

Origin of information-limiting noise correlations

Ingmar Kanitscheider^{a,b,1,2}, Ruben Coen-Cagli^{a,1}, and Alexandre Pouget^{a,c,d}

^aDepartment of Basic Neuroscience, University of Geneva, 1211 Geneva, Switzerland; ^bCenter of Learning and Memory and Department of Neuroscience, The University of Texas at Austin, Austin, TX 78712; ^cDepartment of Brain and Cognitive Sciences, University of Rochester, Rochester, NY 14627; and ^dGatsby Computational Neuroscience Unit, University College London, London W1T 4JG, United Kingdom

Edited by Terrence J. Sejnowski, Salk Institute for Biological Studies, La Jolla, CA, and approved September 28, 2015 (received for review May 4, 2015)

The ability to discriminate between similar sensory stimuli relies on the amount of information encoded in sensory neuronal populations. Such information can be substantially reduced by correlated trial-to-trial variability. Noise correlations have been measured across a wide range of areas in the brain, but their origin is still far from clear. Here we show analytically and with simulations that optimal computation on inputs with limited information creates patterns of noise correlations that account for a broad range of experimental observations while at same time causing information to saturate in large neural populations. With the example of a network of V1 neurons extracting orientation from a noisy image, we illustrate to our knowledge the first generative model of noise correlations that is consistent both with neurophysiology and with behavioral thresholds, without invoking suboptimal encoding or decoding or internal sources of variability such as stochastic network dynamics or cortical state fluctuations. We further show that when information is limited at the input, both suboptimal connectivity and internal fluctuations could similarly reduce the asymptotic information, but they have qualitatively different effects on correlations leading to specific experimental predictions. Our study indicates that noise at the sensory periphery could have a major effect on cortical representations in widely studied discrimination tasks. It also provides an analytical framework to understand the functional relevance of different sources of experimentally measured correlations.

noise correlations | information theory | neural computation | efficient coding | neuronal variability

The response of cortical neurons to an identical stimulus varies from trial to trial. Moreover, this variability tends to be correlated among pairs of nearby neurons. These correlations, known as noise correlations, have been the subject of numerous experimental as well as theoretical studies because they can have a profound impact on behavioral performance (1–7). Indeed, behavioral performance in discrimination tasks is inversely proportional to the Fisher information available in the neural responses, which itself is strongly dependent on the pattern of correlations. In particular, correlations can strongly limit information in the sense that some patterns of correlations can lead information to saturate to a finite value in large populations, in sharp contrast to the case of independent neurons for which information grows proportionally to the number of neurons. However, the saturation is observed for only one type of correlations known as differential correlations. If the correlation pattern slightly deviates from differential correlations, information typically scales with the number of neurons, just like it does for independent neurons (7). These previous results clarify how correlations impact information and consequently behavioral performance but fail to address another fundamental question, namely, Where do noise correlations, and in particular information-limiting differential correlation, come from? Understanding the origin of information-limiting correlation is a key step toward understanding how neural circuits can increase information transfer, thereby improving behavioral performance, via either perceptual learning or attentional selection.

Several groups have started to investigate sources of noise correlations such as shared connectivity (2), feedback signals (8), internal dynamics (9–11), or global fluctuations in the excitability of

cortical circuits (12–16). Global fluctuations have received a lot of attention recently as they appear to account for a large fraction of the measured correlations in the primary visual cortex. Correlations induced by global fluctuations, however, do not limit information in most discrimination tasks (with the possible exception of contrast discrimination for visual stimuli). Therefore, if cortex indeed operates at information saturation, the source of information-limiting correlations is still very much unclear.

In this paper, we focus on correlations induced by feedforward processing of stimuli whose information content is small compared with the information capacity of neural circuits. Using orientation selectivity as a case study, we find that feedforward processing induces correlations that share many properties of the correlations observed *in vivo*. Moreover, we also show feedforward processing leads to information-limiting correlations as a direct consequence of the data processing inequality. Interestingly, these information-limiting correlations represent only a small fraction of the overall correlations induced by feedforward processing, making them difficult to detect through direct measurements of correlations. Finally, we demonstrate that correlations induced by global fluctuations cannot limit information on their own, but can reduce the level at which information saturates in the presence of information-limiting correlations. Despite our focus on orientation selectivity, our results can be generalized to other modalities, stimuli, and brain areas.

In summary, this work identifies a major source of noise correlations and, importantly, a source of information-limiting noise correlations, while clarifying the interactions between information-limiting correlations and correlations induced by global fluctuations.

Significance

Populations of neurons encode information in activity patterns that vary across repeated presentation of the same input and are correlated across neurons (noise correlations). Such noise correlations can limit information about sensory stimuli and therefore limit behavioral performance in tasks such as discrimination between two similar stimuli. Therefore it is important to understand where and how noise correlations are generated. Most previous accounts focused on sources of variability inside the brain. Here we focus instead on noise that is injected at the sensory periphery and propagated to the cortex: We show that this simple framework accounts for many known properties of noise correlations and explains behavioral performance in discrimination tasks, without the need to assume further sources of information loss.

Author contributions: I.K., R.C.-C., and A.P. designed research; I.K. and R.C.-C. performed research; and I.K., R.C.-C., and A.P. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

¹I.K. and R.C.-C. contributed equally to this work.

²To whom correspondence should be addressed. Email: I.Kanitscheider@mail.clm.utexas.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1508738112/-DCSupplemental.

Results

Information-Limiting Noise Correlations Induced by Computation. We first characterize the noise correlations produced by a simple feedforward network and study whether they limit information in the population. We consider a population of model V1 neurons with spatially overlapping receptive fields extracting orientation from an image presented to the retina (Fig. 1A). We model V1 neurons as standard linear–nonlinear Poisson (LNP) units with Gabor receptive fields followed by a rectifying nonlinearity and Poisson spike generation. The image consists of an oriented Gabor corrupted by independent pixel noise in the retina (e.g., due to noisy photoreceptors). The noise is additive Gaussian white noise with SD equal to 20% of the range of retinal response levels (estimated from refs. 17 and 18), except where otherwise noted. For simplicity, our model lumps the retina and the lateral geniculate nucleus (LGN) into a single layer. Dividing up this layer in multiple stages would not affect our conclusions qualitatively, while significantly complicating the analytical approximations. As a measure of information, we use linear Fisher information, because it quantifies the performance (i.e., the inverse variance) of the optimal linear readout in fine discrimination tasks (19), which are commonly studied experimentally (20). To quantify the behavioral discrimination performance that could be achieved based on such a network, we assume throughout the text that all of the information encoded in the network responses will be

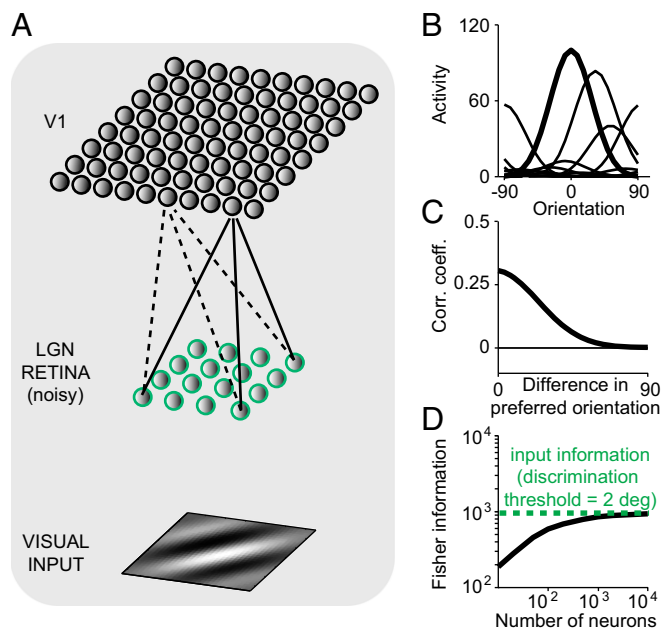


Fig. 1. A feedforward orientation network accounts for heterogeneous tuning, realistic correlations, and limited information. (A) Schematic of the orientation network. (A, Bottom) The visual input is a Gabor image. (A, Middle) Neurons in the sensory periphery add noise to the visual inputs. (A, Top) Due to cortical expansion, V1 neurons share part of their noisy inputs, introducing noise correlations and limiting information about any stimulus. Note that we model only one cortical hypercolumn: The receptive fields of all neurons in the population fully overlap in space. (B) Orientation tuning curves for a subset of neurons. (C) Average pairwise noise correlations in the network are positive and decay with the difference in preferred orientation. (D) Information as a function of the number of neurons in the V1 layer (solid black line). Information in V1 saturates for a large number of neurons. In this particular case, the network is optimal as indicated by the fact that information saturates at the upper limit of information imposed by the input information (dashed green line). The noise in the sensory periphery was set to yield a discrimination threshold of 2°. Simulation parameters are specified in *SI Appendix, Table S2*.

extracted and used to guide behavior. Equations for the network and information analysis are reported in *Methods*, and further derivations are reported in *SI Appendix*. Details of the simulation parameters used for each figure are provided in *SI Appendix, Tables S2–S5*.

For this simple network, it is possible to derive analytically the mean and covariance of the neural response in the V1 layer (*Methods*, Eq. 3 and *SI Appendix, section 1*). This analysis reveals that the V1 orientation tuning curves and noise correlation patterns are qualitatively similar to those observed experimentally (Fig. 1B and C). Noise correlations in our model arise because receptor noise propagating from the retina into V1 is shared between neurons according to their filter overlap. Because the filter of a neuron determines its tuning, it follows that neurons with similar tuning will be more correlated than neurons with dissimilar tuning. Thus, correlations that emerge from computation naturally reproduce the limited-range structure that has been observed experimentally in visual cortex (1, 21–25). Furthermore, pairwise correlations in the network decay with the distance between the centers of the filters and do so more rapidly for small than for large filters (*SI Appendix, Fig. S1*), as found in a recent comparison of V1 and V4 (25).

Pairwise correlations measured in V1 experimentally are on average small, ~ 0.1 for similarly tuned neurons (26). In contrast, the network of Fig. 1C exhibits much higher correlations. The reason for this discrepancy is that, in the feedforward model, the correlation coefficient is effectively a measure of the overlap between the filters representing two neurons. In Fig. 1C all filters are perfectly overlapping in space and spatial frequency and differ only in the preferred orientation and amplitude (parameter details in *SI Appendix, Table S2*). On the other hand, in typical recordings neurons display a heterogeneity of receptive field sizes, positions, and spatial frequency preferences (27). Fig. 2A shows that a network with heterogeneous filter properties drawn from the typical distributions found in V1 experiments (parameter details in *SI Appendix, Table S3*) indeed produces much lower correlation coefficients (mean correlation coefficient 0.09 for neurons with similar orientation preference and 0.03 across the entire population).

Our feedforward network can reproduce the pattern of correlations observed in vivo but it is not clear yet whether these correlations limit the information in the V1 layer of the network. A simple intuitive argument shows that information-limiting correlations must be present. Given the noise corrupting the visual input in our model, the information about orientation in the image is finite as long as the number of pixels is finite. Moreover, according to the data-processing inequality (28), any amount of posterior processing in the cortex can at most preserve the amount of information present in the input. Consequently, if we keep the size of the image constant, the information conveyed by the network as a function of the number of V1 units must eventually saturate to a value equal to or smaller than the information available in the retina.

This is indeed the case in our model. Using analytical expressions for tuning curves and covariance matrices (*Methods*, Eq. 3), we can calculate the linear Fisher information in the V1 layer. As illustrated in Fig. 1D, information in V1 increases with population size initially, but saturates for larger populations. Note that this is also the case for the population with heterogeneous filters (Fig. 2B), for which the resulting tuning curves for orientation have heterogeneous width and amplitude, although saturation occurs at a larger number of neurons. In ref. 7, it was shown that only a very special type of correlations known as differential correlations can limit information in large population codes. Because the correlations induced by the feedforward network limit information, they must contain differential correlations for large networks.

These results generalize to other stimuli besides orientation. Qualitatively, any stimulus that is displayed to a noisy receptor

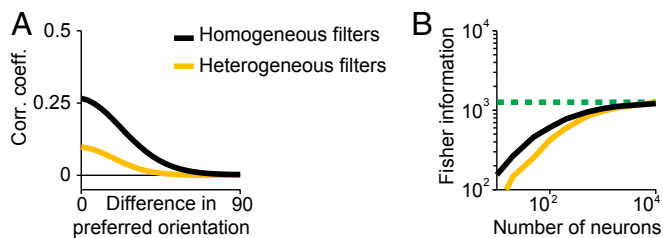


Fig. 2. Effects of tuning heterogeneity on correlations and information. (A and B) Noise correlations (A) and information (B) in the same network as in Fig. 1 (black) and in a network with heterogeneous filters (yellow). The same conventions as in Fig. 1 are used. Simulation parameters are specified in *SI Appendix, Table S3*.

with finite capacity will have finite information and therefore induce differential correlations in subsequent layers. Similarly, neurons with more overlap in their linearized receptive fields will share more feedforward-induced noise and therefore have larger noise correlation, giving rise to the experimentally observed noise profile. The quantitative results of tuning curves, covariance matrices, and Fisher information (*Methods, Eq. 3*) rely on the assumption that the input pattern can be parameterized as $\mathbf{I}(s)$, where s is the manipulated stimulus. When this is true, a linear decoder is sufficient to perform optimal fine discrimination, in the sense that the linear decoder would recover all of the Fisher information.

Therefore, our analysis shows that a simple feedforward model, which also explicitly considers the noise injected by the sensory periphery into sensory signals, accounts for realistic noise correlations as well as for limited information, without the need to postulate any additional internal source of correlated noise.

Network Optimization Cannot Increase Information Indefinitely. The data-processing inequality guarantees that information saturates in large networks, but it does not fully determine the value at which information saturates; it provides only an upper bound. The saturation value depends on whether the network performs optimal computation. If there are suboptimal steps, the information will saturate below the upper bound.

When this is the case, it is possible to improve the efficiency of the code by optimizing the tuning and noise correlations. Behavioral improvement achieved by perceptual learning and attention is typically accompanied by increased sensitivity in single neurons (i.e., tuning sharpening or amplification) (29, 30) and reduced noise correlations (24, 31, 32), suggesting that these might be efficient strategies to increase information. Many previous theoretical studies have derived more general solutions for the optimal shape of tuning curves under the assumption of independent response variability (33–36). Others have considered the effects of correlated variability on coding accuracy for a fixed set of tuning curves (3, 4, 6, 37). However, in these studies, the authors considered population codes consisting of isolated units whose tuning curves and correlations can be manipulated independently. In a network model like the one we considered, this is not an option: The tuning curves and the correlations cannot be independently adjusted (19, 38). Instead, the only way to optimize the network consists in modifying the connectivity that determines in turn both the tuning curves and the correlations.

In the case of our model, it is possible to determine analytically the optimal connectivity profile for orientation discrimination (i.e., the one that allows the network to recover all of the input information), along with the resulting correlation structure. In the above example the oriented Gabor stimuli can be formalized by a function $\mathbf{I}(\theta)$ parameterized by orientation θ , where the vector notation denotes image pixels. Given the linear filters \mathbf{F}_i of the neural population, the tuning curves are given by the responses to Gabor images parameterized by orientation, i.e., the rectified dot

product between filter and image (*Methods, Eq. 3*). Similarly, the noise covariance matrix is given by the product of the filter matrix (a matrix containing each filter as a row) and its transpose, plus a stimulus-dependent diagonal matrix for the Poisson step (*Methods, Eq. 3*).

As we show in *Methods* and *SI Appendix, section 2*, the linear Fisher information in a neural population without Poisson noise is given by $\text{FI}_{\text{neural}} = \text{FI}_{\text{input}} \cos^2(\alpha)$, where α is the angle between the derivative $\mathbf{I}'(s)$ and the vector space spanned by the filters. If $\mathbf{I}'(s)$ can be written as a linear combination of the filters, the angle $\alpha = 0$ and all input information is preserved, corresponding to optimal connectivity. In the case of an oriented Gabor stimulus, this is the case if a subset of the neural filters is also given by Gabors of the same size and spatial frequency as the stimulus: The input information is preserved because $\mathbf{I}'(s)$ can be approximated by the difference of two Gabor filters of nearby orientations. Note that whereas this is the case for a homogeneous population where all neural filters have the same frequency as the stimulus, it is also true for heterogeneous filters (Fig. 2B) as long as the stimulus frequency is not outside the range covered by the filters. On the other hand, if, for instance, all of the neural Gabor filters share the same spatial frequency, and such frequency differs from that of the stimulus, information will be lost (Fig. 3C). This is intuitively clear if one thinks for instance of a discrimination task between two signals: It is a well-known result of signal processing that the optimal template is the difference between the patterns to be discriminated (39). These results carry over for neurons that are additionally corrupted by independent Poisson noise, if the population is large enough, as illustrated in Fig. 3C.

Therefore, although a suboptimal network contains less information than an optimal one, information does not grow indefinitely in either case. Importantly, the optimal tuning curves predicted by our approach are very different from the one predicted by efficient coding approaches that ignore that both arise as a consequence of connectivity and limited input information. It is well known for instance that for a population code for a 1D variable such as orientation, corrupted by independent Poisson noise, the information is optimized when the tuning curves are infinitely sharp (33). This is not the case in our network: Tuning width depends on the shape of the filters (*SI Appendix, section 3*); e.g., if the suboptimal Gabor filter has a higher spatial frequency (lower spatial wavelength) than the optimal one, tuning curves will be sharper for suboptimal than optimal neurons (Fig. 3A). Second, in our network optimality does not imply decorrelation either.

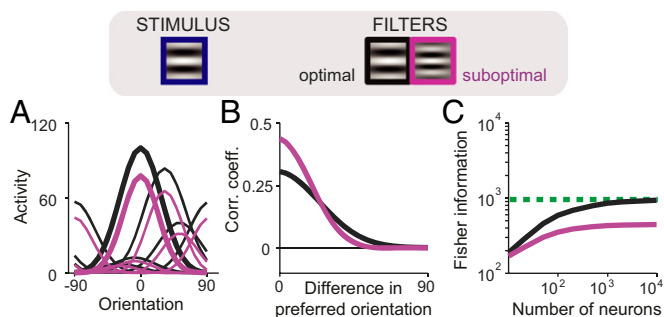


Fig. 3. Network optimization does not imply sharpening or decorrelation. (Top) A population of filters with size and spatial frequency matched to the stimulus is optimal. (A) Tuning curves for optimal filters (black) are not sharper than for suboptimal filters (purple). (B) Noise correlations in the optimal network are not uniformly lower than in the suboptimal network. (C) Suboptimal filters cannot extract asymptotically all of the information contained in the input. Simulation parameters are specified in *SI Appendix, Table S2*.

For instance, because noise correlations are determined by filter overlap, if suboptimal Gabor filters have a higher spatial frequency than the optimal one, optimizing the connectivity will lead to a broadening of the limited-range structure, rather than a uniform decrease in magnitude (Fig. 3B). Conversely, the optimal homogeneous population requires fewer neurons than a heterogeneous population to reach saturation (Fig. 2B), but the former has uniformly larger correlations (Fig. 2A).

A Quadratic Nonlinearity Yields Contrast-Independent Correlations and Decoding. So far, we have considered a model with a rectifying nonlinearity but experimental data suggest that the nonlinearity in vivo is closer to a quadratic function [as is used in standard models of simple cells in V1 (40)]. Importantly, the results we have presented so far do not depend substantially on this choice: In both cases, linear rectified and quadratic, correlations decline with difference in preferred tuning, information is limited by the input, and the maximal input information can be recovered using optimal filters (*SI Appendix*, Fig. S2). However, there are two important aspects of this network that are critically dependent on the choice of nonlinearity.

First, whereas correlations in the linear-rectifying model depend on both stimulus and contrast, they are approximately invariant in the quadratic model (*SI Appendix*, sections 4 and 5, and Fig. 4A). It has been reported that noise correlations in visual cortex are largely invariant to both stimulus value and stimulus strength on behaviorally relevant timescales (i.e., a few hundred milliseconds) (21, 22). This observation thus favors the quadratic over the linear-rectifying nonlinearity. Second, the quadratic model leads to simpler decoding, because both the tuning curves and the covariance matrix scale approximately homogeneously with contrast. Because the locally optimal linear decoder of neural activity is proportional to the product of the inverse covariance matrix with the vector of tuning curve derivatives (19), the optimal linear decoder scales homogeneously as well (*SI Appendix*, section 5). This has the advantage that if contrast fluctuates from trial to trial, as is the case in natural

circumstances, the optimal decoder does not need to be adapted as the same relative decoding weights can be used on different trials (Fig. 4B). Furthermore, because the product of the inverse covariance with the tuning curves is invariant to contrast, the feedforward model with quadratic nonlinearity implements a linear probabilistic population code that is invariant to contrast (41).

Although the optimal linear decoder is independent of contrast, it is not independent of the level of input noise for both nonlinearities (Fig. 4C and *SI Appendix*, sections 4 and 5). The intuition is that, for purely Poisson variability, the most informative neurons (i.e., those with large decoding weights) are tuned away from the stimulus to be decoded, because their tuning curve has maximal slope, leading to a broad decoder profile. Conversely, if the input noise is much larger than the Poisson variability, then it is sufficient to assign large weight to the two neurons closest to the stimulus to be decoded, one on each side (as explained in the previous section, the difference between such two neurons is essentially the output of the optimal template); therefore, when the noise is mostly due to input variability, the decoder's profile is narrow. Varying the relative contribution of input noise and Poisson variability results in a gradual broadening or narrowing of the decoder's profile. Thus, our framework predicts that, if the input noise level is manipulated by adding pixel noise, perceptual learning on one particular noise level should improve performance for that particular noise level but not necessarily for higher or lower levels. Lu and Doshier (42) indeed reported that training on noisy stimuli does not transfer to noiseless stimuli. Note that in addition to affecting the optimal decoder profile, the level of input noise determines also the input information level and therefore the level at which the network information saturates. We found that cortical information does not simply scale with input noise level, but the population size required for saturation changes as well. Thus, in a population with homogeneous filters, 95% of the input information was recovered by ~500 neurons at 50% input noise level but more than 10,000 neurons are needed at 5% noise level (*SI Appendix*, Fig. S3). This prediction can also be tested experimentally by recording from large cortical populations while adding different amounts of pixel noise to the stimuli.

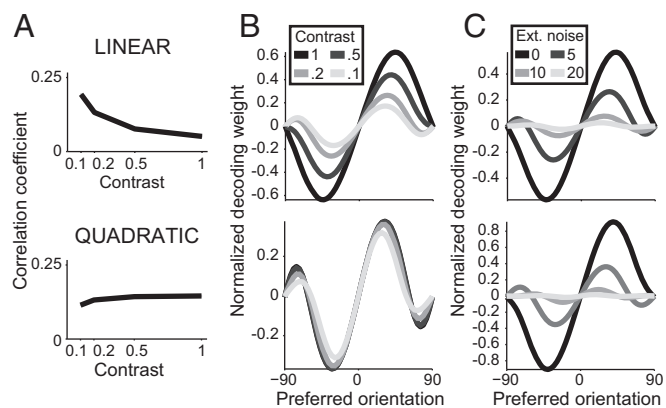


Fig. 4. Optimal decoding with linear and quadratic V1 models. (A) Noise correlations depend on image contrast in the linear model (Top), but are approximately contrast invariant in the quadratic model (Bottom). (B) Weights of the locally optimal decoder of orientation around 0° plotted as a function of the preferred orientation of each neuron. The optimal linear decoder depends on image contrast in the linear model (Top), but is contrast independent in the quadratic model (Bottom). (C) The locally optimal linear decoder depends on the amount of noise added at the sensory periphery in both models. In B and C the normalization convention for the decoding weights is $\mathbf{w}_{\text{norm}} = \mathbf{w}_{\text{opt}} \text{FI} = \Sigma^{-1} \mathbf{f}'$, where Σ is the covariance matrix and \mathbf{f}' is the derivative of the tuning curves (*Methods* and *SI Appendix*, section 5). Note that the decoding weights in both models are smooth despite heterogeneous tuning. Simulation parameters are specified in *SI Appendix*, Table S4.

Global Fluctuations Reduce, but Do Not Limit, Information. So far we have discussed properties of the response variability that arises from variability in the sensory periphery, and we have shown that a simple feedforward model is sufficient to account both for limited information in large populations and for many properties of experimental noise correlations. However, recent studies have revealed that noise correlations in visual areas also reflect slow internal fluctuations that are shared between groups of neurons and proposed that such fluctuations are the main source of experimentally measured noise correlations (12–16). However, the impact of this prominent source of correlation on information is largely unknown, except for two empirical studies of auditory (43) and visual (44) cortex whose results are based on small groups of neurons and cannot be extrapolated to large populations. Extrapolation from small populations to large populations is indeed prone to very large errors (7). Therefore, we asked how global fluctuations affect feedforward-generated noise correlations and what impact they have on information.

Following ref. 15, we model global fluctuations resulting from the product of a fluctuating gain factor with an underlying stimulus-driven rate. In our case, the latter is determined by the feedforward filtering of the noisy stimulus. In *SI Appendix*, section 6, we show that such global fluctuations modify the feedforward covariance by rescaling it (i.e., expanding the noise covariance equally in all directions) and adding a one-dimensional perturbation (Fig. 5A). As a result, noise correlations are amplified for all neuronal pairs and regardless of whether the neuronal filters are homogeneous (Fig. 5B) or heterogeneous (Fig. 5C).

Because global fluctuations amplify noise correlations, do they also destroy information? Intuitively, a global scaling of the population response affects only the height of the population hill of activity, leaving the location of the peak (and therefore the stimulus value extracted by a linear readout) unchanged. Therefore, one could expect that global fluctuations do not change information. In *SI Appendix, section 6* we show that this is only partly correct: Global fluctuations do decrease the asymptotic information in a large population, to an extent that depends on the level of the fluctuations. Intuitively, this is because multiplicative global fluctuations expand the original noise covariance, and the expansion is equally large in all directions: Therefore, noise is amplified also in the direction of the signal, thus increasing the magnitude of differential correlations. However, global fluctuations by themselves cannot limit information in large populations, consistent with the original intuition (Fig. 5 *D* and *E*).

Furthermore, global fluctuations reduce asymptotic information only if the decoder does not have access to the fluctuating gain factor. If the latter is known, one can easily divide it out and recover the full input information. For example, if global fluctuations in a neural population arise due to a fluctuating top-down signal but the same top-down signal is also available to the downstream circuits reading out the population, the information loss due to global fluctuations would not be relevant for behavior even though it is measured in neural recordings. Goris et al. (15) and Ecker et al. (14) conjectured nonsensory signals such as arousal, attention, and adaptation to be among the major causes of global fluctuations in awake animals; whether these signals are shared across sensory and decision areas and removed at the decision stage is currently unknown.

If the fluctuating gain parameter is not shared with decision areas and asymptotic information is affected by global fluctuations, attention could also be interpreted as a mechanism that improves behavioral performance by reducing the variance of global fluctuations. Such an interpretation would be consistent

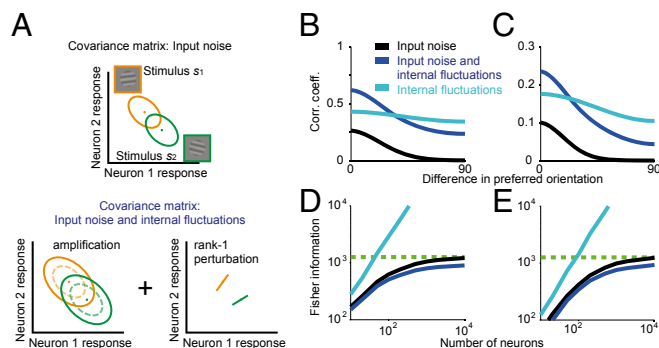


Fig. 5. Effects of internal global fluctuations on correlations and information. (A) Illustration of the effects of global fluctuations on the noise covariance matrix. (A, Top) Covariance matrices (ellipses) due to input noise around two stimuli (dots represent the mean responses to each stimulus). (A, Bottom) Global fluctuations produce a uniform amplification of the covariances (Left) and an additive, rank-1 perturbation (Right). See *SI Appendix, section 6* for derivation. (B) Internal fluctuations increase pairwise noise correlations approximately uniformly (dark blue) compared with the model with peripheral noise alone (black). In the absence of peripheral noise, internal fluctuations produce correlations that depend only weakly on the difference in preferred orientation when averaged over the entire population (light blue). Note that in these simulations the neurons have completely overlapping receptive fields and hence the large correlations. (C) Same as B, but with heterogeneous instead of homogeneous filters. (D) Internal fluctuations can reduce asymptotic information when considered in a model with peripheral noise (compare dark blue and black). In a model without peripheral noise, internal fluctuations do not limit asymptotic information (light blue). (E) Same as D, but with heterogeneous instead of homogeneous filters. Simulation parameters are specified in *SI Appendix, Table S3*.

with the finding that attention is often accompanied by uniform decorrelation (31, 32) and with the recent observation that directing attention to a target decreases the variance of global fluctuations (45). Rather than by decorrelating an existing correlation pattern as originally suggested (31), attention in this scenario would increase information by suppressing gain fluctuation as an additional, detrimental source of variance.

Computation-Induced Noise Correlations Contain a Tiny Amount of Differential Correlations. The tuning curve pattern, correlation pattern, and information saturation behavior of our model (Figs. 1 and 2) seem in apparent contradiction with the results of ref. 6. Both their and our models have heterogeneous tuning curves and positive correlations proportional to tuning curve similarity, but although they report that information grows unboundedly with network size, it saturates in our case. What is the reason for this discrepancy?

The discrepancy comes from the fact that our model, unlike the one of Ecker et al. (6), contains information-limiting correlations. Still, the overall pattern of correlations is dominated by non-information-limiting correlations; this is why overall correlations look similar in both models. Moreno et al. (7) already suggested that differential correlations might be small compared with other correlations in vivo but our model allows us to test this idea quantitatively in a realistic model of cortical computation.

As we show in *SI Appendix, section 7*, any information-limiting covariance matrix can be split into a positive definite non-information-limiting part and an information-limiting part, which is proportional to the product of the tuning curve derivatives. The size of the information-limiting part is determined by the relative size of the behavioral threshold and the tuning curve width, squared. Recall that, under the assumption that all of the information encoded in V1 will be extracted and used to guide behavior, behavioral thresholds are inversely related to linear Fisher information via Eq. 7 in *Methods*. If the behavioral threshold is much smaller than the tuning curve width, the magnitude of the information-limiting part is very small, while having a significant impact on information. For instance, in the case of orientation discrimination at high contrast, with a threshold of roughly 2° and tuning width of roughly 20° , the square of their ratio is 0.01; i.e., information-limiting correlations contribute to 1% of the total correlations. Given the size and measurement error of experimentally observed correlations, this implies that it would be very hard to detect differential correlations directly in experimental data.

In Fig. 6 we compare the information, tuning curves, and correlations of our model and a model in which the differential correlations have been artificially removed. For a behavioral threshold of 2° , the contribution of the differential correlations is so small that it is virtually impossible to discern the difference in the correlations plot (Fig. 6*B*). This is true not only on average, but also on a pair-by-pair basis: Even ignoring measurement errors, the difference in correlation coefficients with vs. without differential correlations is larger than 0.05 in only about 5% of the pairs and larger than 0.1 in only 0.2% of the pairs (*SI Appendix, Fig. S4*). The impact on information, however, is dramatic: When the tiny amount of differential correlations is removed, saturating information turns into nonsaturating information (Fig. 6*C*).

This resolves the apparent paradox: Having a set of tuning curves and correlations that look similar to experimental data does not determine the information content in the population, because a tiny error in the measured correlation can have a huge impact on measured information for large populations.

Paradoxes of Cortical Models That Do Not Operate at Saturation. So far we have shown that a simple model based on noisy inputs and feedforward computation accounts for realistic and information-limiting correlations. We also saw that in a large network tiny differential correlations will have a huge impact on information because they are solely responsible for saturation. But is the

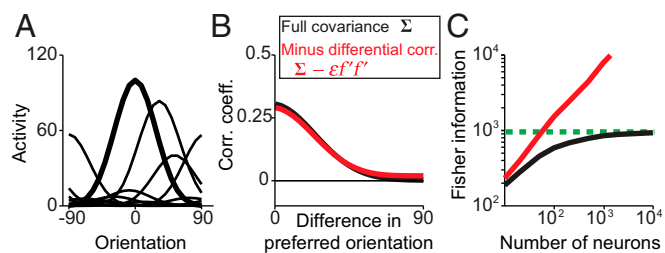


Fig. 6. Limited input information introduces differential correlations. (A) Tuning curves are identical before or after removing differential correlations. (B) Only minimal changes in the pattern of average noise correlations result from removing differential correlations. (C) Information is not asymptotically limited after removing differential correlations. Simulation parameters are specified in *SI Appendix, Table S2*.

number of neurons involved in a given computation large enough to make the black curve in Fig. 7 saturate? In the specific example of V1 and orientation, this means assuming that the number of relevant V1 neurons for the discrimination (i.e., those that are orientation tuned and project to higher cortex) is much larger than the number of peripheral inputs. This is a biologically realistic assumption: For instance, based on anatomy considerations we estimated that the experimental stimuli used in typical psychophysics studies (20) would activate $\sim 300,000$ relevant V1 neurons (46–48) but only 10,000 LGN neurons (48). Similar ratios of inputs to outputs have been reported previously for V1 (49) and other sensory cortices (50–52).

Whereas saturation of information for large populations is an inevitable necessity given limited input information, there also exists the possibility that the brain operates at a point of the saturation where information still grows if more neurons are added. However, the behavioral thresholds that would be expected from independent variability (or non-information-limiting correlations alone) and realistic numbers of neurons are an order of magnitude smaller than what is found experimentally (Fig. 7) (53).

Furthermore, it is unclear how this scenario can be made consistent with two key experimental observations on the relation between single-neuron activity and perceptual choices (assuming that the population responses are read out optimally and disregarding feedback effects, assumptions we discuss below). First, neurometric thresholds (i.e., the minimal stimulus difference that can be discriminated reliably from the responses of a single neuron) are comparable to psychometric thresholds (i.e., those measured in the perceptual discrimination task) (54, 55) (Fig. 8A). This result is hard to explain in the nonsaturated regime: If information grows linearly with the number of neurons, the single-neuron information is negligible compared with the full population, and therefore the neurometric threshold is large compared with the psychometric threshold (Fig. 8B and C). Second, the activity of single neurons can be used to predict behavioral responses on a trial-by-trial basis, as quantified by choice probabilities (a measure of the correlation between neural activity and choice) (Fig. 8D) (54, 56–58). However, in the absence of saturation the contribution of each single neuron to behavior is minimal and vanishes in the limit of large populations (Fig. 8E and F). These results assume that the behavioral performance of the animal corresponds to that of the optimal decoder and that choice probabilities are not influenced by feedback effects. Indeed, one way to rescue a model with nonsaturating information is to assume that in areas like V1 there is much more information than subjects are able to extract in a simple task like orientation discrimination. Such a suboptimal readout of V1 stands in conflict with the ability of human subjects (53) to approximate ideal observer performance in orientation discrimination in the presence of external noise, at least after perceptual learning, as can be shown by a

simple calculation (*SI Appendix, section 8*). For instance, in the experiments of Doshier and Lu (53), when the stimuli were not corrupted by external noise, human subjects needed between 0.75% and 2.2% signal contrast to achieve 80% correct performance, whereas the ideal observer needs between 0.6% and 1.2% signal contrast; similarly, for the largest amount of external noise tested in the experiment, subjects needed between 4.5% and 7% signal contrast and the ideal observer between 3.5% and 7% (Fig. 9). Also a recently developed formalism of relating choice probabilities to stimulus preferences gives support to the scenario of saturated information and optimal readout as opposed to non-saturated information and suboptimal readout (59). The above arguments address the case of computationally simple tasks like orientation discrimination; for hard tasks like object recognition under naturalistic viewing conditions the readout of V1 is likely suboptimal due to extensive computational approximations of downstream circuitry (60).

It is also possible that choice probabilities are high not because of saturating information but because they are enhanced by feedback from areas related to decision making (57, 61). However, a time-course analysis of the decision signal (8, 57, 58) and inactivation studies (62) suggest that feedback alone is not sufficient to explain the observed above-chance choice probability. Therefore, this scenario remains speculative but certainly deserves to be investigated in the future.

Discussion

Noise correlations have attracted considerable attention over the last few years because of their impact on the information of neural representations. Of particular importance is to understand the origin of one type of correlations known as differential correlations, the only kind of correlations that can make information saturate to a finite value for a large neuronal population. Here, using orientation selectivity in V1 as a case study, we show that noise in peripheral sensors and potentially suboptimal computation are sufficient to account for a significant fraction of noise

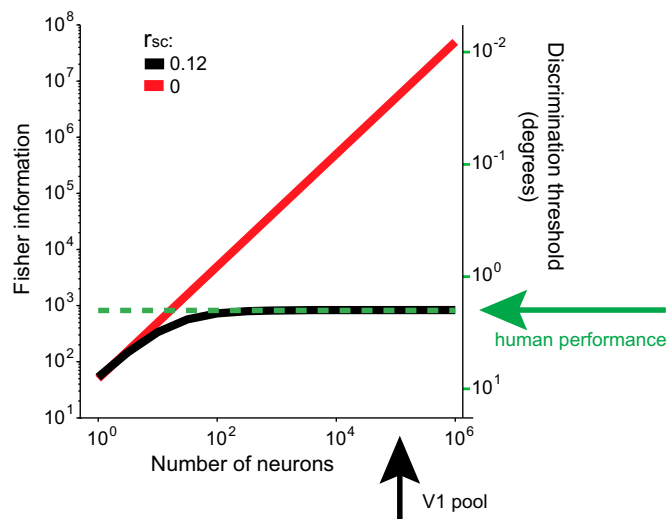


Fig. 7. Information-limiting noise correlations account for realistic psychophysical thresholds. A population of independent Poisson neurons (red line) roughly the size of the V1 pool that can be used downstream for discrimination in typical psychophysics experiments (black arrow; see text for estimation) leads to psychophysical thresholds that are orders of magnitude smaller than human performance (green arrow). A population with information-limiting correlations of 0.12 on average (black line) accounts for realistic thresholds. Simulations are based on a simplified model with synthetic covariance and tuning curves, with average amplitude of 20 spikes per second (*SI Appendix, Table S5*).

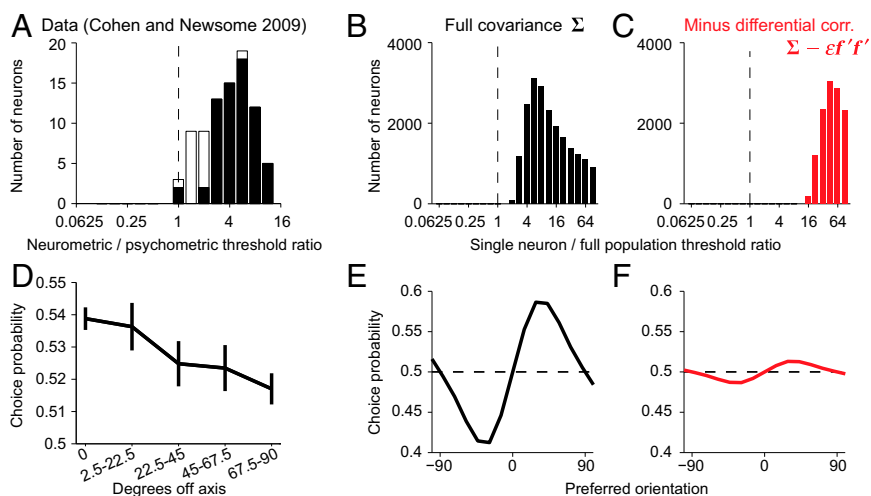


Fig. 8. Information-limiting noise correlations account for neurometric thresholds and choice probabilities. (A) Distribution of the ratio of neurometric to psychometric thresholds (54). The neurometric threshold is defined as the minimal stimulus difference that can be discriminated reliably from the responses of a single neuron to the two stimuli. Ratios close to 1 imply that the performance of the animal in the discrimination task can be closely approximated by an ideal observer of the single-neuron responses. Black bars, cases with ratio significantly different from 1. Note that the histogram plotted here is displaced by one bin to the right compared with the original one published by Cohen and Newsome (54). This is because the analysis of Cohen and Newsome (54) overestimated the single-neuron information by a factor of 2 (Methods, Eq. 8) due to their neuron/antineuron assumption that effectively turns one neuron into two independent neurons. Republished with permission from ref. 54. (B and C) Same as A, for the network model before (B) and after (C) removing differential correlations. (D) Average choice probabilities in a motion discrimination task, as a function of the difference between neuronal preferred direction and the discrimination axis, as in ref. 54. Choice probability measures the correlation between single-neuron activity and perceptual choice on a trial-by-trial basis; values close to 0.5 indicate no correlation, and values larger than 0.5 indicate that the behavior can be predicted from the single-neuron response. Republished with permission from ref. 54. (E and F) Average choice probabilities as a function of preferred orientation for a discrimination around 0°, before (E) and after (F) removing differential correlations. Note that in D the stimulus is motion direction and the task is a coarse discrimination, whereas E and F are a fine discrimination for orientation, so the abscissae and stimulus dependences are not directly comparable. Simulation parameters are specified in *SI Appendix, Table S2*.

correlations and in particular differential correlations. In contrast, previous influential accounts have emphasized the potential role of internal noise in the emergence of noise correlations, due for instance to shared cortical input (2) or chaotic dynamics in balanced networks (9, 10, 63). This correlated internal noise has been suggested to limit information and therefore behavioral performance (2, 11). The problem with this approach is that internal noise is unlikely to limit information, unless it is fine-tuned to contain differential correlations (7). This is particularly unlikely given that differential correlations are stimulus specific. Thus, if cortex happens to generate noise that contains differential correlations for, say, direction of motion, these correlations could not make information saturate for speed of motion or many other perceptual variables. One might argue that even if internally generated correlations do not contain differential correlations, they might reduce information substantially relative to independent variability, such as the correlation structure explored by ref. 6. This corresponds to a red line with lower slope in Fig. 7 and narrows the gap between decoding performance and behavioral performance. However, it is unclear whether a noise correlation structure like that of ref. 6 could actually be generated purely by internal noise. Another recently proposed alternative is that differential correlations are induced via feedback from decision-related areas (61). However, such differential correlations do not necessarily impact information because they are known to the decision area and could be easily factored out.

It is therefore unclear whether internal variability contributes significantly to information-limiting differential correlations. Instead, we have argued elsewhere that the main sources of information limitation are variability in peripheral sensors and suboptimal computation (60). This study shows that the two sources do indeed generate information-limiting correlations and lead to overall patterns of correlations very similar to those observed in vivo. Interestingly, we also found that only a tiny fraction of correlations in our model are differential correlations. Moreno et al. (7) previously suggested that this might explain why

the overall patterns of correlations in vivo bear no resemblance to the pattern predicted by pure differential correlations. We have shown here that this is indeed the case in a realistic model of orientation selectivity. More specifically, we found that when tuning curve widths are larger than psychophysical thresholds, differential correlations will be tiny compared with nonlimiting ones (Fig. 6). This has important ramifications for experimentalists because it suggests that detecting differential correlations directly will be highly challenging: A tiny measurement error can give rise to drastically different population information. It might be easier to measure information directly from simultaneous recordings of large neuronal populations but this might require one to record for a population of several thousand active neurons, which is currently beyond what is technically possible.

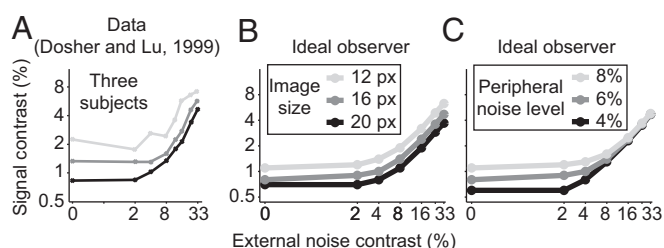


Fig. 9. Assessing optimality of orientation discrimination performance in trained human subjects. (A–C) Threshold vs. contrast curves for three human subjects as in ref. 53 (A) and the ideal observer (B and C). The ordinate represents the grating contrast required to achieve a discrimination performance of 79.3% correct, at each noise level (abscissa). The derivation of stimulus parameters and ideal observer are in *SI Appendix, section 8*. Different shades of gray correspond to different experimental subjects (A), different assumptions about the number of pixels in the image (B), and different assumptions about the peripheral noise level (C). Data from ref. 53.

An interesting question to ask is how the differential correlations induced by feedforward connectivity are modified by recurrent connectivity. In ref. 64, it was shown that a recurrent network of balanced excitatory and inhibitory connections can decorrelate weak feedforward-induced correlations to be arbitrarily small, as the network size increases. However, the feedforward-induced correlations in this study did not contain differential correlations, because the size of the input layer grew together with the output layer. Moreno et al. (7) found that a balanced recurrent network can reduce average correlations even if differential correlations are present in the input, but this decorrelation had only a minimal effect on information and only for small networks. The reason is twofold: First, not the average level of correlations, but only the specific pattern of differential correlations is responsible for information saturation in large networks. Second, differential correlations must be present, because the input information is limited, which by the data-processing inequality implies that also the output information must be limited. In our case, the input information is limited because the sensory periphery is noisy and has a small capacity compared with cortex. The results of Moreno et al. (7) suggest that although recurrent connectivity might affect the overall size and shape of noise correlations, it will not affect information saturation or substantially affect the size of the network necessary for saturation.

This work allowed us also to understand how global fluctuations influence information in population codes. Global fluctuations have been identified as a major source of correlations in vivo. However, global fluctuations on their own do not generate differential correlations for orientation and, as such, do not strongly impact information for orientation and any other sensory variables except perhaps contrast. Nonetheless, if differential correlations are already present in a code, we showed that global fluctuations can lower the level at which information saturates. This suggests an intriguing theory of attentional control: Attention might increase information in population codes, and thus improve behavioral performance, by reducing global fluctuations. This would predict that attention should lead to an overall drop of correlations (Fig. 5), which is indeed what has been observed in vivo.

Recent work has addressed the question of how sensory noise affects coding and optimal receptive fields of single neurons (65). More closely related to our work, how correlations induced by feedforward connectivity affect information was explored previously by refs. 19, 38, and 66–68. As in our model, it was found that feedforward-induced correlations have a big impact on information and affect the choice of optimal tuning curves. Our model goes beyond these earlier studies in that it explicitly limits information at the input and in that it allows quantifying analytically the information loss relative to the input, comparing neural information to psychophysical thresholds and separating information-limiting and nonlimiting correlations.

Importantly, our model makes testable experimental predictions. It predicts in particular correlations of individual pairs given the filters associated with each neuron. Moreover, the relation between tuning and noise correlation has an explicit dependence on the level of input noise. This could be readily tested by recording from neurons with closely overlapping receptive fields while manipulating the input noise.

Even though we focused our discussion on orientation selectivity and V1, our model readily generalizes to other modalities, stimuli, and brain areas. Instead of a Gabor image parameterized by orientation, $\mathbf{I}(s)$ can be made to denote any high-dimensional input parameterized by a low-dimensional stimulus, such as auditory or tactile stimuli parameterized by frequency. Similarly the neural filters \mathbf{F}_i can be defined with respect to the appropriate high-dimensional input space. In all such cases, noise in the input will limit information and induce correlations in the neural representation. Similarly the analytical results on optimality in the case of matched filters and the relationship between filters, tuning curves, and covariance generalize to this case.

What is not yet addressed in our model is the case of a high-dimensional input $\mathbf{I}(s, t)$, which depends not only on a well-controlled stimulus s but also on a fluctuating nuisance variable t . Such a case arises, for example, in object recognition under naturalistic viewing conditions where s corresponds to object categories in a discrimination task and t corresponds to irrelevant image features that need to be marginalized out (69). In such a situation, sophisticated nonlinear processing is required to reformat information in such a way that it can be read out linearly. A specific example was studied in the coding of visual disparity in binocular images, with the stimulus given by a random stereogram (70). However, that study addressed only single neuron rather than population information. It would be interesting to generalize our model for such processing and predict correlations in higher visual areas.

Methods

Orientation Network Model. The model we used for the simulations is based on a network comprising a noisy input layer followed by a bank of orientation-selective filters with static nonlinearity and Poisson spike generation. Here we provide a brief summary of the model details. Full mathematical derivations and network parameters for the simulations are presented in *SI Appendix*.

The inputs to the network are Gabor patches corrupted by additive white noise with variance σ_0^2 ,

$$\tilde{\mathbf{I}}|\theta \sim \mathcal{N}(\mathbf{I}(\theta), \sigma_0^2 \mathbf{1}_P), \quad [1]$$

where P is the number of pixels in the image, $\mathbf{1}_P$ is the identity matrix of size $P \times P$, and $\mathbf{I}(\theta)$ is a noiseless Gabor with orientation θ .

The V1 model neurons filter the sensory input by a Gabor $\mathbf{F}_i \in \mathbb{R}^P$ oriented according to their preferred orientation θ_i , followed by a rectifying nonlinearity and a Poisson step. The neural response of neuron i is then sampled from

$$r_i \sim \text{Poisson}\left(\left[\mathbf{F}_i \cdot \tilde{\mathbf{I}}\right]_+\right), \quad [2]$$

where $[\cdot]_+$ denotes half rectification. The tuning curves and covariance matrix are given by

$$f_i(\theta) = \langle r_i \rangle \approx [\mathbf{F}_i \cdot \mathbf{I}(\theta)]_+, \quad [3]$$

$$\Sigma_{ij}(\theta) = \text{Cov}(r_i, r_j) \approx \sigma_0^2 [\mathbf{F}_i \cdot \mathbf{F}_j]_+ + \delta_{ij} f_i(\theta),$$

where the approximation is valid for small input noise and for neurons that are well above the firing threshold.

The correlation coefficient for two neurons $i \neq j$ is given by

$$\rho_{ij} = \frac{\sigma_0^2 [\mathbf{F}_i \cdot \mathbf{F}_j]_+}{\sqrt{\sigma_0^2 |\mathbf{F}_i|^2 + f_i(\theta)} \sqrt{\sigma_0^2 |\mathbf{F}_j|^2 + f_j(\theta)}}. \quad [4]$$

From expression [4] we see that the correlations between the neurons will be proportional to the overlap of their filters \mathbf{F}_i and \mathbf{F}_j , giving rise to the limited-range structure of correlations. In particular, the result that neurons with similar receptive fields share more feedforward-induced noise correlation is not specific to Gabor filters and orientation and can be generalized to other stimuli and modalities.

More details and an analytical approximation to tuning curves and covariance matrix for Gabor filters can be found in *SI Appendix, section 1*.

The Fisher information about orientation in the image is given by

$$F_{\text{input}}(\theta) = \frac{1}{\sigma_0^2} |\mathbf{I}'(\theta)|^2,$$

where $' = d/d\theta$ is the derivative. For a finite number of pixels, information is limited. The linear Fisher information in the neural population without Poisson noise is given by

$$F_{\text{neural}} = \sum_{ij} f_i(\theta) \Sigma_{ij}^{-1} f_j(\theta) = \frac{1}{\sigma_0^2} \mathbf{I}'(\theta) \cdot \sum_{ij} \mathbf{F}_i (\mathbf{F} \cdot \mathbf{F})_{ij}^{-1} \mathbf{F}_j \cdot \mathbf{I}'(\theta) = F_{\text{input}} \cos^2 \alpha,$$

where α is the angle between the derivative $\mathbf{I}'(s)$ and the vector space spanned by the filters. If $\mathbf{I}'(s)$ can be written as a linear combination of the filters, this angle is zero and the input information is preserved. More details can be found in *SI Appendix, section 2*.

For the simulations in Fig. 4, we also considered a quadratic nonlinearity rather than half rectification. The tuning curves and covariance matrix can also be computed in closed form and are given by

$$f_i^{(NL)}(\theta) = \left(f_i^{(L)}(\theta)\right)^2 + \Sigma_{ii}^{(L)}$$

$$\Sigma_{ij}^{(NL)}(\theta) = 4\Sigma_{ij}^{(L)}f_i^{(L)}(\theta)f_j^{(L)}(\theta) + 2\left(\Sigma_{ij}^{(L)}\right)^2,$$

where $f_i^{(L)}(\theta)$ and $\Sigma_{ij}^{(L)}(\theta)$ correspond to tuning curves and covariance matrix, respectively, of the linear model. This allows us to calculate the correlation coefficients in Fig. 4A. The normalized decoding weights for both linear and nonlinear models are defined as

$$\mathbf{w}_{\text{norm}} = \Sigma^{-1}\mathbf{f}.$$

More details can be found in *SI Appendix, section 5*.

For the simulations in Fig. 5, we extended the network by introducing an internal source of shared variability, modeled as a positive random variable that multiplies the rate of the neurons before the Poisson step, as in ref. 15. In this case, the neural response of neuron i is then sampled from

$$r_i \sim \text{Poisson}\left(g\left[\mathbf{F}_i \cdot \bar{\mathbf{I}}\right]_+\right)$$

$$g \sim \Gamma\left(\frac{1}{\sigma_g^2}, \frac{1}{\sigma_g^2}\right); \langle g \rangle = 1, \quad [5]$$

where σ_g^2 is the variance of the internal fluctuations. This extension does not change the tuning curve of the neurons, but it changes the covariance matrix to

$$\Sigma^{(g)} = \left(\sigma_g^2 + 1\right)\Sigma + \sigma_g^2\mathbf{f}\mathbf{f}^T + \text{diag}(\mathbf{f}), \quad [6]$$

- Zohary E, Shadlen MN, Newsome WT (1994) Correlated neuronal discharge rate and its implications for psychophysical performance. *Nature* 370(6485):140–143.
- Shadlen MN, Newsome WT (1998) The variable discharge of cortical neurons: Implications for connectivity, computation, and information coding. *J Neurosci* 18(10):3870–3896.
- Abbott LF, Dayan P (1999) The effect of correlated variability on the accuracy of a population code. *Neural Comput* 11(1):91–101.
- Sompolinsky H, Yoon H, Kang K, Shamir M (2001) Population coding in neuronal systems with correlated noise. *Phys Rev E Stat Nonlin Soft Matter Phys* 64(5 Pt 1):051904.
- Averbeck BB, Latham PE, Pouget A (2006) Neural correlations, population coding and computation. *Nat Rev Neurosci* 7(5):358–366.
- Ecker AS, Berens P, Tolias AS, Bethge M (2011) The effect of noise correlations in populations of diversely tuned neurons. *J Neurosci* 31(40):14272–14283.
- Moreno-Bote R, et al. (2014) Information-limiting correlations. *Nat Neurosci* 17(10):1410–1417.
- Wimmer K, et al. (2015) Sensory integration dynamics in a hierarchical network explains choice probabilities in cortical area MT. *Nat Commun* 6:6177.
- Ben-Yishai R, Bar-Or RL, Sompolinsky H (1995) Theory of orientation tuning in visual cortex. *Proc Natl Acad Sci USA* 92(9):3844–3848.
- Litwin-Kumar A, Doiron B (2012) Slow dynamics and high variability in balanced cortical networks with clustered connections. *Nat Neurosci* 15(11):1498–1505.
- Ly C, Middleton JW, Doiron B (2012) Cellular and circuit mechanisms maintain low spike co-variability and enhance population coding in somatosensory cortex. *Front Comput Neurosci* 6:7.
- Arieli A, Sterkin A, Grinvald A, Aertsen A (1996) Dynamics of ongoing activity: Explanation of the large variability in evoked cortical responses. *Science* 273(5283):1868–1871.
- Harris KD, Thiele A (2011) Cortical state and attention. *Nat Rev Neurosci* 12(9):509–523.
- Ecker AS, et al. (2014) State dependence of noise correlations in macaque primary visual cortex. *Neuron* 82(1):235–248.
- Goris RLT, Movshon JA, Simoncelli EP (2014) Partitioning neuronal variability. *Nat Neurosci* 17(6):858–865.
- Schölvinck ML, Saleem AB, Benucci A, Harris KD, Carandini M (2015) Cortical state determines global variability and correlations in visual cortex. *J Neurosci* 35(1):170–178.
- Croner LJ, Purpura K, Kaplan E (1993) Response variability in retinal ganglion cells of primates. *Proc Natl Acad Sci USA* 90(17):8128–8130.
- Uzzell VJ, Chichilnisky EJ (2004) Precision of spike trains in primate retinal ganglion cells. *J Neurophysiol* 92(2):780–789.
- Seriés P, Latham PE, Pouget A (2004) Tuning curve sharpening for orientation selectivity: Coding efficiency and the impact of correlations. *Nat Neurosci* 7(10):1129–1135.
- Dosher BA, Lu Z-L (1998) Perceptual learning reflects external noise filtering and internal noise reduction through channel reweighting. *Proc Natl Acad Sci USA* 95(23):13988–13993.
- Bair W, Zohary E, Newsome WT (2001) Correlated firing in macaque visual area MT: Time scales and relationship to behavior. *J Neurosci* 21(5):1676–1697.

where, in the right-hand side, Σ is the same as in Eq. 3, and $\text{diag}(\mathbf{f})$ is a diagonal matrix with entries given by the tuning curves of the neurons. More details can be found in *SI Appendix, section 6*.

Analysis of Model Neural Responses. We computed linear Fisher information directly from its definition, using the analytical expressions derived for the tuning curves and covariance,

$$\text{FI} = \mathbf{f}^T \Sigma^{-1} \mathbf{f}, \quad [7]$$

where the symbol ' denotes the derivative with respect to the stimulus.

In Fig. 8, we plot neurometric discrimination thresholds and choice probabilities (CP). The discrimination threshold for a neural population with linear Fisher information FI is

$$\delta_{\text{population}} = \frac{1}{\sqrt{\text{FI}}} \quad [8]$$

and the discrimination threshold for a single neuron is

$$\delta_i^{\text{neuron}} = \sqrt{\frac{\Sigma_{ii}}{f_i^2}}. \quad [9]$$

We computed CPs using the analytical expression derived in ref. 71:

$$\text{CP}_i = 0.5 + \arctan \sqrt{2\xi_i^2 - 1}^{-1}$$

$$\xi_i = \frac{\delta_{\text{population}}}{\delta_i^{\text{neuron}}}. \quad [10]$$

ACKNOWLEDGMENTS. We thank M. Carandini, S. Ganguli, R. Goris, A. Kohn, and P. Latham for helpful discussions and M. Cohen for providing us with Fig. 8 A and D. A.P. was supported by a grant from the Swiss National Science Foundation (31003A_143707) and a grant from the Simons Foundation. R.C.-C. was supported by a fellowship from the Swiss National Science Foundation (PAIBA3-145045).

- Kohn A, Smith MA (2005) Stimulus dependence of neuronal correlation in primary visual cortex of the macaque. *J Neurosci* 25(14):3661–3673.
- Huang X, Lisberger SG (2009) Noise correlations in cortical area MT and their potential impact on trial-by-trial variation in the direction and speed of smooth-pursuit eye movements. *J Neurophysiol* 101:3012–3030.
- Gu Y, et al. (2011) Perceptual learning reduces interneuronal correlations in macaque visual cortex. *Neuron* 71(4):750–761.
- Smith MA, Sommer MA (2013) Spatial and temporal scales of neuronal correlation in visual area V4. *J Neurosci* 33(12):5422–5432.
- Cohen MR, Kohn A (2011) Measuring and interpreting neuronal correlations. *Nat Neurosci* 14(7):811–819.
- Foster KH, Gaska JP, Nagler M, Pollen DA (1985) Spatial and temporal frequency selectivity of neurones in visual cortical areas V1 and V2 of the macaque monkey. *J Physiol* 365(1):331–363.
- Cover TM, Thomas JA (1991) *Elements of Information Theory* (Wiley, New York).
- Schoups A, Vogels R, Qian N, Orban G (2001) Practising orientation identification improves orientation coding in V1 neurons. *Nature* 412(6846):549–553.
- Yang T, Maunsell JHR (2004) The effect of perceptual learning on neuronal responses in monkey visual area V4. *J Neurosci* 24(7):1617–1626.
- Cohen MR, Maunsell JH (2009) Attention improves performance primarily by reducing interneuronal correlations. *Nat Neurosci* 12(12):1594–1600.
- Mitchell JF, Sundberg KA, Reynolds JH (2009) Spatial attention decorrelates intrinsic activity fluctuations in macaque area V4. *Neuron* 63(6):879–888.
- Zhang K, Sejnowski TJ (1999) Neuronal tuning: To sharpen or broaden? *Neural Comput* 11(1):75–84.
- Bethge M, Rotermund D, Pawelzik K (2002) Optimal short-term population coding: When Fisher information fails. *Neural Comput* 14(10):2317–2351.
- Wang Z, Stocker A, Lee D (2013) Optimal neural tuning curves for arbitrary stimulus distributions: Discrimax, infomax and minimum Lp loss. *Adv Neural Inf Process Syst* 25:2177–2185.
- Ganguli D, Simoncelli EP (2014) Efficient sensory encoding and Bayesian inference with heterogeneous neural populations. *Neural Comput* 26(10):2103–2134.
- Shadlen MN, Britten KH, Newsome WT, Movshon JA (1996) A computational analysis of the relationship between neuronal and behavioral responses to visual motion. *J Neurosci* 16(4):1486–1510.
- Bejjanki VR, Beck JM, Lu Z-L, Pouget A (2011) Perceptual learning as improved probabilistic inference in early sensory areas. *Nat Neurosci* 14(5):642–648.
- Green DM, Swets JA (1966) *Signal Detection Theory and Psychophysics* (Wiley, New York).
- Heeger DJ (1992) Normalization of cell responses in cat striate cortex. *Vis Neurosci* 9(2):181–197.
- Ma WJ, Beck JM, Latham PE, Pouget A (2006) Bayesian inference with probabilistic population codes. *Nat Neurosci* 9(11):1432–1438.
- Dosher BA, Lu Z-L (2005) Perceptual learning in clear displays optimizes perceptual expertise: Learning the limiting process. *Proc Natl Acad Sci USA* 102(14):5286–5290.

43. Pachitariu M, Lyamzin DR, Sahani M, Lesica NA (2015) State-dependent population coding in primary auditory cortex. *J Neurosci* 35(5):2058–2073.
44. Lin I-C, Okun M, Carandini M, Harris KD (2015) The nature of shared cortical variability. *Neuron* 87(3):644–656.
45. Rabinowitz N, Goris R, Cohen M, Simoncelli EP (2015) Modulators of V4 population activity under attention. *eLife* 2015:10.7554.
46. O’Kusky J, Colonnier M (1982) A laminar analysis of the number of neurons, glia, and synapses in the adult cortex (area 17) of adult macaque monkeys. *J Comp Neurol* 210(3):278–290.
47. Van Essen DC, Newsome WT, Maunsell JH (1984) The visual field representation in striate cortex of the macaque monkey: Asymmetries, anisotropies, and individual variability. *Vision Res* 24(5):429–448.
48. Malpeli JG, Lee D, Baker FH (1996) Laminar and retinotopic organization of the macaque lateral geniculate nucleus: Magnocellular and parvocellular magnification functions. *J Comp Neurol* 375(3):363–377.
49. Olshausen BA, Field DJ (2004) Sparse coding of sensory inputs. *Curr Opin Neurobiol* 14(4):481–487.
50. Mombaerts P, et al. (1996) Visualizing an olfactory sensory map. *Cell* 87(4):675–686.
51. DeWeese MR, Wehr M, Zador AM (2003) Binary spiking in auditory cortex. *J Neurosci* 23(21):7940–7949.
52. Brecht M, Sakmann B (2002) Dynamic representation of whisker deflection by synaptic potentials in spiny stellate and pyramidal cells in the barrels and septa of layer 4 rat somatosensory cortex. *J Physiol* 543(Pt 1):49–70.
53. Doshier BA, Lu Z-L (1999) Mechanisms of perceptual learning. *Vision Res* 39(19):3197–3221.
54. Cohen MR, Newsome WT (2009) Estimates of the contribution of single neurons to perception depend on timescale and noise correlation. *J Neurosci* 29(20):6635–6648.
55. Britten KH, Shadlen MN, Newsome WT, Movshon JA (1992) The analysis of visual motion: A comparison of neuronal and psychophysical performance. *J Neurosci* 12(12):4745–4765.
56. Uka T, DeAngelis GC (2004) Contribution of area MT to stereoscopic depth perception: Choice-related response modulations reflect task strategy. *Neuron* 42(2):297–310.
57. Nienborg H, Cumming BG (2009) Decision-related activity in sensory neurons reflects more than a neuron’s causal effect. *Nature* 459(7243):89–92.
58. Britten KH, Newsome WT, Shadlen MN, Celebrini S, Movshon JA (1996) A relationship between behavioral choice and the visual responses of neurons in macaque MT. *Vis Neurosci* 13(1):87–100.
59. Pitkow X, Liu S, Angelaki DE, DeAngelis GC, Pouget A (2015) How can single sensory neurons predict behavior? *Neuron* 87(2):411–423.
60. Beck JM, Ma WJ, Pitkow X, Latham PE, Pouget A (2012) Not noisy, just wrong: The role of suboptimal inference in behavioral variability. *Neuron* 74(1):30–39.
61. Haefner RM, Berkes P, Fiser J (2014) Perceptual decision-making as probabilistic inference by neural sampling. *ArXiv:1409.0257* [q-bio.NC].
62. Smolyanskaya A, Haefner RM, Lomber SG, Born RT (2015) A modality-specific feed-forward component of choice-related activity in MT. *Neuron* 87(1):208–219.
63. Ponce-Alvarez A, Thiele A, Albright TD, Stoner GR, Deco G (2013) Stimulus-dependent variability and noise correlations in cortical MT neurons. *Proc Natl Acad Sci USA* 110(32):13162–13167.
64. Renart A, et al. (2010) The asynchronous state in cortical circuits. *Science* 327(5965):587–590.
65. Karklin Y, Simoncelli EP, eds (2011) *Efficient Coding of Natural Images with a Population of Noisy Linear-Nonlinear Neurons* (MIT Press, Cambridge, MA), Vol 24.
66. Klam F, Zemel RS, Pouget A (2008) Population coding with motion energy filters: The impact of correlations. *Neural Comput* 20(1):146–175.
67. Beck J, Bejjanki VR, Pouget A (2011) Insights from a simple expression for linear Fisher information in a recurrently connected population of spiking neurons. *Neural Comput* 23(6):1484–1502.
68. Renart A, van Rossum MC (2012) Transmission of population-coded information. *Neural Comput* 24(2):391–407.
69. DiCarlo JJ, Zoccolan D, Rust NC (2012) How does the brain solve visual object recognition? *Neuron* 73(3):415–434.
70. Haefner R, Bethge M (2010) Evaluating neuronal codes for inference using Fisher information. *Adv Neural Inf Process Syst* 23:1–9.
71. Haefner RM, Gerwinn S, Macke JH, Bethge M (2013) Inferring decoding strategies from choice probabilities in the presence of correlated variability. *Nat Neurosci* 16(2):235–242.