Feedback Dynamics Determine the Structure of Spike-Count Correlation in Visual Cortex

Adrian G. Bondy[1] & Bruce G. Cumming[1]

[1] Laboratory of Sensorimotor Research, National Eye Institute, NIH
49 Convent Drive, Rm. 2A50
Bethesda, MD 20892

Correspondence:
Adrian Bondy
49 Convent Drive, Room 2A50
Bethesda, MD 20892
adrian.bondy@gmail.com
301-451-4926

The variable spiking discharge of sensory neurons in response to a fixed stimulus tends to be weakly correlated (spike-count correlation, r_{sc}), an observation with profound implications for neural coding of sensory information. However, the source of r_{sc} is unclear. It is widely thought to reflect "bottom up" stochastic noise in shared sensory afferents. However, it may also be generated by changes over time in feedback from higher-order brain regions. Here we test this alternative directly by measuring spiking activity in populations of primary visual cortical (V1) neurons in rhesus monkeys while the animals performed different visual discrimination tasks on the same set of visual inputs. We found that the structure of r_{sc} (the way r_{sc} varied with neuronal stimulus preference) changed dramatically with task instruction despite identical retinal input, directly implying that r_{sc} structure primarily reflects feedback dynamics engaged by the task, not noise in sensory afferents. These results fundamentally alter our view of the origin and function of r_{sc} in sensory neurons, suggesting that correlated variability may be best described as a signature of neural computation rather than stochastic sensory encoding.

Neurons in sensory cortices fire at a rate that is highly dependent on sensory input. However, they also display pronounced response variability that tends to be weakly correlated between pairs of sensory neurons[1]. Because perceptual states are thought to be generated by pooling responses of many sensory neurons, correlated variability can have a profound effect on the reliability of sensory processing: while independent variability can be averaged away by pooling enough neurons, correlated variability cannot necessarily. As a result, correlated variability is thought to profoundly influence the fidelity of sensory information in the brain[2–6] and, relatedly, to generate trial-to-trial correlations between single-neuron variability and perceptual reports in psychophysical tasks (choice-related activity)[7–9].

Unfortunately, little is currently known about the origin of r_{sc}. Presumably r_{sc} is derived from variability in common inputs. The predominant view is that stochastic noise in the afferent pathway is the primary source of this common variable input[10–12]. Consistent with this, r_{sc} correlates with physical proximity and similarity in stimulus preference, both of which are predictive of greater shared afferent input[2,10,11,13–15]. However, sensory cortical areas receive only a minority of their inputs from the upstream brain regions conveying sensory information from the periphery[16,17]. Consequently, variation over time in shared inputs from downstream areas (i.e. "top-down"; "feedback"), for instance in signals related to attention, reward, arousal, or decision-making, may make a significant contribution to r_{sc}. It would not necessarily follow that this source of correlated variability impacts downstream sensory decoding, since downstream areas may have access to the state of these other inputs. Rather, correlated variability in a sensory area may best be seen as an effect of downstream computation rather than noise in the brain's sensory representations.

Here, we directly investigate the relative contribution of feedforward and feedback sources of correlated variability in sensory neurons. We recorded spiking activity in populations of primary visual cortical (V1) neurons in macaque monkeys performing different orientation discrimination tasks using the same set of stimuli. The only difference between the tasks was the pair of orientations being discriminated. If r_{sc} primarily reflects noise in sensory afferents, it should be invariant to changes in the task given fixed retinal input. Instead, we hypothesized a task-dependent pattern of correlated variability across the V1 population generated by feedback specifically targeting neurons tuned to the orientations being discriminated. To test this, we measured the r_{sc} structure (how r_{sc} varies as a function of pairwise orientation preference) of V1 under the different task contexts. We found that this structure was highly task-dependent, in precisely the manner predicted by our hypothesis. Strikingly, we were unable to identify any r_{sc} structure that did not change with the task, entirely inconsistent with noise in sensory afferents as a major source.

We go on to show how these results fundamentally change our view of the role of r_{sc} in decision making and information coding. First, we show that the feedback dynamics introduce a pattern of correlated variability that degrades the accuracy of pooled sensory signals that can be extracted from V1 by a traditional linear decoder. However, our discovery of its feedback origin points to the possibility that the brain can, in principle, outperform such a decoder by exploiting knowledge of the changing state of downstream brain areas when decoding V1 activity. Next, we show quantitatively that these feedback dynamics are the primary source of the choice-related activity we observe in V1, clarifying an ongoing debate about the origin of choice-related signals in sensory neurons.

Results

We trained two rhesus monkeys (*Macaca mulatta*) to perform different versions of a two-alternative forced choice (2AFC) coarse orientation discrimination task (Fig. 1), used previously[18]. On a given trial, the subject was shown a dynamic, 2D filtered noise stimulus for a fixed duration of 2 seconds, after which it had to make a saccade to one of two choice targets to report the stimulus orientation. The discriminanda were always an orthogonal pair of orientations (for instance, horizontal and vertical). This pair defined the “task context” and was explicitly cued using two oriented Gabor patches as the choice targets. The stimuli were bandpass filtered in the Fourier domain to include only orientations within a predetermined range. On a given trial, the stimulus filter was centered on one of the two discriminandum orientations and its orientation bandwidth was varied to modulate task difficulty. The stimulus was placed over the joint receptive field of the population of V1 neurons being recorded. We included 0%-signal trials for which the stimuli were unfiltered for orientation. These were statistically identical across task contexts, allowing us to examine the effect of task context on r_{sc} in the presence of a fixed retinal input.

Our approach relied on being able to detect changes in r_{sc} structure that depended on the task context. Thus it was critical that the subjects in fact based their choices on the presence of the cued set of orientations. To ensure this, we used psychophysical reverse correlation[18–20] (PRC; see Methods) to directly measure the influence of different stimulus orientations on the subject’s choices (the “psychophysical kernel”). We found that subjects required multiple days of retraining after a change in the task context to fully update their psychophysical kernel. For this reason, we kept the task context fixed for the duration of each recording session, and excluded sessions from the analysis if subjects failed to demonstrate an appropriate kernel (Supplementary Fig. 1).

We hypothesized that the pattern of correlated variability in V1 is generated by task-related feedback dynamics. We specifically considered the possibility that a feedback signal is alternatingly targeted towards neurons that represent the two orientations the animal must choose between, introducing a particular pattern of correlated fluctuations in V1 that depends on orientation preference. A key motivation for this hypothesis is the observation of correlations between neuronal variability and choice in V1 neurons during performance of this task. While the initial interpretation of choice-related activity was that it reflects the feedforward influence of neuronal variability on choice[21–26], several recent studies suggest choice-related activity may be an effect of choice on sensory neurons via feedback[27,28]. A feedback origin of choice-related activity necessarily implies the presence of task-related feedback signals, such that on trials when the subject reports orientation 1, feedback excites V1 neurons preferring orientation 1, and so forth. However, similar predictions also arise from considering the effect of fluctuations across trials in the allocation of feature-based attention[27] or Bayesian priors[29] during a 2AFC task. Therefore the presence of task-dependent r_{sc} structure would, on its own, be consistent with a range of feedback mechanisms with potentially diverse functional roles, as we discuss later.

To make quantitative predictions for the effect of task-related feedback on correlated fluctuations in V1, we parameterize the hypothesized feedback in V1 as a sinusoidal function of preferred orientation with a peak and trough at the discriminandum orientations, such that the effect on V1 firing rates at a given instant is simply a scalar multiple of this function[30]. Examples for two task contexts (cardinal and oblique discrimination) are shown in Fig. 2d,e. This parameterization makes a specific, testable prediction for the way r_{sc} will vary as a function of pairwise preferred orientation (i.e. the "r_{sc} matrix"). In particular it defines a mode of covariability in V1 which is equivalent to an eigenvector of the r_{sc} matrix. Assuming it is the only eigenvector, the predicted r_{sc} matrix exhibits a lattice-like pattern (Fig. 2b,c) characterized by high r_{sc} values for pairs preferring the same discriminandum orientation (within-pool pairs), low values for pairs preferring opposite discriminandum orientations (between-pool pairs), and average values for pairs that are not task-relevant. Because any additional source of global V1 fluctuations (such as arousal) would introduce positive correlations amongst all neurons, the average magnitude of r_{sc} is unconstrained. Crucially, the predicted pattern changes systematically with the task, such that the peaks and troughs in correlation are always aligned to the discriminandum orientations. This amounts to a shift in the r_{sc} structure between the two matrices along the diagonal by an amount matching the change in the task context.

The predominant view that the structure of r_{sc} in a sensory area is primarily determined by noise in common afferent inputs yields a different prediction for the r_{sc} matrix: a diagonal banded structure (X), in which r_{sc} depends only on the similarity in orientation preference between a given pair. This pattern (known as a "limited-range" correlation structure) is thought to be due to the anatomical convergence of afferent inputs to neurons with similar stimulus preferences[10–12] and has been postulated as the critical organizational logic of r_{sc} in sensory neurons[3–5,31]. Crucially, limited-range correlations should be invariant to the changes in task context since they depend only on noise in the afferent pathway. A main motivation for belief in limited-range correlations is the positive correlation between r_{sc} and similarity in stimulus tuning of neuronal pairs, including in V1[2,10,11,13–15]. However, this observation is also consistent with the pattern predicted by task-dependent feedback—on average, neurons with more similar orientation preference produce higher r_{sc} values under that prediction, as well (Fig. 2f). Therefore, only by measuring the full r_{sc} matrix across multiple task contexts, as the present study is the first to do, can these two hypotheses be distinguished.

We recorded extracellular spiking activity from populations of single V1 neurons using multi-electrode arrays. Neurons were excluded from analysis if they were not well orientation tuned, as determined in separate blocks of trials (see Methods). Yields varied across sessions, with a mean of 20 simultaneously recorded pairs (from 5 cells) per session. The final dataset includes 811 simultaneously recorded pairs from 200 cells. For each pair, we calculated its r_{sc} value as the Pearson correlation between trial-length spike counts in response to identical stimuli. Measuring r_{sc} using 0%-signal trials isolates any changes due to the task context. In practice, we found that measuring r_{sc} across all trials (after normalizing spike counts to remove rate changes due to the stimulus) did not qualitatively alter our main results (Supplementary Fig. 2) and increased statistical power, so this is what we report.

R_{sc} Structure Changes Systematically with the Task Context

The results were striking: r_{sc} structure changed dramatically with task context, closely matching the prediction based on task-related feedback and inconsistent with r_{sc} structure primarily driven by afferent noise. To show this, we first divided the recording sessions into two groups based on the task context used (Fig. 3a). To estimate the r_{sc} matrix for a given subset of sessions, we combined data from pairs recorded across sessions. The location of a given r_{sc} value in the matrix was determined by the preferred orientations of the pair. We then applied a von Mises smoothing kernel to obtain a continuous, smooth measure of the rsc matrix.

We predicted that both matrices would contain a similar lattice-like pattern with peaks and troughs in r_{sc} that matched the within- and between-pool regions of the matrix. However, because the task context was different, this predicts patterns of r_{sc} shifted along the diagonal by an offset matching the change in the discriminandum orientations (Fig. 3b,c). Furthermore, the largest eigenvector of the two predicted matrices should resemble sinusoids with a peak and trough matching the two discriminandum orientations (Fig. 3d,e). The observed data matched both of these predictions. Peaks and troughs in r_{sc} were markedly different among the two subsets of data, with both patterns resembling the prediction based on task-dependent feedback (Fig. 3f,g). Furthermore, the largest eigenvector of the two matrices resembled a sinusoid with a circular mean orientation that was not significantly deviated from the discriminanda (Fig. 3h,i).

The task-dependent r_{sc} structure in the data is most clearly evident by combining r_{sc} values from all 811 simultaneously recorded pairs. To illustrate this, we generated a task-aligned average r_{sc} matric by expressing each neuron's preferred orientation relative to the two discriminandum orientations on its respective recording session. In this task-aligned coordinate frame, 0^{o} and 90^{o} always index the discriminandum orientations. We found that this task-aligned matrix contained a lattice pattern that almost precisely replicated the prediction based on the task-related feedback hypothesis (Fig. 4a,b). Furthermore, the first eigenvector (Fig. 4d) resembled a sinusoid with a peak not significantly different from 0^{o}, demonstrating a striking degree of alignment across sessions between the structure of correlated variability in V1 and the subject's task. These features were also present in the task-aligned r_{sc} matrix when computed separately for each subject (Supplementary Fig. 3), so we performed further analyses on the combined dataset. We found that the first eigenvalue of the task-aligned matrix was much greater than chance, suggesting the r_{sc} matrix is largely explained by its first eigenvector, again consistent with a single source of covariability that depended on the task (Fig. 4e). The chance distribution was obtained by randomly translating each individual r_{sc} measurement along the diagonal, ruling out the possibility that we observed a lattice pattern simply due to a diagonal ridge and sampling error ($p<0.005$, permutation test). Taken together, these results show that the V1 r_{sc} matrix closely matched the prediction based on task-related feedback and does not primarily reflect limited-range correlations introduced by a fixed source of noise in shared sensory afferents.

We observed a different result during separate blocks of trials interleaved in the same recording sessions, during which the subject fixated passively for reward but the same set of stimuli was shown. During these blocks, the r_{sc} matrix could not be distinguished from a diagonal ridge (Fig. 5). This demonstrates that the dynamic changes observed during task performance depended on active task engagement, and could not be explained, for instance, simply as an effect of recent task experience. This is further evidence that trial-by-trial variation in feedback is responsible for the r_{sc} structure, rather than, for instance, slow time scale changes in local V1 circuitry. We also ruled out a number of potential confounds related to the retinal input that could in principle have produced a task-dependent r_{sc} matrix, such as fixational eye movements and the effect of trial history, strengthening our interpretation that centrally-generated signals reflecting task engagement are responsible (see Supplementary Discussion §1 and Supplementary Figs. 4-7).

Segregating Fixed and Task-Aligned Components of R_{sc} Structure

The preceding analysis established that a substantial component of r_{sc} structure changes with task context. To better quantify this, and to determine if there is also a component that remains fixed, we turned to a statistical model. The model contained two components: a fixed component (an r_{sc} matrix for orientation that did not change with the task), and a task-aligned component (an r_{sc} matrix that did change). Each r_{sc} measurement was described as the sum of values at the appropriate points in the two matrices. For the fixed component, this location was determined by the raw orientation preferences of each pair. For the task-aligned component, it was determined using the preferences expressed relative to the discriminandum orientations. By construction, if r_{sc} depends only on the raw orientation preferences of neuronal pairs, with no effect of task context, then the model assigns large coefficients to the fixed component and coefficients of zero to the task-aligned component. If r_{sc} is entirely task-dependent, the reverse is true.

When fitted to the observed r_{sc} measurements, the task-aligned component of the model explained most of the explainable variance in the data (79%. Fig. 6a). Not surprisingly, its shape recapitulated the task-dependent lattice pattern in the observed data. The fixed component had a markedly smaller amplitude, with a less organized structure that did not clearly resemble a diagonal ridge (Fig. 6b). Removing the fixed component from the model altogether had little effect, while removing the task-aligned component dramatically impaired model performance (Fig. 6c). (We were able to reproduce these model results individually for data from one subject). This demonstrates that the majority of r_{sc} structure in V1 changes dynamically with task instruction. We failed to reliably identify a fixed source of r_{sc} structure, such as the limited-range correlations postulated in prior studies, during task performance.

Effect of Task-Dependent R_{sc} Structure on Neural Coding

Spike-count correlations in sensory neurons are typically studied with the view that they are a source of noise that impacts the ability of a downstream brain area to decode a sensory input. Our results show that the predominant source of r_{sc} structure in V1 is top-down in origin V1, at least during performance of a discrimination task. Therefore, the impact on decoding depends crucially on whether the spikes in V1 generated by feedback act as an additional source of sensory noise or can in some way be taken into account by the decoder. To clarify this distinction, consider two decoding schemes. The first ("purely sensory decoder") is applied to the activity of V1 neurons only, and so cannot differentiate different sources of correlation. The second ("extended decoder") is applied to the activity of V1 neurons and the activity of the feedback connections arriving in V1. If there are variations in the activity of feedback connections, the extended decoder can perform much better than a purely sensory decoder by effectively discounting the spikes in V1 that are generated by feedback. Since our results imply that that there is variation in the activity of feedback connections during psychophysical tasks, it follows that current analyses, almost exclusively based on the assumption of purely sensory decoders, may be misleading.

Since we do not know whether the brain has access to extended decoders, we investigated the impact of the observed r_{sc} structure under the assumption that extended decoding is not used. Assuming the brain uses "purely sensory" decoders and optimal linear readout weights, it is known[32] that a particular r_{sc} structure places a strict upper bound on decoding accuracy. These so-called "differential" correlations are those that resemble the correlations produced by changes in the stimulus along the axis defining the task. When stimulus-independent correlations share this structure, they are indistinguishable from task-relevant stimulus changes and hence confound subjects' judgments. We examined the structure of differential correlations in our data set by comparing the mean responses of each neuronal pair to the various stimuli used in the task. Specifically the differential correlation for a pair is given by the product of the slopes of the mean responses as a function of signal strength (Fig. 7a)[32]. We then plotted these values as a matrix indexed by task-aligned preferred orientation (Fig. 7b). This showed a lattice-like pattern strikingly similar to the observed r_{sc} matrix (Fig. 4b). Confirming this similarity, the task-aligned component of r_{sc} structure identified by the regression model was highly correlated on a pair-by-pair basis with the differential correlations (r=0.62, Fig. 7c). In other words, the structure of stimulus-independent covariability in the V1 population was closely similar to the structure of covariability introduced by stimulus variation. In a sense this is not suprising since the r_{sc} structure reflects feedback that is matched to the task. However, this implies that task-related feedback may have the consequence of contaminating the sensory representation in V1. Critically, this implication depends on the assumption that the decoder knows nothing about the feedback (a "purely sensory" decoder). While this assumption is currently widespread, out results suggest it may be highly misleading.

Relationship between R_{sc} Structure and Perceptual Choice

We found significant choice-related activity in the V1 population during performance of the task, similar to a recent study. For each neuron, we calculated its Choice Probability (CP), a metric which quantifies the probability with which an ideal observer could correctly predict the subject's choice from that neuron's trial-by-trial variability[33,34]. Across the population, we found an average CP of 0.54 for task-relevant neurons, significantly above chance level (Fig. 8a), and similar in magnitude to reports from another study using the same task[18].

Currently, the source of CP in sensory neurons is not well understood, despite its ubiquity. The traditional view is that it reflects the influence of neuronal variability on choice[21–26]. However, recent studies suggest that at least some of CP is due to a feedback effect of choice on sensory neurons[27,28]. Theoretical studies have emphasized that CP in a population of sensory neurons is closely related to the structure of spike-count correlation[8,9,34,35]. Simply put, if many sensory neurons have variability that is correlated with choice, then this implies the variability of individual neurons is also correlated. It follows that there exist correlated fluctuations in the V1 population that relate to the subject's choices in our task. However, the relationship to choice could reflect one or both of two directions of causality: 1) correlated fluctuations directly affect the choices a subject makes trial to trial (a feedforward source of CP); or 2) the correlated fluctuations reflect variation across trials in a feedback signal related to the upcoming choice (a feedback source).

Our detailed measures of r_{sc} structure during performance of a discrimination task allow us to make significant progress in addressing the origin of choice-related activity. We reasoned that a feedback effect of CP would predict an attenuated r_{sc} structure across trial in which the subject made the same choice. We found a significant, but very modest, attenuation (about 3%). Similarly, the r_{sc} structure was also attenuated on high-signal trials relative to 0% signal trials, in a manner which depended weakly, but systematically, with the strength of the signal (Supplementary Fig. 2). This is also consistent with a feedback source of CP, since there is naturally less variability in choice on trials with high signal. These data suggest an involvement of feedback related to choice in generating the r_{sc} structure in V1, but also rule out a simple post-decisional mechanism in which the state of feedback is perfectly correlated with the final report. Consistent with this interpretation, we found that the rsc structure, when calculated using spikes from 200-ms bins during the trial, showed a relatively stable timecourse that did not grow in amplitude with decision formation (Supplementary Fig. 8). We conclude that the task-related feedback relays information related to but not determined by the final report, such as biases, the decision variable, or attention to orientation.

Next, we considered the possibility that CP partially reflects the feedforward influence of correlated fluctuations on choice. We made use of the known, but untested, relationship between the spike-count correlations for a pool of sensory neurons, their (linear) readout weights, and their CPs that follows from assuming a feedforward "purely sensory" decoder[8]. Our novel approach of measuring the entire r_{sc} matrix during task performance allows us to test these predictions for the first time. Furthermore, our separation of the r_{sc} structure into putatively feedforward and feedback components allows us to distinguish components of CP due to the influence on choice of each of these potential sources of correlated variability. If the fixed component alone could explain CP, this would be compatible with the standard view that CP arises via the influence of sensory afferent noise on choice. If a task-aligned component is necessary, it indicates that CP requires the presence of task-related feedback.

To perform this analysis, we made the assumption that choices are based on a V1 population of infinite size and that neurons are defined only by their preferred orientation. We can then define neuronal CP as a continuous function of task-aligned preferred orientation (Fig. 8b), similar to how we define the r_{sc} matrix in Fig. 4e. To obtain this full CP profile for the V1 population, we again combine measures across sessions. We found that the relationship between CP and preferred orientation followed a predictable pattern: CPs were highest for task-relevant neurons (those with preferred orientation close to one of the discriminandum orientations) and were at chance level for neurons that were not task-relevant.

We then made a quantitative prediction for the CP profile that would be observed given a linear readount of the V1 population characterized by the r_{sc} matrix observed during task performance (Fig. 8b). We found that the predicted profile closely matched what we observed empirically, the first quantitative demonstration that feedforward pooling is consistent with experimentally observed measurements of r_{sc} structure and CP. There were two sets of free parameters: the readout weights applied to neurons at each preferred orientation and a uniform scaling factor reflecting the possibility of decision noise after the pooling stage. The readout weights were unobserved, but all results we report were insensitive to the readout weights. The pooling noise term was fit to the data. (see Method, Supplementary Discussion and Supplementary Fig. 10).

Next, we repeated these predictions using each component of the r_{sc} structure ("fixed" and "task-aligned") independently, keeping the pooling noise term fixed. We found that we could predict most of the observed CP (82% of its overall magnitude) using only the task-aligned component, while the fixed component alone could explain only a small fraction (Fig. 8c,d). (This does not follow necessarily from the larger amplitude of the task-aligned component, but depends crucially on the pattern of correlation.) This finding rules out the possibility that CP is primarily generated by the feedforward effect of noise in sensory afferents on perceptual decisions. Instead, CP is primarily generated by the task-related feedback signals that generate the dynamic pattern of r_{sc} in the population. The success of the feedforward predictions for explaining the observed CP shows that we cannot exclude the possibility that CP arises partially through the feedforward effect on choice of the correlated fluctuations introduced by feedback, although the dependence of the r_{sc} structure on variability in choice suggests that feedback already contains information about the upcoming report. Thus our results shed new light on the origin of choice-related activity while also, in a sense, muddying the distinction between feedforward and feedback pathways that are causally responsible. One possibile interpretation of the data, consistent with other reports[29,36], is that CP reflects a self-reinforcing loop in which decision-related feedback generates correlated fluctuations in V1 that in turn influence the decision. Assessing this possibility will require future experiments to better understand the information conveyed by task-related feedback signals and may require knowing whether the brain in fact has access to "extended" decoders.

Discussion

Correlations in the variability of sensory neurons may have profound implications for the way sensory information is encoded in the brain. However, attempts to study this typically take the view that correlated variability reflects stochastic noise accumulating in afferent pathways. In the present study, we show that the pattern of correlated variability in V1 is almost entirely dependent on the task context, independent of any changes in retinal input. This demonstrates that r_{sc} structure reflects an important source of feedback that is not well characterized as “noise”. This finding has profound implications for the way we think about sensory processing in the brain. First, it points to the importance of considering the interconnectedness of cortical areas as a crucial aspect of the way sensory information is encoded and decoded by the brain. As a result, strict hierarchical notions of “feedforward” and “feedback” do not adequately capture how sensory input is processed. Strikingly, this appears to be the case even at the earliest stage of cortical processing for vision in the primate.

Furthermore, studies examining the impact of correlated variability on the fidelity of the brain’s sensory representations have typically assumed that the brain applies decoders in a purely feedforward fashion: sensory input is encoded in the activity of a given population of neurons and subsequently read out downstream. However, this approach is ill-posed if r_{sc} is generated by variation in feedback signals, as the decoder may take into account the activity of neurons carrying the feedback signals (an “extended” decoder) to improve its performance. Our results demonstrate that such feedback modulation is a primary source of r_{sc} structure in V1, and therefore that the brain may benefit from “extended” decoding. In principle, this could be straightforward to implement: since we observed that the effect of feedback variability on V1 resembled the effect of changing the stimulus along the task axis (“differential correlations”), the subject could simply adjust its criterion trial-by-trial to remove the impact of the correlated variability on performance.

Our results may at first appear to contrast with reports about the influence of spatial attention on r_{sc}[37–39]. In those studies, context-dependent changes in r_{sc} appeared to improve the accuracy with which a model (“purely sensory”) decoder could reconstruct the stimulus, suggesting that top-down input can act as a control signal that modulates r_{sc} in a sensory brain area specifically to improve the signal-to-noise ratio of relevant sensory signals. By contrast, the data presented here suggest that r_{sc} can be, in the first place, a consequence of variation in feedback inputs that change with a subject’s task. Bridging this apparent gap, a recent reanalysis[40] of neuronal data from one of those studies[38] suggests this may also be the mechanism by which spatial attention operates. The new study revealed that the reduction in r_{sc} under spatial attention could be explained by an attenuation of the ongoing variability of a small number of common gain-modulating inputs, presumably feedback in origin. This is consistent with the view that spatial attention defines a condition under which downstream computation changes, and therefore may not be aimed at improving the sensory representation in sensory cortex.

Finally, we ask: what function do the task-dependent changes in r_{sc} reported here serve, particularly as they appear not to provide any improvement in the sensory representation contained in V1? While the form of these modulations is consistent with known feedback mechanisms—e.g. fluctuations in the allocation of feature-based attention and/or choice-related feedback—the potentially problematic effect of the correlated fluctuations they introduce speaks to the need for a principled, normative explanation, beyond reference simply to “downstream computation”. We argue that taking the view of perception as probabilistic inference may provide such an explanation. A companion paper in this issue[41] lays out the proposal that this computation is implemented in sensory neurons themselves, such that activation of a sensory neuron reflects both sensory input and prior beliefs about the presence of its preferred feature in the scene. In the context of our 2AFC task, a subject’s prior belief of a particular orientation being presented is concentrated bimodally around the two discriminandum orientations. It follows that trial-to-trial fluctuations in the prior will introduce task-dependent patterns of correlated variability in V1 similar to what we have reported. Impressively, this framework succeeds both in generating quantitative predictions that match many of the results reported here with few free parameters, and in providing a normative account for these data. Thus, we conclude that an intriguing interpretation of our findings deserving further study is that the structure of correlated variability in sensory cortex reflects an adaptive integration of incoming sensory input with prior knowledge about the structure of the world.

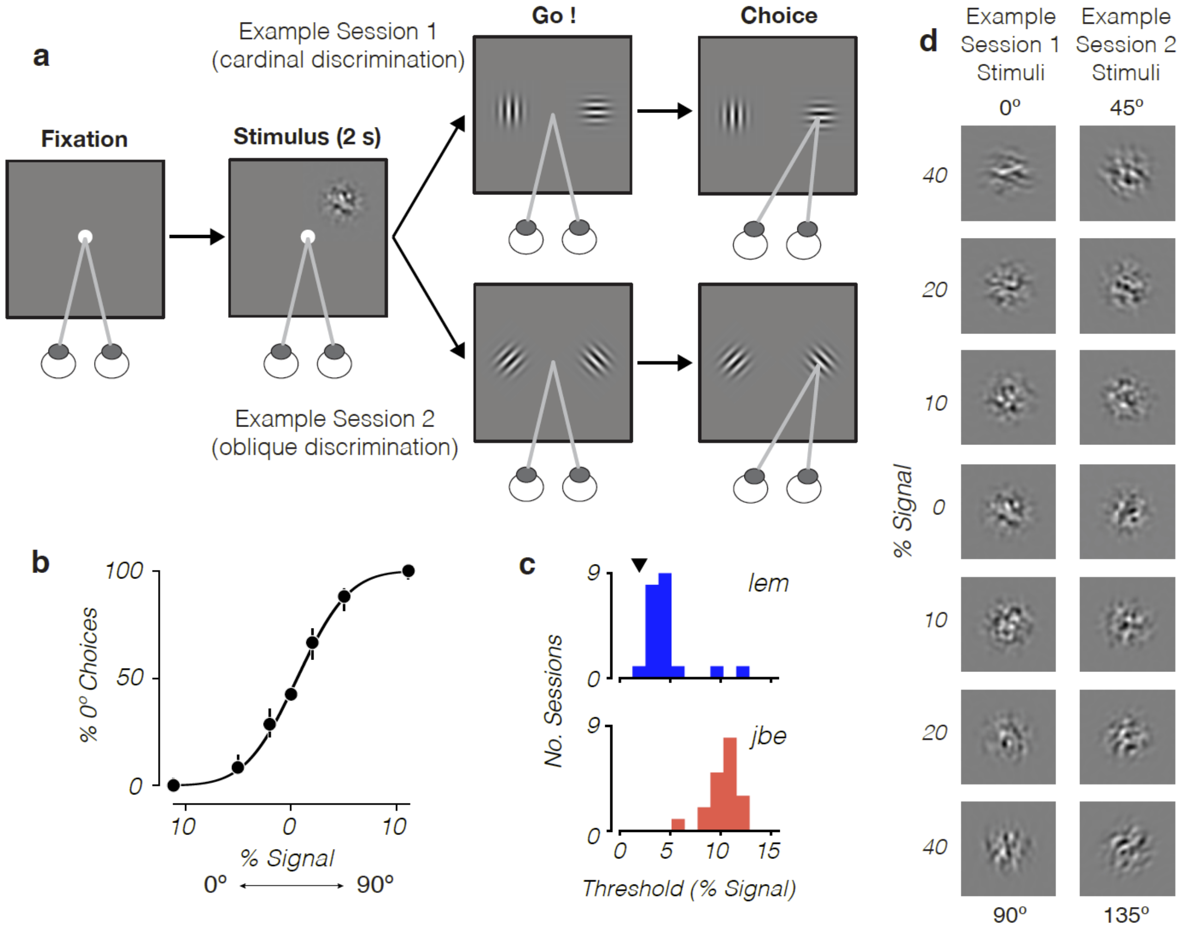

Figure 1. Orthogonal orientation discrimination task. **a**. Schematic illustration of the task. After the subject acquired fixation, a dynamic, filtered noise stimulus appeared for a fixed duration of 2 s. Then the subject had to saccade to the one of two orthogonal choice targets (Gabor patches) whose orientation matched the stimulus. Two example task contexts shown (cardinal and oblique discriminations). The task context was fixed in a given recording session, but varied across sessions. **b**. Psychometric function for monkey *'lem'*, example session. Black curve is a probit fit, and error bars are 95% confidence intervals. **c**. Histograms showing the distribution of psychometric thresholds across sessions for the two subjects. Thresholds were defined as the signal level eliciting 75% correct performance. Black triangle indicates the threshold corresponding to the example session in (b). **d**. Example single stimulus frames corresponding to the two example task contexts in (a). The stimuli consisted of dynamic, white noise filtered in the Fourier domain for orientation (see Methods). The filter was centered on one of the two discriminandum orientations and its bandwidth determined signal strength. A given trial consisted of many frames of

independent noise with a fixed filter. 0% signal stimuli were unfiltered for orientation and were statistically identical across task contexts.

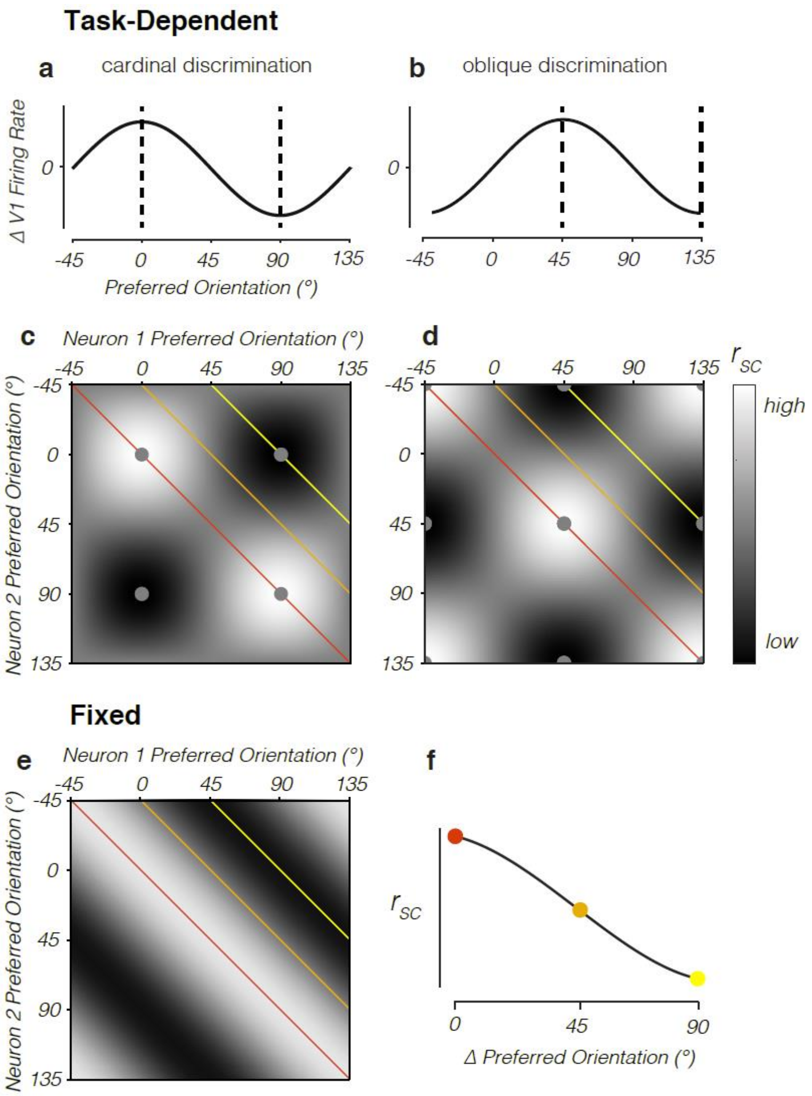

Figure 2. Predictions for Task-Dependent and Fixed Sources of r_{sc} structure. a-b. We hypothesize the presence of task-related feedback that selectively targets the two subpopulations of neurons tuned for the discriminandum orientations, alternating in its allocation across trials. These feedback dynamics are parameterized using a sinusoidal function of preferred orientation, with a peak and trough at the discriminanda. The effect on V1 firing rates on a given trial is a scalar multiple of this function. Examples for cardinal and oblique discrimination shown. Dashed lines indicate the discriminanda. **c-d**. Task-related feedback, as illustrated in (a) and (b), introduces correlated fluctuations amongst pairs of V1 neurons that depend on pairwise orientation preference, and which change systematically with the task context. We illustrate these dependencies using r_{sc} matrices indexed by neuronal prefered orientation. The structure in the matrices is given simply as the outer product of the sinusoidal functions in (a) and (b). In both cases, the result is a lattice-like pattern in the r_{sc} matrix, offset along the diagonal such that peaks and troughs in r_{sc} are aligned to the regions of the matrix indicating pairs preferring the same and opposite discriminandum orientations. These regions are marked by the black and white circles, respectively. Colored lines indicate a constant difference in neuronal preferred orientation (Δ preferred orientation: 0 (red), 45 (orange), and 90 (yellow)). **e.** The traditional view is that r_{sc} structure reflects a fixed source of stochastic afferent noise. Given the known anatomy of afferent input,

this is thought to generate “limited-range” correlations that decrease as a function of difference in stimulus preference (in this case for orientation), as illustrated here. Colored lines are as in (c-d). **f**. All predicted r_{sc} matrices (c-e) contain an identical downward-sloping relationship between r_{sc} and Δ preferred orientation. Thus, they cannot be distinguished using existing experimental observations relating r_{sc} and similarity in stimulus preference.

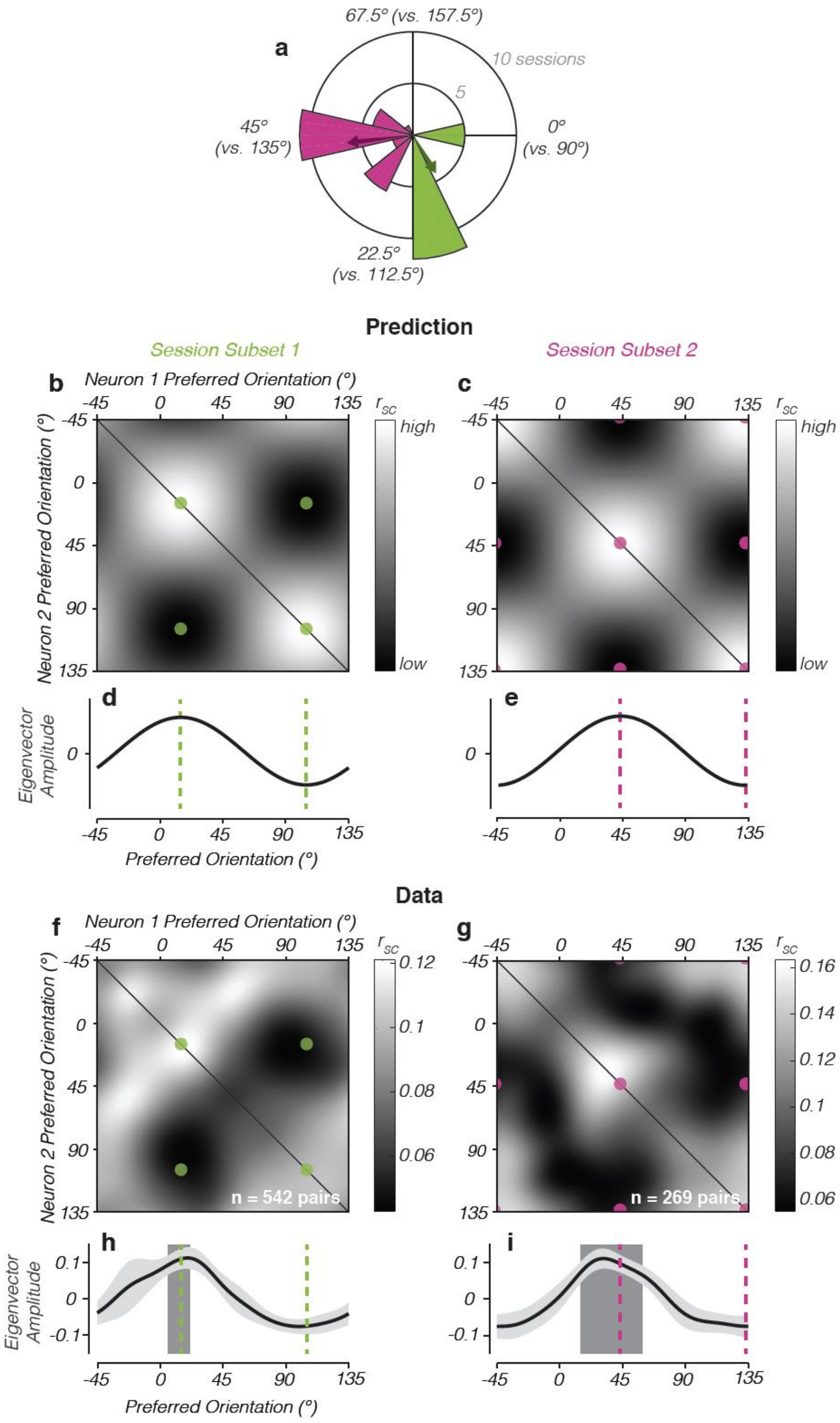

Figure 3. R_{sc} structure in V1 depends systematically on task context. a. To test for task-dependent changes in r_{sc} structure, we divided the set of recording sessions into two groups based on the set of discriminanda (task context) used. Polar histogram shows the distribution of task contexts used across sessions, with color indicating the division into two subsets. Note that the period of these angular variables is 90° because of the inherent orthogonality of the task. Colored arrows indicate the mean task context associated with each subset. **b-c**. The hypothesis that task-related feedback introduces task-dependent r_{sc} structure predicts a distinct r_{sc} matrix associated with the two subsets of sessions. The locations in the matrix where peaks and troughs in r_{sc} are predicted are highlighted with colored circles. These correspond to the mean discriminandum orientations indicated with arrows in (a). **d-e**. The sole eigenvectors of the matrices in (b) and (c) have peaks and troughs aligned to the mean set of discriminandum orientations associated with each subset (vertical lines). **f-g**. Observed r_{sc} matrices for the two subsets of sessions. These are obtained by combining r_{sc} measurements made across

the set of sessions in each subset and applying smoothing. The observed structure is distinct across the two matrices, corresponding to the predictions in (b-c). Peaks and troughs in r_{sc} closely match the predicted locations highlighted by the colored dots. **h-i**. The rank-1 eigenvector for the two matrices in (f) and (g) in each case closely resembled a sinusoid whose phase matches the task context. Light gray shaded region indicates +/- 1 boostrap s.e. Dark gray vertical bar indicates the peak in the eigenvector (the angle of its mean resultant vector) +/- 1 bootstrap s.e. In neither case was the peak significantly different from one of the discriminandum orientations.

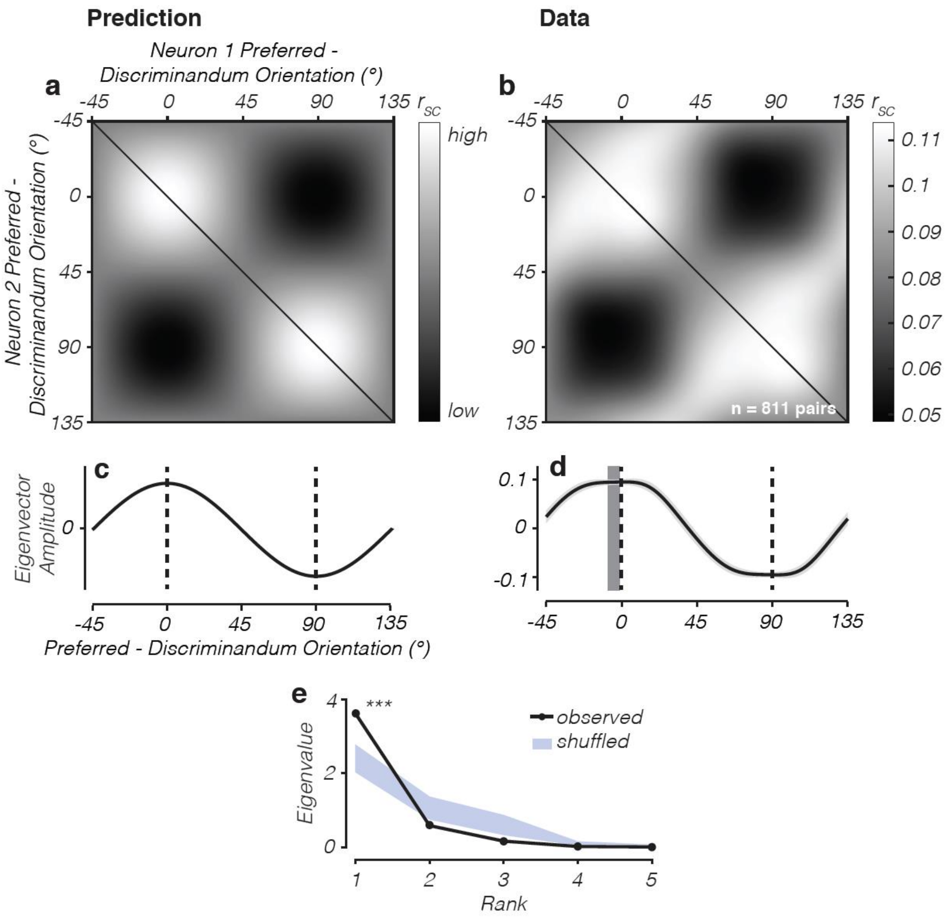

Figure 4. Task-Aligned Summary *R_{sc}* Matrix. a. To illustrate the task-dependent pattern of r_{sc} structure, we combined data from all recorded sessions and generated an r_{sc} matrix in which each pair's preferred orientations were expressed relative to the discriminandum orientations on the session they were recorded. In this task-aligned coordinate frame, 0º and 90º always index the discriminandum orientations. (a) shows the prediction for the form of this task-aligned r_{sc} matrix based on the hypothesis of task-related feedback (data identical to Fig. 2c). (b) shows the observed task-aligned r_{sc} matrix, which included all recorded pairs, and was smoothed identically to the data in Fig. 3. (c) and (d). The rank-1 eigenvectors of the predicted and observed r_{sc} matrices in (a) and (b) are closely similar. Shaded regions in (d) indicate +/- 1 bootstrap s.e. as in Fig. 3. The peak in the eigenvector in (d) was not significantly different from 0, indicating good alignment between the dynamic pattern of r_{sc} in V1 and the subject's task. (e) Eigenspectrum for the observed matrix in (b). Most of the variance in the matrix was explained by its rank-1 eigenvector (shown in d), significantly more than would be predicted by chance ($p<0.001$,

resampling test). Chance level was determined by randomly permuting the preferred orientations of the neurons. This demonstrates that the r_{sc} structure in V1 can be largely explained by a single mode of covariability that is determined by the task context, confirming the prediction based on task-related feedback.

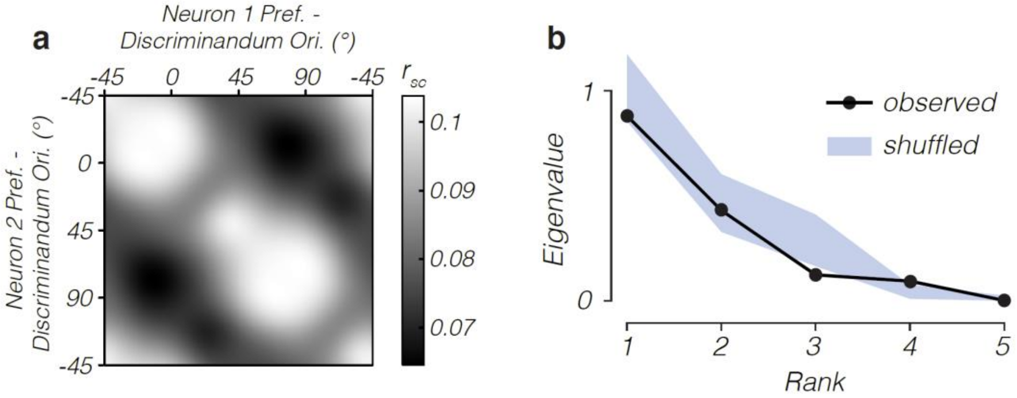

Figure 5. Task-dependent r_{sc} structure absent during passive fixation. a. Lattice pattern was absent in the task-aligned r_{sc} matrix during separate blocks of trials used to measure neuronal orientation tuning, during which the animal fixated passively for reward (matrix obtained as in Fig. 4b). Instead, the pattern in the matrix more closely resembled a fixed diagonal ridge, demonstrating that the task-dependent changes in r_{sc} structure observed during task performance depend on active task engagement. R_{sc} values were calculated only using interleaved presentations of the 0%-signal orientation-filtered noise stimulus, to facilitate comparison with the r_{sc} matrix observed during task performance. Data from only 556 pairs are shown, as not all recording sessions included fixation blocks with the 0%-signal filtered noise stimuli interleaved. **b**. The first eigenvalue of the r_{sc} matrix in (a) was not significantly above chance, unlike during task performance (Fig. 4) and inconsistent with the prediction based on task-dependent feedback. (Note that the eigenvalues are smaller than in Fig. 4e because the variance within the matrix in (a) is smaller.)

501

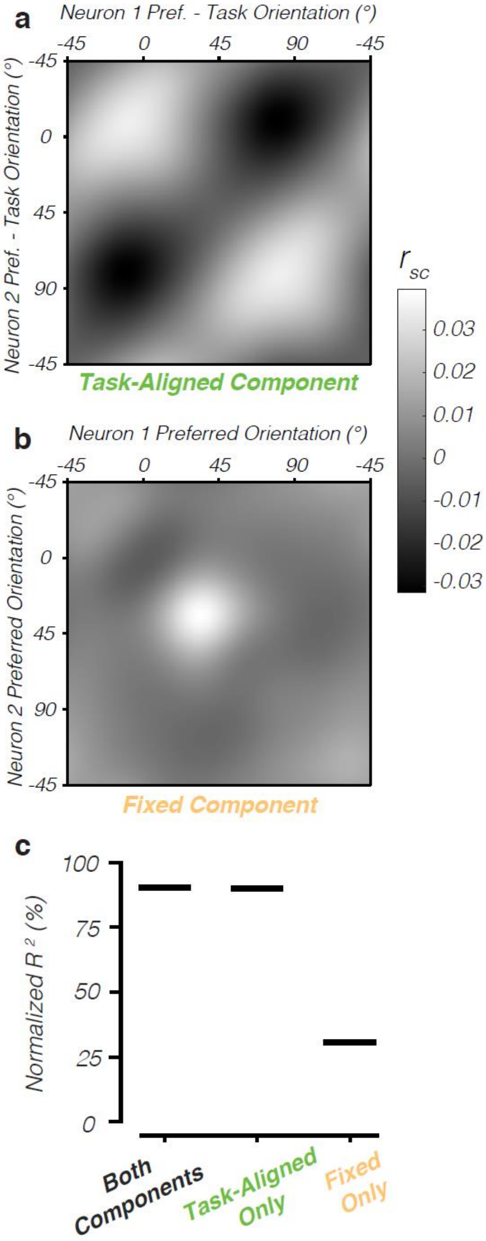

Figure 6. Segregating "fixed" and "task-aligned" components of r_{sc} structure. a-b. The two components of observed r_{sc} structure, jointly estimated using multilinear regression. The amplitude of the "task-aligned" component is larger than the "fixed" component, showing the majority of r_{sc} structure changes with the subject's task. Note that preferences for the task-aligned component, but not the fixed component, are expressed relative to the discriminandum orientations. Mean values are close to 0 due to the inclusion of a model constant. **c**. Goodness-of-fit for the joint model and two reduced models that included only one of the two components. Values are expressed relative to an estimate of the explainable variance in the data (see Methods). Removing the "task-aligned" component (but not the "fixed" component) dramatically reduced goodness-of-fit.

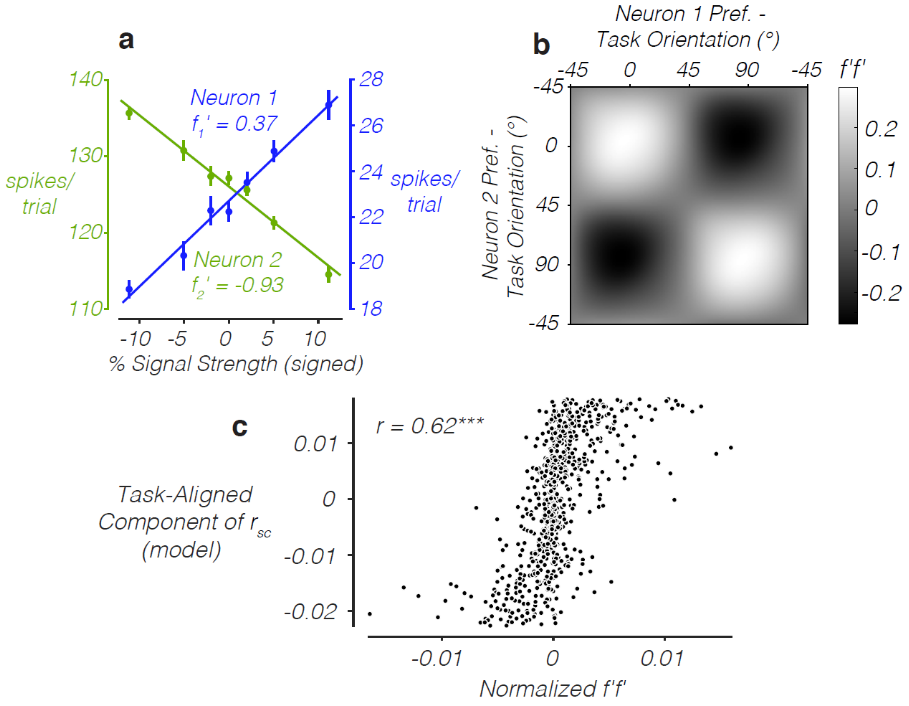

Figure 7. Task-Related Feedback Introduces Differential Correlations. **a**. Responses (mean +/- 1 s.e.) to the stimuli used in the task at various signal strengths for two example neurons. Calling these response functions f_1 and f_2, the differential correlation for this pair is proportional to the product of the derivatives $(f_1'f_2')$[32]. This product can be viewed as a metric of similarity in tuning for the task. Therefore, differential correlations are those that resemble the effect of changes in the stimulus along the axis defining the task. **b**. The matrix of $f'f'$ values, as a function of task-aligned pairwise orientation preference, obtained using kernel smoothing as in Fig. 4b. The lattice pattern is extremely similar to the structure of task-dependent r_{sc} we observed during task performance (Fig. 4b), suggesting task-related feedback introduces a source of differential correlation to the V1 population. **c**. Scatter plot of the task-aligned (putatively top-down) component of r_{sc} (Fig. 6a) against $f'f'$ values (normalized; see Methods) for each recorded neuronal pair. The two were highly correlated across the population ($r=0.62$, $p<10^{-5}$).

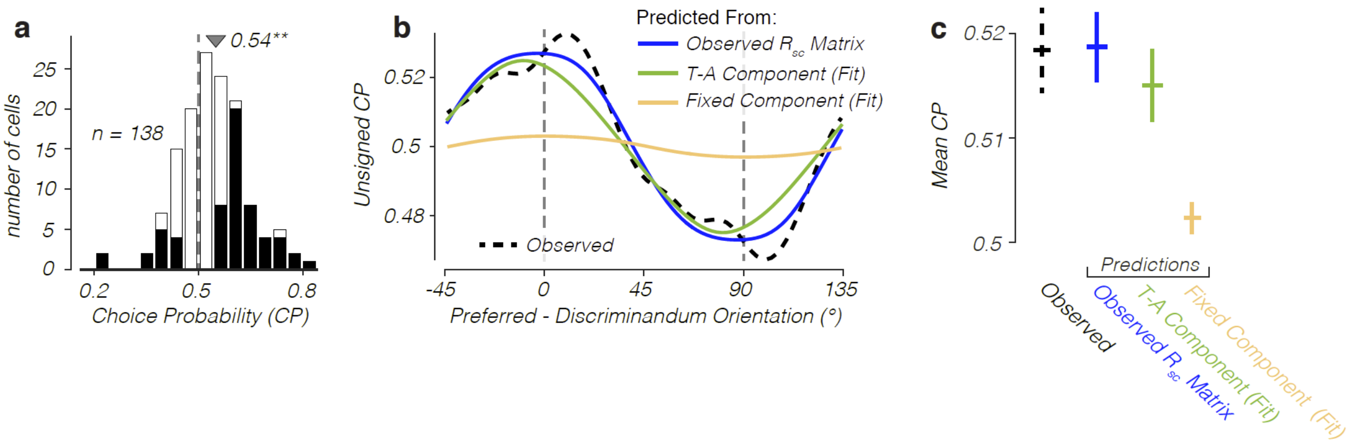

Figure 8. The task-dependent component of r_{sc} structure accounts for choice-related activity. a. Histogram of observed CPs, from neurons significantly preferring one of the two discriminandum orientations (d′>0.9 at highest signal level). The mean of 0.54 was significantly above chance (bootstrap test, cell resampling, p<0.01). CPs that were individually significant (p<0.05; bootstrap test, trial resampling) are shown in black. **b**. Observed (dashed) and predicted (solid) CP profiles, signed arbitrarily with respect to the 0° choice. (CP>0.5 indicates greater firing on 0°-choice trials, and CP<0.5 indicates greater firing on 90°-choice trials). The observed profile was obtained using data from all recorded neurons and smoothed with a von Mises kernel that approximated a wrapped Gaussian with 10° s.d. The predicted profiles were generated by applying a linear readout to V1 assuming different r_{sc} structures, as labeled (see Methods). The profiles shown are averages of a large set generated from different assumed readout weight profiles (see Supplementary Fig. 10), which introduced relatively little variability. The prediction using the observed r_{sc} matrix (Fig. 4b) closely match the observed CP profile. This could be replicated with predictions using only the task-aligned component (Fig. 6a), but not the fixed component (Fig. 6b), demonstrating that the feedforward effect of fixed sources of r_{sc} structure was insufficient to explain the observed magnitude of CP. **c**. Mean (rectified) CPs associated with the profiles in (b), +/- 1 bootstrap s.e. obtained by resampling from the data (gray bars). Note that the mean observed CP is lower here than in (a) because all neurons are included, regardless of their orientation preference.

Methods

Electrophysiology

We recorded extracellular spiking activity from populations of V1 neurons in two awake, head-fixed rhesus monkeys (*Macaca mulatta*). Both monkeys were implanted with a head post and scleral search coils under general anaesthesia[42]. In monkey '*lem*', a recording chamber was implanted over a craniotomy above the right occipital operculum, as described previously[43], by which we introduced linear microelectrode arrays (U- and V-probes, Plexon; 24-contacts, 50 or 60 µm spacing) at an angle approximately perpendicular to the cortical surface with a custom micro-drive. We positioned the linear arrays so that we roughly spanned the cortical sheet, as confirmed with current-source density analysis, and removed them after each recording session. In monkey '*jbe*', a planar "Utah" array (Blackrock Microsystems; 96 electrodes 1mm in length inserted to target supragranular layers, 400 um spacing) was chronically implanted, also over the right occipital operculum. All procedures were performed in accordance with the U.S. Public Health Service Policy on the humane care and use of laboratory animals and all protocols were approved by the National Eye Institute Animal Care and Use Committee.

Broadband signals were digitized at 30 or 40 kHz and stored to disk. Spike sorting was performed offline using custom software in MATLAB®. First, spikes were detected using a voltage threshold applied to high-pass filtered signals. Next, triggered waveforms were projected into spaces defined either by principal components or similarity to a template. Clusters boundaries were finally estimated with a Gaussian mixture model, and then rigorously verified and adjusted by hand when needed. In the linear array recordings, spike sorting yield and quality was substantially improved by treating sets of three or four neighboring contacts as "n-trodes". As this was not possible with the Utah array due to the greater interelectrode spacing, we excluded pairs of neurons recorded on the same electrode to avoid contamination by misclassification. Neurons from separate recording sessions were treated as independent. To reduce the possibility that a single neuron from the Utah array contributed to two datasets, we included only sessions that were separated by at least 48 hours (with a median separation of

5 days). We excluded from analysis those neurons whose mean evoked firing rate did not exceed 7 spikes/second.

Visual Stimuli

All stimuli were presented binocularly on two gamma-corrected cathode ray tube (CRT) monitors viewed through a mirror haploscope, at 85 or 100Hz. The monitors subtended 24.1° x 19.3° of visual angle (1280 x 1024 pixels). The stimuli presented during performance of the discrimination task consisted of bandpass filtered dynamic white noise, as described previously[18]. Briefly, stimuli were filtered in the Fourier domain with a polar-separable Gaussian. The peak spatial frequency was optimized for the recorded neuronal population (1 and 4 cpd medians for *'lem'* and '*jbe*', respectively) while the peak orientation could take one of two orthogonal values the animal had to discriminate in a given session. The angular s.d. of the filter modulated the orientation bandwidth and was varied trial to trial. A 2D Gaussian contrast envelope was applied to the stimulus so that its spatial extent was as small as possible while still covering the minimum response fields of the neuronal populations being recorded. The median envelope s.d. was 0.6 degrees for both animals. The median stimulus eccentricity was 5.4 degrees for *'lem'* and 0.5 degrees for '*jbe*'. In Fig. 1, we quantify orientation bandwidth as % signal strength. This was calculated as $100 * R$, where R is the length of the resultant vector associated with the angular component of the stimulus filter.

We estimated neuronal orientation preferences in separate blocks of trials, using 420-ms presentations of the following types of stimuli, presented at a range of orientations: 1) an orientation narrowband version of the stimulus described above (10° angular s.d.); 2) sinusoidal gratings; and 3) circular patches of dynamic 1D noise patterns (random lines). The preferred orientation of a neuron was calculated as the circular mean of its orientation tuning curve. For each neuron, from among the set of tuning curves elicited by the different stimulus types described above, we chose as the final estimate of preferred orientation the one with the smallest standard error, obtained by resampling trials. We excluded from further analysis all neurons where this exceeded 5°. On a subset of sessions, we also used these orientation-tuning blocks to present examples of the 0%-signal orientation-filtered noise stimuli. These were presented at the same location and size as during task performance, allowing us to calculate r_{sc} structure in the absence of task engagement but with identical retinal input.

Orientation Discrimination Task

The animals performed a coarse orientation discrimination task using the orientation-filtered noise stimuli, as described previously[18]. To initiate a trial, the subject had to acquire a central fixation square. After a delay of 50 ms, the stimulus appeared for a fixed duration of 2 seconds. The trial was aborted if the subject broke fixation at any point during the stimulus presentation, defined as either 1) making a microsaccade covering a distance greater than a predefined threshold (typically 0.5°) or 2) a deviation in mean eye position from the center of the fixation point of more than a predefined threshold, typically 0.7°. At the end of the stimulus presentation, two choice targets appeared. These were Gabor patches of 2-3° in spatial extent, oriented at each of the two discriminandum orientations. The locations of the choice targets depended on the task. For discriminandum pairs near horizontal and vertical (-22.5° – +22.5° and 67.5° – 112.5°), the choice targets were positioned along the vertical meridian, at an eccentricity of about 3°, with the more vertically-oriented target appearing always in the upper hemifield. For orientation pairs near the obliques (22.5° – 67.5° and 112.5° – 157.5°), the choice targets were positioned along the horizontal meridian, at the same range of eccentricities, with the smaller of the two orientations always appearing in the left hemifield. (We use the convention that horizontal is 0° and that orientation increases with clockwise rotation.) To penalize random guessing, the volume of liquid reward delivered after correct choices was doubled with each consecutive correct choice, up to a maximum of four times the initial amount. Since we were primarily interested in the effect of task engagement on neuronal activity, we applied a behavioral criterion to our data, excluding sessions where the subject's psychophysical threshold (defined as the signal level eliciting 75% correct performance) exceeded 14% signal. A two-pass presentation procedure was used. Each instance of a stimulus (generated with a given noise seed) was shown twice per experimental block. This allowed us to account for any effect of fluctuations in the stimulus on r_{sc} (see Supplementary Discussion §1.1 and Supplementary Fig. 4).

Psychophysical Reverse Correlation (PRC)

We performed PRC to objectively measure the weights subjects applied to different stimulus orientations when making their choices. To do this, we first summarized the stimulus on each trial as the radial sum of its 2D Fourier amplitude spectrum, averaged across frames. This isolates the orientation content while removing any information about spatial frequency and phase. Psychophysical kernels were calculated as the difference between the two choice-conditioned radial sums. This was performed separately for each signal level and the resulting kernels were averaged.

Spike-Count Correlation Measurements

Spike-count correlations were calculated as the Pearson correlation between spike counts, counted over the entire duration of the stimulus, with a 50-ms delay to account for the typical V1 response latency. Spike counts were first z-scored separately within each experimental block (typically a set of 100-200 trials lasting about 10 minutes) and each stimulus condition. This removed correlations related to long-term firing rate nonstationarities and allowed us to combine trials at different signal levels without introducing correlations related to similarity in stimulus preference. We used a balanced z-scoring method proposed recently to prevent bias related to differences in choice distributions across signal levels[44]. We excluded pairs that were not simultaneously isolated for at least 25 trials total. The median number of trials per pair during task performance was 752. For the analysis of r_{sc} during passive fixation blocks (Fig. 5), the median was 120.

A main goal of the study was to measure how spike-count correlation varies with pairwise orientation. We describe this relationship as a smoothed function estimated from measures of r_{sc} combined across multiple recording sessions, which we then sampled discretely with 1° resolution. The smoothed estimates were obtained using a bivariate von Mises smoothing kernel. A point in the correlation matrix **C** was given as:

$$\mathbf{C}(x,y)=\tanh\frac{\sum_{i=1}^{n} z_i K(x,y,\theta_i,\phi_i)}{\sum_{i=1}^{n} K(x,y,\theta_i,\phi_i)}\text{, where } K(x,y,\theta_i,\phi_i)=e^{\kappa\,(\cos(\theta_i-x)+\cos(\phi_i-y))}, \tag{1}$$

z_i is the i^{th} (Fisher z-transformed) r_{sc} measurement, θ_i and ϕ_i are the preferred orientations of the i^{th} pair, and κ is the von Mises dispersion parameter. We set $\kappa = 1.3\pi$, yielding a smoothing kernel closely approximating a bivariate wrapped Gaussian with 15° s.d. In some cases, we expressed the r_{sc} matrix in a task-aligned coordinate frame (e.g. Fig. 4e), for which the preferred orientations of the i^{th} pair relative to the task orientation were used for θ_i and ϕ_i. Since there were always two orthogonal task orientations, we averaged across both possible alignments, such that $\mathbf{C}(x,y)=\mathbf{C}(x+90°,y+90°)$. All angular quantities were doubled for the calculations, as orientation has a period of 180°. To generate the kernel-smoothed profile of CP (Fig. X), we used a one-dimensional equivalent of the procedure above, in which preferred orientations were parameterized only by a single parameter.

Regression Model

We used a multilinear regression model to identify “fixed” and “task-aligned” components of the structured correlations we observed. Our approach was to describe the set of observations (811 individual pairwise r_{sc} measurements, Fisher z-transformed to produce normal error) in terms of a set of two underlying correlation structures: one defining r_{sc} as a function of pairwise preferred orientation alone (“fixed”) and the other defining r_{sc} as a function of pairwise preferred orientation relative to the task (“task-aligned”). In order to provide a continuous and smooth description of the data, each component was parameterized as the sum of an array of $n\ x\ n$ evenly spaced basis functions. Each observation, y_i, was expressed as:

$$y_i = x_i^{fixed}\cdot\beta^{fixed} + x_i^{task}\cdot\beta^{task} + \beta_0 + \varepsilon_i \tag{2}$$

x_i^{fixed} and x_i^{task} are length-n^2 vectors of loadings onto the basis functions, which were given by evaluating the basis functions at the location corresponding to the pairwise orientation preference of the *i*th pair. β^{fixed} and β^{task} are the length-n^2 vectors of amplitudes of the basis functions (coefficients to be fit), β_0 is a model constant, and $\cdot$ is the element-wise product. For the basis functions, we used bivariate von Mises functions, with no correlation and equal dispersion in both dimensions. Thus the *k*th loading ($x_i^{fixed}(k)$ or $x_i^{task}(k)$) was given by:

$$x_i(k) = \frac{e^{\kappa\left(\cos\left(\theta_i - \mu_k^1\right) + \cos\left(\phi_i - \mu_k^2\right)\right)}}{Z} \qquad (3)$$

where θ_i and ϕ_i are the preferred orientations of the *i*th pair (relative to the task orientation in the case of the task-aligned loadings), μ_k is a pair of orientations defining the location of the *k*th basis function, *Z* is a normalization constant such that the sum of all loadings for observation *i* ($x_i^{fixed} + x_i^{task}$) is 1, and κ is the von Mises dispersion parameter. Again, angular quantities were doubled and κ was set to 1.3π. We found that arrays of 8x8 were sufficient to describe the structure of the two components. Because the observations were pairwise correlations, it was sufficient only to fit the upper triangular portion of the array of basis functions. Thus, the two-component model contained 73 parameters (36 for each component, plus the model constant).

We fit the model by finding the parameters (β^{fixed}, β^{task} & β_0) that minimized the L1 error (to encourage sparseness) plus two additional terms that encouraged smoothness and symmetric positive semi-definiteness, as the two components were meant to represent correlation matrices. The solution was obtained as:

$$\hat{\beta}^{fixed}, \hat{\beta}^{task}, \hat{\beta}_0 = \underset{\beta^{fixed}, \beta^{task}, \beta_0}{\operatorname{argmin}} \sum_i |\varepsilon_i| + \alpha_1 \Gamma\left(\beta^{fixed} + \beta^{task}\right) + \alpha_2 D_{SPD}\left(\beta^{fixed} + \beta^{task}\right) \qquad (4)$$

where Γ is the discrete 2D Laplace operator, corresponding to circular convolution with the kernel:

$\begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix}$ and $D_{SPD}(X)$ is the 2-norm between X and the nearest symmetric positive semidefinite matrix $\hat{X}$, which is given by $(B+H)/2$ where H is the symmetric polar factor of $B = \frac{(A+A')}{2}$ [45]. The α's controlled the strength of regularization and were chosen to produce the best fit (as measured with R^2 under 50-fold cross-validation). The solution was obtained by gradient descent using the MATLAB function *fminunc*.

While this model did not explain more than a small percentage (3.2%) of the variance of the raw observed r_{sc} values, this is not surprising as the raw correlation data do not vary smoothly with preferred orientation (reflecting both noise, and the fact that r_{sc} is known to depend on parameters other than orientation.[1,14,15]). For this reason, we measured goodness-of-fit relative to the variance explained simply by a smoothed version of the raw data (sum of values in fixed and task-aligned matrices was 3.6%).

Choice Probability Predictions

Choice Probability was calculated in the standard way[33]. We only used 0%-signal trials, as the uneven choice distributions elicited by signal trials yield noisier CP measurements. Assuming feedforward pooling with linear pooling weights, the relationship between the covariance matrix for a population of neurons, the pooling weight of each neuron, and the Choice Probability (*CP*) of each neuron is:

$$CP_k = \frac{1}{2} + \frac{2}{\pi}\operatorname{sgn}(\xi_k)\arctan\sqrt{2{\xi_k}^{-2} - 1}^{\,-1} \quad \text{with} \quad \xi_k = \frac{(\mathbf{C}\beta)_k}{\sqrt{C_{kk}\beta^{\mathrm{T}}\mathbf{C}\beta}} \tag{5}$$

where CP_k is the CP of neuron k with respect to choice 1, β is the vector of pooling weights and **C** is the covariance matrix[8]. We used this equation to quantify the CPs that would be associated with the correlation structure we observed and the fixed and task-aligned components we identified. CPs, correlations, and pooling weights were described as functions of preferred orientation, relative to the task. (For the fixed component of r_{sc}, which was indexed relative to raw orientation preferences, we generated a task-aligned version by substituting the observed r_{sc} values with model fits, using only the fixed component of the model, and then generating a smoothed task-aligned matrix as in Fig. 4e). We assumed a population of infinite size spanning this space, that was homogeneous at a given orientation, and that we sampled discretely with 1° resolution. Since the pooling weights were unknown, we generated a random distribution of plausible pooling weights that could support task performance. To do this, we started with a vector of randomly assigned weights (drawn from a normal distribution) and applied the 90° symmetry inherent in the task, such that $\beta_\theta = -\beta_{\theta+90}$, where β_θ is the weight assigned to neurons with task-relative preferred orientation θ. Then, we smoothed the pooling weight profiles with a wrapped Gaussian kernel with 15° s.d. and excluded profiles which did not have a circular mean within 22.5° of choice 1 (0°). To guarantee real-valued CPs on [0,1], we performed the calculations using a symmetric positive definite approximation[45] of the correlation matrices, which introduced negligible error.

Estimating mean covariance for a population of neurons is necessarily more error-prone than estimating mean correlation, as the former is sensitive to sampling error in measurements of average spike-count variance (and therefore firing rate), so for this reason we preferred to perform the calculations using correlations (see Supplemental Discussion §3). We can use correlations interchangeably with covariances in equation 1, under the simplifying assumption that the variance is uniform as a function of preferred orientation. If $\mathbf{\Sigma}$ is the correlation matrix for a population with uniform variance α, then it follows that:

$$\xi_k = \frac{a(\mathbf{\Sigma}\beta)_k}{\sqrt{a\Sigma_{kk}\beta^{\mathrm{T}}(a\mathbf{\Sigma})\beta}} = \frac{(\mathbf{\Sigma}\beta)_k}{\sqrt{\Sigma_{kk}\beta^{\mathrm{T}}\mathbf{\Sigma}\beta}} \tag{6}$$

where $\Sigma_{kk} \equiv 1$ for all k. We felt that spike-count variance that depended systematically on preferred orientation was unlikely to be a feature of the V1 representation, and thus that the advantages of using correlations outweighed the cost.

Noise in the decision process after pooling (pooling noise) has the effect of uniformly scaling down CPs, such that ξ_k in Eq. 5 is substituted with: $\frac{(\mathbf{C}\beta)_k}{\sqrt{C_{kk}(\beta^{\mathrm{T}}\mathbf{C}\beta+\sigma_{pool}^2)}}$, where σ_{pool}^2 is the variance of the pooling noise[6]. We found that non-zero pooling noise was needed to avoid overestimating the magnitude of CP from the observed correlation structure. We used a fixed value of pooling noise in our predictions such that the average squared difference between the CP profile predicted from the observed correlation matrix and the observed CP profile was minimized. Empirically, we found that pooling noise variance of 0.6 was optimal. Since our spike counts were normalized to have unit variance, this implies pooling noise whose variance is 60% of the average spike-count variance of single neurons. This should be interpreted with care, as overestimation of CPs may also be an artefact related to the assumption of a homogeneous population[8]. Alternatively, the need to invoke pooling noise may be due to nonuniform sensory integration across the trial, which is distinct but which would also have an attenuating effect on CP when measured over the entire trial.

Calculating Differential Correlations

The information capacity of a sensory population, assuming a linear read out, is bounded when the spike-count covariances sufficiently match the differential correlations[32]. Since we made use of spike-count correlations, rather than covariances, in the present study, we normalized the measurements of differential correlations by the product of the standard deviations of the stimulus-independent variability of each pair.

1. Cohen, M. R. & Kohn, A. Measuring and interpreting neuronal correlations. *Nat. Neurosci.* **14,** 811–9 (2011).

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

3. Abbott, L. F. & Dayan, P. The effect of correlated variability on the accuracy of a population code. *Neural Comput.* **11,** 91–101 (1999).

4. Sompolinsky, H., Yoon, H., Kang, K. & Shamir, M. Population coding in neuronal systems with correlated noise. *Phys. Rev. E* **64,** 51904 (2001).

5. Snippe, H. & Koenderink, J. Information in channel-coded systems: correlated receivers. *Biol. Cybern.* **190,** 183–190 (1992).

6. Averbeck, B. B., Latham, P. E. & Pouget, A. Neural correlations, population coding and computation. *Nat. Rev. Neurosci.* **7,** 358–66 (2006).

7. Shadlen, M. N., Britten, K. H., Newsome, W. T. & Movshon, J. A. A computational analysis of the relationship between neuronal and behavioral responses to visual motion. *J. Neurosci.* **16,** 1486–510 (1996).

8. Haefner, R. M., Gerwinn, S., Macke, J. H. & Bethge, M. Inferring decoding strategies from choice probabilities in the presence of correlated variability. *Nat. Neurosci.* **16,** 235–42 (2013).

9. Nienborg, H. & Cumming, B. G. 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 (2010).

10. Bair, W., Zohary, E. & Newsome, W. T. Correlated firing in macaque visual area MT: time scales and relationship to behavior. *J. Neurosci.* **21,** 1676–97 (2001).

11. Mastronarde, D. N. Correlated firing of cat retinal ganglion cells. I. Spontaneously active inputs to X- and Y-cells. *J. Neurophysiol.* **49,** 303–24 (1983).

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

13. Lee, D., Port, N. L., Kruse, W. & Georgopoulos, a P. Variability and correlated noise in the discharge of neurons in motor and parietal areas of the primate cortex. *J. Neurosci.* **18,** 1161–70 (1998).

14. Smith, M. A. & Kohn, A. Spatial and temporal scales of neuronal correlation in primary visual cortex. *J. Neurosci.* **28,** 12591–603 (2008).

15. Kohn, A. & Smith, M. a. Stimulus dependence of neuronal correlation in primary visual cortex of the macaque. *J. Neurosci.* **25,** 3661–73 (2005).

16. Callaway, E. M. Feedforward, feedback and inhibitory connections in primate visual cortex. *Neural Netw.* **17,** 625–32 (2004).

17. Sillito, A. M., Cudeiro, J. & Jones, H. E. Always returning: feedback and sensory processing in visual cortex and thalamus. *Trends Neurosci.* **29,** 307–16 (2006).

18. Nienborg, H. & Cumming, B. G. Decision-Related Activity in Sensory Neurons May Depend on the Columnar Architecture of Cerebral Cortex. *J. Neurosci.* **34,** 3579–3585 (2014).

19. Nienborg, H. & Cumming, B. G. Psychophysically measured task strategy for disparity

discrimination is reflected in V2 neurons. *Nat. Neurosci.* **10,** 1608–1614 (2007).

20. Ahumada Jr, A. J. Perceptual classification images from Vernier acuity masked by noise. in *Perception ECVP abstract* **25,** 0 (Pion Ltd, 1996).

21. Newsome, W. T., Britten, K. H., Movshon, J. A. & Shadlen, M. N. in *Neural Mechanism of Visual Perception* 171–197 (1989).

22. Uka, T. & DeAngelis, G. C. Contribution of Area MT to Stereoscopic Depth Perception. *Neuron* **42,** 297–310 (2004).

23. Cook, E. P. & Maunsell, J. H. R. Dynamics of neuronal responses in macaque MT and VIP during motion detection. *Nat. Neurosci.* **5,** 985–94 (2002).

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

25. Gu, Y., Angelaki, D. E. & DeAngelis, G. C. Neural correlates of multisensory cue integration in macaque MSTd. *Nat. Neurosci.* **11,** 1201–1210 (2008).

26. Price, N. S. C. & Born, R. T. Timescales of sensory- and decision-related activity in the middle temporal and medial superior temporal areas. *J. Neurosci.* **30,** 14036–14045 (2010).

27. Nienborg, H. & Cumming, B. G. Decision-related activity in sensory neurons reflects more than a neuron's causal effect. *Nature* **459,** 89–92 (2009).

28. Wimmer, K. *et al.* Sensory integration dynamics in a hierarchical network explains choice probabilities in cortical area MT. *Nat. Commun.* **6,** 6177 (2015).

29. Haefner, R. M., Berkes, P. & Fiser, J. Perceptual Decision-Making as Probabilistic Inference by Neural Sampling. *Neuron* **90,** 649–660 (2016).

30. Ecker, A. S., Denfield, G. H., Bethge, M. & Tolias, A. S. On the Structure of Neuronal Population Activity under Fluctuations in Attentional State. *J. Neurosci.* **36,** 1775–1789 (2016).

31. Shamir, M. & Sompolinsky, H. Implications of neuronal diversity on population coding. *Neural Comput.* **18,** 1951–1986 (2006).

32. Moreno-Bote, R. *et al.* Information-limiting correlations. *Nat. Neurosci.* **17,** 1410–1417 (2014).

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

34. Crapse, T. B. & Basso, M. A. Insights into decision making using choice probability. *J. Neurophysiol.* **114,** 3039–3049 (2015).

35. Nienborg, H., Cohen, M. R. & Cumming, B. G. Decision-Related Activity in Sensory Neurons: Correlations Among Neurons and with Behavior. *Annu. Rev. Neurosci.* **35,** 463–483 (2012).

36. Kwon, S. E., Yang, H., Minamisawa, G. & O'Connor, D. H. Sensory and decision-related activity propagate in a cortical feedback loop during touch perception. *Nat. Neurosci.* **19,** (2016).

37. Mitchell, J. F., Sundberg, K. a & Reynolds, J. H. Spatial attention decorrelates intrinsic activity fluctuations in macaque area V4. *Neuron* **63,** 879–88 (2009).

38. Cohen, M. R. & Maunsell, J. H. R. Attention improves performance primarily by reducing interneuronal correlations. *Nat. Neurosci.* **12,** 1594–600 (2009).

39. Ruff, D. A. & Cohen, M. R. Attention can either increase or decrease spike count correlations in visual cortex. *Nat. Neurosci.* **17,** 1591–7 (2014).

40. Rabinowitz, N. C., Goris, R. L., Cohen, M. & Simoncelli, E. Attention stabilizes the shared gain of V4 populations. *Elife* **4,** e08998 (2015).

41. Lange, R. D. & Haefner, R. M. Inferring the brain’s internal model from sensory responses in a probabilistic inference framework. (2016). doi:10.1101/081661

42. Judge, S. J., Richmond, B. J. & Chu, F. C. Implantation of magnetic search coils for measurement of eye position: An improved method. *Vision Res.* **20,** 535–538 (1980).

43. Cumming, B. G. & Parker, A. J. Binocular neurons in V1 of awake monkeys are selective for absolute, not relative, disparity. *J. Neurosci.* **19,** 5602–18 (1999).

44. Kang, I. & Maunsell, J. H. R. Potential confounds in estimating trial-to-trial correlations between neuronal response and behavior using choice probabilities. *J. Neurophysiol.* **108,** 3403–15 (2012).

45. Higham, N. J. Computing a nearest symmetric positive semidefinite matrix. *Linear Algebra Appl.* **103,** 103–118 (1988).

Supplemental Information available online.

Acknowledgements We thank Michael Shadlen and James McFarland for useful discussions; Ralf Haefner for advice about the analysis; Beth Nagy, Irina Bunea, and Denise Parker for veterinary care; and Richard Krauzlis, Bevil Conway, and Ali Ghazizadeh for comments on an earlier version of the manuscript.

Author Contribution A.G.B. and B.G.C. conceived and designed the experiments. A.G.B. performed the experiments and all aspects of the analysis. A.G.B. and B.G.C. wrote the paper. B.G.C. advised at all stages.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to A.G.B. (adrian.bondy@gmail.com).