nature
neuroscience

Tuning curve sharpening for orientation selectivity: coding efficiency and the impact of correlations

Peggy Seriès[1], Peter E Latham[1] & Alexandre Pouget[2]

Several studies have shown that the information conveyed by bell-shaped tuning curves increases as their width decreases, leading to the notion that sharpening of tuning curves improves population codes. This notion, however, is based on assumptions that the noise distribution is independent among neurons and independent of the tuning curve width. Here we reexamine these assumptions in networks of spiking neurons by using orientation selectivity as an example. We compare two principal classes of model: one in which the tuning curves are sharpened through cortical lateral interactions, and one in which they are not. We report that sharpening through lateral interactions does not improve population codes but, on the contrary, leads to a severe loss of information. In addition, the sharpening models generate complicated codes that rely extensively on pairwise correlations. Our study generates several experimental predictions that can be used to distinguish between these two classes of model.

Many cortical neurons encode variables in the external world via bell-shaped tuning curves. In this coding scheme, the mean firing rate of a neuron is a gaussian function of some variable such as the orientation of a bar, the speed of a moving object or the frequency of a tone, among others. Of particular importance is determining how much information is contained in a population of such neurons, because this is a prerequisite to understanding how downstream neurons compute efficiently with population activity.

For a single neuron, the information about a variable can be characterized by the discrimination threshold, in other words, by the smallest change in the variable that can be reliably detected on the basis of single trial responses. The inverse of the threshold is proportional to the square root of what is known as Fisher information[1,2]. This information depends on two quantities: the noise in the response and the slope of the tuning curve. Specifically, the information is inversely proportional to the level of noise (the higher the noise, the less information) and directly proportional to the slope of the tuning curve (the larger the slope, the larger the change in the neuronal response for a given change in the encoded variable, and thus the easier the discrimination task).

For bell-shaped tuning curves, the slope increases as the width of the tuning curve decreases, which has led to the notion that sharpening always improves the quality of a code. The idea that 'sharper is better' has been invoked in many different contexts in neuroscience, including orientation selectivity[3], attention[4–6], perceptual learning[7–9] and auditory processing[10]. For example, behavioral improvements in the ability of monkeys to discriminate small changes in orientation have been related to sharpening of orientation tuning curves in V1 (ref. 7).

From the point of view of Fisher information, populations of neurons are very similar to single neurons: sharpening increases information and

noise decreases it, as long as the noise is independent among neurons[2]. Spikes from cortical neurons are, however, not independent[11,12]. In particular, for a given stimulus, the trial-to-trial fluctuations in the spike count of one neuron partially predict the spike count fluctuations of another. In addition, the extent of these 'noise' correlations almost always depends on the width of the tuning curves. This is because the correlations arise primarily through common input, and the level of common input is determined by connections from the thalamus and other cortical areas, and from local intracortical lateral connections. Because sharpening requires changes in at least one of these, it is generally accompanied by changes in correlations. Consequently, sharpening in realistic networks does not guarantee an increase in information.

Here we address this issue in the context of orientation selectivity, as it has been studied widely in this setting experimentally[13–16]. Several orientation selectivity models have been proposed in the past decade[14,17], and these can be grouped into two main classes: sharpening models and no sharpening models. In the no-sharpening model, the selectivity of cortical cells is due primarily to the convergence of lateral geniculate nucleus (LGN) afferences[13,14], and cortical lateral connections are used for amplification and/or contrast invariance. In the sharpening models, the LGN afferences provide broad tuning curves that are subsequently sharpened through cortical interactions mediated by lateral connections[3,16,18]. In the sharpening model, the assumption is that the broadly tuned input from the LGN provides a poor representation of orientation. This problem is fixed, in principle, by sharpening the tuning curves through the cortical interactions in V1. Indeed, with sufficient sharpening the tuning curves can be as narrow as those obtained in the no-sharpening model, at which point one might expect the two models to provide equivalent representations of orientation.

[1]Gatsby Computational Neuroscience Unit, Alexandra House, 17 Queen Square, London WC1N 3AR, UK. [2]Department of Brain and Cognitive Sciences, University of Rochester, Rochester, New York, USA. Correspondence should be addressed to A.P. (alex@bcs.rochester.edu).

Figure 1 Properties of V1 neurons in response to a bar oriented at 90 deg. (**a**) Mean (each point represents an average over 1,008 trials) pooled LGN afferences for the sharpening model (blue) and no-sharpening model (red). Neurons are ranked by their preferred orientation, which was assigned *a priori* to set up the thalamocortical connectivity. The pooled LGN afferences of a neuron is defined as the sum of thalamocortical spikes converging on that neuron. As expected, the orientation tuning of the pooled LGN afferences is broader for the sharpening model. (**b**) Mean population activity of the cortical neurons, colored as in **a**. The slight shift of the population patterns of activity away from 90 deg is not due to noise; it is a systematic shift caused by the heterogeneity of the network, which induces a mismatch between the preferred orientation assigned to each neuron and its actual preferred orientation. The shift, however, is taken into account by our estimator (which effectively relabels neurons) and thus does not lead to biased estimates. (**c**) Circular normal functions fit to the data in **b**. The population patterns are closely matched overall. For the purposes of comparison, the curves have been shifted so that the peaks are at the same place. (**d**,**e**) Contrast sensitivity (fit of the data) for the sharpening (**d**) and no-sharpening (**e**) models. The patterns of activity are roughly invariant with respect to contrast in both models. (**f**) Log of the variance versus the log of the mean spike count over 1 s. Both models show statistics that are consistent with what has been reported for actual V1 neurons: the log of the variance is proportional to the log of the mean, and the constant of proportionality is close to 1.
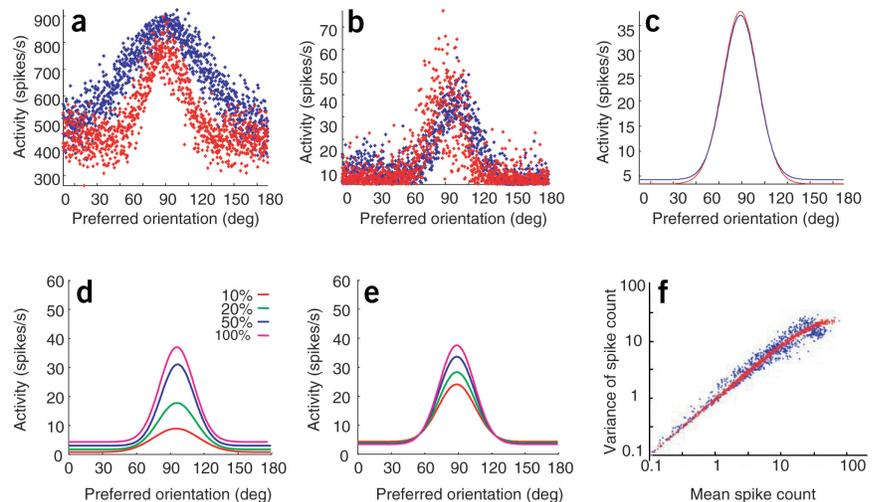
We show here that this is not the case. We have implemented both models in networks of spiking neurons and have estimated the information in the population activity. Our results show that the sharpening model conveys far less information than does the no-sharpening model. Thus, the sharpening model does not lead to an increase in information, but rather to a large loss. This loss can be traced to the pattern of correlations generated by the sharpening process—a finding that illustrates the effect of computation-induced correlations on population codes. In addition, the code in the sharpening model contains correlations that make it particularly inefficient for computing and learning.

Although our analysis focuses on orientation selectivity, the results apply to sharpening in general (for example, in auditory cortex or somatosensory cortex). Notably, the fact that correlations depend on the extent of sharpening leads to two experimental pre-dictions that can be used to establish whether the cortex uses a sharpening or no-sharpening architecture.

RESULTS
Models
The models consist of three stages: retina, LGN and V1. The retinal stage corresponds to grids of ON and OFF ganglion cells modeled by difference-of-gaussian filters. The output of each filter is passed through a saturating nonlinearity and used to drive the LGN cells, which generate Poisson spikes. The output of the LGN cells, which is pooled by using Gabor function receptive fields, is used as input to V1. The V1 stage, which represents a hypercolumn of layer IV simple cells, consists of 1,260 conductance-based integrate-and-fire neurons, 80% of which are excitatory regular spiking cells and 20% of which are inhibitory fast-spiking cells. These cells are coupled through lateral connections.

The sharpening model is very similar to a published model[3] that we refer to here as S^{low}_{MH} because its lateral connectivity follows a Mexican hat (MH) profile and the thalamo-cortical conductance is low (relative to the other models). The no-sharpening network implements a classical Hubel and Wiesel model with feedforward inhibition based on published work[19,20]. We refer to this network as $NS^{high}_{\infty I}$, where 'high' refers to the strength of the thalamocortical connections, 'I' indicates that the only active cortical connections are from inhibitory to excitatory cells, and '∞' indicates that the connections extend throughout the network. For details, see Methods and the **Supplementary Note** online. For definiteness, we initially focused on these two models. To ensure that our results were not specific to any one particular property, however, we considered several variations as described below.
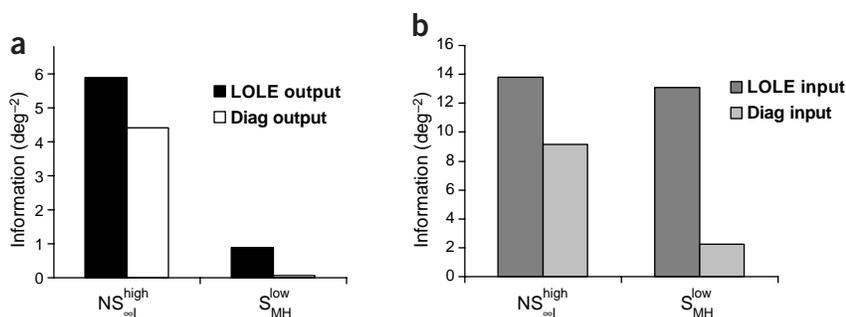

Figure 2 Information comparison across models. (**a**) Information in the cortical neurons recovered by a locally optimal estimator (I_{LOLE}) and by a decoder that ignores correlations (I_{diag}) for the sharpening (S^{low}_{MH}) and no-sharpening ($NS^{high}_{\infty I}$) models. The no-sharpening model recovers six times as much information as the sharpening model, despite identical tuning widths (**Fig. 1c**). The ratio I_{diag}/I_{LOLE} is also very small (7%) for the sharpening network. Because I_{diag} corresponds to the information recovered by a decoder that ignores correlations, the small ratio of I_{diag}/I_{LOLE} for the sharpening model indicates that the code relies heavily on correlations. (**b**) Same as in **a** but for the pooled LGN afferences, that is, the input onto the cortical layer. For both models, the information in the pooled LGN input is about the same, despite very different tuning widths (**Fig. 1a**). Similar to the output code, the input code in the sharpening model relies extensively on correlations, as indicated by the small value of I_{diag}/I_{LOLE}.
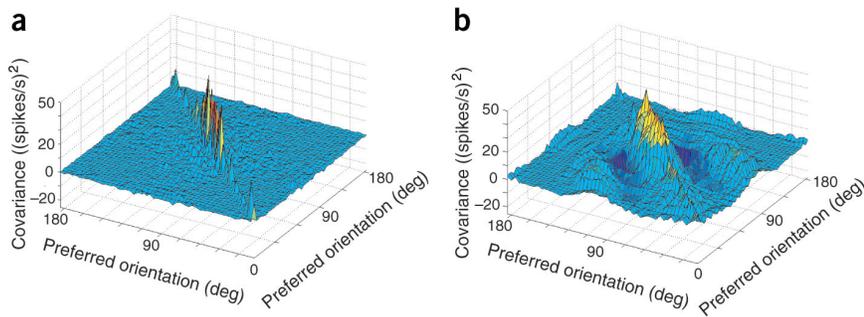
Figure 3 Covariance matrices of the V1 cells in both models. (**a**) In the no-sharpening network, correlations are mostly positive and confined to cells with similar preferred orientations. (**b**) In the sharpening model, correlations tend to be longer range and are both negative and positive.

Single-cell properties

To make a fair comparison between the sharpening and no-sharpening models, it is important that both show the same single-cell properties. The parameters of the models were thus adjusted so that their responses to oriented bars matched each other's, as well as the response properties of V1 neurons *in vivo*, as closely as possible. We found that the input from the LGN was broader for the sharpening model (the half width at half height of the population activity, averaged over 1,008 trials, was 19.3 degrees (deg) for the no-sharpening models as compared with 40.2 deg for the sharpening model; **Fig. 1a**). By contrast, the output patterns—the population activity across the cortical neurons—had nearly identical widths (18.1 deg for the no-sharpening models as compared with 17.8 deg for the sharpening model; **Figs. 1b,c**). Both models also showed approximate contrast invariant tuning curves, with contrast gain curves, similarly to ones reported *in vivo*[21] (**Figs. 1d,e**). The variance of the spike count had a power law dependence on the mean, with an exponent of 0.88 and a multiplier of 0.98 for the no-sharpening model, and values of 0.85 and 1.04, respectively, for the sharpening model (**Fig. 1f**). These values are close to what has been reported *in vivo* and to what is expected for a near-Poisson process[22]. The coefficient of variation (CV) of the interspike interval distribution was also similar across the models (sharpening model, $CV = 0.8 \pm 0.23$; no-sharpening model, $CV = 0.85 \pm 0.17$) and within biological range[23].

Given how close these numbers are, it would essentially be impossible to distinguish between the sharpening and no-sharpening models on the basis of the variance versus mean relationship, the CV of the interspike interval distribution, or the mean activity.

Information from a locally optimal linear estimator

To determine which model is more efficient at encoding orientation, we decoded the responses of each model to bar stimuli and analyzed the amount of information that they provided about the stimulus orientation. The information that we report here is I_{LOLE}, a lower bound on Fisher information obtained by using a locally optimal linear estimator (LOLE; see Methods). We also considered several nonlinear methods and found almost no improvement over I_{LOLE}, indicating that it is a tight bound (see Discussion and Methods).

The no-sharpening model extracted more than six times as much information as the sharpening model ($I_{LOLE} = 5.88$ deg^{-2} versus 0.88 deg^{-2} respectively, corresponding to discrimination thresholds of 0.56 deg versus 1.43 deg, respectively; **Fig. 2a** and see Methods, equation (3)). It might be wondered whether this difference in information is due to a difference in information conveyed by the LGN afferences. Indeed, we found that the orientation tuning of the

pooled LGN afferences (the sum of the thalamocortical afferences onto each cell) is broader for the sharpening network than for the no-sharpening network (**Fig. 1a**). It is therefore conceivable that the input to the sharpening network is less informative than is the input to the no-sharpening network. To address this, we computed the information in the pooled LGN afferences. We found, in fact, that there is little difference in information in the pooled LGN afferences across models (**Fig. 2b**). The tuning curves specified by the LGN afferences were broader in the sharpening models, but the correlational structure was such that there was very little reduction in the information in comparison to the no-sharpening model. This illustrates the importance of taking into account the correlations among neurons when computing information.

Because the tuning curves are matched and there is essentially no difference in information in the LGN input, the discrepancy in information between the output of the two models can be due only to the structure of the second order statistics of the cells; that is, the covariance matrix. Examination of the covariance matrices showed that there are indeed clear differences (**Fig. 3**). In the no-sharpening model, correlations tended to be mostly positive and restricted to pairs of cells with similar orientation preferences. In the sharpening model, there was much more structure to the covariance matrix: both positive and negative correlations were seen, and strong correlations existed even among cells with very different preferred orientations. This is, indeed, the type of covariance matrix that is expected for a sharpening model (see refs 18,24). We next explored in greater detail the impact of these correlations on the structure of the codes.

I_{LOLE} versus $I_{shuffled}$

To relate our work to experimental studies[25], we computed the information in the population patterns of activity generated by the cortical layer of our model after shuffling the data across trials to remove correlations (denoted $I_{shuffled}$; see Methods). $I_{shuffled}$ corresponds to the information that would be measured by a neurophysiologist making several single-cell recordings, because such recordings are
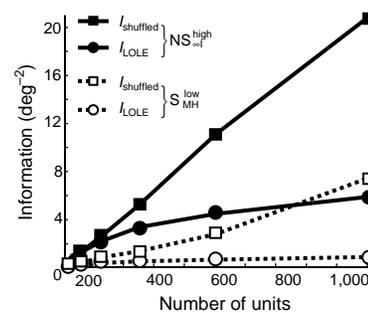

Figure 4 Shuffled information. Comparison of I_{LOLE} and $I_{shuffled}$ as a function of the number of neurons. The neurons were evenly sampled with respect to their preferred orientation. The information estimated from single-cell recordings overestimates the true information in both models. Unbroken lines, no-sharpening model (NS$^{high}_{\infty l}$); broken lines, sharpening model (S$^{low}_{MH}$).

necessarily uncorrelated. A comparison of $I_{shuffled}$ and I_{LOLE} indicates the extent to which single-cell recordings over- or underestimate the true information contained in a population of neurons. In theoretical studies[11,26,27], whether $I_{shuffled}$ is greater or less than I_{LOLE} depends crucially on the structure of the covariance matrix.

We found that $I_{shuffled}$ vastly overestimates I_{LOLE}; in other words, the correlations induced by the computation in both of our models are the kind that reduce the overall information (**Fig. 4**). This result is not unexpected when the number of neurons is large: in the output layer of a multilayer architecture, the information is guaranteed to saturate at or below the information available in the input layer, regardless of the number of neurons in the output layer (this explains why I_{LOLE} saturates with the number of neurons in **Fig. 4**). By contrast, the information for independent neurons (which is what $I_{shuffled}$ computes) is proportional to the number of neurons, even when the input information is fixed and finite. Consequently, $I_{shuffled}$ is guaranteed to be larger than I_{LOLE} for a sufficiently large number of neurons. Notably, we found that it is larger even for a small number of neurons, suggesting that the correlations in our models always decrease the information contained in the population.

Note that for both models, $I_{shuffled}$ is roughly five times larger than I_{LOLE}. This is unexpected for the no-sharpening model because the correlations seem weak in **Figure 3a**. With such weak correlations, shuffling might be expected to have little effect on information. This is clearly not the case: these weak correlations add up to a large effect on the structure of the code. The fact that $I_{shuffled}$ is necessarily larger than I_{LOLE} for a sufficiently large number of neurons might seem to contradict previous studies that found that the relative magnitude of $I_{shuffled}$ versus I_{LOLE} depends on the covariance matrix and, in particular, that $I_{shuffled}$ can be less than I_{LOLE} (refs 26,27). In those studies, however, I_{LOLE} was allowed to grow indefinitely with the number of neurons. This cannot happen in the networks that we are considering here, or in realistic networks in general, because I_{LOLE} cannot be greater than the information provided by the LGN, regardless of the number of cortical neurons.

Diagonal approximation

In the cortex, orientation-selective neurons project to downstream neurons involved in tasks such as shape analysis and sensorimotor transformations. The computations associated with these tasks are efficient only if the downstream neurons have some knowledge of the statistics of the activity generated by the orientation-selective neurons. Given the strong impact of correlations on the information of population codes, as measured by comparing I_{LOLE} to $I_{shuffled}$, it is tempting to conclude that correlations must be known by the downstream neurons if these computations are to be carried out efficiently. In fact, this is not necessarily true. It is possible for codes to have high correlations and still be decoded optimally without knowledge of those correlations[28,29]. This is a crucial issue, because it affects the kind of computations that must be performed by downstream neurons; in particular, it affects how complex those computations must be.

To determine whether knowledge of correlations is necessary to decode the V1 responses in our networks, we computed the information when ignoring correlations, a quantity that we call I_{diag}. This quantity is less than or equal to I_{LOLE}, because it corresponds to the information recovered by an LOLE trained on the shuffled data but tested on the actual data (as opposed to I_{LOLE}, which is obtained by training and testing an LOLE on the actual data; see Methods).

We found that for the no-sharpening model I_{diag} was 75% of I_{LOLE}, whereas it was 7% for the sharpening model (**Fig. 2a**). In
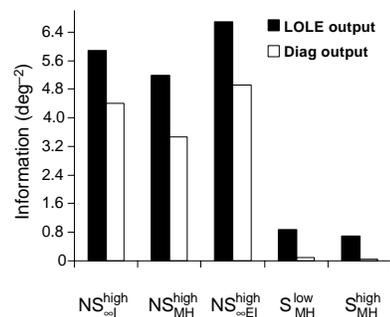

Figure 5 Information (I_{LOLE} and I_{diag}) in the output layer of additional models. $NS^{high}_{\infty I}$ corresponds to the model shown in **Fig. 2**; that is, a no-sharpening model with long-range inhibition and no cortical excitation. NS^{high}_{MH} is a no-sharpening model with short-range excitation and long-range inhibition. $NS^{high}_{\infty EI}$ is a no-sharpening model with long-range excitation and long-range inhibition. S^{low}_{MH} is the same sharpening model as in **Fig. 2**. S^{high}_{MH} is very similar to NS^{high}_{MH} but receives a widely tuned input from the LGN. Regardless of the details of the cortical connectivity, the no-sharpening models outperform the sharpening models by a minimum factor of five.

other words, 93% of the information is embedded in the pairwise correlations in the sharpening model as compared with 25% in the no-sharpening model. The code generated by the no-sharpening model therefore conveys more information than does the sharpening model, and it does not rely heavily on correlations. The fact that correlations can be largely ignored implies that the no-sharpening code is efficient for learning, because algorithms based on variational approximations in which correlations are ignored—a common approach to intractable learning problems—are closer to being exact[30]. In addition, we found that learning through gradient descent was slower when the ratio I_{diag}/I_{LOLE} was small: it took more than six times as long to train the LOLE on the output of the sharpening model than on the output of the no-sharpening model. This slower learning is related to the fact that, for the sharpening model, pairwise correlations must be known to recover more than 7% of the information available; this involves estimating more than 500,000 parameters (for 1,008 neurons).

Codes in which correlations can be ignored are also easier to use for computations. For example, it is possible to design biologically plausible network architectures that can perform optimal Bayesian inferences when the code is independent, or can be treated as such (such as when I_{diag} is close to I_{LOLE}), whereas no general biologically plausible networks exist when the codes rely extensively on correlations[31].

It might be argued that downstream neurons are not involved in estimating orientation and, consequently, may not care about the accuracy with which orientation can be estimated. The problem, however, is that Fisher information provides a bound on estimating not only orientation but also any function of orientation. Therefore, regardless of what downstream neurons do with orientation, a code with little information and a complicated format imposes severe limitations on the accuracy with which downstream computations are made.

We also computed I_{diag} in the pooled LGN afferences and found the same result: much of the information (66%) was encoded in the correlations for the sharpening model, but less (17%) was encoded in the correlations for the no-sharpening model (**Fig. 2b**). Thus, the problem with I_{diag} is, to a large extent, inherited from the thalamocortical afferences.

Parametric study

To determine whether our results were robust to network parameters, we varied the thalamocortical and intracortical conductances one at a time by ±10% (see Methods) and recomputed the information. We found that the no-sharpening network always outperformed the sharpening network, even though the tuning curve amplitudes and response variabilities were no longer matched. This was true even when we compared the best sharpening network across all simulations to the worst no-sharpening network: here, the sharpening network recovered only 22% of the information of the no-sharpening network.

Given the differences in contrast sensitivity, we considered whether our results depended on contrast. We found this not to be the case: whether we compared the networks at 5% or 100% contrast, the no-sharpening network was still better than the sharpening network ($I_{LOLE} = 1.35$ deg^{-2} versus 0.06 deg^{-2} at 5% contrast).

We also explored whether our results were due to differences in the relative strength of the thalamocortical afferences or in the pattern of the cortical connections. Indeed, both parameters varied across networks: the conductances of the thalamocortical afferences were larger in the no-sharpening model, and the no-sharpening model contained only inhibitory cortical connections. To address this issue, we implemented new sharpening and new no-sharpening models with similar thalamocortical conductances and cortical connectivity (NS$^{high}_{MH}$ and S$^{high}_{MH}$ in **Fig. 5**; see **Supplementary Note**). In the NS$^{high}_{MH}$ model, the width of the input from the LGN was 19.4 deg as compared with 16.2 deg for the cortical output. In the S$^{high}_{MH}$ model, these values were 40.12 deg and 18.36 deg, respectively. As above, the models were tuned to generate near-Poisson spike trains. The results were essentially the same as those from the original networks. Because NS$^{high}_{MH}$ and S$^{high}_{MH}$ share nearly identical connectivity, the poor performance of the sharpening model is unlikely to be due to its connectivity (**Fig. 5**). Rather, it would seem that the sharpening model is unable to process the widely tuned input that it receives from the LGN efficiently, a point that we discuss below.

To test further the influence of the cortical connectivity, we also implemented a no-sharpening model with global excitatory and global inhibitory connections (connections independent of the difference in preferred orientation), as opposed to the 'Mexican hat' connectivity (short-range excitation and long-range inhibition) used for NS$^{high}_{MH}$ and S$^{high}_{MH}$. This manipulation had no effect on our results (NS$^{high}_{\infty EI}$ in **Fig. 5**).

Last, we tested the robustness of our results to changes in the correlational structure of the inputs to V1. We constructed sharpening and no-sharpening models in which all cortical cells received independent inputs from the LGN, as opposed to the above studies in which any pair of neighboring V1 cells shared significant common inputs. We found that even in this extreme situation, the no-sharpening model outperformed the sharpening model by a factor of 11 ($I_{LOLE} = 10.51$ deg^{-2} versus 0.93 deg^{-2}).

DISCUSSION

Our results challenge two commonly held notions about orientation selectivity. The first is that narrow tuning curves convey more information than do wide ones. We have shown that this is not necessarily true: in our model, the information in the pooled LGN afferences was similar whether the tuning was narrow or wide (**Fig. 2b**). The second is that one can improve the quality of a code in a broadly tuned population by transmitting it to a second population and then sharpening it. We found the opposite: that is, that sharpening results in a loss of information.

This second result implies that sharpening is not necessarily beneficial—a fact that has implications well beyond models of orientation selectivity. Indeed, many models rely on the idea that sharpening is necessarily better. For example, it has been proposed that attention improves discrimination by sharpening tuning curves in early visual areas[5]. A similar argument has been used by to explain how discrimination improves after extensive training[8]. These models assume, however, that the noise distribution is independent of the tuning curve width. As we have shown in our simulations, and as we argued initially, this assumption is very unlikely to hold with a realistic noise model. As a result, sharpening does not have to be better. In fact, because sharpening through lateral connections is a form of postprocessing, it is guaranteed to lead to an information loss or at best no change (because of the data processing inequality[32]). The unexpected finding is the very large information loss over a wide range of sharpening networks.

In our models, the information loss is related to the pattern of correlations introduced by the sharpening process (**Fig. 3b**). It should not be concluded, however, that the information in a population code is low whenever correlations follow the overall pattern shown in **Figure 3b**. In fact, correlations of this type can be associated with optimal information transmission[24]. Therefore, it is the fine structure of the correlations in **Figure 3b** that are responsible for the information loss, not the overall profile. Several studies have already identified some of the factors that are essential for determining the amount of information in a population of neurons[27,33], but more work is needed in this area.

We emphasize that we are not arguing against sharpening in general; our results only concern sharpening with lateral connections within a hypercolumn, a feature that is commonly used in models of orientation selectivity[3]. There might be other forms of sharpening that are not subject to the problems that we report here. For example, tuning curves can sharpen as a result of training on an orientation discrimination task[7,9]. In this case, the sharpening might be due to changes in the feedforward connectivity and could indeed result in higher information transmission. Sharpening may also help when the lateral connections connect cells processing nearby spatial locations (in our model, all units see the same patch of the LGN and therefore the same spatial location). Notably, some data indicate that when sharpening occurs, it involves such long-range interactions (Xing, D., Shapley, R., Hawken, M. & Ringach, D., *Soc. Neurosci. Abstr.* 910.4, 2003). In this case, it is possible that sharpening is not crucial to orientation selectivity but might be involved in other tasks, such as decorrelation across space[34] or contour completion[35–38]. Last, sharpening may be beneficial if significant noise is injected during processing by downstream areas, although whether or not this is so depends crucially on exactly how the noise is injected.

Here we have used the word 'information' to refer to a lower bound on Fisher information: specifically, the information recovered by an LOLE. If this bound is not tight, the differences that we have reported across models may not reflect differences in actual information. In particular, it is conceivable that the sharpening model has information beyond what the LOLE recovers[33]. However, we do not think that this is the case: we have tried several nonlinear methods and have found at most a 9% increase in information (see Methods), suggesting that our lower bound on Fisher information is in fact very close to the true value.

In addition to our theoretical findings, our work makes two experimental predictions. First, pairwise correlations differ across models (**Fig. 3**). This prediction is consistent with what had been suggested by simpler models based on analog units[18,24] and stochastic spiking neu-

rons[39]. Second, in the sharpening model a large fraction of the information is in the correlations, whereas in the no-sharpening model only a small fraction is in correlations. Both of these predictions can be tested with multielectrode recordings. To our knowledge, current data does not seem to support the sharpening model. For example, correlations tend to be positive for units with similar stimulus preferences (ref. 11 and Kohn, A., Smith, M.A., Bair, W. & Movshon, J.A., *Soc. Neurosci. Abstr.* 557.14, 2002) and close to zero otherwise (which is consistent with the covariance matrix of the non-sharpening model; **Fig. 3**). In addition, I_{diag} is typically about 80–90% of the total information in the retina and cortex[40–43]. Those studies, however, often involved different information measures and were mostly obtained outside V1. Therefore, it is too early to tell whether the experimental data favors one model over the other.

Why should sharpening reduce information transmission? Our previous work with networks of analog units suggests an explanation. We found that lateral connections in a recurrent network can be used to embed expectations about the statistics of the input patterns[44]. The network is most efficient when the input statistics match those expectations. The same mechanism seems to be at work here; in other words, the wide input to the sharpening model is simply suboptimal. If this is correct, driving a sharpening model with a sharp input (as sharp as the one used for the no-sharpening model) should improve performance to a level comparable to the performance of the no-sharpening model. This is indeed what we found: NS^{high}_{MH} in **Figure 5** is essentially the second sharpening model (S^{high}_{MH}) driven by a narrowly tuned input, and its performance is very close to that of the original no-sharpening network.

METHODS

Orientation models. The models simulate the circuits involved in one hypercolumn of primary visual cortex. They consist of three stages: retina, LGN and V1. The retinal and LGN stages are identical in the sharpening and no-sharpening models. These stages closely follow those described in ref. 3. In brief, the retinal stage corresponds to grids of ON and OFF ganglion cells modeled by difference-of-gaussian filters. The activity of the filters is passed through a saturating nonlinearity to account for stimulus contrast sensitivity and is used at the LGN level to generate spikes by a Poisson process.

The V1 stage represents a hypercolumn of layer IV simple cells. It comprises 1,260 conductance-based integrate-and-fire neurons, 80% of which are excitatory regular spiking cells and 20% of which are inhibitory fast-spiking cells. The organization of the thalamo-cortical afferents controls the initial receptive field structure and initial orientation selectivity[13]. This initial receptive field structure is established by segregation of ON and OFF LGN afferences into ON and OFF subfields, and is modeled using a Gabor function.

In the sharpening model, S^{low}_{MH}, the parameters of the Gabor function are such that the LGN afferences are weakly anisotropic and broadly tuned to orientation (**Fig. 1a**). The superscript 'low' refers to the strength of the thalamo-cortical conductance. The subscript 'MH' indicates that the lateral connections follow a Mexican hat profile, in which excitatory interconnections are strongest among cortical cells with similar orientation preference, and inhibitory projections can target cells with more wide-ranging orientation preferences[3,17]. The effect of these connections is to amplify and to sharpen the input tuning curves.

In the no-sharpening model, $NS^{high}_{\infty I}$, the parameters of the Gabor function are such that the pooled LGN afferences are more selective to orientation (**Fig. 1a**). The subscript '∞' indicates that the cortical connections extend throughout the network, regardless of the difference in preferred orientation, and the subscript 'I' indicates that only the inhibitory to excitatory connections are active. The inhibition is assumed to be nonspecific with respect to the preferred phase of the neurons[16]. This model thus implements a pure 'feedforward inhibition' that counteracts LGN afferences at non-preferred orientations and produces contrast invariant orientation tuning curves[14,19].

We implemented three other versions of this model, denoted S^{high}_{MH}, NS^{high}_{MH} and $NS^{high}_{\infty EI}$. The superscripts and subscripts have the meaning

described above; the subscript 'EI' indicates that both excitatory and inhibitory interconnections are active (see **Supplementary Note** online for more details).

Oriented bar stimuli were presented to the retinal cells for 500 ms. The models were constrained so that the responses showed Poisson-like variability. Variability was quantified in two ways: by plotting the variance of the spike count over 500 ms versus its mean, and computing the slope and intercept; and by computing the CV, that is, the ratio of the standard deviation of the interspike interval to its mean. The CVs were computed for all neurons simultaneously by presenting a bar at an orientation of 90 deg for 10 s and by estimating the interspike interval histogram. Neurons that fired fewer than two spikes were removed from this analysis. Details of the implementation are given in the **Supplementary Note**.

I_{LOLE}. To estimate the information contained in the V1 population activity, we trained an LOLE of orientation for each model. The LOLE was trained on the spike counts of the 1,008 excitatory neurons in response to an oriented bar presented for 500 ms (inhibitory neurons were not decoded because they do not project out of V1). This estimator has the form

$$\hat{\theta} = \mathbf{w}\mathbf{r} + b \tag{1}$$

where $\hat{\theta}$ is the estimate of orientation, $\mathbf{w}$ and b are adjustable parameters, and $\mathbf{r}$ is a vector containing the spike counts of all neurons on a given trial. We optimized the weight $\mathbf{w}$ and the constant b to estimate two values of orientation differing by a small angle: 89.5 deg and 90.5 deg. A LOLE is equivalent to a two layer network in which the input layer encodes the spike counts of all neurons, $\mathbf{r}$, and the output layer consists of one unit trained to estimate the orientation encoded by $\mathbf{r}$. We did not test whether the timing of the spikes conveyed additional information.

For each model, 1,008 population patterns of activity were generated in response to a bar oriented at 89.5 deg, and another 1,008 trials were done for a bar oriented at 90.5 deg. Half of those trials (504 at 89.5 deg and 504 at 90.5 deg) were used to train the weights of a LOLE; training was done by using a gradient descent. Generalization performance was obtained by testing the weights on the 1,008 trials not used for training (504 for each angle). To prevent overfitting, we stopped the gradient descent when the generalization performance started decreasing.

Once training was completed, we computed the mean and variance of the estimates for both orientations ($\theta_1 = 89.5$ deg and $\theta_2 = 90.5$ deg) on the 504 test trials, denoted

$$\left\{ \langle \hat{\theta}_i \rangle, \sigma_{\hat{\theta}_i} \right\}_{i=1,2}$$

to obtain an estimate of Fisher information,

$$I_{LOLE} = \frac{\left(\left(\langle \hat{\theta}_2 \rangle - \langle \hat{\theta}_1 \rangle \right) / \delta\theta \right)^2}{\left(\sigma^2_{\hat{\theta}_2} + \sigma^2_{\hat{\theta}_1} \right)^{1/2}} \tag{2}$$

where $\delta\theta$ corresponds to the difference between the two angles (1 deg). I_{LOLE} is a lower bound on the Fisher information (see below). From equation 2, one can compute an upper bound on the discrimination threshold of an ideal observer—that is, the change in orientation, $\Delta\theta$, that can be detected 75% of the time[45]. This upper bound is

$$\Delta\theta = \frac{1.35}{\sqrt{I_{LOLE}}} \tag{3}$$

Instead of using a lower bound, we could have estimated Fisher information directly. For neuronal responses characterized by tuning curves $\mathbf{f}(\theta)$ and corrupted by multivariate gaussian noise with covariance matrix $\mathbf{Q}(\theta)$ (a good approximation for our data), Fisher information is given by

$$I(\theta) = \mathbf{f}'(\theta)^T \mathbf{Q}^{-1}(\theta) \mathbf{f}'(\theta) + \frac{1}{2} \text{Tr}\left[\mathbf{Q}'(\theta) \mathbf{Q}^{-1}(\theta) \mathbf{Q}'(\theta) \mathbf{Q}^{-1}(\theta) \right] \tag{4}$$

where $\mathbf{f}'(\theta)$ and $\mathbf{Q}'(\theta)$ are the derivatives of the tuning curves and covariance matrix, respectively, with respect to θ (ref. 27). Estimating this expression directly is difficult because it requires estimates of $\mathbf{Q}(\theta)$ and $\mathbf{Q}'(\theta)$, both of which require a very large amount of data. Using a LOLE provides a way to estimate the first term without estimating $\mathbf{Q}(\theta)$ explicitly[31]. We have shown previously that this term can be also recovered by a biologically plausible architecture known as a line attractor network[31,44]. Occasionally the trace term in equation (4) can be large, in which case our lower bound would be different from the actual information. This does not seem to be the case for our models. We have estimated Fisher information with several nonlinear methods, including K nearest neighbors[46], backpropagation[47], support vector machine[48] and another published method[33]. Support vector machine was the only method that outperformed LOLE, resulting in, at most, an improvement of 9%.

$I_{\mathbf{shuffled}}$. We denote I_{shuffled} as the information available in an artificial data set in which the activity of the units was shuffled over trials to remove all correlations among cells. This shuffling operation is analogous to making single-cell recordings and generating artificial population patterns of activity by grouping the activity of the different cells collected under the same stimulus conditions. To compute I_{shuffled}, we trained and tested a LOLE on the shuffled data set using the same cross-validation technique described above.

$I_{\mathbf{diag}}$. We denote I_{diag} as the information recovered from the actual population patterns of activity when correlations are ignored (diag stands for the diagonal of the covariance matrix). We computed I_{diag} by first training a LOLE on the shuffled data and then using the weights on the original (nonshuffled) data.

Note that I_{diag} cannot be greater than I_{LOLE}. Indeed, the weights used to compute I_{LOLE} were designed to be optimal for the original data, whereas the weights used to compute I_{diag} were obtained from the shuffled data. The latter weights were thus suboptimal (or at best the same as the I_{LOLE} weights); as a result, the ratio $I_{\text{diag}}/I_{\text{LOLE}}$, which is the fraction of the total information that can be recovered when correlations are ignored, is always less than or equal to 1.

Note: Supplementary information is available on the Nature Neuroscience website.

1. Paradiso, M. A theory of the use of visual orientation information which exploits the columnar structure of striate cortex. *Biol. Cybern.* **58**, 35–49 (1988).
2. Seung, H. & Sompolinsky, H. Simple model for reading neuronal population codes. *Proc. Natl. Acad. Sci. USA* **90**, 10749–10753 (1993).
3. Somers, D.C., Nelson, S.B. & Sur, M. An emergent model of orientation selectivity in cat visual cortical simple cells. *J. Neurosci.* **15**, 5448–5465 (1995).
4. Spitzer, H., Desimone, R. & Moran, J. Increased attention enhances both behavioral and neuronal performance. *Science* **240**, 338–340 (1988).
5. Lee, D.K., Itti, L., Koch, C. & Braun, J. Attention activates winner-take-all competition among visual filters. *Nat. Neurosci.* **2**, 375–381 (1999).
6. Murray, S.O. & Wojciulik, E. Attention increases neural selectivity in the human lateral occipital complex. *Nat. Neurosci.* **7**, 70–74 (2004).
7. Schoups, A., Vogels, R., Qian, N. & Orban, G. Practising orientation identification improves orientation coding in V1 neurons. *Nature* **412**, 549–553 (2001).
8. Teich, A.F. & Qian, N. Learning and adaptation in a recurrent model of V1 orientation selectivity. *J. Neurophysiol.* **89**, 2086–2100 (2003).
9. Yang, T. & Maunsell, J.H. The effect of perceptual learning on neuronal responses in monkey visual area V4. *J. Neurosci.* **24**, 1617–1626 (2004).
10. Fitzpatrick, D.C., Batra, R., Stanford, T.R. & Kuwada, S. A neuronal population code for sound localization. *Nature* **388**, 871–874 (1997).
11. Zohary, E., Shadlen, M. & Newsome, W. Correlated neuronal discharge rate and its implication for psychophysical performance. *Nature* **370**, 140–143 (1994).
12. 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**, 1161–1170 (1998).
13. Hubel, D. & Wiesel, T. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J. Physiol. (Lond.)* **160**, 106–154 (1962).
14. Ferster, D. & Miller, K.D. Neural mechanisms of orientation selectivity in the visual cortex. *Annu. Rev. Neurosci.* **23**, 441–471 (2000).
15. Monier, C., Chavane, F., Baudot, P., Graham, L.J. & Frégnac, Y. Orientation and direction selectivity of synaptic inputs in visual cortical neurons: a diversity of combinations produces spike tuning. *Neuron* **37**, 663–680 (2003).
16. Shapley, R., Hawken, M. & Ringach, D.L. Dynamics of orientation selectivity in the primary visual cortex. *Neuron* **38**, 689–699 (2003).
17. Sompolinsky, H. & Shapley, R. New perspectives on the mechanisms for orientation selectivity. *Curr. Opin. Neurobiol.* **7**, 514–522 (1997).
18. Ben-Yishai, R., Bar-Or, R.L. & Sompolinsky, H. Theory of orientation tuning in visual cortex. *Proc. Natl. Acad. Sci. USA* **92**, 3844–3848 (1995).
19. Troyer, T.W., Krukowski, A.E., Priebe, N.J. & Miller, K.D. Contrast-invariant orientation tuning in cat visual cortex: thalamocortical input tuning and correlation-based intracortical connectivity. *J. Neurosci.* **18**, 5908–5927 (1998).
20. McLaughlin, D., Shapley, R., Shelley, M. & Wielaard, D.J. A neuronal network model of macaque primary visual cortex (V1): orientation selectivity and dynamics in the input layer 4Cα. *Proc. Natl. Acad. Sci. USA* **97**, 8087–8092 (2000).
21. Sclar, G. & Freeman, R. Orientation selectivity in the cat's striate cortex is invariant with stimulus contrast. *Exp. Brain Res.* **46**, 457–461 (1982).
22. Gershon, E.D., Wiener, M.C., Latham, P.E. & Richmond, B.J. Coding strategies in monkey V1 and inferior temporal cortices. *J. Neurophysiol.* **79**, 1135–1144 (1998).
23. Softky, W.R. & Koch, C. The highly irregular firing of cortical cells is inconsistent with temporal integration of random EPSPs. *J. Neurosci.* **13**, 334–350 (1993).
24. Pouget, A., Zhang, K., Deneve, S. & Latham, P.E. Statistically efficient estimation using population codes. *Neural Comput.* **10**, 373–401 (1998).
25. Newsome, W.T., Britten, K.H. & Movshon, J.A. Neuronal correlates of a perceptual decision. *Nature* **341**, 52–54 (1989).
26. Yoon, H. & Sompolinsky, H. The effect of correlations on the Fisher Information of Populations codes. in *Advances in Neural Information Processing Systems* (eds. Kearns, M. S., Solla, S. & Cohn, D. A.) 167–173 (MIT Press, Cambridge, Massachusetts, USA, 1999).
27. Abbott, L. & Dayan, P. The effect of correlated variability on the accuracy of a population code. *Neural Comput.* **11**, 91–101 (1999).
28. Wu, S., Nakahara, H. & Amari, S. Population coding with correlation and an unfaithful model. *Neural Comput.* **13**, 775–797 (2001).
29. Nirenberg, S. & Latham, P.E. Decoding neuronal spike trains: how important are correlations? *Proc. Natl. Acad. Sci. USA* **100**, 7348–7353 (2003).
30. Jordan, M.I., Ghahramani, Z., Jaakkola, T.S. & Saul, L.K. An introduction to variational methods for graphical models. *Mach. Learn.* **37**, 183–233 (1999).
31. Latham, P.E., Deneve, S. & Pouget, A. Optimal computation with attractor networks. *J. Physiol. (Paris)* **97**, 683–694 (2003).
32. Cover, T.M. & Thomas, J.A. *Elements of Information Theory* (John Wiley & Sons, New York, 1991).
33. Shamir, M. & Sompolinsky, H. Nonlinear population codes. *Neural Comput.* **16**, 1105–1136 (2004).
34. Simoncelli, E.P. & Olshausen, B.A. Natural image statistics and neural representation. *Annu. Rev. Neurosci.* **24**, 1193–1216 (2001).
35. Zucker, S.W., Dobbins, A. & Iverson, L. Two stages of curve detection suggest two styles of visual computation. *Neural Comput.* **1**, 68–81 (1989).
36. Li, Z. A neural model of contour integration in the primary visual cortex. *Neural Comput.* **10**, 903–940 (1998).
37. Hess, R. & Field, D. Integration of contours: new insights. *Trends Cogn. Sci.* **3**, 480–486 (1999).
38. Seriès, P., Lorenceau, J. & Frégnac, Y. The silent surround of V1 receptive fields: theory and experiments. *J. Physiol. (Paris)* **97**, 453–474 (2004).
39. Spiridon, M. & Gerstner, W. Effect of lateral connections on the accuracy of the population code for a network of spiking neurons. *Network* **12**, 409–421 (2001).
40. Rolls, E.T., Franco, L., Aggelopoulos, N.C. & Reece, S. An information theoretic approach to the contributions of the firing rates and the correlations between the firing of neurons. *J. Neurophysiol.* **89**, 2810–2822 (2003).
41. Nirenberg, S., Carcieri, S.M., Jacobs, A.L. & Latham, P.E. Retinal ganglion cells act largely as independent encoders. *Nature* **411**, 698–701 (2001).
42. Averbeck, B.B. & Lee, D. Neural noise and movement-related codes in the macaque supplementary motor area. *J. Neurosci.* **23**, 7630–7641 (2003).
43. Panzeri, S., Petroni, F., Petersen, R.S. & Diamond, M.E. Decoding neuronal population activity in rat somatosensory cortex: role of columnar organization. *Cereb. Cortex* **13**, 45–52 (2003).
44. Deneve, S., Latham, P. & Pouget, A. Efficient computation and cue integration with noisy population codes. *Nat. Neurosci.* **4**, 826–831 (2001).
45. Green, D.M. & Swets, J.A. *Signal Detection Theory and Psychophysics* (John Wiley & Sons, Los Altos, California, USA, 1966).
46. Duda, R.O. & Hart, P.E. *Pattern Classification and Scene Analysis* (John Wiley & Sons, New York, 1973).
47. Rumelhart, D., Hinton, G. & Williams, R. Learning internal representation by error propagation. in *Parallel Distributed Processing* (eds. Rumelhart, D., McClelland, J. & Group, P.R.) 318–362 (MIT Press, Cambridge, Massachusetts, USA, 1986).
48. Cortes, C. & Vapnik, V. Support-vector networks. *Mach. Learn.* **20**, 273–297 (1995).