Optimal representation of sensory information by neural populations

Mehrdad Jazayeri & J Anthony Movshon

Sensory information is encoded by populations of neurons. The responses of individual neurons are inherently noisy, so the brain must interpret this information as reliably as possible. In most situations, the optimal strategy for decoding the population signal is to compute the likelihoods of the stimuli that are consistent with an observed neural response. But it has not been clear how the brain can directly compute likelihoods. Here we present a simple and biologically plausible model that can realize the likelihood function by computing a weighted sum of sensory neuron responses. The model provides the basis for an optimal decoding of sensory information. It explains a variety of psychophysical observations on detection, discrimination and identification, and it also directly predicts the relative contributions that different sensory neurons make to perceptual judgments.

The ability to detect, discriminate and identify sensory signals is limited by how efficiently information in sensory representations is put to use in the control of behavior. A stimulus activates a population of neurons in various areas of the brain. To guide behavior, the brain must correctly decode this population response and extract the sensory information as reliably as possible. Two important factors make this problem a challenging one. First, each neuron's response is inherently variable: repeated presentations of the same stimulus elicit different responses. Second, sensory neurons represent moment-to-moment changes in sensory input by rapidly changing their firing patterns. The neural machinery that decodes these responses must compute the most reliable solution for the relevant perceptual behavior, and it must do so quickly in order to faithfully reflect the changing patterns of sensory inputs.

For the problem of sensory identification, the standard theoretical framework holds that the brain reads the activity of a population of neurons and collapses it down to a single value to represent the 'best estimate' of the stimulus. Several decoding strategies have provided more or less optimal solutions for this problem[1]. Examples include the "winner-takes-all" and "population vector" models[2], which are suboptimal under most conditions of interest. Recently a neural network model with recurrent architecture has been proposed that under certain conditions can compute the most likely estimate from a population sensory response[3], providing an account of perceptual identification tasks. But extracting a single best estimate is often a questionable strategy. Many perceptual tasks are better viewed as statistical inference problems, for which the brain needs to compute and represent the probability of all of the different stimuli that are consistent with the sensory response. The optimal decoding strategy that generalizes across all these conditions is to compute the likelihood function, which represents the likelihood of the different stimuli that could have given rise to the observed sensory population response.

If the full likelihood function is available, there are natural solutions for most perceptual problems. For identification, the most likely stimulus is the best estimate; for discrimination, the alternative with the highest likelihood must be chosen, and for detection the likelihood of the most likely stimulus is compared to a criterion. When asked to estimate a stimulus from multiple cues, humans are able optimally to combine likelihoods, and not merely separate estimates derived from individual cues[4,5]. Similarly, Bayesian theories propose combining sensory data with prior beliefs about the stimulus[6], which requires the brain to work with the likelihood function. In these cases and others, representing sensory likelihoods is the optimal way for the brain to guide perceptual decisions.

Although the importance of representing likelihood is well understood in theory, it has not been clear in practice how neurons can compute and combine sensory likelihoods. We have developed a simple and neurally plausible model that computes the full likelihood function quickly and continuously. Our design is similar to the population vector model[7] in that it pools the activity of sensory neurons in a simple additive feedforward architecture. However, the model differs from the population vector model and many other models of readout in two important ways: first, it computes not just a single estimate of the stimulus but the full likelihood function, and second, the contribution of each neuron is determined by its own tuning properties and firing statistics. The computation of likelihood thus derives directly and naturally from the properties of the relevant sensory neurons.

We first develop the model for an abstract population of sensory neurons, and then take the specific example of decoding information about motion from the population activity of neurons in area MT/V5 of the visual cortex. We show that this model predicts performance on a wide variety of perceptual tasks that involve judgments of motion. Finally, we use the model to make testable predictions of the way in

Center for Neural Science, 4 Washington Place, Room 809, New York University, New York, New York 10003, USA. Correspondence should be addressed J.A.M. (movshon@nyu.edu).
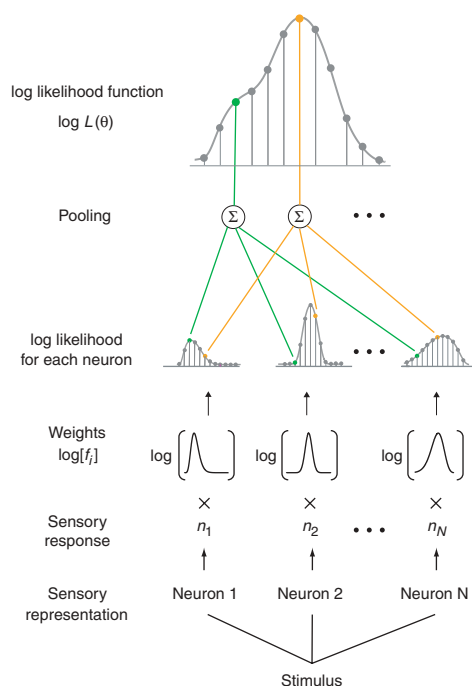
Figure 1 Computing the log likelihood function in a feedforward network. At its input (bottom), a stimulus, elicits $n_1, n_2, \ldots, n_N$ spikes in the sensory representation. The response of each neuron multiplied by the logarithm of its own tuning curve, $\log[f_i]$, gives the contribution of that neuron to the log likelihood function. Adding the contribution of individual neurons (shown for two example stimulus values in orange and green) gives the overall log likelihood function, $\log L(\theta)$ for all values of θ that could have elicited this pattern of responses. Here, the orange point at the peak of the log likelihood function indicates the most likely stimulus.

which individual sensory neurons contribute to perceptual judgments in these tasks.

RESULTS

Computing the likelihood function with neurons

Imagine a sensory stimulus activating a population of neurons in a cortical sensory area. These neurons are often broadly tuned, and the response of each one is noisy. As a result, every stimulus evokes a noisy population response at the level of the sensory neurons; the task is to infer the stimulus from this response. To solve this problem, we ask how likely it is that each possible stimulus elicited the observed response. To determine how likely a given stimulus is, one strategy is to ask each neuron the likelihood that its response was elicited by that stimulus, and then combine the likelihoods to determine the overall likelihood of that stimulus. By repeating the same procedure for all stimuli, one can compute the likelihood of every stimulus for the particular observed population response. This is what the likelihood function represents.

Consider a stimulus, denoted θ_0, causing a neuron tuned to stimulus θ_i to fire n_i spikes in any given time window. If this neuron's firing statistics are described by a Poisson process, then n_i is Poisson-distributed and has a mean of $f_i(\theta_0)$, where $f_i(\theta)$ represents this neuron's tuning function. The likelihood of the stimulus θ_0, denoted $L_i(\theta_0)$, is simply the probability that this neuron would fire n_i spikes in response to that stimulus; that is, $p(n_i|\theta_0)$. Without loss of generality, we can compute this likelihood in the log space. The log likelihood of θ_0 becomes:

$$\log L_i(\theta_0) = \log p(n_i|\theta_0) = \log \frac{f_i(\theta_0)^{n_i}}{n_i!} e^{-f_i(\theta_0)}$$

$$= n_i \log f_i(\theta_0) - f_i(\theta_0) - \log(n_i!) \qquad (1)$$

While this formulation is familiar from previous work[7], we now put it to novel use. The likelihood of every stimulus θ can be computed in the same fashion for each neuron. To compute the overall log likelihood, $\log L(\theta)$, from the population of neurons tuned to different θ_i's, we sum the

individual $\log L_i(\theta)$'s (the summation is a consequence of working in log space). The overall log likelihood of any stimulus θ is then:

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

$$= \sum_{i=1}^{N} n_i \log f_i(\theta) - \sum_{i=1}^{N} f_i(\theta) - \sum_{i=1}^{N} \log(n_i!) \qquad (2)$$

The last two terms in equation (2) can be safely ignored. The last term is clearly independent of θ, and for a homogenous representation, the population response to any stimulus of a given strength sums to a constant and makes the second term also independent of θ. Therefore, the log likelihood of a stimulus at any θ can be computed as a simple weighted sum of the responses of the neurons, where the activity of each neuron is weighted by the log of its own tuning function:

$$\log L(\theta) = \sum_{i=1}^{N} \log L_i(\theta) = \sum_{i=1}^{N} n_i \log f_i(\theta) \qquad (3)$$

Each neuron's contribution to the measurement of the log likelihood of stimulus θ is thus determined by the product of its firing rate and the logarithm of its tuning curve at θ ($\log f_i(\theta)$). The overall $\log L(\theta)$ is simply the sum of the contributions of individual neurons. This computation can be carried out in a single feedforward step: the model receives the sensory responses at its input and, after weighting each neuron's response by the log of its own tuning function, pools these responses into an ensemble of output neurons, where each output neuron gives a measure of the likelihood of a particular stimulus (**Fig. 1**). The recoded profile of activity across the output neurons represents the full log likelihood function.

Our model is quite general and relies only on a few reasonable assumptions. First, we do not assume any specific form for individual tuning curves, so the model can compute the likelihood function from the activity of neurons with widely varying tuning properties as long as the weights take the heterogeneity of tuning functions into account (that is, each neuron is weighted by the log of its own tuning function). The computation is applicable to a variety of sensory parameters with the constraint that the tuning curves for different stimulus conditions (for example, different orientations or directions) should sum to a constant when the stimulus has constant intensity. This assumption is plausible but needs experimental verification; in **Supplementary Methods** online, we discuss the variety of biologically realistic tuning functions that permit the model to be optimal.

Second, while equation (3) presents the case of neurons with Poisson firing statistics for which the feedforward weights are the logs of the tuning functions, we can easily deal with other firing statistics. Poisson variability reasonably describes the firing statistics of cortical neurons[8,9], but it is not the only candidate and it is not required by the model: many other distributions that approximate the firing rate statistics of cortical neurons are also compatible with linear pooling.
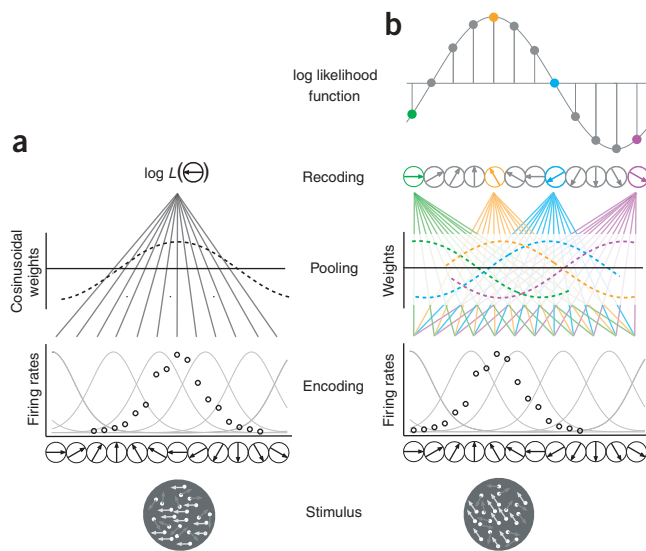
Figure 2 Computing likelihood for the direction of motion. (**a**) A random-dot stimulus (bottom) activates a set of directionally tuned neurons in area MT. The smooth curves represent neuronal tuning curves, and small circles show the noise-perturbed population response on a particular trial. To represent likelihood, we recoded the sensory signals by weighting the inputs from the population of tuned 'encoding' neurons. For the example shown, the correct weighting function has a cosinusoidal form, and the weighted signals converge to an output neuron representing log likelihood for a leftward direction. (**b**) Same as **a**, except here the output layer consists of an ensemble of neurons. The weighted signals converge to this output layer where the neurons represent the log likelihood for all possible directions, the likelihood function. Here, at the output, the average likelihood profile is shown; the colored points represent the average likelihoods of four example directions. The peak of the average likelihood function—the expected maximum-likelihood estimate of the stimulus direction—is shown as orange.

In **Supplementary Methods**, we detail how log L can be computed in a feedforward architecture for a large family of exponential distributions; variations in the firing statistics of neurons simply call for changes in the feedforward weights.

Third, we assume that the encoding neurons are statistically independent, which is usually not correct. Interneuronal correlations can impoverish the quality of pooled signals[9–11], and so one might want to include information about these correlations in the decoding network. But is it reasonable to try to deal with the input correlation structure? We believe not. Interneuronal correlations are not fixed, and they vary from stimulus to stimulus[12]; for a decoder model that does not "sneak a peek" at the stimulus, it is implausible to take such correlations into account[13]. Moreover, the approximately constant variability of neuronal firing across multiple stages of cortical processing may reflect the propagation and not the removal of the correlated noise[9]. We will return to the issue of interneuronal correlations and how they affect the optimality of the model's behavior when we consider the specific case of direction of motion; as we will see, the deviations from optimality that result are consistent with experimental observation.

As we discussed above, representing likelihoods has a number of well-recognized benefits for sensory decoding, and because it computes the full likelihood function our model inherits those benefits. But because our model provides a recipe for computing likelihoods from sensory neurons, it makes specific and testable predictions about how the activity and the tuning properties of each sensory neuron determine its influence on the outcome of perceptual decisions. To illustrate this, we next examine how the properties of neurons in area MT control their contribution to perceptual judgments about the direction of motion.

The likelihood function for the direction of motion

Consider the example of direction of motion in a field of moving dots[14]. Judgments about the direction of motion in such stimuli seem to depend on neural activity in area MT/V5, where most cells are tuned for the direction of motion[14–17]. Neurons in area MT have bell-shaped tuning functions, which can be approximated with the circular Gaussian (the von Mises function[18,19]; κ in equation (4) determines the tuning bandwidth). The tuning function does not change shape with motion strength (coherence)[20]; cells increase their firing rate for favored directions roughly in proportion to coherence,

and their firing statistics are approximately Poisson[21]. Rewriting equation (3) for this case, log L becomes a cosinusoidally weighted sum of a neuron's responses:

$$\log L(\theta) = \kappa \sum_{i=1}^{N} n_i \cos(\theta - \theta_i) \qquad (4)$$

Neurons most effectively increase the likelihood of the direction for which they are tuned; neurons with more remote preferences make smaller contributions; neurons that prefer opposite directions are pooled with negative weights and decrease the likelihood (**Fig. 2a**). The model can therefore be seen as a generalized motion-opponent mechanism of the kind proposed previously[22]. As a consequence of the unequal weighting, neurons tuned optimally to a particular direction, which have a better signal-to-noise ratio (SNR), have more influence than those with a lower SNR that are farther away from the center of the pool. The full likelihood function is represented by an array of output neurons in which each neuron measures the likelihood of a particular direction by pooling the MT responses with a cosinusoidal weighting profile centered at that direction (**Fig. 2b**).

Detecting, identifying and discriminating motion

The properties of incoming sensory stimuli determine the responses of sensory neurons. Our model pools these responses to compute the likelihood function, which in turn determines the performance of the model in perceptual tasks. Let us follow the flow of information in the model from the stimulus to the behavioral performance for the example case of motion in a field of moving dots. After the onset of the stimulus, the number of spikes each MT neuron fires is determined by its tuning function and is subject to Poisson variability. Every spike from every neuron changes the likelihood function. A spike from a neuron tuned to, say, leftward, increases the likelihood of leftward motion, decreases the likelihood of rightward motion and has no influence on the likelihood of upward or downward motion. Accordingly, the variability of MT responses introduces variability into the computed likelihood function and limits the model's performance. We can use the well-documented responses of MT neurons to random-dot stimuli to examine the performance of our model as a function of stimulus characteristics in the classic psychophysical tasks of detection, discrimination and identification.

In a detection task, the subject views a field of randomly moving dots and has to judge whether a given, known direction of motion is present. The optimal strategy is to compare the likelihood for that motion to a criterion; the usual formulation for this problem is in terms of signal detection theory, which uses the receiver operating characteristic (ROC) to distinguish the influence of the choice of criterion from sensitivity (d'). We used our model to make predictions for the hit and
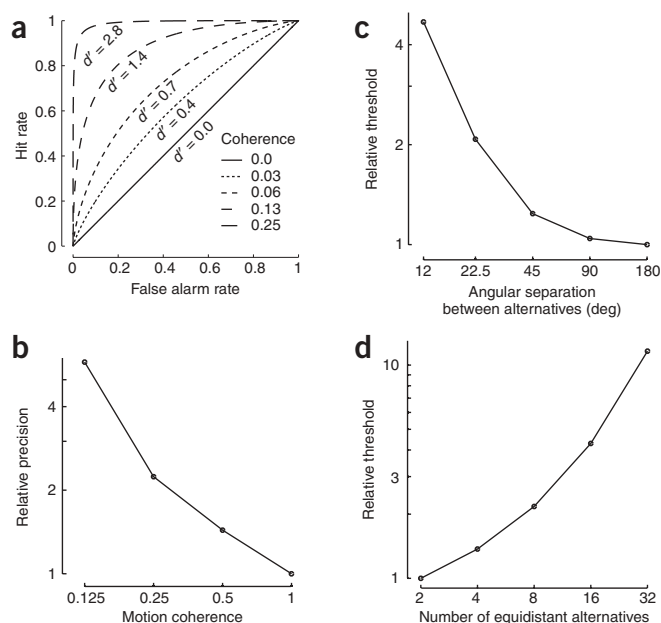
Figure 3 Predictions of the model for behavioral performance in psychophysical tasks. (**a**) The receiver operating characteristic (ROC) for a motion-detection task. The detectability (d') of the stimulus and the area under the ROC curve (that is, percent correct) increase with motion signal strength (coherence). High (low) criterion values correspond to the lower-left (upper-right) part of each curve where both the hit and false alarm rates are low (high). (**b**) The precision of the model's estimate in a direction-of-motion identification task as a function of coherence. The ordinate shows the relative increase of circular standard deviation of the estimate compared to the highest motion strength (coherence = 1). (**c**) Coherence threshold of the model in two-alternative motion discrimination tasks differing in the angular difference between the two discriminanda. The ordinate shows the relative change in coherence threshold as a function of the angular difference between the two alternatives, compared to the easiest condition when the alternatives are 180° apart. (**d**) Coherence threshold of the model in various multiple-alternative motion discrimination tasks with equidistant alternatives (180° for two alternatives, 90° for four alternatives, and so on). The ordinate shows the relative increase in coherence threshold with increasing number of alternatives compared to the easiest condition with only two alternatives.

false alarm rate in a yes–no motion detection task for a range of motion strengths and the associated values of d' (**Fig. 3a**).

In an identification task, the subject views a field of dots and, without knowing in advance which direction is to be presented, must identify the true direction. The optimal strategy is now to find the most likely direction, which is done in our model by applying a simple winner-takes-all rule to the likelihood function (that is, by finding the output neuron with the greatest activity). This strategy predicts how the model becomes more precise at identifying the direction of stronger motion signals (**Fig. 3b**).

When discriminating between two known alternative directions of motion in a random-dot stimulus, the optimal strategy is to choose the alternative with the larger likelihood. Functionally, this is done by comparing the likelihoods of the two alternatives, for which we use the likelihood ratio, or equivalently the difference of log likelihoods. Because the output of our model represents the likelihoods of all stimuli, it easily handles the case of optimal discrimination, which becomes a special case in which we need only compute the likelihood of the two alternatives. Again, each alternative's likelihood is computed by pooling the MT population activity weighted by a cosinusoidal profile centered at that alternative; the difference between the two log likelihoods, log LR, is now the key quantity that optimally discriminates the two alternatives. Our model provides quantitative predictions for how the discrimination threshold would rise as the two alternative directions of motion become closer (**Fig. 3c**).

The importance of having access to the full likelihood function becomes more apparent when one considers cases where more than two alternatives are to be discriminated. Optimal discrimination among multiple alternatives requires the brain to compute and compare several sensory likelihoods. Imagine a case in which a subject is asked to discriminate among N known alternative directions of motion, where the number and the directions of the alternatives change from trial to trial. For optimal discrimination, the subject needs to compute a different set of N sensory likelihoods and compare them on each trial. Such flexibility can be achieved parsimoniously by having access to the full likelihood function. Our model predicts how the coherence threshold would rise with increasing number of alternatives (**Fig. 3d**).

These results (**Fig. 3**) present specific and testable predictions for monkey performance on a range of standard psychophysical tasks, based on the known properties of MT cells and our specific implementation of the likelihood model. By extension, these predictions should also be valid for human observers, and we are now making detailed psychophysical measurements to examine these predictions. While any likelihood-based model can in principle offer a theoretical solution for detection, discrimination and identification of stimuli, our model also provides a plausible neural computation to calculate the likelihood function, and thereby makes a direct link between the activity of sensory neurons and behavioral performance in those tasks.

Neuronal contributions to perceptual judgments of motion

The specific architecture of our model explicitly captures the way that individual sensory neurons contribute to various forms of perceptual behavior. Each neuron's input to the log likelihood is determined by the product of its activity and the log of its tuning curve (equation (3)). An immediate consequence of this computation is that neurons, depending on their level of activity and tuning characteristics, differentially influence any given perceptual behavior. Here, using a simplified homogeneous population of MT neurons, we show how our model reveals the contribution of individual neurons with different direction preferences to motion detection and discrimination tasks.

In a detection task, the log likelihood for the presence of a particular direction of motion is compared to a criterion. The log likelihood is a weighted sum of neuronal responses, where the weight of each neuron is determined by the log of its own tuning function; for the case of motion, this is a cosinusoidal profile (equation (4)). Therefore, the activity of MT neurons tuned to the expected direction should covary positively with the subject's responses; the contribution of neurons tuned away from that direction decreases cosinusoidally, approaches zero for neurons tuned to directions orthogonal to the expected direction, and changes sign for neurons preferring the opposite directions (**Fig. 2a**). These predictions are based on the simplifying assumption that neurons, except for their direction preference, have similar tuning functions. However, in an experimental setting where the tuning bandwidth and firing rates of a neuron under study are readily available, the model makes a specific prediction for how the activity of that particular neuron should covary with the detection behavior.

The problem of decoding neural responses for the discrimination of opposite directions of motion has been well studied. This was first formulated as an opponent process, in which the two directions could be discriminated by subtracting the activity of a neuron tuned to one
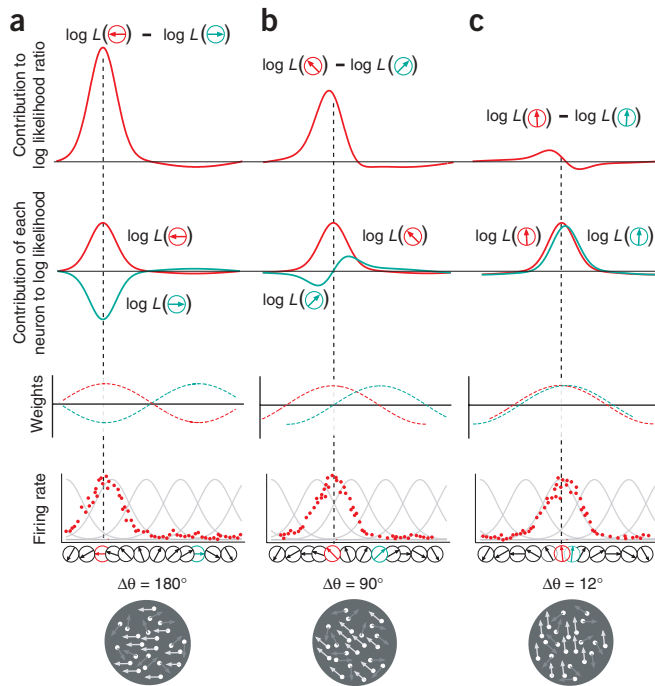
Figure 4 Contributions of MT signals to two-choice motion discrimination. The lower panels show example profiles of activity in area MT in response to a strong motion stimulus, in one of the two directions that are being discriminated. The dashed black line marks the neuron most responsive to this stimulus. The alternatives are 180°, 90° and 12° apart in **a**, **b** and **c**, respectively. The second panels from bottom ("Weights") show the cosinusoidal weighting profiles called for by the model. The contribution of each neuron to the two log likelihoods is computed by multiplying the activity of that neuron by its own weight. Panels in the third row from bottom show the average contribution of each neuron to each of the two log likelihoods (that is, neuron's average firing rate multiplied by its own weight). The top panels show the average contribution of each neuron to the log likelihood ratio. For each neuron, this is computed as the difference between the contribution of that neuron to the two log likelihoods, that is, the difference between the blue and red curves in the third row from bottom. For all three conditions, neurons preferring directions halfway between the two alternatives have similar weights and therefore do not on average contribute to the log LR. For finer discriminations, the overall log likelihood ratio is smaller and, because of the overlap between the weighting profiles, the log likelihood ratio is more strongly determined by neurons with preferences that are shifted away from the two alternatives.

direction from its "anti-neuron" tuned to the opposite direction[14]. This strategy would be optimal if the brain had access to these two neurons only, but it is clearly suboptimal in that it ignores information from neurons that are tuned to directions other than the two alternatives. A later population-decoding model considered the possibility of including such neurons by widening the range of neuronal preferred directions contributing to the decision[23]. But how wide should the range be? Simply adding neurons tuned away from the two alternatives increases both signal and noise, which might either help or harm performance. Our model specifies an exact and optimal form of the pooling profile that would compute the log likelihoods, and it specifies how each neuron contributes to the log likelihood ratio.

To demonstrate this point, we use equation (4) to compute the log likelihood ratio for two alternatives, say θ_1 and θ_2. The log likelihood ratio is simply the difference between the two log likelihoods and can be written:

$$\log \mathrm{LR} = \log L(\theta_1) - \log L(\theta_2)$$
$$= \kappa \sum_{i=1}^{N} n_i [\cos(\theta_1 - \theta_i) - \cos(\theta_2 - \theta_i)] \quad (5)$$

This formulation shows that the contribution of each neuron to the log likelihood ratio is determined by its activity n_i and its preferred direction θ_i relative to the two alternatives. Neurons with similar weights in each of the log likelihoods cancel and do not contribute strongly to the discrimination, whereas neurons with more dissimilar weights in the two log likelihoods have a stronger influence on the model's discrimination behavior.

We implemented this operation in our model for three cases where the two alternative directions are 180°, 90° and 12° apart (**Fig. 4a–c**). Importantly, the readout rule is the same regardless of the two directions that are to be discriminated: for each alternative, the population activity is weighted by a cosine profile centered at that alternative and the difference between the two log likelihoods, log LR, is computed.

When discriminating opposite directions, since the two weighting profiles are opposite cosines (**Fig. 4a**), the MT neurons tuned to the two alternatives that are maximally activated would also have the most dissimilar weights, and would therefore maximally contribute to the measurement of the log LR. This contribution decreases for neurons away from the two alternatives and is zero for neurons tuned to the direction orthogonal to the two alternatives.

In contrast, when computing the log LR for two alternative directions that are 12° apart, the overall contribution of neurons tuned to the two alternatives, despite their high firing rates (**Fig. 4c**, bottom row), would be weakened because they have similar weights (**Fig. 4c**, third row) and will cancel (**Fig. 4c**, top row). More generally, the similarity between the two weighting profiles reduces the contribution of neurons with preferences near the two alternatives and enhances the contribution of neurons tuned away from the two alternatives (**Fig. 4b,c**, top row). In other words, although for any two alternatives the readout rule remains unchanged, for finer discriminations, our model predicts that log LR should be more strongly determined by the activity of neurons tuned to the flanking regions of the two alternatives. This behavior is consistent with the widely assumed role of these "off-optimal" neurons in fine discrimination[24,25], but differs from earlier ideas in making clear that the important influence of these flanking neurons is an automatic consequence of how log likelihoods are computed with neurons.

We show three example cases in which the two alternative directions are 180°, 90° and 12° apart. More generally, as the two alternatives get closer, the contribution of neurons tuned to the alternatives weakens and neurons farther in the flanks become more and more important in the computation of the likelihood ratio. Furthermore, the overall magnitude of the log likelihood ratio is largest for opposite directions where discrimination is easiest (**Fig. 4a**) and becomes progressively smaller as the two alternatives get closer and discriminating between them is more difficult (**Fig. 4b,c**). This change in the magnitude of the log likelihood ratio directly determines the model's performance for the different conditions (**Fig. 3c**).

We have focused here on the model's behavior for the case of direction of motion, and we have presented its predictions for various psychophysical and neurophysiological studies of motion perception. As with the psychophysical predictions in the preceding section, our model makes specific experimental predictions that can be tested against data. In this case, the predicted relationship between neuronal

activity and perceptual performance can be measured by assessing how the response of individual MT neurons covaries with the perceptual judgments[23,26,27] under suitable experimental conditions.

DISCUSSION

Perceptual decision making requires that neural responses in cortical sensory representations be transformed into decision-related variables that can optimally guide behavior. It has been appreciated for some time that sensory likelihoods provide an optimal currency for making perceptual decisions[1], but it has remained controversial whether and how the brain can compute likelihoods from sensory signals[7,28,29]. Our work shows that a linear transformation of sensory cortical responses can compute the needed likelihoods. The computation follows a simple recipe: the likelihood function is computed as the sum of each neuron's response multiplied by the logarithm of its own tuning curve. The model has two important features: it can compute the likelihood for a variety of sensory inputs, and it can account for a variety of perceptual behaviors.

We developed the model for the example case when the sensory signals come from a homogeneous MT population with tuning curves of a particular shape. But the universe of inputs that the model can handle does not depend on the homogeneity or the exact shape of the tuning functions. As discussed in the **Supplementary Methods**, the model can compute sensory likelihoods for different sensory parameters and for a variety of tuning functions. The architecture of the model is such that it combines the overall likelihoods by simple addition of signals from single neurons. Although we have not extended our model beyond the case of a single sensory representation, a particular virtue of the additive computation of likelihood is that the model can combine input from more than one sensory area and compute the likelihood function when multiple cues are present. It does so by simply adding the likelihoods from neurons in more than one sensory representation. It therefore generalizes naturally to a likelihood-based model of cue combination that is called for by human psychophysical data[4,5]. In all cases, the full likelihood function is computed as a weighted sum of neural responses where the weights are directly derived from the firing statistics of neural responses.

The model also provides a single simple neural readout strategy that can support a wide variety of perceptual judgments, including detection, discrimination and identification. In contrast, previous models of sensory decoding were for the most part designed to account for a particular task. For example, in the analysis of motion, one model accounts for discrimination of opposite directions of motion[23,26], another for discrimination of nearby directions[27] and yet others for identification of the direction of motion[30–32]. None of these straightforwardly generalizes to other situations, such as discrimination of multiple alternative directions. The generality of our model, on the other hand, derives directly from the fact that, at its output, it computes the full likelihood function; that is, the likelihood of all stimuli that could have given rise to a given sensory response.

The model solves a sensory detection task by comparing the likelihood of the expected sensory parameter (specified by one output neuron in our model) to a criterion. We show the performance of the model for the case of detecting motion (**Fig. 3a**). In addition, since the model specifies the contribution of each neuron to likelihood, it predicts the relation between the responses of individual neurons in sensory representations to detection behavior. For instance, to detect motion in a given direction, neurons tuned to the expected direction have the highest positive weight[25,33]; the contribution of neurons tuned away from that direction progressively decreases and changes sign for neurons preferring the opposite directions. These predictions are in agreement with data from area MT of monkeys engaged in a

motion detection task (W.H. Bosking and J.H.R. Maunsell, *Soc. Neurosci. Abstr.* 935.7, 2004).

For sensory discrimination tasks, our model computes and compares the likelihoods of the expected alternatives. For the case of motion, it predicts how performance changes with stimulus parameters and task design (**Fig. 3c,d**). More distinctively, it also predicts the contribution of individual sensory neurons in different perceptual situations. When discriminating opposite directions, monkeys seem to rely most on neurons that are tuned to the two alternatives[26]. A recent neurophysiological study of motion discrimination with awake monkeys suggests that for fine direction discrimination, neurons within area MT are pooled with unequal weights[27] so that neurons with preferred directions that are quite remote from the two alternatives have a larger weight in the monkey's choice than ones tuned to the two alternatives. Similarly humans are known to read more effectively the activity of the off-optimal neurons when engaged in a fine sensory discrimination[24,25]. These results suggest that the readout strategy must be adaptable—for discrimination of opposite directions the pooled signals must be centered at the two alternatives[23], but when discriminating nearby directions, the readout mechanism should automatically give the neurons with flanking preferences the highest weight. This is exactly what our model does (**Fig. 4**). When computing the log likelihood ratio for the discrimination of opposite directions, the neurons tuned to the alternatives have the largest weight, whereas for two nearby alternatives, the effective weighting profile has maxima far out on the flanks of the population activity.

We have concentrated on the problem of optimally pooling sensory inputs, but it is also of interest to consider the properties of the neurons that do the pooling. For example, how would one recognize a neuron representing likelihood, such as those at the output of our model? The first issue concerns dynamic range. Because likelihoods can be both increased and decreased by sensory evidence, the pooled signals can assume both positive and negative values. The output neurons should therefore have a baseline firing rate and should operate in an approximately linear manner. The second issue concerns timing. One of the attractive features of our model is that it is purely feedforward, so it can, in principle, represent moment-to-moment changes in likelihood, limited only by the temporal precision of its output neurons. But there is also good reason to believe that neurons involved in perceptual judgment should be able to accumulate information across time when that is desirable[34,35]. This accumulation can either be accomplished by feeding the pooled sensory signals of our model into an integrator network[36,37] or by deferring the integration stage until after the computation of the instantaneous likelihoods by a network like the one we propose.

Given the properties required of these output neurons, we can speculate about where they might be found. Examining the connectional architecture of cortex does not reveal an area in which sensory likelihood might be represented separately from the planning of specific movements: signals from areas thought to have predominantly sensory functions, like area MT, project directly to areas with clear motor planning functions, like the lateral intraparietal area (LIP). The absence of a candidate 'decision area' suggests that the computation of likelihood might actually take place in areas we think of as sensory. For example, with respect to the computation of motion direction, existing models and recent experimental evidence suggest that a population of neurons in area MT, "pattern direction selective cells"[38] might obtain their properties by pooling directionally selective inputs using an architecture similar to the one in our model[39] (N.C. Rust, E.P. Simoncelli, & J.A. Movshon, *Soc. Neurosci. Abstr.* 591.11, 2005). Alternatively, the instantaneous likelihoods might be computed by suitably combining sensory signals directly in association areas such as area LIP[40] or the parietal

reach region (PRR)[41], implying that likelihoods might be calculated separately for behaviors that engage different motor systems.

Whether or not the likelihood calculation is performed in sensory areas, by the time signals reach areas such LIP or PRR, we expect them to reflect the integrated sensory likelihoods. For instance, while discriminating opposite directions of motion, presentation of one of the alternatives should increase the activity of the output neuron that measures the likelihood of the presented alternative while the neuron representing the likelihood of the opposite direction should decrease its activity. This prediction is in agreement with data from area LIP of awake, behaving monkeys[35,40,42]. Moreover, when discriminating two or more equiprobable alternatives, before the presentation of the stimulus the likelihood of all alternatives is the same and depends on the number of alternatives. For two alternatives, the likelihoods are 50% each, whereas for eight alternatives, the likelihoods are 12.5%. This characteristic behavior of the likelihood function makes a clear prediction: before the onset of the stimulus, the baseline firing rate of the output neurons should decrease with the increasing number of alternatives. This effect has recently been observed in the activity of individual LIP neurons in monkeys engaged in a motion discrimination task (A.K. Churchland, M. Tam, J. Palmer, R. Kiani and M.N. Shadlen, *Soc. Neurosci. Abstr.* 16.8, 2005).

Finally, we draw attention to how the properties of sensory neurons relate to the architecture of our model. Why are sensory neurons broadly tuned? Why do they have Poisson-like variability? What is the functional significance of tuning functions that are independent of stimulus intensity? Notably, the simplicity and the flexibility of our model follow directly from these characteristic properties of sensory neurons. Is it a coincidence that these properties are just the ones that allow a simple feedforward architecture to recode the activity of sensory neurons to sensory likelihoods? As a final speculation, we suggest the value of likelihood-based decision making, long recognized in psychophysics, means that sensory representations should be organized to allow a biologically feasible transformation of sensory responses to sensory likelihoods. This is exactly what we think the cerebral cortex does.

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

COMPETING INTERESTS STATEMENT
The authors declare that they have no competing financial interests.

1. Pouget, A., Dayan, P. & Zemel, R.S. Inference and computation with population codes. *Annu. Rev. Neurosci.* **26**, 381–410 (2003).
2. Georgopoulos, A.P., Kalaska, J.F., Caminiti, R. & Massey, J.T. On the relations between the direction of two-dimensional arm movements and cell discharge in primate motor cortex. *J. Neurosci.* **2**, 1527–1537 (1982).
3. Deneve, S., Latham, P.E. & Pouget, A. Reading population codes: a neural implementation of ideal observers. *Nat. Neurosci.* **2**, 740–745 (1999).
4. Ernst, M.O. & Banks, M.S. Humans integrate visual and haptic information in a statistically optimal fashion. *Nature* **415**, 429–433 (2002).
5. Hillis, J.M., Watt, S.J., Landy, M.S. & Banks, M.S. Slant from texture and disparity cues: optimal cue combination. *J. Vis.* **4**, 967–992 (2004).
6. Weiss, Y., Simoncelli, E.P. & Adelson, E.H. Motion illusions as optimal percepts. *Nat. Neurosci.* **5**, 598–604 (2002).
7. Seung, H.S. & Sompolinsky, H. Simple models for reading neuronal population codes. *Proc. Natl. Acad. Sci. USA* **90**, 10749–10753 (1993).
8. 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).
9. Shadlen, M.N. & Newsome, W.T. The variable discharge of cortical neurons: implications for connectivity, computation, and information coding. *J. Neurosci.* **18**, 3870–3896 (1998).
10. Zohary, E., Shadlen, M.N. & Newsome, W.T. Correlated neuronal discharge rate and its implications for psychophysical performance. *Nature* **370**, 140–143 (1994).
11. Bair, W., Zohary, E. & Newsome, W.T. Correlated firing in macaque visual area MT: time scales and relationship to behavior. *J. Neurosci.* **21**, 1676–1697 (2001).
12. Kohn, A. & Smith, M.A. Stimulus dependence of neuronal correlation in primary visual cortex of the macaque. *J. Neurosci.* **25**, 3661–3673 (2005).
13. Shamir, M. & Sompolinsky, H. Nonlinear population codes. *Neural Comput.* **16**, 1105–1136 (2004).
14. Britten, K.H., Shadlen, M.N., Newsome, W.T. & Movshon, J.A. The analysis of visual motion: a comparison of neuronal and psychophysical performance. *J. Neurosci.* **12**, 4745–4765 (1992).
15. Maunsell, J.H. & van Essen, D.C. The connections of the middle temporal visual area (MT) and their relationship to a cortical hierarchy in the macaque monkey. *J. Neurosci.* **3**, 2563–2586 (1983).
16. Salzman, C.D., Murasugi, C.M., Britten, K.H. & Newsome, W.T. Microstimulation in visual area MT: effects on direction discrimination performance. *J. Neurosci.* **12**, 2331–2355 (1992).
17. Salzman, C.D. & Newsome, W.T. Neural mechanisms for forming a perceptual decision. *Science* **264**, 231–237 (1994).
18. Mardia, K.V. *Statistics of Directional Data* (Academic Press, London, New York, 1972).
19. Swindale, N.V. Orientation tuning curves: empirical description and estimation of parameters. *Biol. Cybern.* **78**, 45–56 (1998).
20. Britten, K.H. & Newsome, W.T. Tuning bandwidths for near-threshold stimuli in area MT. *J. Neurophysiol.* **80**, 762–770 (1998).
21. Britten, K.H., Shadlen, M.N., Newsome, W.T. & Movshon, J.A. Responses of neurons in macaque MT to stochastic motion signals. *Vis. Neurosci.* **10**, 1157–1169 (1993).
22. Adelson, E.H. & Bergen, J.R. Spatiotemporal energy models for the perception of motion. *J. Opt. Soc. Am. A* **2**, 284–299 (1985).
23. 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**, 1486–1510 (1996).
24. Regan, D. & Beverley, K.I. Postadaptation orientation discrimination. *J. Opt. Soc. Am. A* **2**, 147–155 (1985).
25. Hol, K. & Treue, S. Different populations of neurons contribute to the detection and discrimination of visual motion. *Vision Res.* **41**, 685–689 (2001).
26. Britten, K.H., Newsome, W.T., Shadlen, M.N., Celebrini, S. & Movshon, J.A. A relationship between behavioral choice and the visual responses of neurons in macaque MT. *Vis. Neurosci.* **13**, 87–100 (1996).
27. Purushothaman, G. & Bradley, D.C. Neural population code for fine perceptual decisions in area MT. *Nat. Neurosci.* **8**, 99–106 (2005).
28. Salinas, E. & Abbott, L.F. Vector reconstruction from firing rates. *J. Comput. Neurosci.* **1**, 89–107 (1994).
29. Gold, J.I. & Shadlen, M.N. Neural computations that underlie decisions about sensory stimuli. *Trends Cogn. Sci.* **5**, 10–16 (2001).
30. Salzman, C.D., Britten, K.H. & Newsome, W.T. Cortical microstimulation influences perceptual judgments of motion direction. *Nature* **346**, 174–177 (1990).
31. Groh, J.M., Born, R.T. & Newsome, W.T. How is a sensory map read out? Effects of microstimulation in visual area MT on saccades and smooth pursuit eye movements. *J. Neurosci.* **17**, 4312–4330 (1997).
32. Nichols, M.J. & Newsome, W.T. Middle temporal visual area microstimulation influences veridical judgments of motion direction. *J. Neurosci.* **22**, 9530–9540 (2002).
33. Sekuler, R. & Ball, K. Mental set alters visibility of moving targets. *Science* **32**, 60–62 (1977).
34. Smith, P.L. & Ratcliff, R. Psychology and neurobiology of simple decisions. *Trends Neurosci.* **27**, 161–168 (2004).
35. Roitman, J.D. & Shadlen, M.N. Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. *J. Neurosci.* **22**, 9475–9489 (2002).
36. Wang, X.J. Probabilistic decision making by slow reverberation in cortical circuits. *Neuron* **36**, 955–968 (2002).
37. Koulakov, A.A., Raghavachari, S., Kepecs, A. & Lisman, J.E. Model for a robust neural integrator. *Nat. Neurosci.* **5**, 775–782 (2002).
38. Movshon, J.A., Adelson, E.H., Gizzi, M.S. & Newsome, W.T. The analysis of moving visual patterns. in *Experimental Brain Research Supplementum* II: *Pattern Recognition Mechanisms* (eds. Chagas, C., Gattass, R. & Gross, C.) 117–151 (Springer, New York, 1986).
39. Simoncelli, E.P. & Heeger, D.J. A model of neuronal responses in visual area MT. *Vision Res.* **38**, 743–761 (1998).
40. Mazurek, M.E., Roitman, J.D., Ditterich, J. & Shadlen, M.N. A role for neural integrators in perceptual decision making. *Cereb. Cortex* **13**, 1257–1269 (2003).
41. Snyder, L.H., Batista, A.P. & Andersen, R.A. Coding of intention in the posterior parietal cortex. *Nature* **386**, 167–170 (1997).
42. Shadlen, M.N. & Newsome, W.T. Neural basis of a perceptual decision in the parietal cortex (area LIP) of the rhesus monkey. *J. Neurophysiol.* **86**, 1916–1936 (2001).

Corrigendum: Optimal representation of sensory information by neural populations

Mehrdad Jazayeri & J Anthony Movshon

In the supplementary information initially published online to accompany this article, equation 5 and the last two equations on page 4 contained typographical errors. The errors have been corrected online.