Neuron
Article

An Optimal Decision Population Code that Accounts for Correlated Variability Unambiguously Predicts a Subject's Choice

Federico Carnevale,[1] Victor de Lafuente,[2] Ranulfo Romo,[3,4,*] and Néstor Parga[1]
[1]Departamento de Física Teórica, Universidad Autónoma de Madrid, Cantoblanco, 28049 Madrid, Spain
[2]Instituto de Neurobiología, Universidad Nacional Autónoma de México, 76230 Querétaro, México
[3]Instituto de Fisiología Celular-Neurociencias, Universidad Nacional Autónoma de México, 04510 México, DF, México
[4]El Colegio Nacional, 06020 México, DF, México
*Correspondence: rromo@ifc.unam.mx
http://dx.doi.org/10.1016/j.neuron.2013.09.023

SUMMARY

Decisions emerge from the concerted activity of neuronal populations distributed across brain circuits. However, the analytical tools best suited to decode decision signals from neuronal populations remain unknown. Here we show that knowledge of correlated variability between pairs of cortical neurons allows perfect decoding of decisions from population firing rates. We recorded pairs of neurons from secondary somatosensory (S2) and premotor (PM) cortices while monkeys reported the presence or absence of a tactile stimulus. We found that while populations of S2 and sensory-like PM neurons are only partially correlated with behavior, those PM neurons active during a delay period preceding the motor report predict unequivocally the animal's decision report. Thus, a population rate code that optimally reveals a subject's perceptual decisions can be implemented just by knowing the correlations of PM neurons representing decision variables.

INTRODUCTION

When decisions are based on sensory evidence, decision-related signals evolve across sensory and frontoparietal cortices (for reviews, see Gold and Shadlen, 2001; Romo and de Lafuente, 2013; Romo and Salinas, 2003). The involvement of single neurons in decision-making processes is usually studied in terms of the choice probability (CP) index, a measure of covariation between a neuron's firing rate activity and the subject's choice (Britten et al., 1996; Green and Swets, 1966). In the brain, however, decisions engage multiple pools of neurons distributed across brain areas (de Lafuente and Romo, 2006; Hernández et al., 2010; Romo and de Lafuente, 2013). Hence, if one is to decode behavioral choices, the relevant measurements must come from population variables constructed from the spiking activity of multiple neuronal pools.

To understand how decisions emerge, one must first define proper measures to quantify how population activity covaries with behavior. It is known that firing rates vary stochastically from trial to trial (Shadlen and Newsome, 1998; Tolhurst et al., 1983) and that pairs of neurons exhibit correlated variability (Gawne and Richmond, 1993; Zohary et al., 1994), often named noise correlation. Such correlations between neurons strongly impact the association between neuronal activity and behavior (Shadlen et al., 1996); in particular, it is known that the CP index depends on the correlation structure of the neuronal network (Cohen and Newsome, 2009; Haefner et al., 2013; Nienborg and Cumming, 2010; Nienborg et al., 2012). In addition, we recently demonstrated that the temporal profile of the noise correlation coefficient changes as the task progresses, reflecting dynamic effects of stimuli and internally generated signals on frontal lobe neurons that might participate in the decision process (Carnevale et al., 2012). Given that a decision evolves over time, we think it is important to detect and describe transient interpool interactions. Some knowledge about the dynamics of a large-scale cortical network during decision making has been obtained by studying macroscopic signals from magnetoencephalographic recordings (Siegel et al., 2011), but the dynamical profile of correlations has rarely been studied at the circuit level (Pesaran et al., 2008).

Motivated by these observations, we developed analytical tools to study the dynamics of neuronal pools and their relation to behavior. We tested these tools with data from simultaneous recordings of neuron pairs obtained while monkeys performed a decision-making task (de Lafuente and Romo, 2005, 2006). Specifically, we have first extended the concept of CP index, which traditionally refers to single neurons, to define measures of covariation between behavior and the firing rates of two or more neurons. We have then derived analytical expressions that explicitly relate these measures to statistical properties of the pools' spiking activity, obtaining a precise description of how noise correlations affect the standard CP index and the generalized indices introduced here. We find that the CP becomes significant when the correlation coefficients depend strongly on the choice outcomes of the trials used to compute them and that the association between population activity and behavior increases notably when the choice-conditioned correlations are small.

To address the issue of how neuronal pools cooperate to form the decision, we reasoned that the concerned pools combine their firing outputs and send the resulting signal downstream for further processing. Since the observed behavior is a consequence of these neural computations, we assumed that an important combination of pool activities would be one that covaries closely with behavior.

To test these ideas, we analyzed simultaneous recordings of pairs of premotor cortex (PM) neurons of distinct functional types and also of neuron pairs in the secondary somatosensory area (S2) (de Lafuente and Romo, 2005, 2006). In the decision-making task, monkeys had to detect a stimulus that often was very weak and was absent in half of the trials. Both the PM and S2 areas contained two types of neurons that exhibited oppositely tuned responses to stimulation (de Lafuente and Romo, 2006). Presumably, these two neuronal pools contribute to the decision-making process. For the detection task analyzed here, we found that sensory-like neurons in PM areas covary strongly with the decision report during the stimulation period, although this covariation does not reach its largest possible value. In contrast, pools of PM neurons exhibiting delay activity during the period between the application of the stimulus and a cue signal that triggers the decision's motor report become fully correlated with the subject's choice. Interestingly, this occurs when the population firing rates of the relevant pools are combined optimally, maximizing the generalized measures of covariance with behavior.

RESULTS

Measures of Covariance between Behavior and the Activity of Neural Pools

Consider a perceptual decision-making task in which the subject has to decide between two possible choices, A or B. Covariation between the activity of single neurons and the subject's choice is often quantified by the CP index. This quantity represents the average probability with which an external observer could predict the subject's decision from the activity of a single neuron, using the accrued knowledge of the firing rate distributions computed over trials in which option A or option B was selected. If the neuron responds identically in trials in which the subject chooses A (A trials) and in trials in which it chooses B (B trials), the prediction performance of the external observer is at chance level (CP = 0.5). Conversely, if the firing rate distributions of the neuron in trials A and B are fully distinct, the external observer could perfectly predict the subject's decision (CP = 1).

The CP index can be computed as the area under the receiver-operating characteristic curve (ROC) of the neuron's firing rate, segregating trials according to the subject's choice (Britten et al., 1996; Green and Swets, 1966). If the neuron's firing rate distributions in trials A and B can be described as Gaussian, one finds the following analytical expression (see Supplemental Information available online for the derivation)

$$\mathrm{CP} = \frac{1}{2}\,\mathrm{erfc}\left(-\frac{\delta}{2}\right),\quad \delta = \frac{\mu^A - \mu^B}{\sqrt{\frac{1}{2}\left(\sigma_A^2 + \sigma_B^2\right)}}, \qquad \text{(Equation 1)}$$

where μ^c and σ_c^2 are the mean and variance of the firing rate over trials in which the subject's choice was $c = \mathrm{A},\ \mathrm{B}$. The quantity δ is the difference between the firing rate means in trials A and B, measured in units of the arithmetic mean of the two variances.

The CP index is a useful measure to study how the activity of a single neuron covaries with behavior. However, the decision-making process is determined by neuronal populations (de Lafuente and Romo, 2006; Heekeren et al., 2004; Hernández et al., 2010; Pesaran et al., 2008; Siegel et al., 2011). Understanding how the decision is formed in the brain requires the use of proper measures to quantify the covariance of population activity variables with the subject's choices. This can be done by extending the concept of CP index to the combined activity of several neurons. Here we start by considering the case in which cells can be sorted into homogeneous pools of similar responses, and in section "Finding the Optimal Decision Code," we study the general case. For the simplest example, two neurons from the same pool, we consider the ROC index of the sum of their firing rates r_1 and r_2, $r_w = r_1 + r_2$, which can be estimated by

$$\mathrm{CP}_{2,w} = \frac{1}{2}\,\mathrm{erfc}\left(-\frac{\Delta_w}{2}\right),\quad \Delta_w = \frac{\mu_w^A - \mu_w^B}{\sqrt{\frac{1}{2}\left(\sigma_{w,A}^2 + \sigma_{w,B}^2\right)}}, \qquad \text{(Equation 2)}$$

where μ_w^c and $\sigma_{w,c}^2$ are the mean and variance of r_w over trials of choice c. This can be expressed in terms of the firing properties of the pair of neurons

$$\Delta_w = \frac{\sqrt{2}\delta_{1,2}}{\sqrt{1+\rho_{12}^w}}. \qquad \text{(Equation 3)}$$

Here $\delta_{1,2}$ is the arithmetic mean of δ_1 and δ_2, defined in Equation 1, and $\rho_{12}^w = 0.5(\rho_{12}^{w,A} + \rho_{12}^{w,B})$ is the arithmetic mean of the correlation coefficients between r_1 and r_2, computed over trials of decision A and B, $\rho_{12}^{w,A}$ and $\rho_{12}^{w,B}$ (see Equation 11 in Experimental Procedures). The superscript w indicates that the two neurons belong to the same pool. For simplicity, in Equation 3 we assumed that the variance of the single neuron's firing rate distributions is equal for both neurons, in both trial types (A and B). Equation 2 relates the subject's choices to the activity of the pool of neurons, but it does so in terms of the properties of the two neurons in the pool and their interaction, as captured by the correlation coefficient. The general expression is given in the Supplemental Information (Equation S16). Notice that $\mathrm{CP}_{2,w}$ will always be higher than CP, except when $\rho_{12}^w = 1$. This is a reasonable result: the averaged activity of two neurons in the same pool covaries with behavior more than that of single neurons, provided that their responses are significantly different, i.e., that their correlation is not too large.

For two neurons in different pools, we consider an arbitrary linear combination of their firing rates, $r_b = C_1 r_1 + C_2 r_2$, and quantify its covariation with the subject's choices by another ROC index, $\mathrm{CP}_{2,b}$. This index can be estimated as (see Supplemental Information for details)

$$\mathrm{CP}_{2,b} = \frac{1}{2}\,\mathrm{erfc}\left(-\frac{1}{2}\Delta_b\right),\quad \Delta_b = \frac{\delta_1 + D\delta_2}{\sqrt{1 + D^2 + 2D\rho_{12}^b}} \qquad \text{(Equation 4)}$$

where $D = C_2/C_1$. Here $\rho^b_{12} = 0.5(\rho^{b,A}_{12} + \rho^{b,B}_{12})$ with $\rho^{b,c}_{12}$ being the correlation coefficient between the firing rates of neurons 1 and 2 in different pools, computed over trials of choice c. Again, we assumed that the variance of the firing rates is equal for both neurons and both types of trials (see general expression in Equation S23). Note that if $\rho^b_{12}<0$ and $D<0$, the $CP_{2,b}$ index increases as $|\rho^b_{12}|$ decreases.

To test the amount of covariation with behavior of the combined activity of different neural populations, we consider a further extension of this procedure. Given a set of P pools each having N neurons and a population firing rate r_α ($\alpha = 1,..., P$), we can quantify the amount of covariation of an arbitrary combination of the pools' firing rates, $r_N = \sum C_\alpha r_\alpha$, defining the CP_N index,

$$CP_N = \frac{1}{2}\mathrm{erfc}\left(-\frac{\Delta_N}{2}\right), \quad \Delta_N = \frac{\mu^A_N - \mu^B_N}{\sqrt{\frac{1}{2}\left(\sigma^2_{N,A} + \sigma^2_{N,B}\right)}}. \qquad \text{(Equation 5)}$$

In applying these calculations to experimental data, we will be interested in linear combinations of rates from two pools, $r_N = C_1 r_+ + C_2 r_-$, with the pools defined as + and −. Δ_N can be expressed in terms of population-averaged firing rates, variances, and correlation coefficients,

$$\Delta_N = \frac{\sqrt{N}\left(\bar{\delta}_+ + D\bar{\delta}_-\right)}{\sqrt{(1+D^2)[1+(N-1)\bar{\rho}^w] + 2DN\bar{\rho}^b}}, \qquad \text{(Equation 6)}$$

where again $D = C_2/C_1$. The bar indicates population average. For simplicity, we took equal population-averaged variances for the two pools and the two choices (see the general expression in Equation S31). Correlation coefficients $\bar{\rho}^w$ and $\bar{\rho}^b$ affect the CP_N index in a manner similar to that for $CP_{2,w}$ and $CP_{2,b}$, respectively. The factor that amplifies the population-averaged δ s increases as $\bar{\rho}^w$ decreases. For D < 0 and $\bar{\rho}^b<0$, Δ_N also increases when $|\bar{\rho}^b|$ decreases. For the particular case in which $(N-1)\bar{\rho}^w >> 1$, Δ_N becomes independent of the number of neurons. If in addition, $\bar{\rho}^b \sim 0$, $\bar{\rho}^w$ still modulates the amplification of this index, $\Delta_N \propto (\bar{\rho}^w)^{-1/2}$.

Finding the Optimal Decision Code

The measures defined above can be used to study the interaction of neural pools during the decision-making process. If two pools cooperate in forming the decision, then combinations of their firing rates must covary with the behavioral response. But the reverse should also be true: maximizing this covariation should lead to the combination of rates that optimally predicts the animal's decision. This can be done by optimizing the CP_N index with respect to the relative contribution of the two pools to the population variable (D), which is equivalent to maximizing the mean difference between choices divided by the choice-conditioned variances (Equation 5).

To study how population activity covaries with behavior, we considered the case in which cells can be assigned to one of two discrete pools. If neurons could be sorted into discrete and homogeneous pools, one would assign equal weights to all neurons within the same pool. However, the assumption of neurons distributed in discrete pools can be relaxed. In a more general case, neurons contribute to the population variable in a graded manner, with their firing rates weighted with different coefficients. A CP_N index associated with this variable can still be defined and expressed in terms of its means, variances, and covariance matrix in the two conditions (see Supplemental Information). The proposed optimality criterion is again equivalent to maximizing the mean difference between choices divided by the choice-conditioned variances. Equivalently, finding the weights amounts to obtaining the Linear Fisher's Discriminant between the two decisions (Equations S37 and S38).

In this more general situation, the CP_N index depends on the covariance between every pair in the population. Thus, searching for optimal population variables in experimental data requires the simultaneous recording of multiple neurons in the population. However, in situations in which neurons can be classified into discrete pools, the CP_N can be computed from Equations 5 and 6. The degrees of freedom involved in the optimization procedure are reduced to one: the coefficient D that combines the pools' activities. Besides this parameter, in this case, the CP_N index only depends on two population-averaged correlation coefficients, one for neural pairs within each pool and another for pairs between pools.

Notice that the optimization procedure does not assume the existence of a decision rule based on the neurons' firing activity. In our formalism, the population variable could be any combination of the firing rates and is not necessarily related to a decision rule. The proposed procedure can be seen as a way to search for population variables that are optimally correlated with behavior. It can be applied to neurons in any area participating in the decision-making process. The covariance between global activity and behavior is determined by the network correlation structure.

The Covariance between Global Activity and Behavior Is Determined by the Network Correlation Structure

We now turn to a more detailed analysis of how the correlation structure affects the covariance between firing activity and behavior. The CP index in Equation 1 is computed from properties of single neurons (means and variances of the firing rate distributions in trials A and B). Although pairwise correlations do not appear explicitly in this equation, the CP index does depend on the correlation structure of the neural population involved in the decision-making process. This is because the firing rate distributions are conditioned to the subject's choice, which is determined by the network state during the trial.

To make this dependence explicit, we must relate the usual correlation coefficient R_{ij} for the neuron pair (i, j) (that is, the correlation coefficient computed using all trial types, regardless of the subject's choice) with the difference in mean firing rates between trials ending in each of the two choices. The latter are essentially the quantities δ, defined for neurons i and j, as in Equation 1. Given an arbitrary partition of the set trials into two different groups (A and B), the correlation coefficient R_{ij} can be expressed as

$$R_{ij} = \frac{\frac{1}{2}\left(\rho^A_{ij} + \rho^B_{ij}\right) + \frac{1}{4}\delta_i\delta_j}{\sqrt{\left[1+\left(\frac{\delta_i}{2}\right)^2\right]}\sqrt{\left[1+\left(\frac{\delta_j}{2}\right)^2\right]}} \qquad \text{(Equation 7)}$$

where, for simplicity, we assumed that the variance of the firing rates of the two neurons is equal in both types of trials (see Equation S48 for the general expression). Equation 7 shows that, apart from a common factor, the correlation coefficient R_{ij} is the sum of two effects: a contribution from the difference in rates between the two choice conditions (δ_i and δ_j, see Equation 1) and a contribution from the choice-conditioned correlation coefficients, ρ^A_{ij} and ρ^B_{ij}.

If the network contains several neural pools, the correlation structure consists of correlation coefficients of pairs of neurons in the same pool and correlations between neurons in different pools. Given a triplet of cells (1, 2, 3), with neurons (1, 2) in the same pool and neuron 3 in a different pool, we can use the equation above for the pairs (1, 2), (1, 3), and (2, 3). This leads to equations for R_{12}, R_{13}, and R_{23} that can be solved for δ_1, δ_2, and δ_3, obtaining the CP indices as a function of correlations between neurons in the same or different pools. However, implementing this procedure to analyze electrophysiological data requires the simultaneous recording of triplets of neurons. Moreover, the mathematical solution in terms of correlation coefficients between the three pairs of neurons becomes rather cumbersome. It is desirable to have a way to estimate the CP index using only data from simultaneous recordings of pairs of neurons. Now we show that it is possible to obtain a simple approximate expression for the population-averaged CP index based only on correlations between pairs of neurons. It is enough to assume that, given two neurons (1, 2) in the same population, $\delta_1 \sim \delta_2 \sim \delta_0$. Using R_{12} we obtain

$$\mathrm{CP} \sim \frac{1}{2}\mathrm{erfc}\left(-\frac{\delta_0}{2}\right),\ \left(\frac{\delta_0}{2}\right)^2 \sim \frac{R^w_{12}-\rho^w_{12}}{1-R^w_{12}}, \qquad \text{(Equation 8)}$$

where $\delta_0 = (\delta_1 + \delta_2)/2$ (see Equations S50–S53 in the Supplemental Information for a discussion of the accuracy of this approximation). Averaging Equation 8 over the population of independent pairs gives the estimate for the population-averaged CP index. This is our main result. It shows that the CP index is different from 0.5 when correlations evaluated using all trial types differ from the correlations conditioned on the subject's choice. Neurons could covary significantly with behavior even if the latter correlations are very small. It has been pointed out that correlated activity is necessary for observing robust covariations between single neuron responses and behavior (Shadlen et al., 1996). On the other hand, pairwise correlations in recurrent networks can be quite small (Ecker et al., 2010; Renart et al., 2010). The equations above show that there is no contradiction between these two statements: decorrelation in the recurrent network makes ρ^w_{12} small, but there is still a contribution to CP coming from R^w_{12}, which is produced by the difference in firing rates between trials A and B (Equation 7; see also Brody, 1999). In fact, Equation 8 shows that the CP index is maximized when the overall correlations R^w_{12} are large but the choice-conditioned correlations are small. Notice that Equation 8 does not assume any model that mechanistically relates the activity of the neurons to the subject's decision.

We can use Equation 8 to draw several conclusions. First, notice that $R^w_{12} - \rho^w_{12} \geq 0$: correlations for neurons in the same pool, computed with fixed-choice trials, are smaller than those obtained with the whole set of trials. Instead, if neurons (1, 3) in two different pools have opposite mean responses in the two conditions: $\mu^A_1 - \mu^B_1 \sim -(\mu^A_3 - \mu^B_3)$ (that is, $\delta_1\delta_3<0$), from Equation 7, we observe that $R^b_{13} - \rho^b_{13} \leq 0$ (see Supplemental Information). In both cases, the sign is determined by the difference in the mean activities in the two trial types. Finally, neurons in a given pool show CP = 0.5 if pairwise correlations in that pool obey $R^w_{12} = \rho^w_{12}$.

Similar considerations apply to the other generalized choice probability indices (Equations 3, 4, and 5). In particular for the $\mathrm{CP_N}$ index (Equations 5 and 6), apart from the population-averaged δ_+ and δ_-, for which the discussion above still holds, there is a factor depending only on the choice-conditioned correlation coefficients. A potential effect of this factor is to amplify Δ_N, thereby pushing $\mathrm{CP_N}$ to saturation, that is, closer to full covariance between firing activity and behavior.

Equation 8 was obtained for a pair (1, 2) of neurons in the same pool. As a more complex example, we now consider a two-pool network satisfying the condition that for a pair (1, 3) of neurons in different pools, $\delta_3 \sim -\delta_1 \sim -\delta_0$. This can be seen as a constraint on the correlation structure of the network. Using this constraint, and replacing pairwise correlations by their population-averaged values ($\overline{R}^w$, $\overline{R}^b$, $\overline{\rho}^w$, $\overline{\rho}^b$), it is interesting to observe that the average CP index can be estimated as

$$\overline{\mathrm{CP}} \sim \frac{1}{2}\mathrm{erfc}\left(-\sqrt{\frac{\left(\overline{R}^w-\overline{R}^b\right)-\left(\overline{\rho}^w-\overline{\rho}^b\right)}{2-\left(\overline{R}^w-\overline{R}^b\right)}}\right). \qquad \text{(Equation 9)}$$

This equation shows an explicit dependence on the difference $\overline{R}^w - \overline{R}^b$ but not on $\overline{R}^w + \overline{R}^b$. In a computational model constrained as in this example, where $\overline{R}^w$ and $\overline{R}^b$ were considered to be free parameters, it was found that the CP index depends only on $\overline{R}^w - \overline{R}^b$ (Nienborg and Cumming, 2010; Nienborg et al., 2012). However, that model used an explicit decision rule. In contrast, Equation 9 does not make any hypotheses about how the decision is made. To compare our prediction with the modeling results, we have determined the choice-conditioned noise correlations by simulating the same model. We found that $\overline{\rho}^w - \overline{\rho}^b$ depends only on $\overline{R}^w - \overline{R}^b$, while $\overline{\rho}^w + \overline{\rho}^b$ depends only on $\overline{R}^w + \overline{R}^b$ (data not shown), confirming the conclusion reached by Nienborg and Cumming (2010).

Analysis of Electrophysiological Data from a Vibrotactile Detection Task

In the remaining sections of the paper, we analyze electrophysiological data recorded in a vibrotactile detection task (Figure 1A, see Experimental Procedures), using the analytical results derived above. We analyze data from S2 and PM. Previous studies of these data showed that in both areas neuronal activity covaries with the animal's behavior (de Lafuente and Romo, 2005, 2006). Importantly, it was found that this covariation is related to the animal's perception of the sensory stimulus rather than to the selection of the motor plan. Trials were classified as hits (H), misses (M), correct rejections (CR) or false alarms (FA), depending on whether the stimulus was present or absent and on the behavioral response (Figure 1B). We analyzed stimulus-present trials, so types A and B (as denoted in all the equations above) correspond to H and M, respectively.

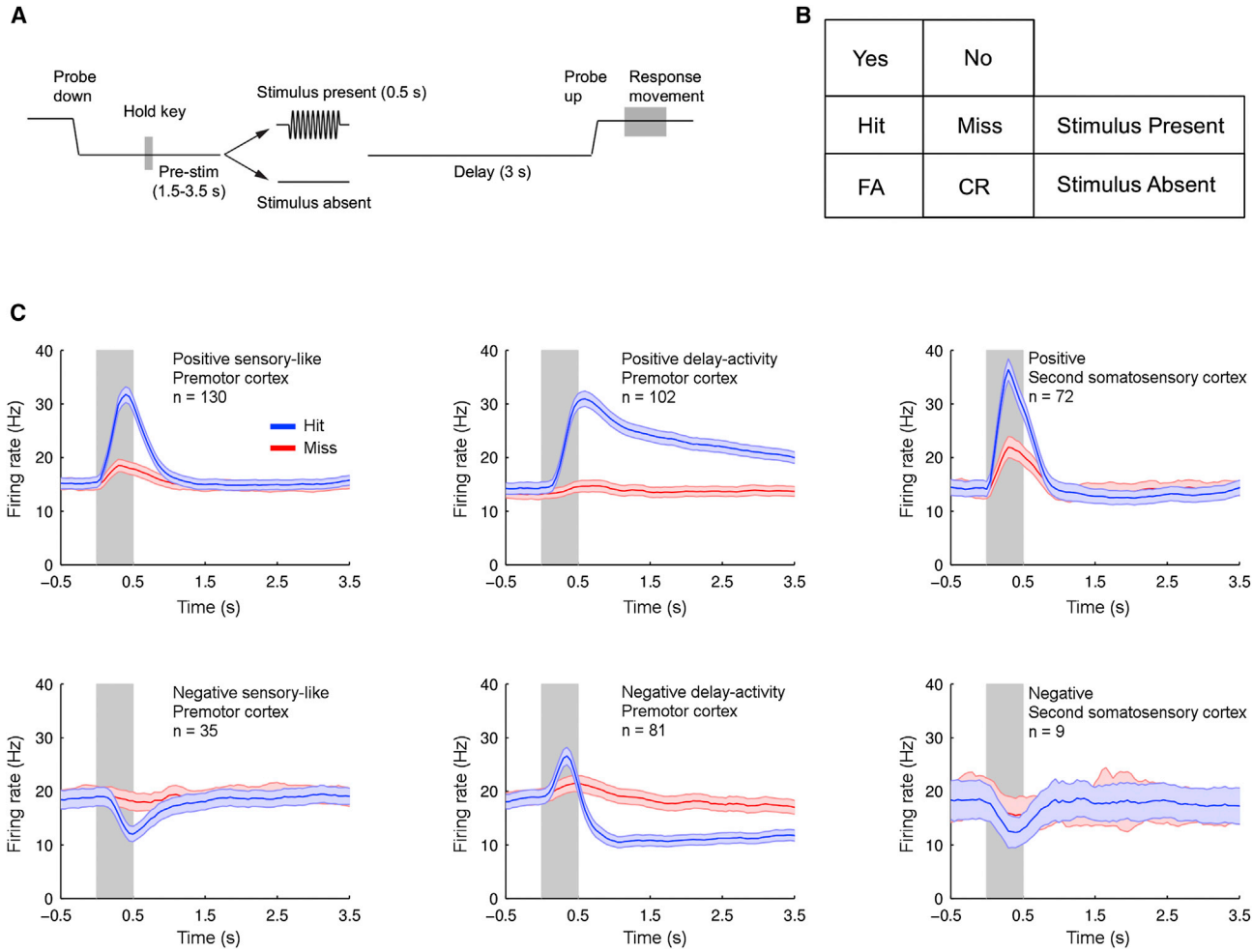

Figure 1. Detection Task and Neural Populations

(A) The mechanical probe indented the skin of one fingertip of the restrained hand (Probe down) and the monkey reacted by placing its free hand on an immovable key (Hold key). After a variable prestimulus period (1.5–3.5 s), a vibratory 0.5 s stimulus was presented on half of the trials. At the end of a fixed delay period, the stimulator probe moved up (Probe up), instructing the monkey to make a response movement to one of two push buttons. The pressed button indicated whether or not the monkey felt the stimulus.

(B) A trial is classified according to stimulus presence or absence and to the subject's response as a hit (H), miss (M), correct rejection (CR), or false alarm (FA). Stimulus amplitude was pseudorandomly chosen. A run was composed of 90 trials (amplitude 0) and 90 stimulus-present trials, with varying amplitudes (nine amplitudes with ten repetitions each; 2.3–34.6 μm).

(C) Temporal profile of mean firing rates during hit (blue traces) and miss (red traces) trials for PM neurons (first two columns) and S2 neurons (third column). Top row shows pools of positively tuned neurons and bottom row negatively tuned neurons (n is the number of neurons). Colored area represents SEM.

PM neurons were classified according to their responses to a strong stimulus in H trials (de Lafuente and Romo, 2006) (Figure 1C). Those that responded only during the stimulation period were labeled as sensory-like neurons and those showing sustained activity during the delay period were classified as delay-activity neurons. In addition, both S2 and PM neurons were labeled as positive if their firing rate transiently increased with the stimulus and as negative if their firing rate decreased in response to the stimulus. These criteria define two oppositely tuned neuronal pools (denoted as positive and negative) for each of the three populations of neurons (S2, sensory-like PM, and delay-activity PM).

We start by showing that employing two neurons to predict the animal's choice increases the level of covariation with behavior. We considered two neurons from the same neural population and compared the $CP_{2,w}$ index with the pairwise averaged CP. Figure 2 shows the temporal evolution of these two quantities for the population of positive sensory-like PM neurons (Figure 2A), positive delay-activity PM neurons (Figure 2B), and positive S2 neurons (Figure 2C). As expected, the sum of activities of two simultaneously recorded neurons from the same population is more predictive of the animal's choice than the activity of single neurons. For S2 (Figure 2C) and PM sensory-like neurons (Figure 2A), this is true during the period of stimulus presentation,

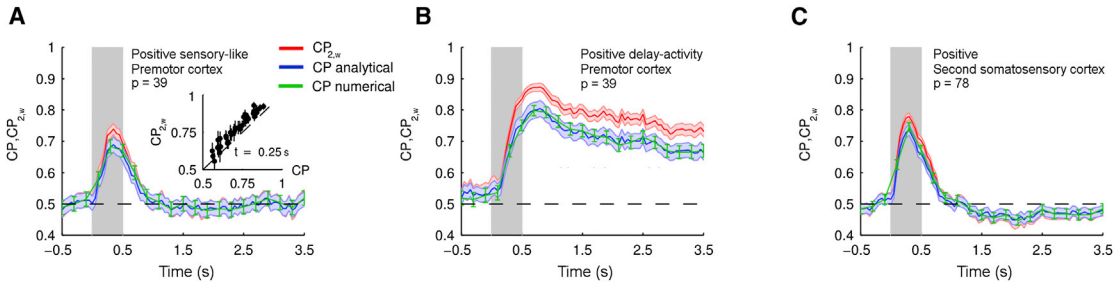

Figure 2. Choice Probability Obtained by Summing Pairs of Neurons

Temporal profile of population-averaged $CP_{2,w}$ index for pairs of positively tuned neurons (red traces) compared with population-averaged CP index for the same neural pool (blue and green traces). As expected, the combined activity of two neurons better predicts the animal's choice than the activity of a single neuron.

(A) Pool of positive PM sensory-like neurons. Inset shows $CP_{2,w}$ versus the pairwise averaged CP for each pair (t = 0.250 s).

(B) Pool of positive PM delay-activity neurons.

(C) Pool of positive S2 neurons. $CP_{2,w}$ was from Equation S16 (see Supplemental Information). A good agreement can be observed between the analytical CP (Equation 1) and its direct numerical evaluation (green traces, see Experimental Procedures). Gray boxes indicate the period of stimulation; error bars and colored areas represent SEM and p the number of neuron pairs (see also Figure S1).

while for the PM delay-activity neurons, the effect is maintained until the end of the delay period (Figure 2B). In addition, the CP indices obtained from Equation 1, which assumes Gaussian distributions of responses and involves only their means and variances, are in good agreement with those obtained by direct evaluation, for which no assumption about the response distributions is made (Figure 2, blue versus green traces). The $CP_{2,w}$ calculated based on the means and variances of individual neurons and their pairwise correlations (Equation S16) also compares well with its direct numerical evaluation (Figure S1 in the Supplemental Information).

This analysis shows that these measures of covariance with behavior can be evaluated accurately using only first- and second-order statistics of the neuronal firing rate activity. The analytical expressions can then be used reliably for studying the more general effects of neuronal correlations in the detection task.

How Correlated Variability Determines Choice Probability

The observation that single neurons covary with the subject's response is usually explained by the existence of correlated variability among the cells in the neuronal population (Shadlen et al., 1996). This argument refers to noise correlations evaluated over the whole set of trials, a quantity that may receive a substantial contribution from the difference in the firing rates in trials of different choices. However, one may wonder whether noise correlations estimated using subsets of similar trials—presumably with similar firing rates—might affect choice probability (or, more generally, any of the indices defined above). Indeed, the results obtained in Equations 7 and 8 indicate that noise correlation decreases when conditioned on the choice and that these choice-conditioned noise correlations could reduce the CP index.

To investigate this issue further, we start by analyzing the relationship between the correlation coefficients and the difference between the mean firing rates in H and M trials (Equation 7). We present this analysis for positive and negative delay-activity neurons. We consider correlation coefficients from pairs within the same neural pool, denoted by the superscript w (that is, R^w and ρ^w), and between neurons from different neural pools, denoted by b (R^b and ρ^b) (see Figure 3A). First, we show in Figure 3B the temporal profile of the population-averaged correlation coefficients R^w and R^b, obtained by direct numerical evaluation from all trial types (green traces). As we have seen, two separate factors contribute to these correlations: (1) the differences in firing rates in H and M trials and (2) the correlations conditioned on the choice (Equation 7). We then compared the same correlations with those obtained analytically by combining these two factors (blue traces, Equation S48). This comparison shows quite similar values, both for pairs of neurons within the same pool (Figure 3B, top) and for pairs of neurons belonging to different pools (Figure 3B, bottom, and Figures S2A and S2B for pools of S2 and sensory-like PM neurons).

Second, we compared the population-averaged correlation coefficients computed with all stimulus-present trials ($\overline{R}^w$ and $\overline{R}^b$) with those obtained using trials with a fixed choice ($\overline{\rho}^w$ and $\overline{\rho}^b$, red traces). Again, the comparison appears in Figure 3B (top) for pairs of neurons within the same pool ($\overline{R}^w$ and $\overline{\rho}^w$) and in Figure 3B (bottom) for pairs of neurons from different pools ($\overline{R}^b$ and $\overline{\rho}^b$). Noise correlations decrease when they are conditioned on the animal's choice: $\overline{R}^w$ exceeds $\overline{\rho}^w$ after the stimulus onset and during the entire delay period. This is explained by our analytic expression, Equation 8: for each pair of neurons, the difference $R^w - \rho^w$ is positive and comes from the difference in mean firing rate in trials of different choice (δ_0). In contrast, when the cells belong to different pools, R^b is lower than ρ^b (Figure 3B, bottom). This is because the firing rate of positively tuned neurons in H trials is larger than in M trials, while the opposite occurs for negatively tuned neurons (Figure 1C).

Notice the rather different temporal profiles of correlations conditioned on the choice and correlations defined over the whole set of trials. Whereas the latter are strongly modulated by the stimulus, the choice-conditioned correlations $\overline{\rho}^w$ and $\overline{\rho}^b$ are not. Only toward the end of the delay period does $\overline{\rho}^w$ decrease significantly below the value that it had before stimulus onset (Figure 3B, top), although during this period the firing rate of PM delay-activity neurons is higher than before stimulus

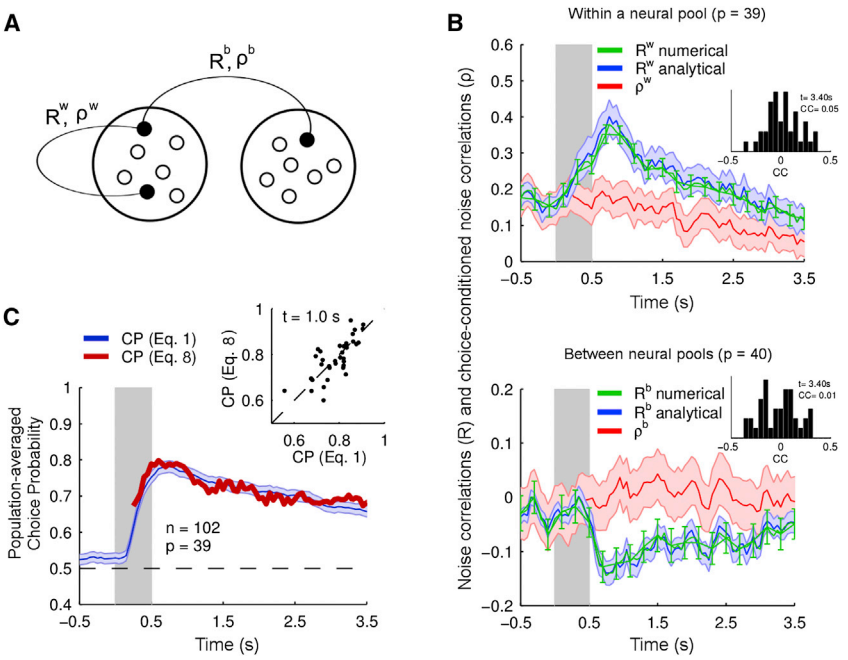

Figure 3. Noise Correlations and Choice-Conditioned Noise Correlations Determine the CP Index

(A) Correlation structure in a two-pool network. R^w, R^b, ρ^w, and ρ^b denote spike-count noise correlation coefficients between neurons in the same (w) or in different (b) pools. R^w and R^b are computed using all stimulus-present trials, while ρ^w and ρ^b are obtained from trials with a fixed subject's choice. Together, they define the correlation structure of this example network.

(B) Temporal evolution of mean correlation coefficients of delay-activity PM neurons computed with all trials (R, blue and green traces) compared with average correlations obtained from hit and miss trials separately (ρ, red traces). Top: pairs within the pool of positive delay-activity PM neurons. Bottom: pairs of positive and negative delay-activity PM neurons. Mean correlation coefficients were obtained by averaging over all pairs from the same functional type. Gray boxes indicate the period of stimulus presentation; error bars and colored areas represent SEM and p the number of pairs. Green traces depict the correlation coefficients computed numerically. Blue traces show predictions from Equation S48 (see Supplemental Information). Insets show the distribution of ρ over the population of pairs for a 250 ms time window centered at t = 3.40 s.

(C) The CP index of delay-activity neurons computed from correlation coefficients (Equation 8) is compared with its evaluation from the mean and variance of the firing rate, in H and M trials (Equation 1). Inset shows the pairwise averaged CP computed with Equation 1, compared with the CP obtained using Equation 8, for each neuronal pair at t = 1.0 s (see also Figure S2).

presentation. This behavior is consistent with the existence of a common, slowly fluctuating signal that correlates neurons in the positive pool and is present before stimulus onset (Carnevale et al., 2012). After stimulus offset, this signal tends to disappear and choice-conditioned correlations thus fall to their smallest value at the end of the delay period. In contrast, correlations evaluated using H and M trials remain high because firing rates in these two trial types are different throughout the delay period (top middle panel in Figure 1C). Correlations between neurons in different pools, $\overline{\rho}^b$, are much smaller than those between neurons in the same pool.

The difference in the temporal profiles of the correlation coefficients $\overline{R}^w$ and $\overline{\rho}^w$ fully explains the temporal evolution of the population-averaged CP index. In fact, we have seen that this index can be approximated in terms of that difference (Equation 8). We studied this prediction using data from the population of PM delay-activity neurons. The average error introduced by this approximation in our data is 11%. The population-averaged CP index, evaluated using only correlation coefficients, is shown in Figure 3C, together with the prediction from Equation 1. This result confirms that the increase of the population-averaged CP index occurring after stimulus presentation and its subsequent slight decrease during the delay period (Figure 2B) are controlled by the transient modulations of the difference $\overline{R}^w - \overline{\rho}^w$ (Figure 3B, top). Although mean choice-conditioned correlations can be rather small ($\overline{\rho}^w$ and $\overline{\rho}^b$ in Figure 3B, insets), the population-averaged CP index can be large (about 0.7 for this example) because of the contribution from $\overline{R}^w$ (Shadlen et al., 1996). In fact, correlations conditioned on the choice tend to decrease the covariation of single neurons and of neuronal populations with behavior (Equation 8). The diminishing value of these correlations during the delay period helps to maintain a large CP until the subject makes a movement. Further tests of the validity of our analytical results are shown in Figures S2C and S2D.

Choice Is Unambiguously Decoded from the Activity of Premotor Cortex Neurons

How well can the population activity of PM neurons predict the subject's choice? How are choices affected by noise correlations? To answer these questions, we considered readout neurons implemented as linear combinations of the outputs of two neural pools (Figure 4A). As mentioned before, we reasoned that when two neural pools cooperate to form the decision, the output signal resulting from their interaction must be the one that optimally predicts the subject's decision report. We now use this idea to investigate the involvement in decision making of two neural populations present in PM: sensory-like neurons and delay-activity neurons. Given that these two neuron types presumably play different roles in the process, we have considered them separately. Each of these two populations includes two different pools (+ and −), according to their response to strong stimuli (de Lafuente and Romo, 2006) (Figure 1C). Hence, for each population (sensory-like or delay-activity neurons), we have optimally combined the firing outputs of neurons taken from oppositely tuned pools and have analyzed how well this linear combination predicts the decision response.

As a first example, we considered only two neurons from the same population but different pools and we linearly combined their firing outputs. Since the neurons are in different pools, the relevant measure of covariation with behavior is the $CP_{2,b}$

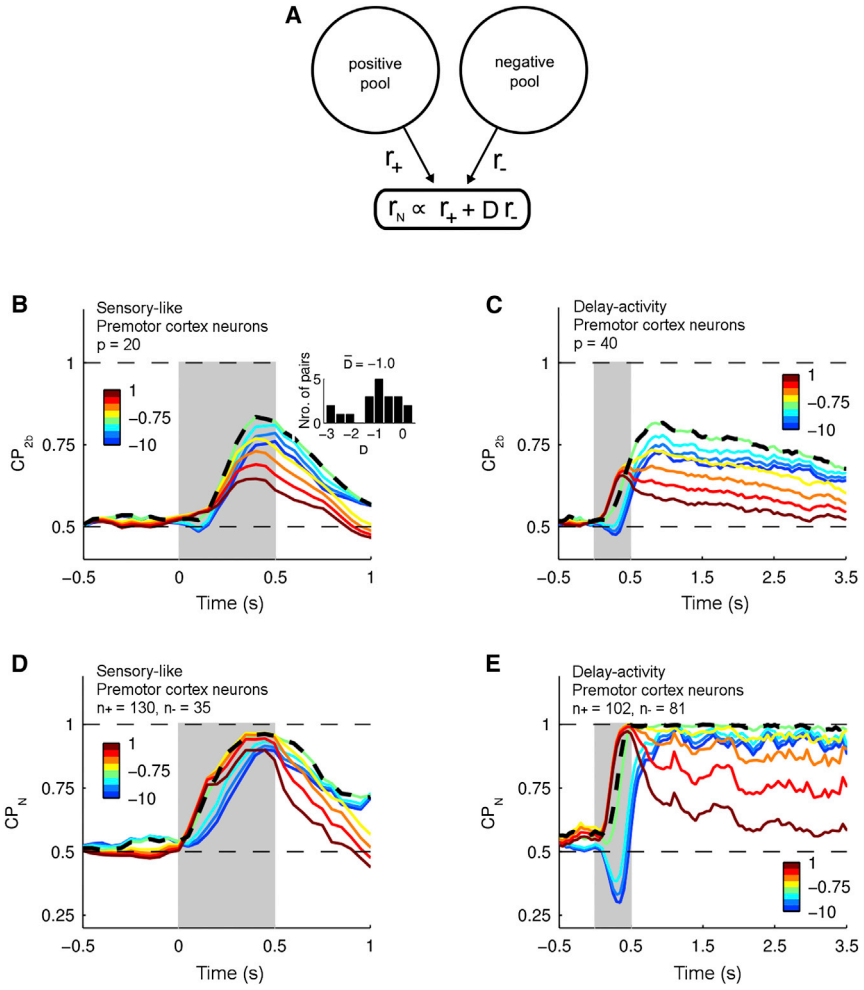

Figure 4. Linear Combinations of Positive and Negative Neurons and Its Covariation with Behavior

(A) The activity of a positive pool is linearly combined with the activity of a negative pool.

(B) Population-averaged $CP_{2,b}$ for pairs of positive and negative sensory-like neurons, computed from Equation S23 for different values of D. Color code corresponds to the value of D. Inset shows the distribution of D that maximizes $CP_{2,b}$ in a 250 ms window centered at the stimulation period. Dashed line indicates the population-averaged $CP_{2,b}$ for the mean value of optimal coefficients D = −1.

(C) Same as (B) for pairs of delay-activity neurons. In this case, the optimal value of D was obtained by averaging over the second half of the delay period resulting in D = −1. Dashed line indicates the population-averaged $CP_{2,b}$ for this value of D.

(D) CP_N for the population of sensory-like PM neurons, computed using Equation S31 for different values of D (color coded). The dashed line corresponds to the CP_N index for the optimal value of D in a 250 ms window centered at the stimulation period, D = −1.2.

(E) Same as panel (D) but for delay activity PM neurons. The dashed line corresponds to the CP_N index for the optimal value D = −0.5, obtained averaging over the second half of the delay period (as it is explained in Figure 5). The number of neurons in the positive and negative pool is denoted by n_+ and n_- respectively. The gray box indicates the period of stimulus presentation (see also Figure S3).

index. Figure 4 shows the results for sensory-like PM neurons (Figures 4B and 4D) and delay-activity PM neurons (Figures 4C and 4E) for different linear combinations of firing rates (characterized by the coefficient D, color coded). These analyses were much more limited for S2 given our experimental database (see Figure S3). For pairs of sensory-like PM neurons, the largest values of $CP_{2,b}$ occur during the presentation of the stimulus. The inset in Figure 4B shows the distribution of D values that maximize $CP_{2,b}$ in a 250 ms window centered at the stimulation period. The mean value of D over the population of pairs is D = −1, which corresponds to the difference of firing rates between the oppositely tuned pairs of neurons. The dashed line in Figure 4B represents the population-averaged $CP_{2,b}$ for this value of D. For delay-activity neurons, large values of $CP_{2,b}$ are observed during the entire delay period (Figure 4C). The most predictive combination of firing rates was again close to the difference (D = −1, averaged over the second half of the delay period) and remained constant until the end of the delay period. Again, the dashed line corresponds to the population averaged $CP_{2,b}$ for D = −1. These results indicate that, as one might intuitively suspect, the perceptual decision about stimulus presence depends on the difference in activity between the responses of oppositely tuned neurons, in agreement with what has been reported in other perceptual decision-making tasks (Gold and Shadlen, 2001; Romo et al., 2003; Heekeren et al., 2004; Romo and de Lafuente, 2013). Although in all the above cases the $CP_{2,b}$ index reaches quite large values, it is still well below its largest possible value. Furthermore, the decision-making process probably involves interactions between pools of multiple neurons (Figure 4A). Hence, we used the CP_N index (Equation 5) to look for linear combinations of mean firing rates of multiple neurons in oppositely tuned pools that would covary maximally with behavior. For sensory-like PM neurons, the optimal combination is obtained during the stimulation period, with D = −1.2 (Figure 4D). The dashed line corresponds to CP_N for this value of D. Although CP_N is larger than $CP_{2,b}$, it remains below 1 and starts to decrease by the end of the stimulus presentation period.

Most remarkably, the combination of pools of delay-activity PM neurons reaches the value $CP_N = 1$ soon after stimulus onset and maintains it during the entire delay period (Figure 4E). Figure 5 (top) shows the temporal profile of the value of D that maximizes the CP_N index at each time window. This optimal value was obtained independently for each shifted time window, as is illustrated in the inset. After a transient modulation, the optimal value of D becomes stationary with a temporal mean of −0.5 until the end of the delay period. Note that, because D in this case depends on the numbers of neurons in each pool (see

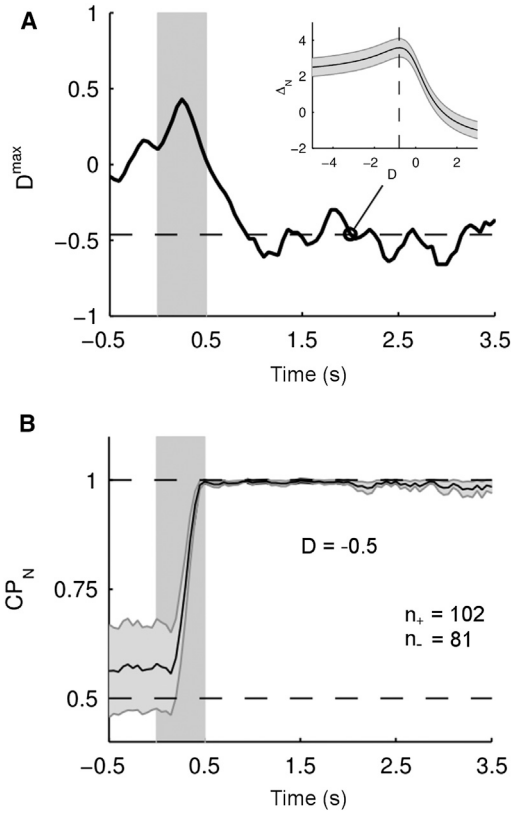

Figure 5. Delay-Activity Neurons Predict Unambiguously the Behavioral Report

(A) Temporal profile of values of D that maximize the CP_N in each 250 ms time window independently. Inset shows the location of the maximum at a 250 ms temporal bin centered at t = 2.0 s as an example. After a transient regime, the optimal value of D reaches the stationary value D = −0.5.

(B) The CP_N index for D = −0.5 is compatible with its maximum possible value of 1 during the entire delay period (see also Figure S4). Shaded area represents SEM.

Equation S32), its temporal modulation is important but its specific value is not necessarily so. The transient positive values during the stimulation period are due to the increased activity of the negative pool at that time (Figure 1, middle bottom), which produces a transient positive difference between the firing rates in hit and miss trials, opposite to the decrease in activity that this pool exhibits during the delay period. Figure 5 (bottom) shows the temporal evolution of the CP_N index when the two pools are combined using the stationary value D = −0.5. The value CP_N = 1 indicates that the population of delay-activity neurons unambiguously predicts the behavioral report during the whole delay period. Notice that because of the difficulties in measuring the entire covariance matrix, we cannot obtain an optimal population variable individually weighting each neuron's firing rate (Equation S38). However, under the assumption of discrete pools, the CP_N index already reaches its maximum possible value, so the conclusion that this population perfectly predicts the animal's behavior still is valid.

We would like to note that the application of our analytical tools does not require any assumption about the role of each pool in the decision process. Even so, a plausible interpretation of the optimal rate combination found for the population of delay-activity neurons is that the activity of the negative pool of PM neurons represents the default decision that the stimulus is absent, because when the stimulus is applied, the activity of these neurons diminishes while the activity of neurons in the positive pool increases.

What factor determines the saturation of the CP_N index? We have noticed before (Equation 6) that this index is affected by the choice-conditioned correlation coefficients. In our data, $(N-1)\overline{\rho}^w >> 1$ so CP_N becomes independent of the number of neurons and the amplification of a single neuron's covariance with behavior is controlled by the inverse of $(1+D^2)\overline{\rho}^w + 2D\overline{\rho}^b$ (the dependence of CP_N on the number of neurons is discussed in the Supplemental Information, Figure S4). We have just seen that for delay-activity neurons D = −0.5. In addition, since $\overline{\rho}^b$ is much smaller than $\overline{\rho}^w$ (Figure 3B), this expression is mainly determined by $\overline{\rho}^w$, the choice-conditioned correlation coefficient of neurons in the same pool. The smaller $\overline{\rho}^w$, the greater the amplification with respect to the single neuron's CP index.

In view of this result, one may wonder if the population of delay-activity neurons can predict the correct choice before the onset of the tactile stimulus, when the presence of such stimulus is indicated by a separate cue at the beginning of the trial. To answer this question, we decoded the animal's choice from the population of neurons with delay activity in a variation of the task in which the correct response button was illuminated at the beginning of the trial (see Experimental Procedures). In this control task, monkeys were not required to attend the vibratory stimuli but just to press the cued button at the end of the trial. We hypothesized that if the correct choice is indicated by the light cue at the beginning of the trial and the same neurons are engaged in this variant of the task, the animal's choice could be decoded from the activity of the neural pools even before the application of the stimulus. To test this hypothesis, we evaluated a CP_N index from the population firing rate during stimulus-present and stimulus-absent control trials. We performed this analysis for delay-activity neurons because they are the only population showing significant covariation with behavior during the delay period of the task. Indeed, before stimulus onset, this index is significantly larger than the CP_N in the detection task (Figure 6). The large value and stationary profile of the CP_N index in control trials indicates that the choice is made during the prestimulation period. After that, the choice is kept in memory in the form of sustained activity until the end of the delay period.

DISCUSSION

In decision-making tasks, the subject's choice results from the coordinated activity of neurons distributed in a large network comprising numerous brain areas. Hence, the decision should be decoded from population variables, based on the spiking activities of the neuronal populations involved. It is then expected that correlated variability between the firing activities of those neurons play a key role in determining the decision. Indeed, the fact that single neurons covary with the subject's report is usually explained by the existence of pairwise

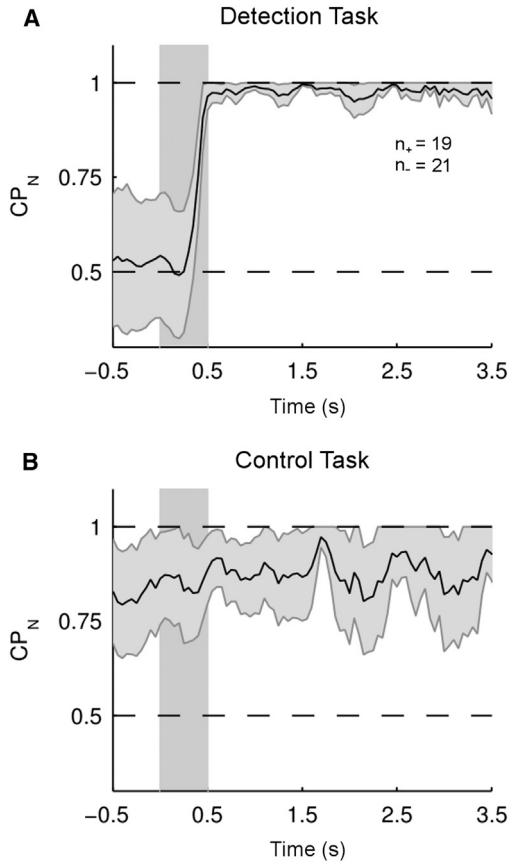

Figure 6. PM Neurons Reflect the Behavioral Choice throughout the Trial when the Correct Response Is Indicated at the Start

(A) Temporal profile of CP_N during the detection task, obtained from a set of delay-activity neurons that were also recorded during cued trials.

(B) Temporal profile of the CP_N index during control trials. In this case, CP_N was computed using stimulus-present and stimulus-absent correct trials. The large value of CP_N before the stimulus presentation indicates that, in contrast to detection-task trials, here the choice was made during the prestimulation period. Shaded area represents SEM.

correlations between neurons in the neuronal population (Shadlen et al., 1996).

Experimental and computational studies have shown that recurrent cortical networks can fire in an uncorrelated fashion, even if they receive significant common inputs (Ecker et al., 2010; Renart et al., 2010). There is no conflict between the ability of cortical networks to decorrelate the responses of pairs of neurons and the observation of significant correlated variability in decision-making tasks. Good performance in these tasks requires nonzero correlations evaluated using trials of both choices, but choice-conditioned correlations are not constrained to be high. To gain further insight into these issues, here we estimated the covariance of firing activity with behavior in terms of pairwise noise correlations and found that choice probability depends in a remarkably simple way on these quantities: it is essentially given by the overall noise correlation coefficient (computed from all trial types together) minus the average choice-conditioned correlation. It is the first term (of Equation 8) that is needed to explain why single neurons have a significant CP, while it is the second term that, according to the decorrelating effect of recurrent cortical networks, has to be small. In fact, the negative contribution of this term shows that nonzero choice-conditioned noise correlations always decrease the covariance between firing activity and behavior. The theoretical understanding of this issue is verified with great accuracy by the analysis of experimental data. CP is indeed explained by the difference between the two correlation types (Figure 3C). Mean choice-conditioned correlations are modulated during the time course of the task, with values in the range between 0.2 and 0.05 (Figure 3B). At least part of these correlations can be explained by the existence of an internally generated signal fluctuating from trial to trial (Carnevale et al., 2012). The smallest observed value could still contain this effect.

So, our study shows that correlations need to be considered if one is to analyze covariations of population firing rate variables with the subject's report. In perceptual decision-making tasks, the perceptual report results from neural processes distributed over several interacting neuronal populations and over a number of brain regions (Hernández et al., 2010; Siegel et al., 2011). The relevant measure of covariance between firing activity and behavior is a generalized CP index defined as a combination of firing rates from appropriate pools of neurons. The hypothesis behind this proposal is that when two or more neuronal populations transiently cooperate in the process of forming the decision, they produce a combined signal that accurately predicts the behavioral report. Hence, the most relevant mixture of the population's firing rates can be obtained by maximizing the covariance between the linear combination of rates and behavior. Following these ideas, we found that populations of PM neurons active during the delay period of a detection task unequivocally predict the animal's decision (Figure 4E). When the animal is cued at the beginning of the trial about the correct choice, full covariance with behavior may be reached even before a (now irrelevant) stimulus is presented (Figure 6B).

In summary, we developed tools to evaluate the CP index from the correlation structure of a network without assuming any decision rule. The tools can be applied to both sensory and frontal areas, as we showed in the analysis of S2 and PM neurons. We then generalized the use of choice probability indices to population variables and found a way to evaluate them based on data from simultaneously recorded pairs of neurons. This allowed us to propose a procedure for determining how the activity of neurons in different populations should be combined to optimally predict the subject's behavior (based on the linear Fisher's discriminant). Finally, we were able to find a population variable that fully covaries with behavior.

These analytical tools may be employed to study the dynamics of cortical networks engaged in keeping relevant information in short-term memory. In the detection task, our results suggest that by the end of the stimulation period, the decision is already made and it is maintained in short-term memory during the entire delay period. In a somatosensory discrimination task with two delay periods (Brody et al., 2003; Hernández et al., 2010; Lemus et al., 2007, 2009; Romo et al., 1999), a first vibratory stimulus is kept in memory during the first of these intervals and this memory is later compared with a second stimulus. After a second delay

period, the subject has to report which of the two stimuli vibrated with the highest frequency. It was found that during this interval, populations of PM neurons maintain in memory the frequency of the two stimuli, even if the decision could already have been made, as in the detection task (Hernández et al., 2010; Lemus et al., 2007). Does this mean that the decision continues to be elaborated, based on the information about the two stimuli maintained in working memory? Analyzing how CP_N evolves in time could answer these dynamical issues.

Neurons recorded in the detection task could be assigned to one of a small set of pools, a property that simplified the study of the CP_N index. In discrimination tasks, however, neuronal firing rates are rather heterogeneous (Jun et al., 2010) and the use of general equations where neurons contribute with different weights could be needed (Equations S38 and S39). In addition to determining whether populations of PM neurons stably maintain a decision after presentation of the second stimulus, estimating the weights with which neurons contribute to the decision would provide a method to rank neurons according to their relevance in the task. This issue could be studied using data sets in which a few tens of neurons have been recorded simultaneously (Hernández et al., 2008).

The ideas developed in this work could be applied to study brain functions other than detection of sensory stimuli. Cohen and Maunsell (2010, 2011) have noticed that attention fluctuations are associated with fluctuations in psychophysical performance. To reach this conclusion, these authors evaluated an ROC index based on a population firing rate variable defined in terms of two attentional states, a quantity somewhat similar to our CP_N index. Noise correlations are reduced by both spatial and feature attention, a fact that is assumed to have a positive effect on stimulus coding (Cohen and Maunsell 2010, 2011). Our results show that noise correlations are also relevant to explain the covariance between neuronal activity and choice and that small choice-conditioned correlated variability is needed to achieve a larger covariance with behavior both for single cells and for neuronal populations (Figures 3C and 4). Whether this is also true for neurons in higher visual areas such as V4 would require an analysis of the time course of correlations.

The simplicity of the approach presented here makes it feasible to study a wide spectrum of problems. From a theoretical viewpoint, we provided an intuitive framework to understand how first- and second-order statistics affect the relationship between network firing activity and behavior, which can be used to further develop computational methods. From a data analysis perspective, our approach could help to reveal how several cortical areas contribute and collaborate in the decision-making process.

EXPERIMENTAL PROCEDURES

Detection Task

Data for this analysis were obtained from two earlier studies (de Lafuente and Romo, 2005, 2006). Stimuli were delivered to the skin of the distal segment of one digit of the restrained hand, via a computer-controlled stimulator (BME Systems; 2 mm round tip). Initial probe indentation was 500 mm. Vibrotactile stimuli consisted of trains of 20 Hz mechanical sinusoids with amplitudes of 2.3–34.6 mm. These were interleaved with an equal number of trials where no mechanical vibrations were delivered to the skin (amplitude = 0). Animals pressed one of two buttons to indicate stimulus present (left button) or stimulus absent (right button). They were rewarded with a drop of liquid for correct responses.

Recordings

The activity of pairs of neurons were simultaneously recorded from the same cortical area including secondary somatosensory cortex (S2), ventral premotor cortex (VPc) on the left hemisphere, and dorsal premotor cortex (DPc) and medial premotor cortex (MPc), bilaterally. Pairs from premotor cortices were not distinguished in this report. Trials in the control light task proceeded exactly as described in Figure 1A, except that at the probe down, the correct target button was illuminated. Vibrotactile stimuli were delivered while the light was kept on; then, at the probe up, the light was turned off. The monkey was rewarded for pressing the previously illuminated button. Detailed description of the experimental techniques was described in de Lafuente and Romo (2005, 2006). Animals were handled in accordance with the standards of the NIH and the Society for Neuroscience.

Data Analysis

Statistical Properties of the Firing Activity

Statistical properties of the firing activity (firing rate, firing rate variance, and correlation coefficient) were computed for each neuron or pair of neurons as a function of time using 250 ms sliding window displaced every 50 ms. Trials were aligned to the time of stimulus onset.

Firing rate, $r(t)$, was calculated as the number of spikes in one sliding window divided by the its temporal length. Mean firing rate in condition c, $\mu_c(t)$, was obtained averaging over all trials of this condition. The SE of the mean firing rate was computed as the SD over trials divided by the square root of the number of trials.

Variance of the firing rate in trials of condition c, $\sigma_c(t)$, was obtained using

$$\sigma_c^2(t) = \left\langle r_i(t)^2 \right\rangle_c - \langle r_i(t) \rangle_c^2, \qquad \text{(Equation 10)}$$

where c indicates average over trials of condition c. The SE of the variance was $\sigma_c^2(t)/\sqrt{2/(N_c-1)}$, where N_c is the number of trials of condition c.

Correlation coefficients of the firing rates of a pair of neurons (i, j), in trials of condition c, were calculated following,

$$\rho^c(t) = \frac{\left\langle r_i(t) r_j(t) \right\rangle_c - \langle r_i(t) \rangle_c \left\langle r_j(t) \right\rangle_c}{\sqrt{\sigma_{i,c}^2(t)\sigma_{j,c}^2(t)}}. \qquad \text{(Equation 11)}$$

Statistical properties of the firing activity were computed only from neural recordings with at least five trials.

Measures of Covariance with Behavior

Choice probability indices were calculated using Equation 1 and computed by direct numerical evaluation (Figure 2). The Complementary error function (erfc) in Equation 1 was computed numerically (MATLAB, MathWorks). The SE of the CP calculated using Equation 1 was obtained by propagation of the SEs of the firing rates and firing rate variances over the formula. Direct numerical evaluation of choice probability was obtained using methods of signal detection theory (Green and Swets, 1966) implemented with custom software written using MATLAB (MathWorks).

Population-averaged $CP_{2,w}$ index in Figure 2 was computed evaluating Equation S16 for each pair of neurons and averaging over all pairs of neurons within the same neural pool. Similarly, population-averaged $CP_{2,b}$ index in Figures 4B and 4C was obtained for each neuronal pair of neurons belonging to different neural pools by evaluating Equation S23 and averaging over the corresponding population of pairs.

CP_N in Figures 4D and 4E and in Figures 5 and 6 was computed from population-averaged statistical properties of the firing activity using Equation S31. Population-averaged quantities were estimated pooling neurons and neuron pairs across different recording sessions. Both $CP_{2,b}$ and CP_N were computed for different linear combinations of pool rates and the optimal one was defined as that with maximum value of $\Delta_{2,b}$ or Δ_N, respectively. This gives the optimal value of the coefficient D.

Full Noise Correlations and Choice-Conditioned Noise Correlations

In Figure 3B, the correlation coefficients of the firing rates computed with all trials (R) were obtained numerically using Equation 11 and analytically by

evaluating Equation S48. Both were computed for each pair of delay-activity PM neurons and then averaged over the population of pairs.

Population-averaged choice probability index as a function of correlation coefficients (Figure 3C) was computed evaluating Equations 8 for each pair within the population of delay-activity PM neurons and then averaged over the population of pairs.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures and four figures and can be found with this article online at http://dx.doi.org/10.1016/j.neuron.2013.09.023.

ACKNOWLEDGMENTS

We warmly thank C. Brody, J. de la Rocha, R. Moreno-Bote, E. Salinas, and M. Tsodyks for their comments and Dorothy Pless for her proofreading. Funding was provided by the Spanish grants FIS 2012-33388 and FIS 2009-09433 (F.C. and N.P.); by an International Research Scholars Award from the Howard Hughes Medical Institute (R.R.) and by grants from the Dirección General de Asuntos del Personal Académico de la Universidad Nacional Autónoma de México and the Consejo Nacional de Ciencia y Tecnología (R.R. and V.d.L.).

Accepted: September 6, 2013
Published: November 21, 2013

REFERENCES

Britten, K.H., Newsome, W.T., Shadlen, M.N., Celebrini, S., and Movshon, J.A. (1996). A relationship between behavioral choice and the visual responses of neurons in macaque MT. Vis. Neurosci. *13*, 87–100.

Brody, C.D. (1999). Correlations without synchrony. Neural Comput. *11*, 1537–1551.

Brody, C.D., Hernández, A., Zainos, A., and Romo, R. (2003). Timing and neural encoding of somatosensory parametric working memory in macaque prefrontal cortex. Cereb. Cortex *13*, 1196–1207.

Carnevale, F., de Lafuente, V., Romo, R., and Parga, N. (2012). Internal signal correlates neural populations and biases perceptual decision reports. Proc. Natl. Acad. Sci. USA *109*, 18938–18943.

Cohen, M.R., and Maunsell, J.H. (2010). A neuronal population measure of attention predicts behavioral performance on individual trials. J. Neurosci. *30*, 15241–15253.

Cohen, M.R., and Maunsell, J.H. (2011). Using neuronal populations to study the mechanisms underlying spatial and feature attention. Neuron *70*, 1192–1204.

Cohen, M.R., and Newsome, W.T. (2009). Estimates of the contribution of single neurons to perception depend on timescale and noise correlation. J. Neurosci. *29*, 6635–6648.

de Lafuente, V., and Romo, R. (2005). Neuronal correlates of subjective sensory experience. Nat. Neurosci. *8*, 1698–1703.

de Lafuente, V., and Romo, R. (2006). Neural correlate of subjective sensory experience gradually builds up across cortical areas. Proc. Natl. Acad. Sci. USA *103*, 14266–14271.

Ecker, A.S., Berens, P., Keliris, G.A., Bethge, M., Logothetis, N.K., and Tolias, A.S. (2010). Decorrelated neuronal firing in cortical microcircuits. Science *327*, 584–587.

Gawne, T.J., and Richmond, B.J. (1993). How independent are the messages carried by adjacent inferior temporal cortical neurons? J. Neurosci. *13*, 2758–2771.

Gold, J.I., and Shadlen, M.N. (2001). Neural computations that underlie decisions about sensory stimuli. Trends Cogn. Sci. *5*, 10–16.

Green, D., and Swets, J.A. (1966). Signal Detection Theory and Psychophysics. (New York: John Wiley and Sons).

Haefner, R.M., Gerwinn, S., Macke, J.H., and Bethge, M. (2013). Inferring decoding strategies from choice probabilities in the presence of correlated variability. Nat. Neurosci. *16*, 235–242.

Heekeren, H.R., Marrett, S., Bandettini, P.A., and Ungerleider, L.G. (2004). A general mechanism for perceptual decision-making in the human brain. Nature *431*, 859–862.

Hernández, A., Nácher, V., Luna, R., Alvarez, M., Zainos, A., Cordero, S., Camarillo, L., Vázquez, Y., Lemus, L., and Romo, R. (2008). Procedure for recording the simultaneous activity of single neurons distributed across cortical areas during sensory discrimination. Proc. Natl. Acad. Sci. USA *105*, 16785–16790.

Hernández, A., Nácher, V., Luna, R., Zainos, A., Lemus, L., Alvarez, M., Vázquez, Y., Camarillo, L., and Romo, R. (2010). Decoding a perceptual decision process across cortex. Neuron *66*, 300–314.

Jun, J.K., Miller, P., Hernández, A., Zainos, A., Lemus, L., Brody, C.D., and Romo, R. (2010). Heterogenous population coding of a short-term memory and decision task. J. Neurosci. *30*, 916–929.

Lemus, L., Hernández, A., Luna, R., Zainos, A., Nácher, V., and Romo, R. (2007). Neural correlates of a postponed decision report. Proc. Natl. Acad. Sci. USA *104*, 17174–17179.

Lemus, L., Hernández, A., and Romo, R. (2009). Neural encoding of auditory discrimination in ventral premotor cortex. Proc. Natl. Acad. Sci. USA *106*, 14640–14645.

Nienborg, H., and Cumming, B. (2010). Correlations between the activity of sensory neurons and behavior: how much do they tell us about a neuron's causality? Curr. Opin. Neurobiol. *20*, 376–381.

Nienborg, H., Cohen, M.R., and Cumming, B.G. (2012). Decision-related activity in sensory neurons: correlations among neurons and with behavior. Annu. Rev. Neurosci. *35*, 463–483.

Pesaran, B., Nelson, M.J., and Andersen, R.A. (2008). Free choice activates a decision circuit between frontal and parietal cortex. Nature *453*, 406–409.

Renart, A., de la Rocha, J., Bartho, P., Hollender, L., Parga, N., Reyes, A., and Harris, K.D. (2010). The asynchronous state in cortical circuits. Science *327*, 587–590.

Romo, R., and de Lafuente, V. (2013). Conversion of sensory signals into perceptual decisions. Prog. Neurobiol. *103*, 41–75.

Romo, R., and Salinas, E. (2003). Flutter discrimination: neural codes, perception, memory and decision making. Nat. Rev. Neurosci. *4*, 203–218.

Romo, R., Brody, C.D., Hernández, A., and Lemus, L. (1999). Neuronal correlates of parametric working memory in the prefrontal cortex. Nature *399*, 470–473.

Romo, R., Hernández, A., Zainos, A., and Salinas, E. (2003). Correlated neuronal discharges that increase coding efficiency during perceptual discrimination. Neuron *38*, 649–657.

Shadlen, M.N., and Newsome, W.T. (1998). The variable discharge of cortical neurons: implications for connectivity, computation, and information coding. J. Neurosci. *18*, 3870–3896.

Shadlen, M.N., Britten, K.H., Newsome, W.T., and Movshon, J.A. (1996). A computational analysis of the relationship between neuronal and behavioral responses to visual motion. J. Neurosci. *16*, 1486–1510.

Siegel, M., Engel, A.K., and Donner, T.H. (2011). Cortical network dynamics of perceptual decision-making in the human brain. Front Hum Neurosci *5*, 21.

Tolhurst, D.J., Movshon, J.A., and Dean, A.F. (1983). The statistical reliability of signals in single neurons in cat and monkey visual cortex. Vision Res. *23*, 775–785.

Zohary, E., Shadlen, M.N., and Newsome, W.T. (1994). Correlated neuronal discharge rate and its implications for psychophysical performance. Nature *370*, 140–143.

Neuron, Volume *80*

Supplemental Information

An Optimal Decision Population Code that Accounts for Correlated Variability Unambiguously Predicts a Subject's Choice

Federico Carnevale, Victor de Lafuente, Ranulfo Romo, and Néstor Parga

SUPPLEMENTAL DATA

$CP_{2,s}$ index: verification of the analytical approximation (related to Figure 2)

In Figure S1 we verify our formula S16 for $CP_{2,w}$ by comparing its prediction with the result obtained by direct numerical evaluation (see Experimental Procedures). The results obtained using our analytical expression are in good agreement with those obtained numerically.

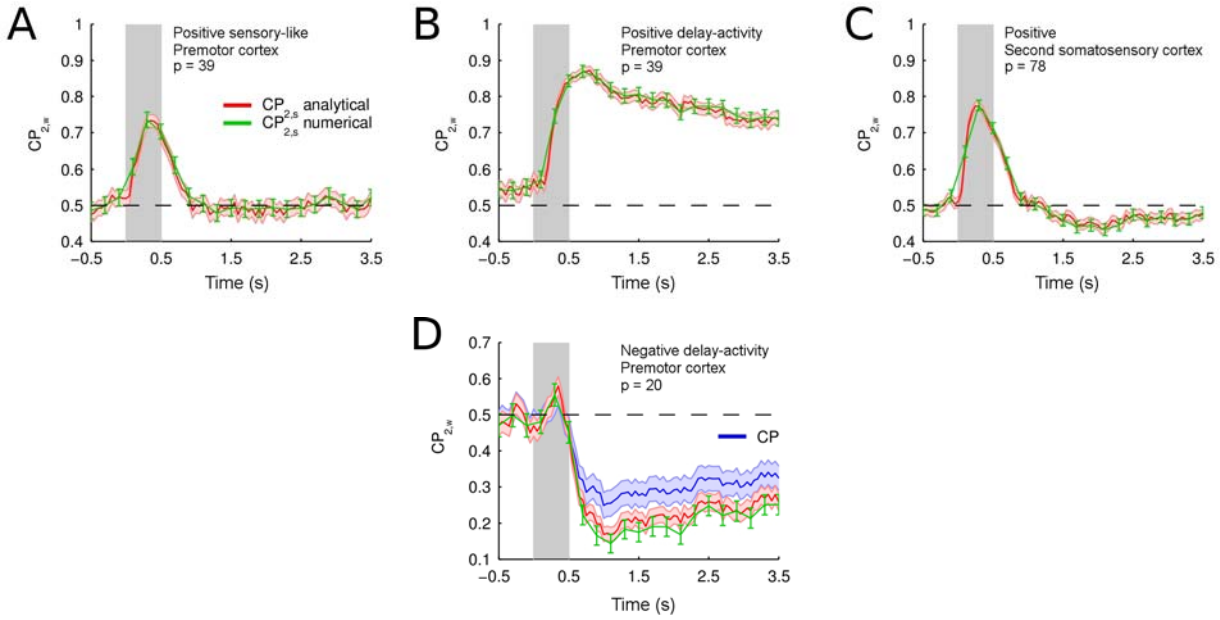

Figure S1: **Numerical and analytical computation of the $CP_{2,w}$ index.** Related to Figure 2. Population-averaged $CP_{2,w}$ index computed with Equation S16 (red traces) compared with the one obtained by direct numerical evaluation (green traces, see Experimental Procedures). Gray boxes indicate the period of stimulation and p is the number of neurons used in this analysis. (A) Pool of positive sensory-like PM neurons. (B) Pool of positive delay-activity PM neurons. (C) Pool of positive S2 neurons. (D) Pool of negative delay-activity PM neurons. For this case we also show the population-averaged CP index in order to verify that the sum of activity of two neurons is more predictive than the activity of single neurons. Due to limitations in the number of simultaneously recorded neural pairs in our database, we cannot perform this analysis in pools of negative sensory-like PM and negative S2 neurons.

Further data analysis of how correlated variability determines choice probability (related to Figure 3)

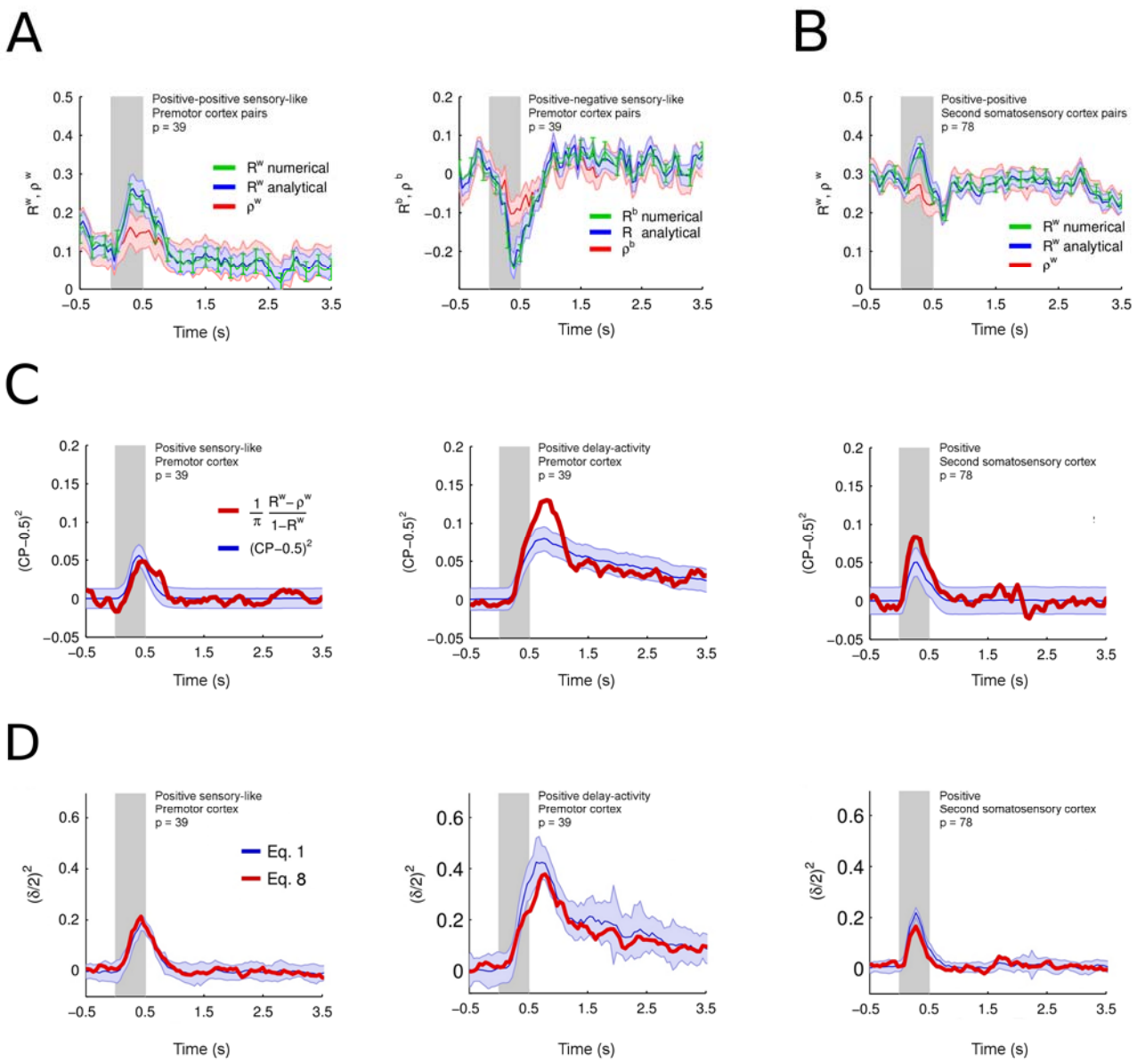

Figure S2: **Noise correlations and choice conditioned noise correlations determine the CP index.** Related to Figure 3. (A)-(B) Temporal evolution of the mean correlation coefficients computed with all trials (R, blue and green traces) compared with the average of correlations obtained using hit and miss trials separately (ρ, red traces). Mean correlation coefficients were obtained averaging over all pairs from the same functional type. Gray boxes indicate the period of stimulus presentation, error bars represent standard error of the mean (SEM) and p is the number of pairs. Green traces are the correlation coefficients computed numerically. Blue traces are the predictions from Equation S48. (A) Pairs within the pool of positive sensory-like PM neurons (Left) and pairs of positive and

negative PM sensory-like neurons (Right). (B) Pairs within the pool of positive S2 neurons. (C) Verification of the linear approximation for the CP in terms of the correlation coefficients (Equation S55). (Left) Pool of positive sensory-like PM neurons. (Middle) Pool of positive delay-activity PM neurons. (Right) Pool of positive S2 neurons. As expected the agreement is good except for large values of the CP index. (D) Population-averaged $\left(\delta_0/2\right)^2$ derived from Equation 1 (blue traces) compared with the one obtained using Equation 8 (red traces). Gray boxes indicate the stimulation period and p the number of pairs. (Left) Pool of positive sensory-like PM neurons. (Middle) Pool of positive delay-activity PM neurons. (Right) Pool of positive S2 neurons.

$CP_{2,b}$ index and optimal rate combination for pairs of S2 neurons (related to Figure 4)

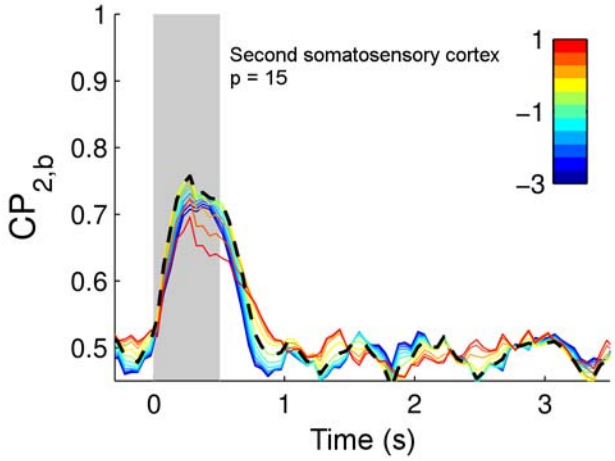

Figure S3: **Optimal rate combination for pairs of S2 neurons.** Related to Figure 4. Population-averaged $CP_{2,b}$ for pairs of positive and negative S2 neurons, computed using Equation S23-S24 for different values of D. Color code corresponds to the value of D. The rate combination with optimal $CP_{2,b}$, during the stimulation period, is obtained for D = -0.7 (black dotted line) .

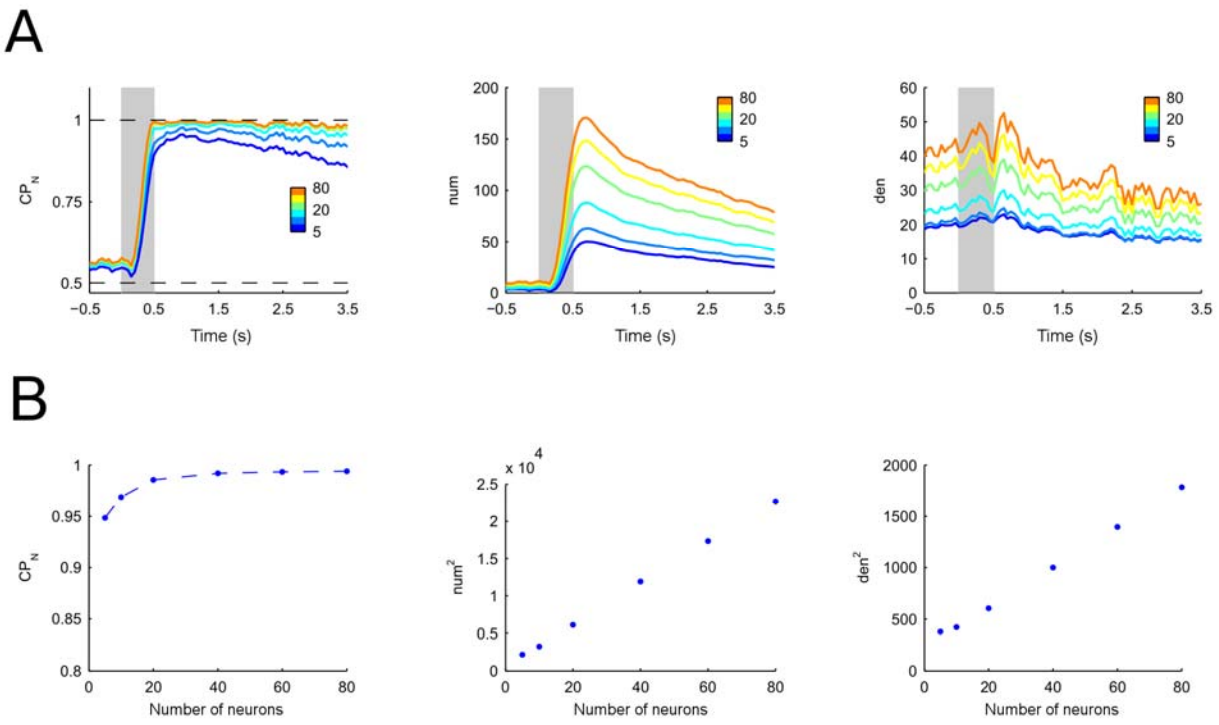

Figure S4: **Dependence of the CP_N index on the number of neurons**. Related to Figure 5. (A) Time course of the CP_N index (Left), numerator (Middle) and denominator (Right) of Equation S61 for N taking values between 5 and 80 neurons per pool (color-coded). For small N the CP_N index increases with N while at larger values of N it saturates. (B) Dependence on N of the CP_N index (Left), the square of the numerator of Equation S61 (Middle) and the square of the denominator (Right) for a fixed time, t = 1 s.

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

Choice Probability Indices

Arbitrary variable

The generalized choice probability index of an arbitrary variable R is defined as the area under the receiver-operating-characteristic (ROC) curve, analogously to the choice probability index [1]. It can be expressed as [2],

$$GCP = \int_0^1 \beta d\alpha \tag{S1}$$

where,

$$\begin{aligned} \alpha(z) &= P[R > z \mid B] \\ \beta(z) &= P[R > z \mid \mathrm{A}] \end{aligned} \tag{S2}$$

A, B are the two possible choices and z is the threshold level. The quantity $\alpha(z)$ is the probability of finding a type B trial with a value of R higher than the threshold z. In the same way, $\beta(z)$ is the probability of finding a type A trial with R higher than the threshold. In Equation S1 we are assuming that the mean value of the distribution of R over A trials is higher than the one in B trials, but the opposite case is analogous.

The probability that a B trial is classified as A, as a function of the threshold z, can be expressed as,

$$\alpha(z) = \int_z^\infty p[R \mid B] dR \tag{S3}$$

from where

$$\frac{d\alpha}{dz} = -p[z \mid B] \ . \tag{S4}$$

Then, Equation S1 can be written as,

$$GCP = \int_{-\infty}^{\infty} p[z \mid B] \beta(z) dz \ . \tag{S5}$$

Note that, because R is an arbitrary variable, it can take any value on the real axis. Therefore, the threshold z can also have any value from $-\infty$ to ∞. If we assume that R has a Gaussian distribution in each of the two set of trials, then we can write,

$$p\left[z \mid B\right] = \frac{1}{\sqrt{2\pi\Sigma_B^2}} \exp\left(-\frac{\left(z-\Theta_B\right)^2}{2\Sigma_B^2}\right) \tag{S6}$$

and

$$\begin{aligned} \beta\left(z\right) &= p\left[R > z \mid A\right] \\ &= \frac{1}{2}\left[1 - erf\left(\frac{z-\Theta_A}{\sqrt{2\Sigma_A^2}}\right)\right] \\ &= \frac{1}{2} erfc\left(\frac{z-\Theta_A}{\sqrt{2\Sigma_A^2}}\right) \end{aligned} \tag{S7}$$

where

$$\begin{aligned} \Theta_A &= \left\langle R \right\rangle_A \\ \Theta_B &= \left\langle R \right\rangle_B \\ \Sigma_A^2 &= \left\langle \left(R-\Theta_A\right)^2 \right\rangle_A \\ \Sigma_B^2 &= \left\langle \left(R-\Theta_B\right)^2 \right\rangle_B \end{aligned} \tag{S8}$$

By replacing both expressions in Equation S5, we can express GCP as

$$\mathrm{GCP} = \frac{1}{2}\mathrm{erfc}\left(-\frac{\Delta}{2}\right), \qquad \Delta = \frac{\Theta_A - \Theta_B}{\sqrt{\frac{1}{2}\left(\Sigma_A^2 + \Sigma_B^2\right)}} \tag{S9}$$

Single neuron firing activity

When the arbitrary variable R is the firing rate r of a neuron, the generalized choice probability reduces to the choice probability index and can be expressed as,

$$\mathrm{GCP} = \frac{1}{2}\mathrm{erfc}\left(-\frac{\delta}{2}\right), \qquad \delta = \frac{\mu_A - \mu_B}{\sqrt{\frac{1}{2}\left(\sigma_A^2 + \sigma_B^2\right)}} \tag{S10}$$

where

$$\begin{aligned} \mu_A &= \langle r \rangle_A \\ \mu_B &= \langle r \rangle_B \\ \sigma_A^2 &= \left\langle \left(r - \mu_A\right)^2 \right\rangle_A \\ \sigma_B^2 &= \left\langle \left(r - \mu_B\right)^2 \right\rangle_B \end{aligned} \tag{S11}$$

that corresponds to Equation 1 of the main text.

Sum of firing activities of two neurons ($CP_{2,w}$)

As the simplest example GCP index, consider the case where R is sum of the firing rates of two neurons from the same neural pool, $r_w = r_1 + r_2$.

$$\begin{aligned} \Delta_{2,w} &= \frac{\langle r_w \rangle_A - \langle r_w \rangle_B}{\sqrt{\frac{1}{2}\left(\Sigma_A^2 + \Sigma_B^2\right)}} \\ &= \frac{\left(\langle r_1 \rangle_A + \langle r_2 \rangle_A\right) - \left(\langle r_1 \rangle_B + \langle r_2 \rangle_B\right)}{\sqrt{\frac{1}{2}\left(\Sigma_A^2 + \Sigma_B^2\right)}} \end{aligned} \tag{S12}$$

The variance Σ_c^2 for c = A,B can be expressed as,

$$\begin{aligned} \Sigma_c^2 &= \left\langle r_w^2 \right\rangle_c - \left\langle r_w \right\rangle_c^2 \\ &= \sigma_{1,c}^2 + \sigma_{2,c}^2 + 2\gamma_{12}^c \end{aligned} \tag{S13}$$

where

$$\begin{aligned} \sigma_{1,c}^2 &= \left\langle r_1^2 \right\rangle_c - \left\langle r_1 \right\rangle_c^2 \\ \sigma_{2,c}^2 &= \left\langle r_2^2 \right\rangle_c - \left\langle r_2 \right\rangle_c^2 \end{aligned} \tag{S14}$$

and $\gamma_{12}^{c} = \left\langle r_1 \, r_2 \right\rangle_c - \left\langle r_1 \right\rangle_c \left\langle r_2 \right\rangle_c$, is the covariance between r_1 and r_2 evaluated over trials of type c = A,B.

From this we have,

$$\mathrm{CP}_{2,w} = \frac{1}{2}\mathrm{erfc}\left(-\frac{\Delta_{2,w}}{2}\right) \tag{S15}$$

$$\Delta_{2,w} = \frac{\left(\mu_1^A - \mu_1^B\right) + \left(\mu_2^A - \mu_2^B\right)}{\sqrt{\frac{1}{2}\left(\sigma_{1,A}^2 + \sigma_{1,B}^2\right) + \frac{1}{2}\left(\sigma_{2,A}^2 + \sigma_{2,B}^2\right) + \left(\gamma_{12}^A + \gamma_{12}^B\right)}} \tag{S16}$$

A simple particular case is obtained if the variance of the firing rate distribution is equal for both neurons and both types of trials. Denoting this variance as σ^2 and defining the correlation coefficient conditioned to choice c, $\rho_{12}^{w,c} = \gamma_{12}^{c} / \sigma^2$, we have

$$\mathrm{CP}_{2,w} = \frac{1}{2}\mathrm{erfc}\left(-\frac{\Delta_{2,w}}{2}\right) \tag{S17}$$

$$\Delta_{2,w} = \frac{\left(\mu_1^A - \mu_1^B\right) + \left(\mu_2^A - \mu_2^B\right)}{\sigma\sqrt{2 + \left(\rho_{12}^{w,A} + \rho_{12}^{w,B}\right)}} = \frac{\sqrt{2}\delta_{1,2}}{\sqrt{1 + \rho_{12}^{w}}} \tag{S18}$$

where $\delta_{1,2}$ is the arithmetic mean of δ_1 and δ_2, and $\rho_{12}^{w} = 0.5\left(\rho_{12}^{w,A} + \rho_{12}^{w,B}\right)$. The superscript *w* indicates that neurons 1 and 2 are within the same neural pool.

Arbitrary combination of the firing activity of two neurons ($\mathrm{CP}_{2,b}$)

If the two neurons belong to different pools, we take R as an arbitrary linear combination of their firing rates, $r_b = C_1 r_1 + C_2 r_2$.

$$\begin{aligned} \Delta_{2,b} &= \frac{\left\langle r_b \right\rangle_A - \left\langle r_b \right\rangle_B}{\sqrt{\frac{1}{2}\left(\Sigma_A^2 + \Sigma_B^2\right)}} \\ &= \frac{C_1\left(\left\langle r_1 \right\rangle_A - \left\langle r_1 \right\rangle_B\right) + C_2\left(\left\langle r_2 \right\rangle_A - \left\langle r_2 \right\rangle_B\right)}{\sqrt{\frac{1}{2}\left(\Sigma_A^2 + \Sigma_B^2\right)}} \end{aligned} \tag{S19}$$

The variance of r_b in trials of type c = A, B is,

$$\Sigma_c^2 = C_1^2\sigma_{1,c}^2 + C_2^2\sigma_{2,c}^2 + 2C_1C_2\gamma_{12}^c \,. \tag{S20}$$

From this we have,

$$\mathrm{CP}_{2,b} = \frac{1}{2}\mathrm{erfc}\left(-\frac{\Delta_{2,b}}{2}\right) \tag{S21}$$

$$\Delta_{2,b} = \frac{\left(\mu_1^A - \mu_1^B\right) + D\left(\mu_2^A - \mu_2^B\right)}{\sqrt{\frac{1}{2}\left(\sigma_{1,A}^2 + \sigma_{1,B}^2\right) + \frac{1}{2}D^2\left(\sigma_{2,A}^2 + \sigma_{2,B}^2\right) + D\left(\gamma_{12}^A + \gamma_{12}^B\right)}} \tag{S22}$$

where we defined $D = C_2/C_1$.

Again, a simple particular case is obtained if the variance of the firing rate distribution is equal for both neurons and both types of trials. Denoting this variance as σ^2 and defining $\rho_{12}^{b,c} = \gamma_{12}^c/\sigma^2$ we have

$$\mathrm{CP}_{2,b} = \frac{1}{2}\mathrm{erfc}\left(-\frac{\Delta_{2,b}}{2}\right) \tag{S23}$$

$$\Delta_{2,b} = \frac{\left(\mu_1^A - \mu_1^B\right) + D\left(\mu_2^A - \mu_2^B\right)}{\sigma\sqrt{1 + D^2 + D\left(\rho_{12}^{b,A} + \rho_{12}^{b,B}\right)}} = \frac{\delta_1 + D\delta_2}{\sqrt{1 + D^2 + 2d\rho_{12}^b}} \tag{S24}$$

where $\rho_{12}^b = 0.5\left(\rho_{12}^{b,A} + \rho_{12}^{b,B}\right)$. The superscript b indicates that neurons 1 and 2 belong to different neural pools.

Arbitrary combination of the firing activity of two neural pools ($\mathrm{CP_N}$)

Consider now the case of two neural pools, denoted by subscripts + and -, having N_+ and N_- neurons respectively. Let us take R as a linear combination of their mean firing activities,

$$r_N = C_1\sum_{j=1}^{N} r_{+,j} + C_2\sum_{j=1}^{N} r_{-,j} \tag{S25}$$

where $r_{\alpha,j}$ is the firing rate of neuron j in the population $\alpha = +,-$. In order to calculate the ROC index we have to express

$$\Delta_N = \frac{\langle r_N \rangle_A - \langle r_N \rangle_B}{\sqrt{\frac{1}{2}\left(\Sigma_A^2 + \Sigma_B^2\right)}} \tag{S26}$$

in terms of single neuron and neuron pair properties. In this case,

$$\Delta_N = \frac{C_1 N_+ \left(\bar{\mu}_+^A - \bar{\mu}_+^B\right) + C_2 N_- \left(\bar{\mu}_-^A - \bar{\mu}_-^B\right)}{\sqrt{\frac{1}{2}\left(\Sigma_A^2 + \Sigma_B^2\right)}} \tag{S27}$$

where the bar indicates population average, i.e.,

$$\bar{\mu}_+^c = \frac{1}{N_+}\sum_{j=1}^{N_+} \mu_{+,j} \tag{S28}$$

$$\bar{\mu}_-^c = \frac{1}{N_-}\sum_{j=1}^{N_-} \mu_{-,j} \tag{S29}$$

The variance Σ_c^2 is now expressed as,

$$\begin{aligned} \Sigma_c^2 &= \left\langle r_N^2 \right\rangle_c - \left\langle r_N \right\rangle_c^2 \\ &= C_1^2 \left(N_+ \overline{\sigma_{+,c}^2} + N_+ \left(N_+ - 1\right) \overline{\gamma_{++}^c} \right) + \\ &+ C_2^2 \left(N_- \overline{\sigma_{-,c}^2} + N_- \left(N_- - 1\right) \overline{\gamma_{--}^c} \right) + \\ &+ 2 C_1 C_2 N_+ N_- \overline{\gamma_{+-}^c} \end{aligned} \tag{S30}$$

Defining $D = C_2 / C_1$ we have,

$$\mathrm{CP}_N = \frac{1}{2}\mathrm{erfc}\left(-\frac{\Delta_N}{2}\right) \tag{S31}$$

$$\Delta_N = \frac{N_+ \left(\bar{\mu}_+^A - \bar{\mu}_+^B\right) + D N_- \left(\bar{\mu}_-^A - \bar{\mu}_-^B\right)}{\sqrt{\frac{1}{2}\left(\frac{\Sigma_A^2}{C_1^2} + \frac{\Sigma_B^2}{C_1^2}\right)}} \tag{S32}$$

with

$$\frac{\Sigma_c^2}{C_1^2} = N_+\overline{\sigma_{+,c}^2} + N_+\left(N_+ - 1\right)\overline{\gamma_{++}^c} + \\ +D^2\left(N_-\overline{\sigma_{-,c}^2} + N_-\left(N_- - 1\right)\overline{\gamma_{--}^c}\right) + \\ +2DN_+N_-\overline{\gamma_{+-}^c} \tag{S33}$$

In the application of Equations S31-S33 to experimental data, we estimated the population-averaged quantities pooling neurons and pairs across different recording sessions. The population-averaged mean and variance of the firing rate in condition c, denoted by $\overline{\mu_i^c}$ and $\overline{\sigma_{\mathrm{i},c}^2}$ respectively, were obtained by pooling neurons from population i = +,- recorded across all the experimental sessions. To estimate the population averaged covariances of firing rates in condition c, referred as $\overline{\gamma_{ij}^c}$, we averaged over all the recorded pairs consisting of one neuron belonging to population i = +,- and the other to population j = +,- across all recording sessions.

In general, the $\mathrm{CP_N}$ index will depend on the number of neurons. In Figure S4 we explore this dependence in our experimental data.

Weighted sum of the activity of a pool of neurons

In a more general context one can study the covariation with behavior of arbitrary linear combinations of the neuron's firing rates. We want to study how the variable

$$r_N = \sum_{i=1}^{N} w_i r_i = \mathbf{w}^T\mathbf{r} \tag{S34}$$

covaries with behavior. The generalized choice probability reads,

$$\mathrm{CP}_N = \frac{1}{2}\mathrm{erfc}\left(-\frac{\Delta_N}{2}\right), \qquad \Delta_N = \frac{\mu_N^A - \mu_N^B}{\sqrt{\frac{1}{2}\left(\sigma_{N,A}^2 + \sigma_{N,B}^2\right)}} = \frac{\mathbf{w}^T\left(\boldsymbol{\mu}^A - \boldsymbol{\mu}^B\right)}{\sqrt{\mathbf{w}^T\boldsymbol{\Gamma}\mathbf{w}}} \tag{S35}$$

where,

$$\mu_N^c = \left\langle \sum_{i=1}^{N} w_i r_i \right\rangle_c = \left\langle \mathbf{w}^T \mathbf{r} \right\rangle_c = \mathbf{w}^T \boldsymbol{\mu}^c$$
$$\sigma_{N,c}^2 = \left\langle \left(\sum_{i=1}^{N} w_i r_i \right)^2 \right\rangle_c - \left\langle \sum_{i=1}^{N} w_i r_i \right\rangle_c^2 = \mathbf{w}^T \boldsymbol{\Gamma}_c \mathbf{w} \tag{S36}$$

The matrix $\boldsymbol{\Gamma}$ is the arithmetic mean of the covariance matrices computed with trials of fixed each choice, $\boldsymbol{\Gamma} = \left(\boldsymbol{\Gamma}_A + \boldsymbol{\Gamma}_B\right)/2$, and $\boldsymbol{\Gamma}_{c,ij} = \left\langle r_i r_j \right\rangle_c - \left\langle r_i \right\rangle_c \left\langle r_j \right\rangle_c$.

The set of weights $\mathbf{w}$ that maximize CP_N can be found by writing Δ_N^2,

$$\Delta_N^2 = \frac{\mathbf{w}^T \left(\boldsymbol{\mu}^A - \boldsymbol{\mu}^B\right)\left(\boldsymbol{\mu}^A - \boldsymbol{\mu}^B\right)^T \mathbf{w}}{\mathbf{w}^T \boldsymbol{\Gamma} \mathbf{w}} \tag{S37}$$

This quantity is the Fisher's Linear Discriminant between classes A and B. It measures the ratio between the squared difference of the mean values of r_N in each class and the average variance within each decision. Assuming that $\boldsymbol{\Gamma}$ has full rank, the optimal vector $\mathbf{w}$ is,

$$\mathbf{w}^{opt} = \boldsymbol{\Gamma}^{-1} \left(\boldsymbol{\mu}^A - \boldsymbol{\mu}^B\right) \tag{S38}$$

and the optimal CP_N is,

$$\mathrm{CP}_N = \frac{1}{2}\mathrm{erfc}\left(-\frac{\Delta_N}{2}\right), \qquad \Delta_N = \sqrt{\left(\boldsymbol{\mu}^A - \boldsymbol{\mu}^B\right)^T \boldsymbol{\Gamma}^{-1T} \left(\boldsymbol{\mu}^A - \boldsymbol{\mu}^B\right)} \tag{S39}$$

Notice that solving this optimization problem requires knowledge of the entire covariance matrix, $\boldsymbol{\Gamma}$, an information that is experimentally very difficult to obtain because it involves the simultaneous recording of the entire population of neurons. In the particular case in which neurons belong to two discrete homogeneous pools, the CP_N reduces to the expression given in S31-S33.

Choice probability and Correlation Structure

The correlation coefficient R_{ij} between a pair of neurons (i,j) computed over trials of both decision types (A and B) is defined as,

$$R_{ij} = \frac{\mathrm{cov}\left(r_i, r_j\right)}{\sigma_i \sigma_j} \tag{S40}$$

$$\begin{aligned} \mathrm{cov}\left(r_i, r_j\right) &= \left\langle \left(r_i - \mu_i\right)\left(r_j - \mu_j\right) \right\rangle \\ \sigma_i &= \left\langle \left(r_i - \mu_i\right)^2 \right\rangle \\ \sigma_j &= \left\langle \left(r_j - \mu_j\right)^2 \right\rangle \end{aligned} \tag{S41}$$

where r_x is the firing rate of neuron x = i,j and μ_x is its mean value evaluated over all trials. The covariance $\mathrm{cov}\left(r_i, r_j\right)$ can be rewritten segregating trials according to the animal's decision,

$$\mathrm{cov}\left(r_i, r_j\right) = \frac{N_A}{N}\left\langle \left(r_i - \mu_i\right)\left(r_j - \mu_j\right) \right\rangle_A + \frac{N_B}{N}\left\langle \left(r_i - \mu_i\right)\left(r_j - \mu_j\right) \right\rangle_B \tag{S42}$$

where N_A and N_B are the number of trials of each type and $N = N_A + N_B$. The mean firing rate over all trials can be expressed for both neurons (x = i, j) as,

$$\begin{aligned} \mu_x &= \frac{N_A}{N}\mu_x^A + \frac{N_B}{N}\mu_x^B \\ &= \mu_x^A - \frac{N_B}{N}\Delta\mu_x^B \\ &= \mu_x^B + \frac{N_A}{N}\Delta\mu_x^B \end{aligned} \tag{S43}$$

with

$$\begin{aligned} \Delta\mu_x &= \mu_x^A - \mu_x^B \\ \mu_x^A &= \left\langle r_x \right\rangle_A \\ \mu_x^B &= \left\langle r_x \right\rangle_B . \end{aligned} \tag{S44}$$

Then, the covariance is

$$\mathrm{cov}\left(r_i, r_j\right) = \frac{N_A}{N}\left\langle \left(r_i - \mu_i^A\right)\left(r_j - \mu_j^A\right) \right\rangle_A + \frac{N_B}{N}\left\langle \left(r_i - \mu_i^B\right)\left(r_j - \mu_j^B\right) \right\rangle_B + \frac{N_A N_B^2 + N_B N_A^2}{N^3}\Delta\mu_i\Delta\mu_j \tag{S45}$$

which can be expressed as,

$$\mathrm{cov}\left(r_i, r_j\right) = \frac{N_A}{N}\gamma_{ij}^A + \frac{N_B}{N}\gamma_{ij}^B + \frac{N_A N_B^2 + N_B N_A^2}{N^3}\Delta\mu_i\Delta\mu_j \tag{S46}$$

where γ_{ij}^{A} and γ_{ij}^{B} are covariance computed segregating trials according to types A and B, respectively. A similar decomposition can be done for variances of the firing rates (x=i, j),

$$\sigma_x^2 = \frac{N_A}{N}\sigma_{x,A}^2 + \frac{N_B}{N}\sigma_{x,B}^2 + \frac{N_A N_B^2 + N_B N_A^2}{N^3}\Delta\mu_x{}^2 \tag{S47}$$

Considering Equations S46 and S47 together,

$$R_{ij} = \frac{\frac{N_A}{N}\sigma_{\mathrm{i,A}}\sigma_{\mathrm{j,A}}\rho_{ij}^{A} + \frac{N_B}{N}\sigma_{\mathrm{i,B}}\sigma_{\mathrm{j,B}}\rho_{ij}^{B} + K\Delta\mu_i\Delta\mu_j}{\sqrt{\frac{N_A}{N}\sigma_{i,A}^2 + \frac{N_B}{N}\sigma_{i,\mathrm{B}}^2 + K\Delta\mu_i{}^2}\sqrt{\frac{N_A}{N}\sigma_{\mathrm{j},A}^2 + \frac{N_B}{N}\sigma_{\mathrm{j,B}}^2 + K\Delta\mu_j{}^2}} \tag{S48}$$

where $K = \frac{N_A N_B^2 + N_B N_A^2}{N^3}$.

For the particular case in which the variances of the two neurons are equal in both type of trials and the number of A and B trials is the same,

$$R_{ij} = \frac{\frac{1}{2}\left(\rho_{ij}^{A} + \rho_{ij}^{B}\right) + \frac{1}{4}\delta_i\delta_j}{\sqrt{\left[1+\left(\frac{\delta_i}{2}\right)^2\right]}\sqrt{\left[1+\left(\frac{\delta_j}{2}\right)^2\right]}} \tag{S49}$$

that is Equation 7 of the main text. Equation S49 expresses the usual correlation coefficient R in terms of correlations coefficients computed segregating trials according to the animal's choice (ρ_{ij}^{A} and ρ_{ij}^{B}) and in terms of δ_i and δ_j (which is the quantity that determines the CP index, Equation S10).

To obtain further insight about the meaning of this expression and to justify the approximate expression for the CP index (Equation 8), let us consider the Taylor expansion of $\delta_{ij} = \left(\delta_i + \delta_j\right)/2$ around $\varepsilon = \left(\delta_i - \delta_j\right)/2 \sim 0$. Under this approximation,

$$\delta_{ij} = \delta_{ij}^{0} + \frac{\varepsilon^2}{2} f\left(R_{ij,}\rho_{ij}^{AB}\right) + O(\varepsilon^4) \tag{S50}$$

with

$$\left(\frac{\delta_{ij}^0}{2}\right)^2 = \frac{R_{ij} - \rho_{ij}}{1 - R_{ij}} \tag{S51}$$

$$f\left(R_{ij}, \rho_{ij}^{AB}\right) = \frac{-1 - 3\left(\delta_{ij}^0\right)^2 + \rho_{ij}^{AB}\left[\left(\delta_{ij}^0\right)^2 - 1\right]}{4(\rho_{ij}^{AB} - 1)\left[\delta_{ij}^0 + \left(\delta_{ij}^0\right)^3\right]} \tag{S52}$$

where $\rho_{ij}^{w} = \left(\rho_{ij}^{w,A} + \rho_{ij}^{w,B}\right)/2$. Comparing the first and second term of the expansion we conclude that the accuracy of this approximation is good if $R_{ij} - \rho_{ij}^{AB} \neq 0$. For example, when $R_{ij} - \rho_{ij}^{AB} > 0.15$, a pair (i,j) with δ_i and δ_j differing in 30% ($\varepsilon = 0.3$) will give a relative error of less than 9%.

When Equation S49 is applied to a pair of neurons (1,2) within the same pool we have,

$$\left(\frac{\delta_0}{2}\right)^2 \sim \frac{R_{12}^{w} - \rho_{12}^{w}}{1 - R_{12}^{w}} \tag{S53}$$

where $\delta_0 = \left(\delta_1 + \delta_2\right)/2$. This is Equation 8 in the main text. Notice from Equation S53 that,

$$R_{12}^{w} - \rho_{12}^{w} \geq 0 \tag{S54}$$

which means that correlations for neurons in the same pool, computed with fixed-choice trials, are smaller than those obtained with the whole set of trials.

For small δ_0 the CP index of neurons in the pair (1,2) can be linearized, obtaining a rather simple relationship between this index and noise correlations,

$$\left(\mathrm{CP} - \frac{1}{2}\right)^2 \sim \frac{1}{\pi}\frac{R_{12}^{w} - \rho_{12}^{w}}{1 - R_{12}^{w}} \tag{S55}$$

and the corresponding average over the population of pairs. The error introduced by this linearization becomes significant at large values of CP (8% for CP = 0.75).

Instead, if we consider that a pair of neurons (1,3) belonging to different pools have opposite sign of the difference of mean responses in the two conditions (that is, $\delta_1\delta_3 < 0$), from Equation S49 we have that

$$R_{13}^{b}=\frac{\frac{1}{2}\left(\rho_{13}^{b,A}+\rho_{13}^{b,B}\right)+\frac{1}{4}\delta_1\delta_3}{\sqrt{\left[1+\left(\frac{\delta_1}{2}\right)^2\right]\left[1+\left(\frac{\delta_3}{2}\right)^2\right]}}$$

$$R_{13}^{b}\le\frac{1}{2}\left(\rho_{13}^{b,A}+\rho_{13}^{b,B}\right)+\frac{1}{4}\delta_1\delta_3 \qquad \text{(S56)}$$

$$R_{13}^{b}\le\rho_{13}^{b}$$

If for a pair of neurons from different pools $\delta_3\sim-\delta_1$, then $\delta_0=\frac{1}{2}\left(|\delta_1|+|\delta_3|\right)$ can be approximated in terms of the correlation coefficients of the pair (1,3) as,

$$\left(\frac{\delta_0}{2}\right)^2\sim\frac{\rho_{13}^{\mathrm{b}}-R_{13}^{b}}{1-R_{13}^{b}}, \qquad \text{(S57)}$$

where we defined $\rho_{13}^{\mathrm{b}}=\left(\rho_{13}^{\mathrm{b},A}+\rho_{13}^{\mathrm{b},B}\right)\big/2$.

As a more complex example, we now consider a two-pool network satisfying both conditions $\delta_1\sim\delta_2\sim-\delta_3$. Equations S53 and S57 can be seen as a constraint on the correlation structure of the network. Using this constraint, and replacing pair-wise correlations by their population-averaged values ($\bar{R}^{w}$, $\bar{R}^{b}$, $\bar{\rho}^{\mathrm{w}}$, $\bar{\rho}^{\mathrm{b}}$) we obtain

$$\overline{\mathrm{CP}}\sim\frac{1}{2}\mathrm{erfc}\left(-\sqrt{\frac{\left(\bar{R}^{w}-\bar{R}^{b}\right)-\left(\bar{\rho}^{w}-\bar{\rho}^{b}\right)}{2-\left(\bar{R}^{w}-\bar{R}^{b}\right)}}\right) \qquad \text{(S58)}$$

which is Equation 9 in the main text.

A similar study can be done for the other generalized indices. For a pair of neurons (1,2) within the same neural pool and a pair (1,3) between different pools, we have

$$\Delta_{2,w}=\frac{\sqrt{2}\delta_0}{\sqrt{1+\rho_{12}^{w}}},\qquad \left(\frac{\delta_0}{2}\right)^2\approx\frac{R_{12}^{w}-\rho_{12}^{w}}{1-R_{12}^{w}}$$

$$\Delta_{2,b}=\frac{\left(1-\mathrm{d}\right)\delta_0}{\sqrt{1+D^2+2D\rho_{13}^{\mathrm{b}}}},\qquad \left(\frac{\delta_0}{2}\right)^2\approx\frac{\rho_{13}^{\mathrm{b}}-R_{13}^{b}}{1-R_{13}^{b}}. \qquad \text{(S55)}$$

The $\mathrm{CP_N}$ index can be related to the choice-conditioned correlation coefficients as,

$$\Delta_N = \frac{\sqrt{N}\left(\bar{\delta}_+ + D\bar{\delta}_-\right)}{\sqrt{\left(1+D^2\right)\left[1+\left(N-1\right)\bar{\rho}^w\right]+2DN\bar{\rho}^b}} \tag{S56}$$

where $\overline{\delta_+}$ and $\overline{\delta_-}$ can be obtained averaging the expression in Equation S51 over the population of positive and negative neuronal pairs, respectively.

Dependence of the CP_N index on the number of neurons

In order to explore the dependence of the CP_N index on the number of neurons, we calculated this index taking from our database subsets of neurons with different sizes. We performed this analysis on the population of delay-activity PM neurons. We randomly selected the same number of neurons from the positive and negative pools and averaged the results over 100 repetitions for each population size.

The CP_N index for a two pools system with N neurons each can be expressed as,

$$\Delta_N = \frac{\sqrt{N}\left[\left(\bar{\mu}_+^A - \bar{\mu}_+^B\right) + D\left(\bar{\mu}_-^A - \bar{\mu}_-^B\right)\right]}{\sqrt{\overline{\sigma_+^2} + \left(N-1\right)\overline{\gamma_{++}^c} + D^2\left(\overline{\sigma_-^2} + \left(N-1\right)\overline{\gamma_{--}^c}\right) + 2DN\overline{\gamma_{+-}^c}}} \tag{S57}$$

Figure S4 (top row) shows the time course of CP_N, the numerator and the denominator of Equation S57, for N taking values between 5 and 80 per pool. For small N the CP_N index increases with N while at larger values of N it saturates. Figure S4 (bottom row) shows the dependence on N of the CP_N index, num^2 and den^2 for a fixed time, t=1s.

The dependence of CP_N on N can be explained in terms of Equation S57. For small N, $\overline{\sigma_{+,\mathrm{AB}}^2} >> \left(N-1\right)\overline{\gamma_{++,AB}^c}$ and $\overline{\sigma_{-,\mathrm{AB}}^2} >> \left(N-1\right)\overline{\gamma_{--,AB}^c}$. Therefore the CP_N index increases with N following the dependence on N of the numerator. For large N, $\overline{\sigma_{+,\mathrm{AB}}^2} << \left(N-1\right)\overline{\gamma_{++,AB}^c}$ and $\overline{\sigma_{-,\mathrm{AB}}^2} << \left(N-1\right)\overline{\gamma_{--,AB}^c}$, therefore Equation S57 can be approximated as,

$$\Delta_N \approx \frac{\sqrt{N}\left[\left(\bar{\mu}_+^A - \bar{\mu}_+^B\right) + D\left(\bar{\mu}_-^A - \bar{\mu}_-^B\right)\right]}{\sqrt{N}\sqrt{\overline{\gamma_{++}^c} + D^2\overline{\gamma_{--}^c} + 2D\overline{\gamma_{+-}^c}}} \tag{S58}$$

which explains the linear dependence of num^2 and den^2 on N (Figure S4, bottom row) and the saturation of CP_N at large N.

SUPPLEMENTAL REFERENCES

1. Green, D. M., and Swet, J. A. (1966). Signal detection theory and pshycophysics (New York, Wiley).

2. Dayan, P. and Abbott L. (2001). Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems (Cambridge, The MIT Press).