Extracting information from neuronal populations: information theory and decoding approaches

Rodrigo Quian Quiroga and Stefano Panzeri‡§*

Abstract | To a large extent, progress in neuroscience has been driven by the study of single-cell responses averaged over several repetitions of stimuli or behaviours. However, the brain typically makes decisions based on single events by evaluating the activity of large neuronal populations. Therefore, to further understand how the brain processes information, it is important to shift from a single-neuron, multiple-trial framework to multiple-neuron, single-trial methodologies. Two related approaches — decoding and information theory — can be used to extract single-trial information from the activity of neuronal populations. Such population analysis can give us more information about how neurons encode stimulus features than traditional single-cell studies.

Spike sorting
The grouping of spikes into clusters based on the similarity of their shapes. Given that, in principle, each neuron tends to fire spikes of a particular shape, the resulting clusters correspond to the activity of different putative neurons. The end result of spike sorting is determining which spike corresponds to which of these neurons.

Decoding
Predicting the most likely stimulus or behaviour eliciting an observed neural response.

**Department of Engineering, University of Leicester, Leicester, LE1 7RH, UK.*
‡Italian Institute of Technology, Department of Robotics, Brain and Cognitive Sciences, 16163 Genoa, Italy.
§Faculty of Life Sciences, University of Manchester, Manchester, M13 9PL, UK.
Correspondence to R.Q.Q.
e-mail: rqqg1@leicester.ac.uk
doi:10.1038/nrn2578

A standard neuroscience technique involves recording the extracellular activity of single neurons to study their responses in different tasks and conditions[1]. This approach has provided invaluable information about the functions of neurons in different areas[2], but it has the caveat that individual neurons are usually studied in isolation, thus giving a limited picture of how neuronal populations give rise to sensation, behaviour or other complex brain processes. Moreover, the responses of single cells to external stimuli are often averaged over several trials to reduce the effects of neuronal variability. However, the brain usually processes information and takes decisions based on single events, making sense of the noisy responses from individual neurons by evaluating the activity of large populations.

Current recording systems and multielectrode probes[3–8] allow recording from dozens or even hundreds of sites simultaneously. However, although the need to study whole neuronal populations is clear[4,9], analysing this type of data involves two major mathematical challenges: the activity of single neurons must be identified from the extracellular recordings using spike-sorting algorithms (BOX 1), and the information from the population of neurons must be extracted with objective measures.

There are two complementary approaches to study how the brain extracts features and deciphers information encoded in the single-trial activity of neuronal populations. The first uses decoding algorithms to predict a given stimulus or behaviour from the pattern of neuronal responses[10–14]. The second uses concepts from information theory[15] to determine how much information neurons carry about the stimuli[12,14,16,17]. Although the two approaches have been developed in parallel and with limited interaction, we aim to show here how they are intrinsically related. We also illustrate with several examples how much more information can be extracted with a population analysis than by studying the firing of individual neurons.

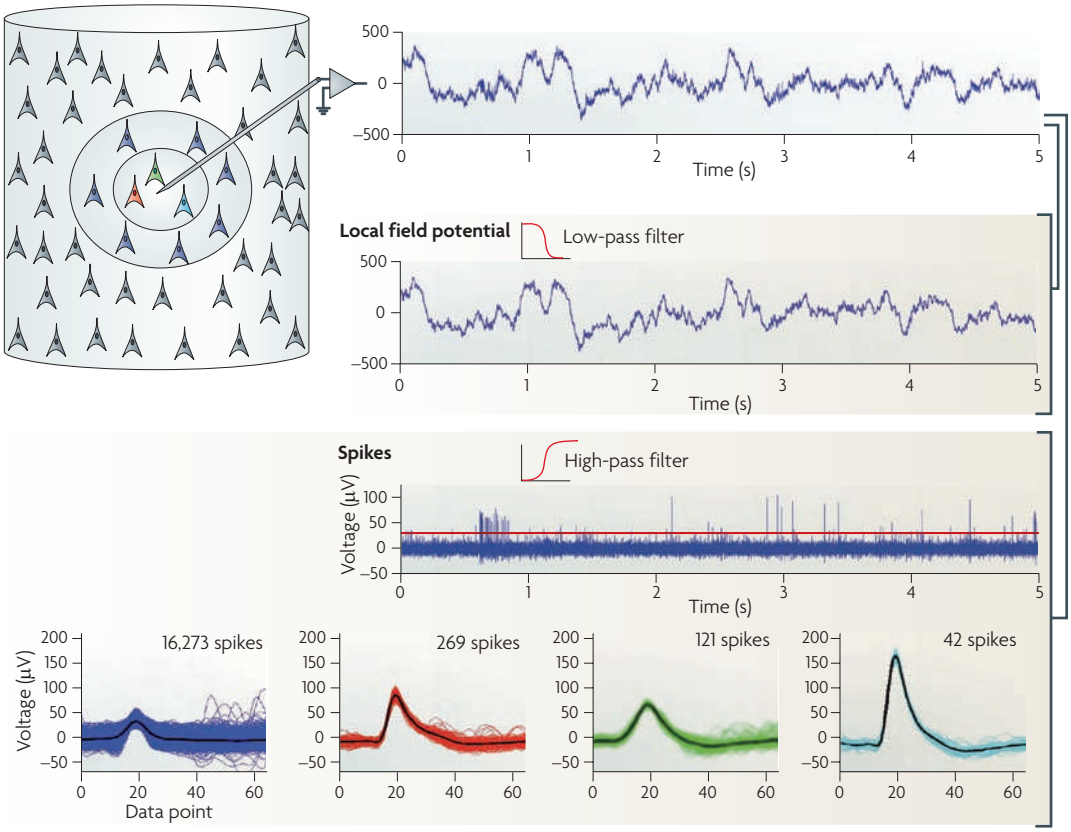
Extracting information from neuronal populations

There are three main steps to analyse multiple single-cell recordings (FIG. 1). First, the neuronal activity is recorded with implanted microwires. Second, the activity of single neurons is extracted from the recorded data using spike-detection and -sorting algorithms[18] (BOX 1). In this respect automatic spike-sorting algorithms are needed for studying large neuronal populations, because they minimize the user's interaction and allow a fast analysis of large numbers of channels[18–21]. The final step is to interpret the resulting patterns of multiple spike trains using decoding or information theory.

Advantages of the population analysis. The power of the population analysis using decoding or information theory relies on several facts. First, these methods consider the information carried by the population of neurons as a whole, and they allow the determination of how each member of the population and the interactions between them contribute to the stimulus representation. For example, the responses of a given neuron might not

© 2009 Macmillan Publishers Limited. All rights reserved

Box 1 | **Extracellular recordings**



Extracellular recordings are usually performed by inserting microwires into the brain[1]. After amplification, the signal is low-pass filtered to obtain the local field potential — the mean field potential generated by neurons in the vicinity of the electrode — and high-pass filtered to identify the activity of single neurons using spike detection and sorting algorithms. The example shown in the figure corresponds to a recording of approximately half an hour in the left hippocampus of an epileptic patient[24,44], of which 5 s of continuous data are shown. After high-pass filtering, the firing of nearby neurons appears as spikes on top of background activity. Spikes are detected using an amplitude threshold (represented by the red horizontal line). Features of the spike shapes are extracted and the spikes are sorted accordingly. For neurons located approximately 50–100 μm from the electrode tip[4,122], the signal-to-noise ratio is good enough to distinguish the activity of each single unit (inner circle; spikes in red, green and cyan). For more distant neurons, up to approximately 150 μm from the tip (outer circle), spikes can be detected but the difference in their shapes is masked by the noise and they are grouped together in a 'multi-unit' cluster (spikes in blue). Spikes from neurons further away from the tip (shown in light grey in the schematic) cannot be detected and contribute to the background noise.

There are several issues that make spike sorting challenging[18]. In particular, some neurons fire very sparsely — for example, the neuron shown in cyan in the figure fired only 42 spikes in approximately half an hour, a mean firing rate of less than 0.05 Hz. These neurons are usually hard to detect. Interestingly, such sparsely firing neurons showed the most selective and interesting responses in human recordings[24].

Information theory
A mathematical theory that deals with measures of information and their application to the study of communication systems. In neuroscience it is used to establish the amount of information about a stimulus or behaviour that is contained in the neural responses.

Local field potential
(LFP). A neurophysiological signal that is obtained by low-pass filtering extracellular recordings. It represents the mean field potential generated by the slow components of synaptic and neural events in the vicinity of the recording electrode.

clearly indicate whether an apple or an orange is being seen, but this uncertainty might be resolved when the activity of other neurons is considered. In fact, the population analysis based on decoding or information theory can reveal how the other neurons resolve this ambiguity — for example, by coordinating their firing to tag particularly salient events[22,23] or by having each neuron represent a particular stimulus feature[24–26]. Second, postsynaptic neuronal systems must usually interpret neuronal responses obtained in only one trial. Both the information-theoretic and the decoding approaches quantify stimulus knowledge obtained with the observation of single-trial population responses, thus providing a framework that is compatible with the strict timescales of online brain processing. Third, the stimulus features encoded by the spike trains can be discovered by assessing whether the population response can discriminate different stimuli containing a particular feature. Fourth, it is possible to systematically evaluate how different features of the spike trains affect the performance of a decoding algorithm or the amount of extracted information. Fifth, the information given by different measures of neuronal activity, such as spike trains and local field potentials (LFPs), can be analysed and combined. Although these two signals have very different characteristics and signal-to-noise ratios, information theory and decoding algorithms allow a direct comparison between LFPs and spikes because they project the two

© 2009 Macmillan Publishers Limited. All rights reserved

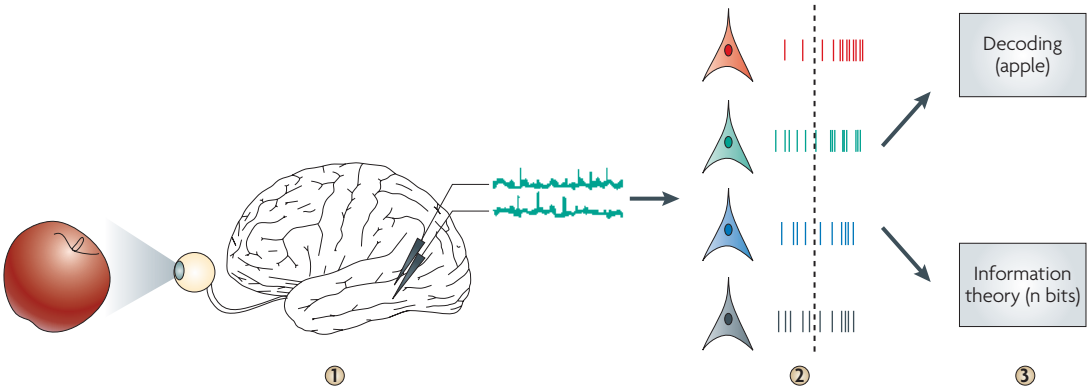


Figure 1 | **Three main steps for the population analysis of neural recordings.** The common steps for analysing how a population of neurons encodes information about visual inputs are shown. First, recordings are taken at different sites with implanted electrodes. Second, the simulated activity of single neurons is extracted from the continuous data using spike-sorting algorithms. Third, information is inferred from the multiple spike trains with decoding algorithms (which can predict that the stimulus was an apple), or information theory (which quantifies the knowledge about the stimulus gained by observing the population response). The population analysis allows the study of the information carried by the different features of the multiple spike trains. For example, it can be established whether the information of the apple is given by an increase in firing (neuron in red), by a particular temporal firing pattern (neuron in green) or by the simultaneous firing of a subset of neurons (neurons in blue and grey). The vertical dotted line marks stimulus onset.

signals on to a common scale. It then becomes possible to assess whether the LFP adds some knowledge about the stimulus that cannot be obtained from spikes alone and vice versa.

Decoding algorithms. Decoding is the prediction of which stimulus or behaviour elicits a particular neuronal response in a single trial. An example is given by Bayesian decoding. Let $P(s)$ denote the probability of presentation of stimulus s (belonging to a set S) and $P(r|s)$ denote the conditional probability of obtaining a population response r (out of a response set R) when stimulus s is presented. Using Bayes' theorem, we obtain[27]:

$$P(s|r) = \frac{P(r|s) \cdot P(s)}{P(r)} \quad (1)$$

with

$$P(r) = \Sigma_s P(r|s) \cdot P(s) \quad (2)$$

Equation 1 gives the posterior probability that, given a response r, stimulus s was presented. Bayesian decoding calculates from this posterior probability distribution a single prediction of the most likely stimulus (s^P) (REFS 13,28–30) — for example, by taking: $s^P = \arg\max_s (P(s|r))$.

Besides the Bayesian approach, there are several other methods to decode the stimulus in a given trial. A thorough discussion of these decoding algorithms has been provided elsewhere[10,11,13,14,31], and a short description of the most common methods is given in BOX 2. The implementation of decoding algorithms is illustrated in FIG. 2.

To validate decoding results, some trials can be used to optimize the decoder (the training set) and the rest to test its performance, a procedure called cross-validation[32]. It is important that trials belonging to the training set are not used to evaluate the decoding performance because this may lead to artificially high values owing to overfitting[33]. Furthermore, both the training and the testing sets should be large enough to avoid underestimating the decoding performance owing to poor optimization of the decoder in the first case and low statistics for testing in the latter. A common procedure is the 'leave-one-out' validation, in which each trial is predicted based on the distribution of all the other trials. This has the advantage that both optimization and testing are based on the largest possible number of trials[33].

Decoding results are usually presented in the form of 'confusion matrices' (FIG. 2c). The values on a given row i and column j of a confusion matrix represent the (normalized) number of times that a presentation of stimulus i is predicted by the decoder to be stimulus j. If the decoding is perfect, the confusion matrix should have entries equal to one along the diagonal and zero everywhere else. For equiprobable stimuli, performance at chance levels should be reflected by a matrix in which each entry has equal probability $1/K$ (with K being the number of stimuli).

Shannon information theory. Another powerful way to study the activity of neuronal populations is to calculate the information about a given stimulus or behaviour contained in the neuronal responses using the formalism of Shannon information theory. As before, suppose that a stimulus s belonging to a set S is presented with a probability $P(s)$. The Shannon entropy $H(S)$ of the distribution of probabilities $P(s)$ for each stimulus is defined as[15,16,27,34]:

$$H(S) = -\sum_s P(s)\log_2 P(s) \quad (3)$$

This quantifies the uncertainty about which stimulus is presented or, conversely, the average amount of information gained with each stimulus presentation. Entropy is measured in bits if the logarithm is taken with base 2

Posterior probability
The posterior probability of a random variable is the conditional probability assigned to the variable given some event. For example, the posterior probability $P(s|r)$ is the conditional probability that stimulus s was presented, given that a response r was observed.

Shannon entropy
A measure of the uncertainty about the value that might be taken by a random variable.

Bit
The unit used to measure reduction of uncertainty. One bit corresponds to a reduction of uncertainty by a factor of two (for example, a correct answer to a yes/no question).

© 2009 Macmillan Publishers Limited. All rights reserved

Box 2 | **Decoding algorithms**

Several decoding algorithms have been developed. Nearest-neighbour decoders assign a given trial to the class of its nearest neighbour[31,61]. Fisher linear discriminant algorithms introduce a dimensionality reduction (by projecting the original space where the classification is performed on to a line with a direction that maximizes the ratio of between-class to within-class distances) that optimally separates the samples of each class[31,44]. Support vector machines project the data into a high-dimensional space in which it is possible to find a hyperplane (a high-dimensional plane) that optimally separates the data[123]. This is particularly useful if points corresponding to two given classes are separated by a nonlinear region. Bayesian decoders assign a given trial to the class that minimizes the probability of an error, given the prior probabilities of each class and the posterior probabilities (see equation 1)[11,28,29]. Decoding can also be performed by training an artificial neural network[65].

Decoding performances are usually quantified by the relative number of hits (that is, the average of the diagonal in the confusion matrix). As the outcomes of the predictions of each stimulus can be regarded as a sequence of Bernoulli trials (independent trials with two possible outcomes: success and failure), the probability of successes in a sequence of trials follows the Binomial distribution[124]. Given a probability p of getting a hit by chance ($p = 1/K$, in which K is the number of stimuli), the probability of getting k hits by chance in n trials is given by

$$P(k) = \binom{n}{k} p^k (1-p)^{n-k} \quad (7)$$

, where

$$\binom{n}{k} = \frac{n!}{(n-k)!k!} \quad (8)$$

is the number of possible ways of having k hits in n trials. From this it is possible to assess statistical significance and calculate a p-value by adding up the probabilities of getting k or more hits by chance[44]:

$$p\text{-value} = \sum_{j=k}^{n} P(j) \quad (9)$$

(as in equations 3–5 and 10–12). The Shannon entropy is zero when the same stimulus is presented each time, and is positive otherwise. It reaches a maximum of $H(S) = log_2 K$ when the probability of presenting each of K different stimuli is equal.

If the neuronal population response r contains information about the stimulus, then its observation will reduce the stimulus uncertainty of equation 3. The residual uncertainty about the stimulus after observing the neuronal response is called equivocation and it is given by the weighted average entropy of the posterior distribution $P(s|r)$ (see equation 1):

$$H(S|R) = -\sum_{s,r} P(r) \cdot P(s|r) \log_2 P(s|r) \quad (4)$$

From this, mutual information is defined as the reduction of uncertainty (or gained information) about the stimulus obtained by knowing the neuronal response. It is given by the difference between the stimulus entropy $H(S)$ and the equivocation $H(S|R)$:

$$I(S;R) = \sum_{s,r} P(r)P(s|r)\log_2 \frac{P(s|r)}{P(s)} \quad (5)$$

$$= \sum_{s,r} P(s,r)\log_2 \frac{P(s,r)}{P(s) \cdot P(r)}$$

$P(s,r)$ denotes the joint probability of observing the response $R = r$ together with the stimulus $S = s$. These probabilities are not known *a priori* but they can be estimated empirically. For example, the range of responses of each neuron can be partitioned into M bins, and the frequency of observing a response in each bin is computed (for reviews of other estimation methods, see REFS 35,36). Information, like entropy, is measured in bits. Every bit of information provided by the neurons reduces the overall uncertainty about the stimulus by a factor of two. If the stimuli and the responses are independent, the mutual information equals zero. Otherwise it takes positive values. Perfect knowledge about the stimulus from the neuronal activity or, in other words, an errorless stimulus reconstruction gives a maximum mutual information of $I(S;R) = H(S)$.

Alternatively, the knowledge about the stimulus contained in the neuronal responses can be measured using Fisher information[34]:

$$FI = \sum_{r} P(r|s)\left[\frac{d}{ds}\ln P(r|s)\right]^2. \quad (6)$$

The inverse of Fisher information is a lower bound to the mean square decoding error obtained with any unbiased decoder. Fisher information can be used only with continuous stimuli and thus its application to experimental neuroscience has remained more limited than that of Shannon information.

Complementarities of decoding and information theory. Both decoding and information theory extract quantitative information from the population responses, by quantifying the knowledge about the stimulus that is gained from observing a neuronal population response on a single trial. However, they each quantify a different aspect of this knowledge. Decoding algorithms predict the stimulus that caused the single-trial neuronal responses, and their performance is typically measured by the percentage of correct predictions (BOX 2). But even with an optimal decoder, the amount of extracted information may be less than the information available in the neuronal responses. Information theory quantifies the overall knowledge about the presented stimulus that is gained with the single-trial neuronal responses. This distinction is important because neurons can convey information by means other than just telling which is the most likely stimulus[37–39]: for example, they can provide information about other stimuli that are unlikely (FIG. 3). The complementarities between decoding and information theory are explicit when considering Bayesian decoders, because in this case both decoding and information theory are just two different computations over the posterior probability $P(s|r)$ of equation 1: Bayesian decoders give the most likely stimulus (arg max $P(s|r)$), whereas information theory gives a smooth integration of information over the whole posterior probability $P(s|r)$. The different amounts of information extracted by decoding algorithms and information theory are detailed in BOX 3.

Only a few studies have quantified the information that is lost when using decoding algorithms. One of them evaluated the information that could be gained from the population activity of head-direction cells (neurons that encode the direction of the animal's head) in the primate

Unbiased decoder
A decoder is said to be unbiased if the expected value of its decoding error (the difference between the true and the estimated stimulus values) is zero.

© 2009 Macmillan Publishers Limited. All rights reserved

and reported that a Bayesian decoding algorithm provided ~95% of the total information available from the neuronal responses[40]. This information loss was higher with predictions using the 'population vector' method (see below)[41], in agreement with previous findings showing that Bayesian decoders performed better than the population vector[13,28,29,42].

There is evidence that animals estimate posterior distributions $P(s|r)$ of stimulus variables[43]; however, the final behavioural outcome of most neuronal calculations is a decision about what is the most probable stimulus or the most appropriate behaviour. Thus, decoding (identifying the most likely stimuli) is well suited for comparing neuronal and behavioural performance. Information theory has the advantage of summarizing all the information contained in the population responses into a single, meaningful number. The mutual information between the stimuli and neuronal responses gives an upper bound to the amount of knowledge that can be provided by a decoding algorithm. This makes information theory a strong tool with which to evaluate the computational capabilities of neuronal codes, as it can consider ways of transmitting information that might not be revealed by decoding algorithms.

A straightforward way to link information theory and decoding is to compute the mutual information between the actual and the predicted stimuli from the decoding outcomes — that is, the mutual information $I(S;S^P)$ between the rows and columns of the confusion matrix[15,34,40,44–46] (BOX 3). This can give more information than just the relative number of hits (correct predictions (BOX 2)). For a given number of hits more information can be obtained if incorrect predictions are concentrated into clusters around the correct stimulus[38,47]. For example, the relative number of hits in FIG. 2 tells us that the presentations of picture 29 were often wrongly decoded, but not that they were mainly confused with picture 26 and not just predicted at random, as would have occurred if the neurons had no information about the stimulus[44]. The combined analysis of decoding and information theory can reveal such systematic errors in the decoder.

It has been shown that populations of neurons can not only compute the most likely stimulus — or the optimal response — but also estimate probability distributions[48–50]. The combination of information theory and decoding could give a valuable measure to clarify how populations of sensory neurons can report the predicted stimulus and other relevant information, such as the uncertainty of these predictions[51–53]. This can be achieved by comparing the information between stimuli and neuronal responses, $I(S;R)$, with the information obtained from the confusion matrix after decoding, $I(S;S^P)$. The difference between these two quantities gives the amount of information available in the neuronal responses that could be gained by means other than decoding the most likely stimulus. It is possible to extend this calculation by computing for each trial the predicted stimulus and another specific aspect U of its uncertainty (for example, the variance of the prediction, or the relative likelihood of the best and second best stimuli) and then evaluate

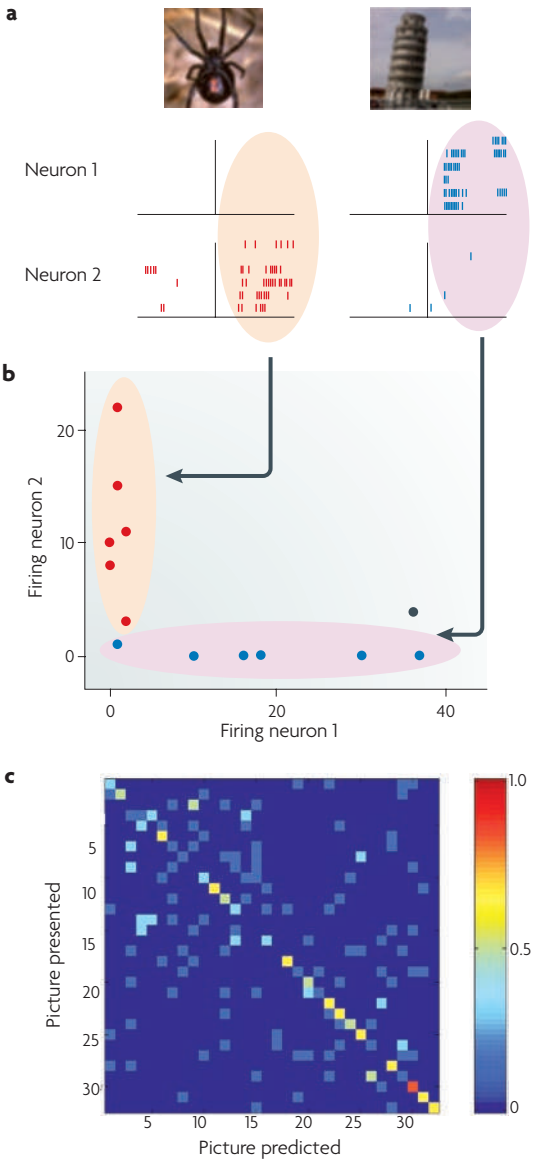


Figure 2 | **Decoding analysis.** The decoding analysis is illustrated with the prediction of which picture presentation elicited single-cell responses recorded from human subjects implanted with intracranial electrodes for clinical reasons. **a** | For simplicity, the responses of two out of a population of N neurons to two out of K stimuli (a picture of a spider and a picture of the Tower of Pisa) are shown. **b** | Each trial is represented as a point in an N-dimensional space. Trials in which the spider was shown are marked in red; those in which the Tower of Pisa was shown are marked in blue. A new trial to be decoded (shown in grey) can, for example, be assigned to the class of its nearest neighbour. **c** | The outcome of the decoding algorithm, in the form of a confusion matrix, for all 32 pictures that elicited responses in this particular recording session. The colour code shows the relative number of times that a presentation of picture i (in the y axis) was predicted to be picture j (in the x axis). A perfect decoding should have all entries in the diagonal. In this case, the overall percentage of hits — the average along the diagonal — was 35.4%, which was significantly better than chance (1/32 = 3.1%) with $p<10^{-49}$ (see BOX 2). Parts **a** and **c** are reproduced, with permission, from REF. 44 © (2007) American Physiological Society.

© 2009 Macmillan Publishers Limited. All rights reserved

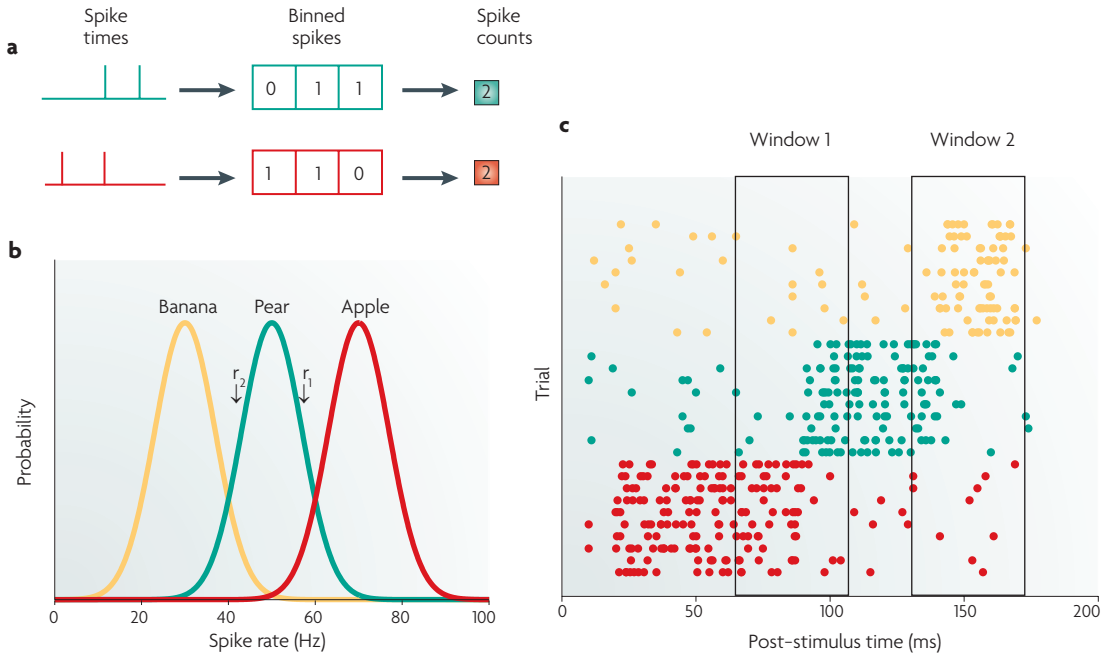


Figure 3 | **Sources of information loss. a** | A first source of information loss in the analysis of spike trains is given by a simplification of the responses — for example, when the original temporal precision is reduced (left), when the spikes are binned in time (centre) or when the total spike counts are used (right). **b** | The use of decoding algorithms gives another source of information loss, because decoders miss out information about unlikely stimuli. Given the response distributions $P(r|s)$ of a (simulated) visual neuron responding to apple, pear and banana stimuli, when the neuron fires a response r_1 just above the average response to pear, a pear will be decoded. However, this particular neural response not only informs us that the pear is the most likely stimulus, it also informs us that a banana is very unlikely. Similarly, when the neuron fires a response r_2 just below the average response to pears, a decoder will again predict that a pear was presented, but it will miss the information that it is very unlikely that an apple was presented. **c** | More generally, information is lost when mistakes about the true neural response probabilities are made. Shown are the spike rasters of a simulated single neuron to the three different stimuli. If the observer has to decode the responses in time window 1, but does not know the precise post-stimulus time at which the considered responses were emitted, it might set the decoder using the wrong response probabilities (for example, those corresponding to window 2 rather than window 1). The stimulus reconstruction will then be flawed and information will be lost.

whether the information given by the combination of the stimulus prediction and its uncertainty, $I(S;S^pU)$, is greater than $I(S;S^p)$ and how much it accounts for all the information $I(S;R)$ carried by the neuronal response.

Limitations of decoding and information theory. The main limitation of decoding is that it does not consider all the potential ways to transmit information. Moreover, decoding algorithms may fail to decode stimuli owing to a high-dimensional response space or the use of incorrect assumptions about the recorded data. In such circumstances it may therefore be dangerous to rule out a candidate neuronal code only because it gives a near-chance performance with a given decoding algorithm.

One disadvantage of information-theoretic measures is that high information values in a neuronal code might not be biologically relevant because neural systems might not be capable of exploiting all of this information. This issue can be addressed by analysing the performance of decoding algorithms that incorporate some of the limitations of the postsynaptic neuronal circuitry.

Another problem of information-theoretic analyses (and, to a lesser degree, of decoding algorithms) is that the stimulus–response probabilities must be estimated empirically. This can also lead to 'limited sampling biases' — that is, pronounced systematic errors caused by limited amounts of data[36,54]. Taking the spike counts from a population of N neurons and partitioning the response of each neuron into M bins gives M^N possible values. The cardinality of the set of all possible responses R increases exponentially with the number of neurons and becomes even larger if we also consider temporal firing patterns. In real experiments, R must be sampled in a finite number of trials and not all possible response outcomes can be sufficiently sampled. Despite recent progress to correct for such biases (by using algorithms that evaluate and remove these systematic errors), when considering simultaneously recorded neurons the minimum number of trials per stimulus that is needed to obtain an unbiased information calculation approximately equals the cardinality of the response set R[36]. This dimensionality curse prevents the application of information theory to very large populations. As decoding algorithms are more data-robust and can be easily applied to recordings of larger populations, a practical way to approximate the information carried by large neuronal populations is to

© 2009 Macmillan Publishers Limited. All rights reserved

Box 3 | **Data-processing inequality and sources of information loss**

There are several sources of information loss when processing neural responses. The first is given by a simplification of the neural responses — for example, by dividing the spike trains into small consecutive time bins or by taking the spike count (FIG. 3a). These operations can be formalized as a transformation of the neural response: $r \rightarrow f(r)$. By analogy with equation 5, the Shannon information about the stimulus carried by the processed response $f(r)$ can be defined as:

$$I(S;f(R)) = \sum_{s,f(r)} P(s,f(r)) \log_2 \frac{P(s,f(r))}{P(s)P(f(r))} \quad (10)$$

The data-processing inequality states that any transformation $f(r)$ can only decrease the Shannon information that was accessible from the original responses; that is, $I(S;f(R)) \leq I(S;R)$[34]. This information loss can, for example, result from ignoring temporal patterns in the spike trains or from ignoring information given by the precise coincident firing of different neurons.

Decoding can be formalized as a second transformation of the neural responses: $r \rightarrow g(f(r)) = s^P$, where s^P denotes stimulus prediction. By analogy with equation 10, the Shannon information about the stimulus carried by s^P is defined as

$$I(S;S^P) = \sum_{s,s^P} P(s,s^P) \log_2 \frac{P(s,s^P)}{P(s)P(s^P)} \quad (11)$$

This is the mutual information between the rows and columns of the confusion matrix (FIG. 2). The data-processing inequality states that the stimulus prediction using decoding algorithms gives a second source of information loss — that is: $I(S,S^P) \leq I(S,f(R)) \leq I(S,R)$, for $s^P = g(f(r))$. An example of this information loss is given in FIG. 3b (for a formal proof, see REFS 37,39).

In general, information losses occur when the observer of the neural responses (the information-extracting algorithm) uses an incorrect probability of stimuli and responses $Q(r,s)$ rather than the correct one $P(r,s)$. For example, the observer may not know the precise post-stimulus time at which current responses were emitted (FIG. 3c). Alternatively, the observer may not know that neurons are correlated and may incorrectly treat them as independent[95].

The information $I'(S;R)$ that can be extracted by an algorithm using the wrong probability $Q(r,s)$ is less than or equal too the information that is encoded by the neural responses: $I'(S;R) \leq I(S;R)$. A lower bound to $I'(S;R)$ is given by:

$$I'(S;R) \geq \sum_{s,r} P(s,r) \log_2 \frac{Q(s,r)}{P(s)Q(r)} \quad (12)$$

A formulation similar to equation 12 was first explored in REF. 32 and later developed in REF. 125. If the error in the probability distribution is made only because of a response simplification $r \rightarrow f(r)$ (that is: $Q(s,r) = P(s,f(r))$; $Q(r) = P(f(r))$), then equation 12 is an equality and becomes identical to equation 10. Similarly, if this simplification is due to decoding (that is: $Q(s,r) = P(s,s^P)$; $Q(r) = P(s^P)$), then equation 12 is equal to equation 11.

compute the mutual information from the confusion matrices. This reduces the dimensionality of R from M^N (the set of all possible responses of N neurons) to K (the number of stimuli). However, this method introduces potentially severe information losses (BOX 3) and its results should be interpreted with care.

Practical applications of population analyses

A pioneering study[41] defined a population vector — the weighted sum of the preferred tuning of each neuron — to show that although neurons in the primate motor cortex are broadly tuned to movement directions, the activity of a population of these neurons could be used to accurately predict the direction of arm movements. Subsequent studies have shown that, in different systems and conditions, it is possible to infer information about a stimulus or behaviour using population vectors or the more comprehensive decoding approaches described in the previous sections[10,14]. We describe a few of these studies, which illustrate how more information can be obtained from a population analysis than from the study of single cells in isolation.

Attention or intention in the posterior parietal cortex? Evidence has shown that neurons in the posterior parietal cortex are involved in movement planning[55]. Reach and delay saccade tasks (FIG. 4) showed that neurons in the lateral intraparietal area (LIP) are involved in saccade plans[56–58] and that neurons in the parietal reach region (PRR) are involved in planning arm reaches[55,58]. However, it has also been argued that these neurons are not involved in movement planning and are exclusively driven by attention to the target location[59,60], and that changes in the responses to reaches and saccades may be due to the different attention loads of each task.

This controversy is difficult to resolve because the information given by single neurons is ambiguous[55,59]. However, a decoding analysis with a population of PRR and LIP neurons revealed that neurons in these areas did encode movement plans. LIP neurons predicted saccades significantly better than reaches, whereas neurons in the PRR predicted reaches significantly better than saccades[61]. Furthermore, predictions of movement intentions were significantly better than predictions of the locus of attention. In no case was a reach confused with a saccade or vice versa, and predictions of both reaches and saccades were nearly perfect. Other studies gave further evidence of the possibility of predicting movement plans from the activity of neurons in the posterior parietal cortex[62–64]. These predictions might be useful in developing prosthetic devices for paralyzed patients or amputees[3,65–67].

Odour identity and concentration in the locust. Further evidence of the value of the population analysis comes from another study that examined whether odour identity and intensity could be determined from the firing pattern of a population of antennal lobe projection neurons in the locust[68]. Single-cell results were inconclusive because different odour concentrations changed the firing patterns of these neurons. However, when the whole neuronal population was considered it was possible to disentangle this ambiguity and establish stimulus identity for different concentrations.

Extracting the stimulus features encoded by the neuronal population. Another example of the use of decoding algorithms to assess what stimulus features are encoded in the neuronal population comes from the analysis of the responses of single cells in the human medial temporal lobe to picture presentations. Neurons in this area have been suggested to encode the identity of the person or object eliciting responses, rather than basic visual features[24]. FIGURE 2 shows the confusion matrix obtained when a decoding algorithm predicted the presentation of 32 different pictures based on the activity of 19 simultaneously recorded units. It was possible to predict which person or object was shown each time (FIG. 2c), but it was not possible to distinguish between different pictures of the same person or object[44].

© 2009 Macmillan Publishers Limited. All rights reserved

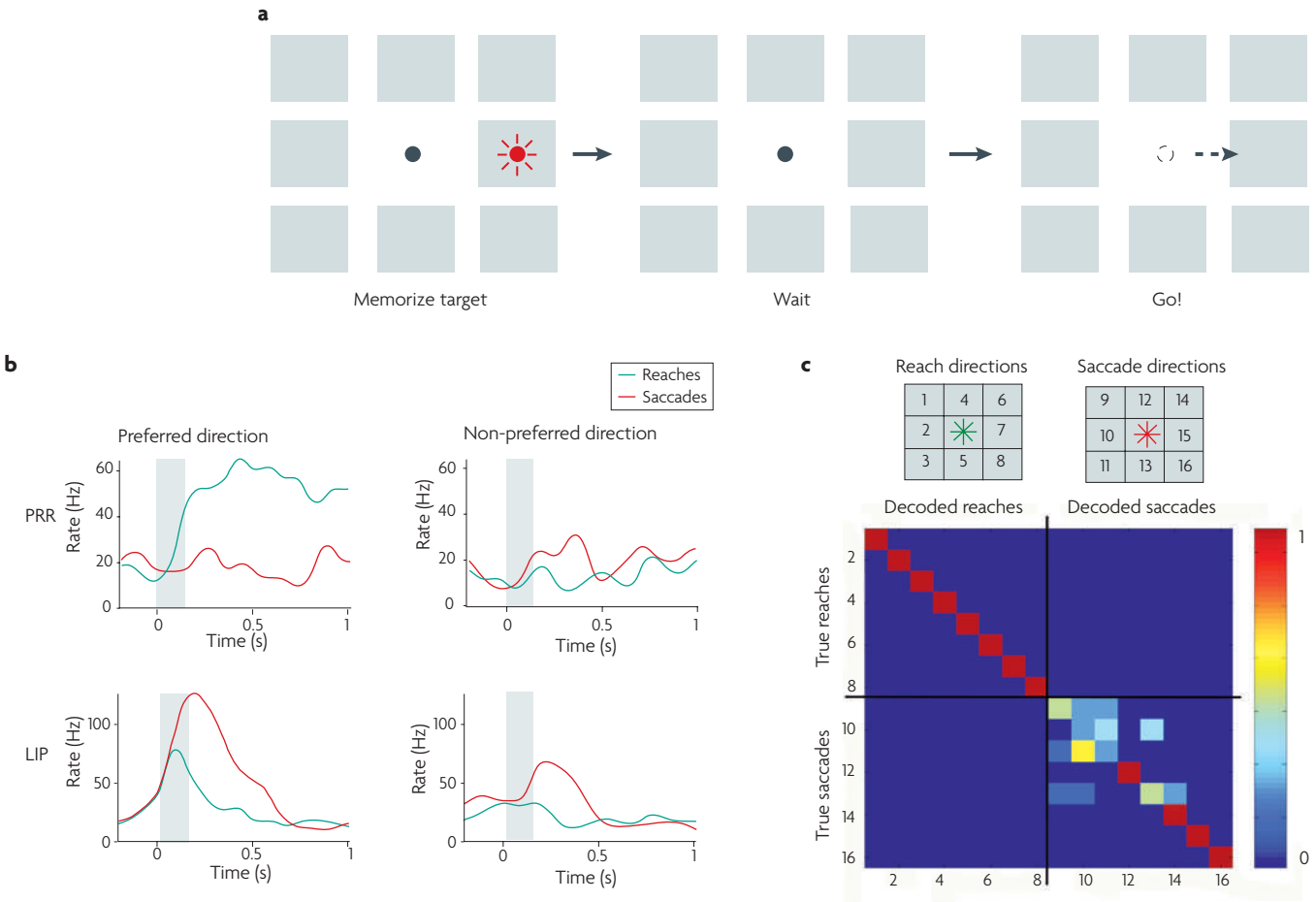


Figure 4 | **Decoding arm reaches and saccades to eight different directions. a** | In a delay reach and saccade task a monkey is shown a briefly flashed target in one of eight possible locations and, after a go signal (the disappearance of the initial fixation point), has to perform either a reach or a saccade depending on the colour of the target (green or red, respectively). **b** | Average firing rates of a neuron in the parietal reach region (PRR) and a neuron in the lateral intraparietal area (LIP) to delay reaches and delay saccades in their preferred and non-preferred directions. The neuron in PRR is tuned to the preferred direction for reaches and has a lower firing rate for saccades. Conversely, the neuron in LIP is tuned to the preferred direction for saccades and has a lower firing rate for reaches. **c** | Decoding of all 8 directions for reaches and saccades using the whole population of 47 PRR and 32 LIP neurons with a nearest-neighbour algorithm. The colour code in the confusion matrix shows the relative number of times that a reach or saccade *i* (along the *y* axis) was predicted to be *j* (along the *x* axis). In no case was a reach confused with a saccade or vice versa. Predictions of both reaches and saccades were nearly perfect. The only exception was for saccades to the left, as recordings were made in the left hemisphere and saccades are usually encoded in the contralateral field. Parts **b** and **c** are modified, with permission, from REF. 61 © (2006) Society for Neuroscience.

The possibility of decoding information from human neuronal population recordings has relevance for the development of brain–machine interfaces[66,67], and such decoding provides information that is not apparent from the analysis of single-cell recordings. In experiments like those shown in FIG. 2 it was found that: first, spike sorting significantly improved predictions[44], stressing the value of optimal processing of the recordings (BOX 1); second, decoding performances were the same when all neurons or only those that produced significant responses were considered, showing that unresponsive neurons did not carry relevant information; third, each responsive neuron carried an average of 0.25 bits of information about the stimulus; and fourth, most of the information about the pictures was given by an average of only 4 spikes fired between 300 and 600 ms after the stimulus onset in a relatively small number of neurons.

The features that are encoded by a population can also be assessed with the information-theory approach. For example, by grouping stimulus features and calculating the amount of information obtained in each case it was found that neurons in the rat barrel cortex encoded a conjunction of the stimulus frequency and amplitude (the product of both, and not each of these parameters independently) when stimulated with sinusoidal whisker vibrations[69].

Temporal coding of sensory information. Two hypotheses have been proposed to explain how neurons encode information. The spike-count hypothesis states that

Brain–machine interface
A direct communication link between a brain (human or animal) and an external device, such as a prosthetic limb or a sensing device.

© 2009 Macmillan Publishers Limited. All rights reserved

neurons represent information only by changing their firing rates over some relevant time window[12,70]. The temporal-coding hypothesis states that the precise timing of spikes adds important information to the information given by the spike counts[45,71–73].

Determining which code best represents information is analogous to evaluating the temporal precision with which spikes provide knowledge about stimuli. High temporal precision favours temporal coding, whereas rough temporal precision favours spike-count codes. The temporal precision can be established by taking transformations of the original response $r \rightarrow f(r)$ that eliminate any information potentially carried in the temporal pattern of the spike trains (BOX 3). This can be achieved by binning the neuronal responses (FIG. 3a) or by randomly jittering the spike times within some time range. Then, using equation 10, we can measure how information, or decoding performance, changes with the timing precision. Temporal information can also be identified by comparing the information given by the total spike counts with that contained in the temporal waveform of the responses — determined, for example, using principal-component analysis[71,74].

Studies in subcortical structures and primary sensory cortices have reported that the timing of spikes adds information to that given by spike counts, even down to a timing precision of the order of a few milliseconds[72,75–79]. Evidence for temporal coding is less established in higher areas[80] (but see REFS 71,81–83 and FIG. 5d,e). Most studies on temporal coding have concentrated on single brain areas and only a few have reported how timing codes are transformed across different stages of a sensory pathway[84–88]. One such study considered how the temporal coding of naturalistic whisker stimuli is transformed across the rat somatosensory pathway, and found that millisecond-precise timing conveyed high amounts of information both in the ganglion and in the somatosensory cortex[89]. In the ganglion the temporal code was supported by precise interspike intervals, whereas in the cortex interspike intervals were variable and the temporal code was based on the times of transient responses measured from the time of the whisker deflections. This 'latency code' has been reported widely at the cortical level[76,90–92]. However, it is unclear whether or how the latency code is used by the brain, given that the measurement of these response latencies requires knowledge of the stimulus time.

The extent to which information in response latencies is accessible to the brain can be addressed with information theory, by considering how much information can be extracted with an imprecise knowledge of the stimulus time. For this, one can use an approximate stimulus–response distribution obtained by shifting the spike trains within the range of the time uncertainty of the stimulus, thus generating a loss of information if stimulus timing is crucial (see equation 12 and FIG. 3c). In the study described above[89], information from the spike timing in the cortex required a precise knowledge of the stimulus time, whereas information from the ganglion did not. This finding can be interpreted by noting that the cortex might estimate stimulus time using the output of the motor system[93] — exploiting the advantages of latency codes, namely scale invariance and simplicity for downstream computations[73] — whereas ganglion cells and their direct targets are unlikely to receive stimulus time information and may be forced to deliver information in other ways.

Assessing the importance of correlations. Given the technical difficulties in acquiring and processing simultaneously recorded data from a large number of neurons (BOX 1), it is common to obtain the population activity by recording one neuron after another under exactly the same experimental paradigm. However, this assumes that neurons fire independently and ignores the role of correlated firing in neuronal information processing. It is therefore important to establish how realistic the information obtained by recording one cell at a time is, compared with that obtained by recording the neurons simultaneously. This issue has been covered in detail elsewhere[13,94–97].

With simultaneous recordings, the role of correlations can be assessed by comparing the information about the stimuli from the recorded population and that obtained after eliminating the correlations across cells. The latter is computed through equation 5 by replacing the true response probability, $P(r,s)$, with $P_{sh}(r,s)$, which is obtained by shuffling the trials corresponding to each stimulus independently for each neuron. This procedure has shown that correlations can either reduce or increase the overall information content of a population. Shuffling the neural responses typically increases the information when positively correlated neurons — those that tend to increase or decrease their firing together — with similar stimulus selectivities are considered[13,94,98]. Alternatively, removing correlations by shuffling can also decrease the information given by a neuronal population when, for example, neurons are positively correlated but have opposite stimulus selectivities or when correlations change from stimulus to stimulus to tag particular sensory features[99]. For decoding, the performance in the absence of correlations can be also estimated by shuffling both training and test data sets[44].

In practice, studies on small populations indicate that noise correlations make a small quantitative difference in stimulus coding[44,95,100–102] (but see REF. 103 for an exception). However, analytical extrapolations suggest that this effect may be important for large neuronal populations[95,98]. Testing these scaling predictions in real data requires the development of improved information-theory or decoding algorithms that are optimized for large populations.

Merging information from different neuronal signals. Measures of neuronal activity such as LFPs reflect a number of subthreshold integrative processes that are not reflected in spike trains, including input and intracortical population synaptic potentials[104,105], spike afterpotentials and voltage-dependent membrane oscillations[106,107]. Thus, the simultaneous recording and analysis of spikes and LFPs (BOX 1) may provide complementary information about the overall network activity.

Principal-component analysis
A linear transformation that projects the data on to an orthogonal base, in which the greatest variance of the data lies on the first coordinate (the first principal component), the second greatest variance on the second coordinate, and so on. It is usually used to reduce the dimensionality of complex data.

Spike afterpotential
A transient hyperpolarization of a neuron following the firing of an action potential. It is caused by K^+ channels, which open during the spike and close a few milliseconds after the neural membrane potential goes back to its resting value.

© 2009 Macmillan Publishers Limited. All rights reserved

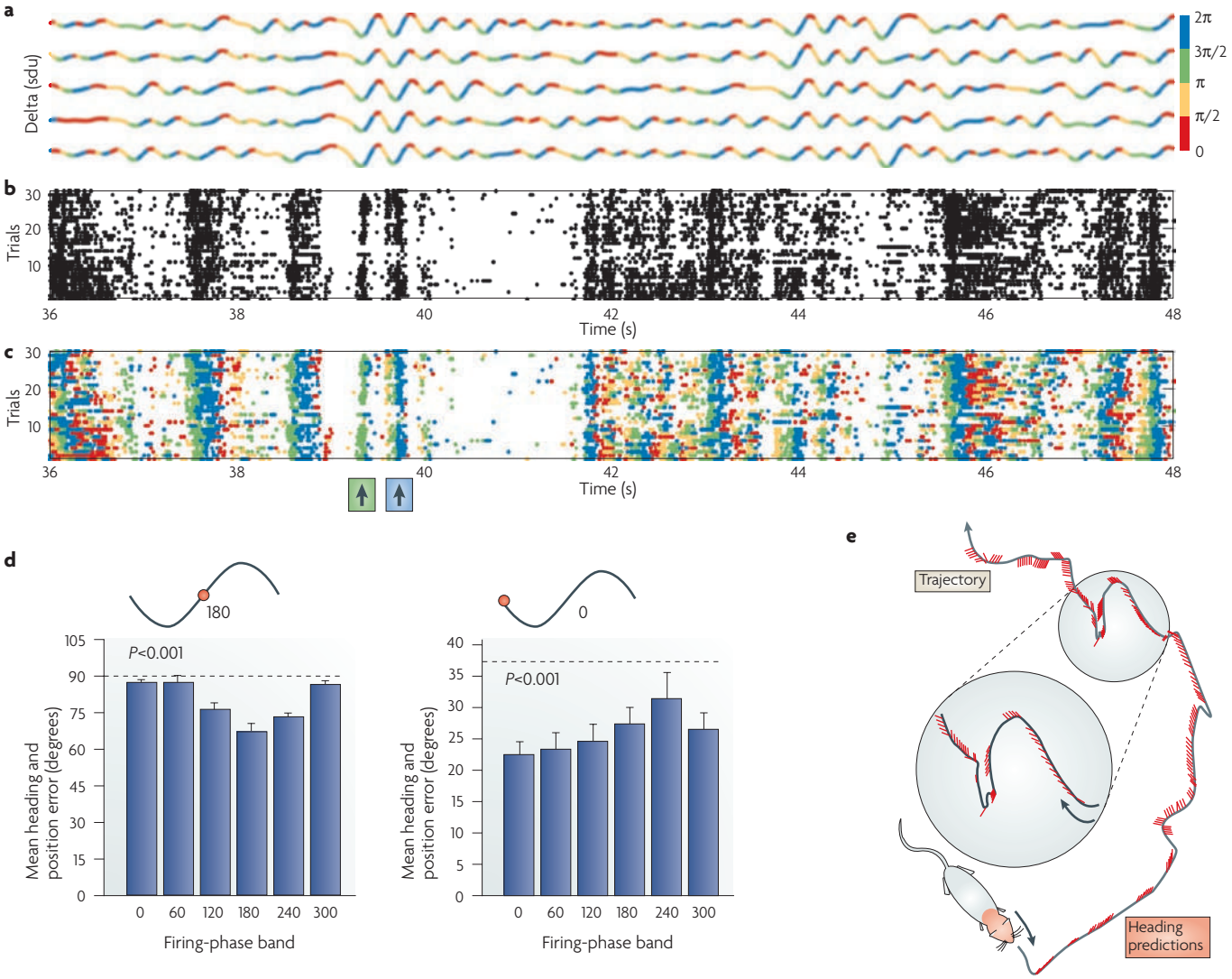


Figure 5 | **Encoding of information by the local field potential phase. a** | Delta band (1–4Hz) local field potential (LFP) traces (measured in standard-deviation units (sdu)) in the monkey primary visual cortex during five presentations of a naturalistic colour movie. The line colour denotes the instantaneous LFP phase (phase range is divided into quarters: 0–π/2, π/2–π, π–3π/2 and 3π/2–2π). **b** | Spike times (indicated by dots) elicited with 30 presentations of the same movie. **c** | The same spike times as in part **b**, but with the spikes coloured according to the concurrent LFP phase. The movie scenes indicated by green and blue arrows can be much better discriminated from each other using the phase of firing (coloured spikes) than using the spike counts (black spikes). **d** | Mean heading and position errors that occur when the trajectory of a rat in a two-dimensional environment is predicted using the spikes fired by place cells at different LFP phases. Heading and position predictions were optimal when considering spikes fired at opposite theta phases (180 and 0 degrees, respectively). **e** | The movement path of the animal over a 10 s period. The heading predictions (marked in red) were based only on spikes fired in the 180-degree phase. The inset shows an enlarged portion of the trajectory in which good predictions were made when the animal turned sharply. Parts **a**–**c** are reproduced, with permission, from REF. 108 © (2008) Cell Press. Parts **d** and **e** are reproduced, with permission, from REF. 81 © (2008) Macmillan Publishers Ltd. All rights reserved.

The population analysis can address the information content of different types of neuronal signals. For example, in the primary visual cortex of anaesthetized macaques[108,109] it was found that the presentation of naturalistic colour movies elicited reliable responses across trials both for the spikes and for the delta (1–4 Hz) LFP phase (FIG. 5a,b). In order to test whether the LFP phase at the time of spike firing conveyed information that could not be extracted from spike rates, the researchers investigated whether it was possible to disambiguate different movie scenes eliciting exactly the same firing rate using the phase of the LFP at which the spikes were fired (FIG. 5c). The phase of firing was indeed found to convey information that was not accessible from the spike rates alone: the overall information about the movie increased by 55% when the LFP phase of firing was considered. It should be noted that this type of phase coding represents another form of spike-timing information discussed in the previous sections.

© 2009 Macmillan Publishers Limited. All rights reserved

Voxel
In MRI research, a voxel refers to the smallest measured volume unit, analogous to a three-dimensional pixel. In functional MRI studies these are typically of the order of 30 mm^3, although much smaller voxel volumes have been achieved in more recent work.

The role of LFP phases in the rat hippocampus has also been debated. Single-cell studies using the global statistical relationship between LFP phase and spike firing could identify a partial dissociation of phase and firing rate in response to the animal's location and speed of movement[82] but could not determine whether these signals encode complementary information about separate external features. Other studies have also shown that both theta phase and rate reflect the amplitude of the cell's input and might convey redundant spatial information[110–112]. Interestingly, a recent population study using decoding techniques[81] clearly demonstrated that the phase of firing of hippocampal place cells encodes features related to the animal's spatial navigation that are not encoded by spike rates (FIG. 5d,e).

The finding of different tunings for LFPs and spikes has also been reported for medial temporal lobe neurons in human subjects[113] and inferotemporal neurons in the macaque cortex[114]. Interestingly, in other cortical areas, such as in the posterior parietal cortex in monkeys, it was found that LFPs and spikes carry similar information about movement plans[63,64]. Moreover, the LFPs carried more information than spikes about the time of the planned movement and the behavioural state. This has significant implications for the development of brain–machine interfaces, given that LFPs are much more stable and easier to record for large periods of time than single-cell activity[115].

Conclusions

We have described the advantages of a single-trial population analysis over traditional single-cell studies of trial-averaged responses. In particular, we have shown how much more knowledge can be extracted using decoding and information-theory methodologies and how information that is ambiguous at the single-cell level can sometimes be clearly interpreted when considering the whole population.

Decoding and information theory describe complementary aspects of knowledge extraction. Decoding has the main advantage of producing an output (the stimulus prediction) that is easy to interpret and is close to behavioural choices, whereas mutual information gives a more comprehensive quantification of the information contained in a neuronal population, by evaluating the reduction of uncertainty about the stimuli that can be obtained from the neuronal responses. The complementary knowledge offered by these two approaches is only now beginning to be investigated in neuroscience[47,54]. A more systematic joint application of both methodologies may offer additional insights.

An interesting question that could benefit from such a joint approach is comparing the information given by the prediction of the stimulus and the uncertainty of these predictions with the total information contained in the population. Representing uncertainty is important for decisions involving risks and may be fundamental for neuronal computations that take place in the presence of both sensory and neural noise[51–53]. As discussed above, combining information-theoretic and decoding approaches provides a rigorous way to quantify how much of the information provided by neuronal populations concerns the prediction of the stimulus, and how much of this information is about specific aspects of the uncertainty of these predictions. A combination of these approaches can therefore provide precise quantitative answers about how the brain deals with intrinsically noisy signals.

Information-theoretic calculations are difficult with neuronal populations because of the curse of dimensionality. Although until recently it was thought to be problematic to compute accurate information measures from the activity of more than one neuron[9], recent progress now permits the computation of the information carried by populations of up to ten neurons[36]. This enables us to examine the details of information processing in local networks, and future work may set this bar even higher. A major and important challenge for neurostatisticians is to find ways to further extend the feasibility of performing information computations with large populations. Two directions are particularly promising. First, it is important to explore the use of optimal dimensionality-reduction techniques that discard dimensions not carrying any relevant information. Second, the information-bias problem for large populations is exacerbated by the fact that the activity of neurons is often correlated. This problem can be diminished by simplifying the correlation structure — for example, by considering pair-wise correlations between neurons and ignoring higher-order interactions[116].

A neuroscientific revolution has been triggered by the use of imaging techniques such as functional MRI. How functional MRI signals correlate with the activity of single neurons and LFPs is far from completely understood[117]. However, all these signals give complementary information, with different coverage and different temporal and spatial resolutions. Decoding and information theory offer elegant ways to combine all of these signals and quantify their information and redundancy. Interestingly, a few studies have already shown the value of decoding for extracting information from populations of voxels[118,119], and other studies have compared the information given by spikes and LFPs. It remains to be seen how much more information can be extracted from a single-trial combined analysis of the neurophysiological signals (spikes and LFPs) and the voxel activity, an enterprise that has only recently been initiated[120,121].

1. Hubel, D. Tungsten microelectrode for recording from single units. *Science* **125**, 549–550 (1957).
2. Kandel, E. R., Schwartz, J. H. & Jessell, T. M. *Principles of Neural Science* (McGraw Hill, New York, 2000).
3. Schwartz, A. B. Cortical neural prosthetics. *Annu. Rev. Neurosci.* **27**, 487–507 (2004).
4. Buzsaki, G. Large-scale recording of neuronal ensembles. *Nature Neurosci.* **7**, 446–451 (2004).
5. Csicsvari, J. *et al.* Massively parallel recording of unit and local field potentials with silicon-based electrodes. *J. Neurophysiol.* **90**, 1314–1323 (2003).
6. Kelly, R. C. *et al.* Comparison of recordings from microelectrode arrays and single electrodes in the visual cortex. *J. Neurosci.* **27**, 261–264 (2007).
7. Rousche, P. J. & Normann, R. A. Chronic recording capability of the Utah intracortical electrode in cat sensory cortex. *J. Neurosci. Methods* **82**, 1–15 (1998).
8. Blanche, T. J., Spacek, M. A., Hetke, J. F. & Swindale, N. V. Polytrodes: high-density silicon electrode arrays for large-scale multiunit recording. *J. Neurophysiol.* **93**, 2987–3000 (2005).
9. Brown, E. N., Kass, R. E. & Mitra, P. P. Multiple neural spike train data analysis: state-of-the-art and future challenges. *Nature Neurosci.* **7**, 456–461 (2004).
10. Abbott, L. F. Decoding neuronal firing and modelling neural networks. *Q. Rev. Biophys.* **27**, 291–331 (1994).

© 2009 Macmillan Publishers Limited. All rights reserved

11. Pouget, A., Dayan, P. & Zemel, R. Information processing with population codes. *Nature Rev. Neurosci.* **1**, 125–132 (2000).
12. Rieke, F., Warland, D., de Ruyter van Steveninck, R. R. & Bialek, W. *Spikes: Exploring the Neural Code* (MIT Press, Cambridge, Massachusetts, 1997).
13. Oram, M. W., Foldiak, P., Perrett, D. I. & Sengpiel, F. The 'ideal homunculus': decoding neural population signals. *Trends Neurosci.* **21**, 259–265 (1998).
14. Dayan, P. & Abbott, L. F. *Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems* (MIT Press, Cambridge, Massachusetts, 2001).
15. Shannon, C. E. A mathematical theory of communication. *Bell Syst. Tech. J.* **27**, 379–423 & 623–656 (1948).
16. Deco, G. & Obradovic, D. *An Information-Theoretic Approach to Neural Computing* (Springer, Berlin, 1997).
17. Borst, A. & Theunissen, F. E. Information theory and neural coding. *Nature Neurosci.* **2**, 947–957 (1999).
18. Quian Quiroga, R. Spike sorting. *Scholarpedia* **2**, 3583 (2007).
A short review describing the steps for processing neural data, basically focused on spike detection and sorting.
19. Quian Quiroga, R., Nadasdy, Z. & Ben-Shaul, Y. Unsupervised spike detection and sorting with wavelets and superparamagnetic clustering. *Neural Comput.* **16**, 1661–1687 (2004).
20. Lewicki, M. A review of methods for spike sorting: the detection and classification of neural action potentials. *Network* **9**, R53–R78 (1998).
21. Harris, K. D., Henze, D. A., Csicsvari, J., Hirase, H. & Buzsaki, G. Accuracy of tetrode spike separation as determined by simultaneous intracellular and extracellular measurements. *J. Neurophysiol.* **84**, 401–414 (2000).
22. Singer, W. & Gray, C. M. Visual feature integration and the temporal correlation hypothesis. *Annu. Rev. Neurosci.* **18**, 555–586 (1995).
23. Engel, A. K. & Singer, W. Temporal binding and the neural correlates of sensory awareness. *Trends Cogn. Sci.* **5**, 16–25 (2001).
24. Quian Quiroga, R., Reddy, L., Kreiman, G., Koch, C. & Fried, I. Invariant visual representation by single neurons in the human brain. *Nature* **435**, 1102–1107 (2005).
25. Reich, D. S., Mechler, F. & Victor, J. D. Independent and redundant information in nearby cortical neurons. *Science* **294**, 2566–2568 (2001).
26. Barlow, H. B., Hill, R. M. & Levick, W. R. Retinal ganglion cells responding selectively to direction and speed of image motion in the rabbit. *J. Physiol. (Lond.)* **173**, 377–407 (1964).
27. MacKay, D. M. *Information Theory, Inference, and Learning Algorithms* (Cambridge Univ. Press, Cambridge, 2003).
28. Foldiak, P. in *Computation and Neural Systems* (eds Eeckman, F. H. & Bower, J.) 55–60 (Kluwer, Norwell, Massachusetts, 1993).
29. Sanger, T. D. Probability density estimation for the interpretation of neural population codes. *J. Neurophysiol.* **76**, 2790–2793 (1996).
30. Paradiso, M. A. A theory for the use of visual orientation information which exploits the columnar structure of striate cortex. *Biol. Cybern.* **58**, 35–49 (1988).
31. Duda, O. H., Hart, P. E. & Stork, D. G. *Pattern Classification* (Wiley & sons, New York, 2001).
32. Kjaer, T. W., Hertz, J. A. & Richmond, B. J. Decoding cortical neuronal signals: network models, information estimation and spatial tuning. *J. Comp. Neurosci.* **1**, 109–139 (1994).
33. Averbeck, B. B. in *Coherent Behavior in Neuronal Networks* (ed. Rubin, J., Matias, M. and Romo, R.) (Springer, New York, in the press).
34. Cover, T. M. & Thomas, J. A. *Elements of Information Theory.* (Wiley & sons, Hoboken, New Jersey, 2006).
35. Victor, J. D. Approaches to information-theoretic analysis of neural activity. *Biol. Theory* **1**, 302–316 (2006).
36. Panzeri, S., Senatore, R., Montemurro, M. A. & Petersen, R. S. Correcting for the sampling bias problem in spike train information measures. *J. Neurophysiol.* **98**, 1064–1072 (2007).
37. Samengo, I. Information loss in an optimal maximum likelihood decoding. *Neural Comput.* **14**, 771–779 (2002).
38. Treves, A. On the perceptual structure of face space. *Biosystems* **40**, 189–196 (1997).
One of the first papers to analyse the information given by the confusion matrix, showing that the distribution of incorrect stimulus predictions is important for understanding neural representations. This work was later extended to the study of how incorrect stimulus predictions relate to the mutual information between stimuli and responses in reference 47.
39. Panzeri, S., Treves, A., Schultz, S. & Rolls, E. T. On decoding the responses of a population of neurons from short time windows. *Neural Comput.* **11**, 1553–1577 (1999).
40. Robertson, R. G., Rolls, E. T., Georges-Francois, P. & Panzeri, S. Head direction cells in the primate pre-subiculum. *Hippocampus* **9**, 206–219 (1999).
41. Georgopoulos, A. P., Schwartz, A. B. & Kettner, R. E. Neuronal population coding of movement direction. *Science* **233**, 1416–1419 (1986).
The first study to implement a population analysis of neuronal responses by using a population vector.
42. Zhang, K., Ginzburg, I., McNaughton, B. L. & Sejnowski, T. J. Interpreting neuronal population activity by reconstruction: unified framework with application to hippocampal place cells. *J. Neurophysiol.* **79**, 1017–1044 (1998).
43. Knill, D. C. & Pouget, A. The bayesian brain: the role of uncertainty in neural coding and computation. *Trends Neurosci.* **27**, 712–719 (2004).
44. Quian Quiroga, R., Reddy, L., Koch, C. & Fried, I. Decoding visual inputs from multiple neurons in the human temporal lobe. *J. Neurophysiol.* **98**, 1997–2007 (2007).
This paper showed that it is possible to correctly predict picture presentations from the firing of neurons in the human medial temporal lobe far above chance. The authors also showed that more information can be extracted from a population analysis than from a single-cell study.
45. Victor, J. D. & Purpura, K. P. Nature and precision of temporal coding in visual cortex: a metric-space analysis. *J. Neurophysiol.* **76**, 1310–1326 (1996).
46. Schnupp, J. W. H., Hall, T. M., Kokelaar, R. F. & Ahmed, B. Plasticity of temporal pattern codes for vocalization stimuli in primary auditory cortex. *J. Neurosci.* **26**, 4785–4795 (2006).
47. Thomson, E. E. & Kristan, W. B. Quantifying stimulus discriminability: a comparison of information theory and ideal observer analysis. *Neural Comput.* **17**, 741–778 (2005).
48. Ma, W. J., Beck, J. M., Latham, P. E. & Pouget, A. Bayesian inference with probabilistic population codes. *Nature Neurosci.* **9**, 1432–1438 (2006).
49. Beck, J. M. *et al.* Probabilistic population codes for Bayesian decision making. *Neuron* **60**, 1142–1152 (2008).
50. Averbeck, B. B., Sohn, J.-W. & Lee, D. Activity in prefrontal cortex during dynamic selection of action sequences. *Nature Neurosci.* **9**, 276–282 (2006).
51. Pouget, A., Zhang, K., Deneve, S. & Latham, P. E. Statistically efficient estimation using population coding. *Neural Comput.* **10**, 373–401 (1998).
52. Sahani, M. & Dayan, P. Doubly distributional population codes: simultaneous representation of uncertainty and multiplicity. *Neural Comput.* **15**, 2255–2279 (2003).
53. Pouget, A., Zemel, R. & Dayan, P. Inference and computation with population code. *Annu. Rev. Neurosci.* **26**, 381–410 (2003).
54. Victor, J. D. & Nirenberg, S. Indices for testing neural codes. *Neural Comput.* **20**, 2895–2936 (2008).
The authors analysed in detail the relative strengths and weaknesses of information theory and Bayesian decoders when they are used to rule out neural codes. They also introduced a set of measures that varied smoothly between information theory and Bayesian decoders.
55. Andersen, R. A. & Buneo, C. A. Intentional maps in the posterior parietal cortex. *Annu. Rev. Neurosci.* **25**, 189–220 (2002).
56. Mountcastle, V. B., Lynch, J. C., Georgopoulos, A., Sakata, H. & Acuna, C. Posterior parietal association cortex of the monkey: command functions for operations within extrapersonal space. *J. Neurophysiol.* **38**, 871–908 (1975).
57. Andersen, R. A., Essick, G. K. & Siegel, R. M. Neurons of area 7 activated by both visual stimuli and oculomotor behavior. *Exp. Brain Res.* **67**, 316–322 (1987).
58. Snyder, L. H., Batista, A. P. & Andersen, R. A. Coding of intention in the posterior parietal cortex. *Nature* **386**, 167–170 (1997).
59. Bisley, J. W. & Goldberg, M. E. Neuronal activity in the lateral intraparietal area and spatial attention. *Science* **299**, 81–81 (2003).
60. Robinson, D. L., Goldberg, M. E. & Stanton, G. B. Parietal association cortex in the primate: sensory mechanisms and behavioral modulations. *J. Neurophysiol.* **41**, 910–932 (1978).
61. Quian Quiroga, R., Snyder, L. H., Batista, A. P., Cui, H. & Andersen, R. A. Movement intention is better predicted than attention in the posterior parietal cortex. *J. Neurosci.* **26**, 3615–3620 (2006).
Using a decoding population analysis, the authors showed that neurons in the posterior parietal cortex encode different movement plans and not just attention to target location.
62. Musallam, S., Corneil, B., Greger, B., Scherberger, H. & Andersen, R. A. Cognitive control signals for neural prosthetics. *Science* **305**, 258–262 (2004).
63. Pesaran, B., Pezaris, J., Sahani, M., Mitra, P. M. & Andersen, R. A. Temporal structure in neuronal activity during working memory in macaque parietal cortex. *Nature Neurosci.* **5**, 805–811 (2002).
64. Scherberger, H., Jarvis, M. & Andersen, R. A. Cortical local field potential encodes movement intentions in the posterior parietal cortex. *Neuron* **46**, 347–354 (2005).
65. Wessberg, J. *et al.* Real-time prediction of hand trajectory by ensembles of cortical neurons in primates. *Nature* **408**, 361–365 (2000).
66. Andersen, R. A., Burdick, J. W., Musallam, S., Pesaran, B. & Cham, J. G. Cognitive neural prosthetics. *Trends Cogn. Sci.* **8**, 486–493 (2004).
67. Nicolelis, M. A. Actions from thoughts. *Nature* **409**, 403–407 (2001).
68. Stopfer, M., Jayaraman, V. & Laurent, G. Intensity versus identity coding in an olfactory system. *Neuron* **39**, 991–1004 (2003).
69. Arabzadeh, E., Panzeri, S. & Diamond, M. E. Whisker vibration information carried by rat barrel cortex neurons. *J. Neurosci.* **24**, 6011–6020 (2004).
70. Adrian, E. D. *The Basis of Sensations* (Norton, New York, 1928).
71. Optican, L. M. & Richmond, B. J. Temporal encoding of two-dimensional patterns by single units in primate inferior temporal cortex. III. Information theoretic analysis. *J. Neurophysiol.* **57**, 162–178 (1987).
The first study to demonstrate that the time profile of cortical spike trains encodes much more information about static visual features than spike counts.
72. de Ruyter van Steveninck, R. R., Lewen, G. D., Strong, S. P. & Koberle, R. Reproducibility and variability in neural spike trains. *Science* **275**, 1805–1808 (1997).
73. Hopfield, J. J. Pattern recognition computation using action potential timing for stimulus representation. *Nature* **376**, 33–36 (1995).
74. Theunissen, F. & Miller, J. P. Temporal encoding in nervous systems: a rigorous definition. *J. Comp. Neurosci.* **2**, 149–162 (1995).
75. Victor, J. D. Temporal aspects of neural coding in the retina and lateral geniculate. *Network* **10**, R1–R66 (1999).
76. Panzeri, S., Petersen, R. S., Schultz, S. R., Lebedev, M. & Diamond, M. E. The role of spike timing in the coding of stimulus location in rat somatosensory cortex. *Neuron* **29**, 769–777 (2001).
77. Reinagel, P. & Clay Reid, R. Temporal coding of visual information in the thalamus. *J. Neurosci.* **20**, 5392–5400 (2000).
78. Berry, M. J., Warland, D. K. & Meister, M. The structure and precision of retinal spike trains. *Proc. Natl Acad. Sci. USA* **94**, 5411–5416 (1997).
79. Engineer, C. T. *et al.* Cortical activity patterns predict speech discrimination ability. *Nature Neurosci.* **11**, 603–608 (2008).
80. Shadlen, M. N. & Movshon, J. A. Synchrony unbound: a critical evaluation of the temporal binding hypothesis. *Neuron* **24**, 67–77; 111–125 (1999).
81. Huxter, J. R., Senior, T. J., Allen, K. & Csicsvari, J. Theta phase–specific codes for two-dimensional position, trajectory and heading in the hippocampus. *Nature Neurosci.* **11**, 587–594 (2008).
The authors recorded spiking activity and local field oscillations in the hippocampus of rats moving through a two-dimensional environment. Using a decoding analysis, they demonstrated that spikes emitted at different oscillation phases represent complementary aspects of the animal's trajectory.
82. Huxter, J. R., Burgess, N. & O'Keefe, J. Independent rate and temporal coding in hippocampal pyramidal cells. *Nature* **425**, 828–832 (2003).

© 2009 Macmillan Publishers Limited. All rights reserved

83. Buracas, G. T., Zador, A., DeWeese, M. & Albright, T. Efficient encoding of rapidly varying stimuli by motion-sensitive neurons in MT of alert monkeys. *Neuron* **20**, 959–969 (1998).
84. Kara, P., Reinagel, P. & Clay Reid, R. Low response variability in simultaneously recorded retinal, thalamic, and cortical neurons. *Neuron* **27**, 635–646 (2000).
85. Theunissen, F. E. From synchrony to sparseness. *Trends Neurosci.* **26**, 61–64 (2003).
86. Perez-Orive, J. *et al.* Oscillations and sparsening of odor representations in the mushroom body. *Science* **297**, 359–365 (2002).
87. Hahnloser, R. H. R., Kozhevnikov, A. A. & Fee, M. S. An ultra-sparse code underlies the generation of neural sequences in a songbird. *Nature* **419**, 65–70 (2002).
88. Salinas, E., Hernandez, A., Zainos, A. & Romo, R. Periodicity and firing rate as candidate neural codes for the frequency of vibrotactile stimuli. *J. Neurosci.* **20**, 5503–5515 (2000).
89. Arabzadeh, E., Panzeri, S. & Diamond, M. E. Deciphering the spike train of a sensory neuron: counts and temporal patterns in the rat whisker pathway. *J. Neurosci.* **26**, 9216–9226 (2006).
This paper developed an analytical formalism that quantifies how much sensory information can be extracted from the spike times when the stimulus time is not known precisely.
90. Reich, D. S., Mechler, F. & Victor, J. D. Temporal coding of contrast in primary visual cortex: when, what, and why. *J. Neurophysiol.* **85**, 1039–1050 (2001).
91. Gawne, T. J., Kjaer, T. W. & Richmond, B. J. Latency: another potential code for feature binding in striate cortex. *J. Neurophysiol.* **76**, 1356–1360 (1996).
92. Muller, J. R., Metha, A. B., Krauskopf, J. & Lennie, P. Information conveyed by onset transients in responses of striate cortical neurons. *J. Neurosci.* **21**, 6978–6990 (2001).
93. Ahrens, K. F. & Kleinfeld, D. Current flow in vibrissa motor cortex can phase-lock with exploratory rhythmic whisking in rat. *J. Neurophysiol.* **92**, 1700–1707 (2004).
94. Abbott, L. F. & Dayan, P. The effect of correlated variability on the accuracy of a population code. *Neural Comput.* **11**, 91–101 (1999).
95. Averbeck, B. B., Latham, P. E. & Pouget, A. Neural correlations, population coding and computation. *Nature Rev. Neurosci.* **7**, 358–366 (2006).
An up-to-date and comprehensive review of the role of noise correlations in population coding.
96. Nirenberg, S. & Latham, P. E. Decoding neuronal spike trains: how important are correlations? *Proc. Natl Acad. Sci. USA* **100**, 7348–7353 (2003).
97. Salinas, E. & Sejnowski, T. J. Correlated neuronal activity and the flow of neural information. *Nature Rev. Neurosci.* **2**, 539–550 (2001).
98. Zohary, E., Shadlen, M. & Newsome, W. Correlated neuronal discharge rate and its implications for psychophysical performance. *Nature* **370**, 140–143 (1994).
99. Panzeri, S., Schultz, S., Treves, A. & Rolls, E. T. Correlations and the encoding of information in the nervous system. *Proc. R. Soc. Lond. B Biol. Sci.* **266**, 1001–1012 (1999).
100. Petersen, R. S., Panzeri, S. & Diamond, M. E. Population coding of stimulus location in rat somatosensory cortex. *Neuron* **32**, 503–514 (2001).
101. Montani, F., Kohn, A., Smith, M. A. & Schultz, S. R. The role of correlations in direction and contrast coding in the primary visual cortex. *J. Neurosci.* **27**, 2338–2348 (2007).
102. Nirenberg, S., Carcieri, S. M., Jacobs, A. L. & Latham, P. E. Retinal ganglion cells act largely as independent encoders. *Nature* **411**, 698–701 (2001).
103. Dan, Y., Alonso, J. M., Usrey, W. M. & Clay Reid, R. Coding of visual information by precisely correlated spikes in the lateral geniculate nucleus. *Nature Neurosci.* **1**, 501–507 (1998).
104. Mitzdorf, U. Properties of the evoked potential generators: current source-density analysis of visually evoked potentials in the cat cortex. *Int. J. Neurosci.* **33**, 33–59 (1987).
105. Juergens, E., Guettler, A. & Eckhorn, R. Visual stimulation elicits locked and induced gamma oscillations in monkey intracortical and EEG-potentials, but not in human EEG. *Exp. Brain Res.* **129**, 247–259 (1999).
106. Harada, Y. & Takahashi, T. The calcium component of the action potential in spinal motoneurones of the rat. *J. Physiol.* **335**, 89–100 (1983).
107. Kamondi, A., Acsady, L., Wang, X. J. & Buzsaki, G. Theta oscillations in somata and dendrites of hippocampal pyramidal cells *in vivo*: activity-dependent phase-precession of action potentials. *Hippocampus* **8**, 244–261 (1998).
108. Montemurro, M., Rasch, M. J., Murayama, Y., Logothetis, N. K. & Panzeri, S. Phase-of-firing coding of natural visual stimuli in primary visual cortex. *Curr. Biol.* **18**, 375–380 (2008).
This study used information theory to demonstrate that, in the macaque primary visual cortex, the phase of slow (< 12 Hz) local field fluctuations encodes information about natural visual scenes that cannot be obtained from spike counts.
109. Belitski, A. *et al.* Low-frequency local field potentials and spikes in primary visual cortex convey independent visual information. *J. Neurosci.* **28**, 5696–5709 (2008).
110. Mehta, M. R., Lee, A. K. & Wilson, M. A. Role of experience and oscillations in transforming a rate code into a temporal code. *Nature* **417**, 741–746 (2002).
111. Harris, K. D. *et al.* Spike train dynamics predicts theta-related phase precession in hippocampal pyramidal cells. *Nature* **417**, 738–741 (2002).
112. Harris, K. D. Neural signatures of cell assembly organization. *Nature Rev. Neurosci.* **6**, 399–407 (2005).
113. Kraskov, A., Quian Quiroga, R., Reddy, L., Fried, I. & Koch, C. Local field potentials and spikes in the human medial temporal lobe are selective to image category. *J. Cogn. Neurosci.* **19**, 479–492 (2007).
114. Kreiman, G. *et al.* Object selectivity of local field potentials and spikes in the macaque inferior temporal cortex. *Neuron* **49**, 433–445 (2006).
115. Andersen, R. A., Musallam, S. & Pesaran, B. Selecting the signals for a brain-machine interface. *Curr. Opin. Neurobiol.* **14**, 720–726 (2004).
116. Schneiman, E., Berry, M. J., Segev, R. & Bialek, W. Weak pairwise correlations imply strongly correlated network states in a neural population. *Nature* **440**, 1007–1012 (2006).
117. Logothetis, N. K. What we can do and what we cannot do with fMRI. *Nature* **453**, 869–878 (2008).
118. Haxby, J. V. *et al.* Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science* **293**, 2425–2430 (2001).
119. Haynes, J.-D. & Rees, G. Decoding mental states from brain activity in humans. *Nature Rev. Neurosci.* **7**, 523–534 (2006).
120. Eichele, T. *et al.* Assessing the spatiotemporal evolution of neuronal activation with single-trial event-related potentials and functional MRI. *Proc. Natl Acad. Sci. USA* **102**, 17798–17803 (2005).
First study using a single-trial analysis to combine the temporal and spatial resolutions of electroencephalogram and functional MRI recordings.
121. Debener, S., Ullsperger, M., Siegel, M. & Engel, A. K. Single-trial EEG-fMRI reveals the dynamics of cognitive function. *Trends Cogn. Sci.* **10**, 558–563 (2006).
122. Gerstein, G. L. & Clark, W. A. Simultaneous studies of firing patterns in several neurons. *Science* **143**, 1325–1327 (1964).
123. Vapnik, R. J. *Statistical Learning Theory* (Wiley & sons, New York, 1998).
124. Soong, T. T. *Fundamentals of Probability and Statistics for Engineers* (Wiley, Sussex, 2004).
125. Latham, P. E. & Nirenberg, S. Synergy, redundancy, and independence in population codes, revisited. *J. Neurosci.* **25**, 5195–5206 (2005).
This study derived a rigorous measure of the information that is lost when using algorithms that make incorrect assumptions about the probabilities of neural responses to stimuli. It provides a useful tool for understanding which features of the stimulus–response relationship are important for information transmission.

Acknowledgements

We are very thankful to F. Montani, K. Whittingstall, J. Csicsvari, A. Mazzoni and G. Kreiman for comments, to P. Dayan for interesting discussions about uncertainty and decoding, and to all our brilliant colleagues that collaborated with us on these topics: R. Andersen, M. Diamond, I. Fried, C. Koch, N. Logothetis, C. Kayser, M. Montemurro, R. Petersen and A. Treves. We acknowledge support from the Engineering and Physical Sciences Research Council, the Medical Research Council, the Royal Society and the Italian Institute of Technology.

FURTHER INFORMATION

Quian Quiroga's homepage:
http://www.le.ac.uk/neuroengineering
Stefano Panzeri's homepage:
http://www.iit.it/stefano_panzeri

ALL LINKS ARE ACTIVE IN THE ONLINE PDF

© 2009 Macmillan Publishers Limited. All rights reserved