and Diamond ME.** Population coding of stimulus location in rat somatosensory cortex. *Neuron* 32: 503–514, 2001.
- Petersen RS, Panzeri S, and Diamond ME.** Population coding in somatosensory cortex. *Curr Opin Neurobiol* 12: 441–447, 2002.
- Pola G, Thiele A, Hoffmann KP, and Panzeri S.** An exact method to quantify the information transmitted by different mechanisms of correlational coding. *Network* 14: 35–60, 2003.
- Poor HV.** *An Introduction to Signal Detection and Estimation.* New York: Springer, 1994.
- Puchalla JL, Schneidman E, Harris RA, and Berry MJ.** Redundancy in the population code of the retina. *Neuron* 46: 493–504, 2005.
- Reich DS, Mechler F, and Victor JD.** Independent and redundant information in nearby cortical neurons. *Science* 294: 2566–2568, 2001.
- Rieke F, Warland D, de Ruyter van Steveninck R, and Bialek W.** *Spikes: Exploring the Neural Code.* Cambridge, MA: MIT Press, 1997.
- Rolls ET, Franco L, Aggelopoulos NC, and Reece S.** 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.
- Romo R, Hernandez A, Zainos A, and Salinas E.** Correlated neuronal discharges that increase coding efficiency during perceptual discrimination. *Neuron* 38: 649–657, 2003.
- Schneidman E, Bialek W, and Berry MJ.** 2nd. Synergy, redundancy, and independence in population codes. *J Neurosci* 23: 11539–11553, 2003.
- Seriès P, Latham PE, and Pouget A.** Tuning curve sharpening for orientation selectivity: coding efficiency and the impact of correlations. *Nat Neurosci* 7: 1129–1135, 2004.
- Shamir M and Sompolinsky H.** Nonlinear population codes. *Neural Comput* 16: 1105–1136, 2004.
- Snippe HP and Koenderink JJ.** Information in channel-coded systems: correlated receivers. *Biol Cybern* 67: 183–190, 1992.
- Sompolinsky H, Yoon H, Kang K, and Shamir M.** Population coding in neuronal systems with correlated noise. *Phys Rev E* 64: 051904, 2001.
- Strong SP, Koberle R, de Ruyter van Steveninck RR, and Bialek W.** Entropy and information in neural spike trains. *Phys Rev Lett* 80: 197–200, 1998.
- Taylor DM, Tillery SI, and Schwartz AB.** Direct cortical control of 3D neuroprosthetic devices. *Science* 296: 1829–1832, 2002.
- Thomson EE and Kristan WB.** Quantifying stimulus discriminability: a comparison of information theory and ideal observer analysis. *Neural Comput* 17: 741–778, 2005.
- Tolhurst DJ, Movshon JA, Thompson ID, and Thompson PG.** The dependence of response amplitude and variance of cat visual cortical neurones on stimulus contrast. *Exp Brain Res* 41: 414–419, 1981.
- Truccolo W, Eden UT, Fellows MR, Donoghue JP, and Brown EN.** A point process framework for relating neural spiking activity to spiking history, neural ensemble, and extrinsic covariate effects. *J Neurophysiol* 93: 1074–1089, 2005.
- Werner G and Mountcastle VB.** The variability of central neural activity in a sensory system, and its implications for the central reflections of sensory events. *J Neurophysiol* 26: 958–977, 1963.
- Wessberg J, Stambaugh CR, Kralik JD, Beck PD, Laubach M, Chapin JK, Kim J, Biggs SJ, Srinivasan MA, and Nicolelis MA.** Real-time prediction of hand trajectory by ensembles of cortical neurons in primates. *Nature* 408: 361–365, 2000.
- Wiener MC and Richmond BJ.** Using response models to estimate channel capacity for neuronal classification of stationary visual stimuli using temporal coding. *J Neurophysiol* 82: 2861–2875, 1999.
- Wilke SD and Eurich CW.** Representational accuracy of stochastic neural populations. *Neural Comput* 14: 155–189, 2002.
- Wu S, Nakahara H, and Amari S.** Population coding with correlation and an unfaithful model. *Neural Comput* 13: 775–797, 2001.
- Zohary E, Shadlen MN, and Newsome WT.** Correlated neuronal discharge rate and its implications for psychophysical performance. *Nature* 370: 140–143, 1994.