Supplementary Materials

General case: computing log likelihood

We first describe the general case of computing the log likelihood of a sensory parameter θ that is encoded by the activity of N neurons. Each neuron's tuning function is described by $\Phi_i(\theta,S)$ where S represents the stimulus strength (e.g. contrast for gratings or coherence for random–dot motion) and fires n_i spikes in response to the stimulus. The average number of spikes elicited is determined by the neuron's mean firing rate (from the tuning function) multiplied by the stimulation time, t, and is subject to Poisson noise (equation (1)). Neurons are assumed to be statistically independent (equation (2)). Equation (3) describes the form of the log likelihood:

$$p(n_i \mid \theta) = \frac{\left(\Phi_i(\theta,S).t\right)^{n_i}}{n_i!} e^{-\Phi_i(\theta,S)t} \tag{1}$$

$$p(n_1, n_2, n_3, \ldots, n_N \mid \theta) = \prod_{i=1}^{N} p(n_i \mid \theta) \tag{2}$$

$$\log L(\theta) = \sum_{i=1}^{N} n_i \log \Phi_i(\theta,S) - t \sum_{i=1}^{N} \Phi_i(\theta,S) - \sum_{i=1}^{N} \log(n_i!) + \log(t) \sum_{i=1}^{N} n_i \tag{3}$$

The last two terms on the right hand side (RHS) are clearly independent of θ and can be dropped. The second term on the RHS is a sum over all tuning curves for a given stimulus. For a homogeneous sensory representation, this term will generally sum to a constant and would therefore be independent of θ. Dropping these terms:

$$\log L(\theta) = \sum_{i=1}^{N} n_i \log \Phi_i(\theta,S)$$

If the stimulus strength only scales the profile of the tuning curves and does not change their shape, $\Phi_i(\theta,S)$ can be written as a product of two functions $f_i(\theta)$ and $g(S)$ and the log likelihood (after removing terms that are independent of θ) becomes:

$$\log L(\theta) = \sum_{i=1}^{N} n_i \log f_i(\theta)$$

To arrive at this solution we made four assumptions:

1. *Homogeneous sensory representation*: This is a common property for many sensory

parameters (e.g. orientation and direction of motion). The uniformity of the representation is also expected for an unbiased cortical representation of sensory parameters. If this requirement were not satisfied the cortical representation would become biased. Log L would reflect this bias. Note that while population profiles need to be self–similar, the individual tuning curves can differ.

2. *Signal strength–invariant tuning functions*: Contrast–independent orientation tuning for grating stimuli in the primary visual cortex (V1) is perhaps the most extensively studied instance of this widespread property[1]. Similarly, direction tuning in area MT is invariant with respect to the coherence of random–dot stimuli[2]. Our model highlights the importance of this property for computing likelihoods. Note that individual tuning curves are not required to have a particular shape so long as assumptions 1 and 2 are satisfied. Examples of the many kinds of tuning curves that could satisfy these requirements include: (1) Gaussian and circular Gaussian tuning curves; (2) a family of low–pass, band–pass and high–pass tuning curves; (3) a family of monotonically increasing and decreasing sigmoidal tuning curves.

3. *Statistical independence*: The key assumption in our model is that the downstream output neurons do not take account of the structure of the interneuronal correlations between the sensory responses. This assumption is based on three factors: (1) interneuronal correlations that are stimulus–dependent[3] cannot be straightforwardly taken into account without prior knowledge about either the stimulus or the full structure of the correlations; (2) the relatively constant variability observed across multiple stages of cortical hierarchy has may arise from the propagation – and not removal – of these correlations[4]; (3) the general structure of these correlations is not known. Available data on pairwise correlations in area MT suggest that neurons preferring similar directions are more correlated than neurons with more remote preferences[5]. The fact that the model ignores correlations does not excuse us from taking them into account in generating predictions. In the next section we describe the model of interneuronal correlations that we used to compute the log likelihood of direction of motion.

4. *Firing rate statistics*: Poisson statistics can describe the variable firing pattern of cortical neurons reasonably well and may be the most widely used model of cortical firing. However, this is not a crucial assumption. The key requirement for our model is to have a log likelihood that is linear with respect to n_i. The likelihood function can include any other term that is independent of the parameter we want to estimate. In particular, considering the family of exponential distribution functions of the form:

$$p(n_i \mid \theta) = \exp\left\{ \frac{f_i(\theta)n_i - b\left[f_i(\theta)\right]}{a(\phi)} + c(n_i, \phi) \right\}$$

When ϕ is treated as a nuisance parameter, it is evident that our formulation holds for all members of the exponential family for which $\sum c(n_i, \phi)$ is independent of n_i.

Example case: decoding motion from MT responses

The encoding model of direction tuned neurons in area MT consists of N similar neurons with uniformly distributed preferred directions, θ_i. The tuning functions of the neurons are self–similar and have a circular Gaussian (von Mises) profile, $f_i(\theta)$ with concentration parameter κ. The mean firing rate of neurons in the model approximates the response of MT neurons to random–dot stimuli where the percentage of coherently moving dots (coherence, C) measures the signal strength. R_{max} is the maximum mean firing rate of the cells in response to a fully coherent stimulus ($C=1$), and n_i is the number of spikes elicited in response to the stimulus and is described by Poisson statistics with a mean that is derived from the coherence of the stimulus, the tuning of the cell, $f_i(\theta)$ and the stimulation time, t. The neurons are assumed to have pairwise correlations with correlation coefficient ρ_{ij}. Neurons preferring the same direction are maximally correlated (correlation coefficient ρ_{max}) and the strength of pairwise correlations monotonically decreases as the preferred directions of the neurons become more different; a circular Gaussian form with concentration parameter δ describes the profile of pairwise correlations (**Fig. S1**).

The elements of the model are:

$$f_i(\theta) = \exp\left\{ \kappa\left[\cos(\theta - \theta_i) - 1\right] \right\}$$

$$p(n_i \mid \theta) = \frac{\left(CR_{max}tf_i(\theta)\right)^{n_i}}{n_i!} e^{-CR_{max}tf_i(\theta)}$$

$$\rho_{ij} = \rho_{max} \exp\left\{ \delta\left[\cos(\theta_i - \theta_j) - 1\right] \right\}$$

The log likelihood function, its mean, μ_L, and variance, σ_L^2 (up to a constant) can be straightforwardly derived:

$$\log L(\theta) = \kappa \sum_{i=1}^{N} n_i \cos(\theta - \theta_i)$$

$$\mu_L = \kappa C R_{max} t \sum_{i=1}^{N} f_i(\theta) \cos(\theta - \theta_i) \tag{4}$$

$$\sigma_L^2 = \kappa^2 C R_{max} t \sum_{i=1}^{N} \sum_{j=1}^{N} \rho_{ij} \sqrt{f_i(\theta) f_j(\theta)} \cos(\theta - \theta_i) \cos(\theta - \theta_j) \tag{5}$$

We used equations (4) and (5) to examine the behavior of the model in a motion detection task (**Fig. 3a**). By assuming that log $L(\theta)$, at the limit of large N, has a Gaussian distribution, the hit and false alarm rates can be computed using standard signal detection theory (SDT). To compute the false alarm rate we assumed that all MT neurons have a mean firing rate of R_{min} imp/s in response to zero coherence motion. We computed the individual ROCs by changing the decision criterion from its largest possible value (hit and false alarms rates of zero) to its smallest value (hit and false alarm rates of 1); we computed d' from these curves in the usual way (**Fig. 3a**).

When discriminating between two alternatives, θ_1 and θ_2, we denote log $L(\theta_1)$ and log $L(\theta_2)$ by $L_{\theta 1}$ and $L_{\theta 2}$ respectively. The log likelihood ratio, log LR, and its first two moments can be written:

$$\log LR = L_{\theta 1} - L_{\theta 2}$$

$$\mu_{LR} = \langle L_{\theta 1} - L_{\theta 2} \rangle = \langle L_{\theta 1} \rangle - \langle L_{\theta 2} \rangle$$

$$\sigma_{LR}^2 = \mathrm{var}\left[L_{\theta 1}\right] + \mathrm{var}\left[L_{\theta 2}\right] - 2\,\mathrm{cov}\left[L_{\theta 1}, L_{\theta 2}\right]$$

Assuming direction θ_1 was presented, we can denote $(\theta_1 - \theta_2)/2$ with $\Delta\theta$ and $(\theta_1 + \theta_2)/2$ with θ_m, and the log LR becomes:

$$\log LR = 2\kappa \sin(\Delta\theta) \sum_{i=1}^{N} n_i \sin(\theta_i - \theta_m)$$

$$\mu_{LR} = 2\kappa C R_{max} \sin(\Delta\theta) t \sum_{i=1}^{N} f_i(\theta_1) \sin(\theta_i - \theta_m)$$

$$\sigma_{LR}^2 = \left[2\kappa \sin(\Delta\theta)\right]^2 C R_{max} t \sum_{i=1}^{N} \sum_{j=1}^{N} \rho_{ij} \sqrt{f_i(\theta_1) f_j(\theta_1)} \sin(\theta_i - \theta_m) \sin(\theta_j - \theta_m)$$

When the neurons are correlated, the pairwise interaction terms increase the variance substantially. To explore how this influences the performance of the model, we assumed that log LR, at the limit of large N, has a Gaussian distribution. The performance of the

Jazayeri & Movshon

model can then be quantified by:

$$P_c = \frac{1}{\sqrt{2\pi\sigma_{LR}^2}} \int_0^\infty \exp\left[\frac{-(x-\mu_{LR})^2}{2\sigma_{LR}^2}\right] dx \qquad (6)$$

We used equation (6) to measure our model's performance for various two–alternative discrimination tasks (**Fig. 3c**). The coherence threshold for 80% correct discrimination was computed as a least square problem using the *fminsearch* unconstrained nonlinear optimization tool from Matlab (Version 7.0.4.352 (R14) Service Pack 2).

To deal with multiple alternative discrimination tasks, we note that, for the case of motion discrimination, the log likelihood of any direction of motion can be readily computed from any other two non–degenerate log likelihoods. Imagine that we know two points on the full likelihood function, say, $L_{\theta 1}$ and $L_{\theta 2}$. We can write:

$$L_{\theta 1} = \kappa \sum_{i=1}^N n_i \cos(\theta_1 - \theta_i) = \kappa \sum_{i=1}^N n_i \left[\cos(\theta_1)\cos(\theta_i) + \sin(\theta_1)\sin(\theta_i)\right]$$

$$L_{\theta 2} = \kappa \sum_{i=1}^N n_i \cos(\theta_2 - \theta_i) = \kappa \sum_{i=1}^N n_i \left[\cos(\theta_2)\cos(\theta_i) + \sin(\theta_2)\sin(\theta_i)\right]$$

In matrix form, we can write:

$$\begin{bmatrix} L_{\theta 1} \\ L_{\theta 2} \end{bmatrix} = \begin{bmatrix} \cos(\theta_1) & \sin(\theta_1) \\ \cos(\theta_2) & \sin(\theta_2) \end{bmatrix} \begin{bmatrix} \sum_{i=1}^N n_i \cos(\theta_i) \\ \sum_{i=1}^N n_i \sin(\theta_i) \end{bmatrix}$$

and therefore, the log likelihood of any other direction, say, θ_3 can be computed as follows:

$$L_{\theta 3} = \kappa \sum_{i=1}^N n_i \cos(\theta_3 - \theta_i) = \begin{bmatrix} \cos(\theta_3) & \sin(\theta_3) \end{bmatrix} \begin{bmatrix} \cos(\theta_1) & \sin(\theta_1) \\ \cos(\theta_2) & \sin(\theta_2) \end{bmatrix}^{-1} \begin{bmatrix} L_{\theta 1} \\ L_{\theta 2} \end{bmatrix} \qquad (7)$$

To compute the performance of the model for multiple alternative motion discrimination (**Fig. 3d**), we assumed that the stimulus was a random–dot pattern moving at the direction θ = 0°. We took 10000 samples from the joint distribution of log likelihoods for θ = 0 and θ = 45°. We then used these values and from equation (7), computed 10000 log likelihood values for all other alternative directions of motion, and asked on what percentage of the trials the log likelihood of θ = 0 exceeded the log likelihood of all other alternative directions. We then used a least-squares optimization (using *fminsearch* from Matlab) to compute the

coherence threshold for 80% correct discrimination.

We used the same method to examine the performance of the model in a direction identification task. For a motion signal moving at the direction $\theta = 0°$, we sampled 10000 pairs of log likelihoods for $\theta = 0$ and $\theta = 45°$ and from there computed the log likelihoods of all directions. On each trial, we extracted the direction with the largest log likelihood (using a max operation) and used the distribution of estimates around the true direction of motion ($\theta = 0$) to compute the circular standard deviation as a function of motion strengths (**Fig. 3b**).

All numerical calculations (**Fig. 3** and **4**) were carried out using Matlab software. The parameters used throughout were:

Name	Value	Comments
N	720	Results largely insensitive to values of $N > 300$
κ	3	Equivalent to a bandwidth of ~ 90°; consistent with reported values in the literature[2]
R_{min}	10 imp/s	Used to measure false alarm rates for motion detection; taken from reported values in the literature[6]
R_{max}	60 imp/s	Results insensitive to its specific value; taken from reported values in the literature[6]
ρ_{max}	0.2	Results insensitive to its specific value; taken from reported values in the literature[5]
δ	0.1	Changes the details of the quantitative predictions; empirical data not available; roughly consistent with literature[5]

References

1. Sclar, G. & Freeman, R.D. Orientation selectivity in the cat's striate cortex is invariant with stimulus contrast. *Exp. Brain Res.* **46**, 457–61 (1982).

2. Britten K.H. & Newsome, W.T. Tuning bandwidths for near-threshold stimuli in area MT. *J. Neurophysiol.* **80**, 762–70 (1998).

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

4. Shadlen, M.N. & Newsome, W.T. The variable discharge of cortical neurons: implications for connectivity, computation, and information coding. *J. Neurosci.* **18**, 3870–96 (1998).

5. Zohary, E., Shadlen, M.N. & Newsome, W.T. Correlated neuronal discharge rate and its implications for psychophysical performance. *Nature* **370**, 140–3 (1994).

6. Shadlen, M.N., Britten, K.H., Newsome, W.T. & Movshon, J.A. A computational analysis of the relationship between neuronal and behavioral responses to visual motion. *J. Neurosci.* **16**, 1486–510 (1996).
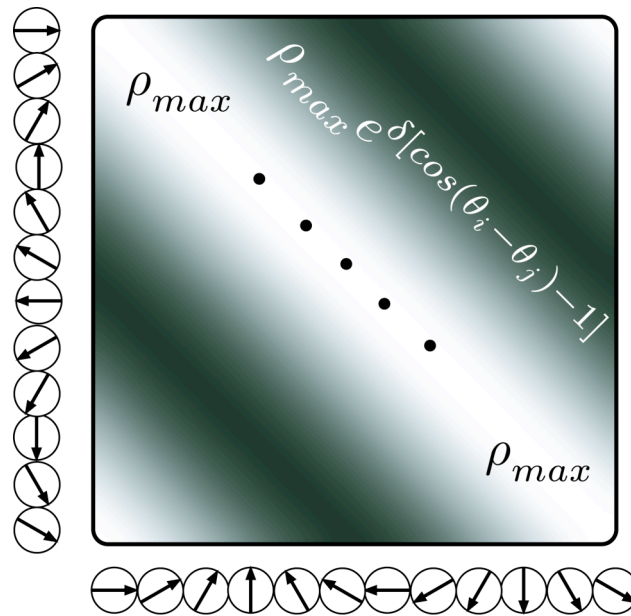
Figure S1. Interneuronal correlation coefficient matrix. A grayscale representation of the structure of the interneuronal pairwise correlation coefficients. The elements on the diagonal are the correlation between pairs of neurons with the same preferred directions, ρ_{max}; off–diagonal elements decrease with a circular Gaussian profile for pairs of neurons with different preferred directions. Higher correlations are represented with lighter grays.