

Neural correlations, population coding and computation

Bruno B. Averbeck*, Peter E. Latham[†] and Alexandre Pouget*

Abstract | How the brain encodes information in population activity, and how it combines and manipulates that activity as it carries out computations, are questions that lie at the heart of systems neuroscience. During the past decade, with the advent of multi-electrode recording and improved theoretical models, these questions have begun to yield answers. However, a complete understanding of neuronal variability, and, in particular, how it affects population codes, is missing. This is because variability in the brain is typically correlated, and although the exact effects of these correlations are not known, it is known that they can be large. Here, we review studies that address the interaction between neuronal noise and population codes, and discuss their implications for population coding in general.

As in any good democracy, individual neurons count for little; it is population activity that matters. For example, as with control of eye^{1,2} and arm³ movements, visual discrimination in the primary visual cortex (V1) is much more accurate than would be predicted from the responses of single neurons⁴. This is, of course, not surprising. As single neurons are not very informative, to obtain accurate information about sensory or motor variables some sort of population averaging must be performed. Exactly how this averaging is carried out in the brain, however, and especially how population codes are used in computations (such as reaching for an object on the basis of visual cues, an action that requires a transformation from population codes in visual areas to those in motor areas), is not fully understood.

Part of the difficulty in understanding population coding is that neurons are noisy: the same pattern of activity never occurs twice, even when the same stimulus is presented. Because of this noise, population coding is necessarily probabilistic. If one is given a single noisy population response, it is impossible to know exactly what stimulus occurred. Instead, the brain must compute some estimate of the stimulus (its best guess, for example), or perhaps a probability distribution over stimuli.

The inability of population activity to perfectly represent variables raises two questions. First, just how accurately can variables be represented? And second, how does the presence of noise affect computations? Not surprisingly, the answer to both depends strongly on the nature of the neuronal noise, and especially on whether or not the noise is correlated (see BOX 1 for definitions). If the noise is uncorrelated, meaning the fluctuations in the response

of one neuron around its average are not correlated with the fluctuations of other neurons, population coding is relatively well understood. Specifically, we know which factors control the amount of information a population code contains⁴⁻⁷, how networks that receive population codes as input can be constructed so that they carry out computations optimally⁸, and even how information in a population code increases as a result of learning or attention⁹⁻¹². Unfortunately, noise in the brain is correlated, and because of this we need to take a second look at the results that have been obtained under the assumption of independent noise. (As discussed in BOX 1, 'correlated' in this article means 'noise correlated'.) For the computational work, this means extending the theories to take into account correlated noise, and for the empirical work this means assessing how attention and learning affect not only single neuron properties, such as tuning curves, but also how they affect correlations in the noise.

For these reasons, it is essential that we gain a thorough understanding of both the correlational structure in the brain and its impact on population coding. Progress has been made on both fronts by adopting two complementary perspectives¹³. One focuses on encoding, and asks whether adding correlations to a population of neurons without modifying single neuron responses (so that the correlated and uncorrelated populations would be indistinguishable on the basis of single neuron recordings) increases or decreases the amount of information in the population. The goal of this approach is to determine whether there are any general principles that relate correlations to increases or decreases in the amount of information, and to assess whether the information

*Department of Brain and Cognitive Sciences and Center for Visual Science, University of Rochester, Rochester, New York 14627, USA.

[†]Gatsby Computational Neuroscience Unit, University College London, 17 Queen Square, London WC1N 3AR, UK. Correspondence to A.P. e-mail: alex@bcs.rochester.edu doi:10.1038/nrn1888

computed from single neurons recorded separately can substitute for the true information in a population (a common technique used in the analysis of experimental data^{9,14–17}). The other focuses on decoding, and more generally on computation. It is driven by the fact that if one wants to extract all the information from a population of correlated neurons (that is, decode optimally), the strategy for doing so can be quite complicated. From this perspective, we can ask how well decoding strategies that ignore correlations, and, therefore, are relatively simple, compare with optimal, but more complex, strategies.

Here, we summarize the empirical and theoretical work that has been carried out in relation to these two perspectives, and discuss what they tell us about the interplay between correlations, population coding and computation.

The encoding perspective: $\Delta I_{\text{shuffled}}$

Perhaps the most straightforward question that can be asked about correlations is whether or not they affect the amount of information in a population code. Although the question is straightforward, the answer is not. For some correlational structures information goes up, for others it goes down, and for still others it stays the same. Although this is somewhat disappointing, because it means that the details of correlations matter, it is important to know for two reasons. First, it raises a cautionary note, as it implies that, in general, the amount of information in a population cannot be computed without knowing the correlational structure (see below). Second, because details matter, we are forced to pay attention to them — general statements such as ‘correlations always hurt so the brain should eliminate them’ or ‘correlations always help so the brain should use them’ can be ruled out.

To determine whether correlations affect the amount of information in a population code, it is necessary to compute the amount of information in the correlated responses, denoted I , and compare this with the amount of information that would be in the responses if they were uncorrelated, denoted I_{shuffled} (the name I_{shuffled} is derived from the fact that, in experiments, responses are decorrelated by shuffling trials). The difference, $\Delta I_{\text{shuffled}}$ ($\equiv I - I_{\text{shuffled}}$), is a measure of the effect of correlations on the amount of information in a population code¹³. An important aspect of this approach is that information is quantifiable, and so can be computed from data.

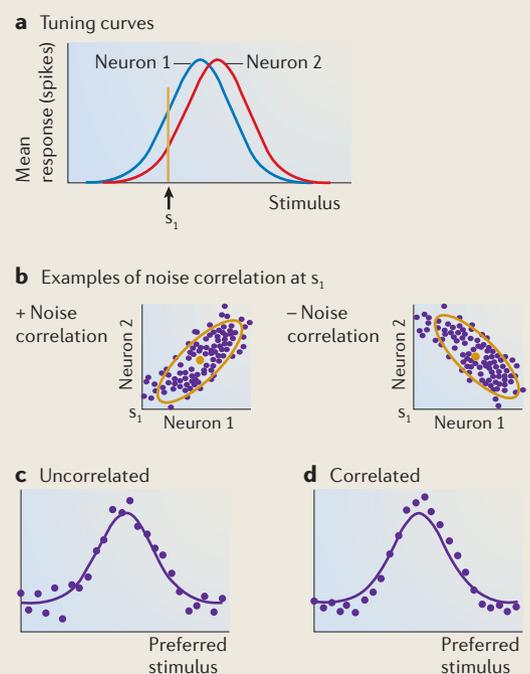
We can develop most of the intuition necessary to understand how correlations affect the amount of information in a population code by considering a two neuron, two stimulus example. Although this is a small population code, it retains several of the features of larger ones. In particular, each stimulus produces a (typically different) set of mean responses, and around those means there is noise in the form of trial-to-trial fluctuations. Because of the noise, any response could be produced by either stimulus, so a response does not tell us definitively which stimulus occurred. Therefore, the noise reduces the information in the responses, with the degree of the reduction depending on both the correlations in the noise and their relationship to the average responses.

To understand the relationship between signal, noise and information in pairs of neurons, we can plot the correlated and uncorrelated response distributions and examine their features. In the left column of FIG. 1 we show a set of correlated responses, and in the right column we show the associated uncorrelated distributions. The response distributions in this figure are indicated schematically by ellipses, which represent 95% confidence intervals (FIG. 1).

Box 1 | Population codes, noise correlation and signal correlation

Population codes are often characterized by the ‘tuning curve plus noise’ model. In this model, the tuning curve represents the average response of a neuron to a set of stimuli, with the average taken across many presentations of each stimulus, and the noise refers to the trial-to-trial variability in the responses. In panel **a**, tuning curves are shown for two neurons that have slightly different preferred stimuli. In panel **b**, we show two hypothetical scatter plots of the single trial responses for this pair of neurons, in response to the repeated presentation of a single stimulus s_1 (arrow in panel **a**). Ellipses represent 95% confidence intervals. The example on the left illustrates positive noise correlation and the example on the right illustrates negative noise correlation. Responses also show a second sort of correlation known as signal correlation⁶¹. These are correlations in the average response. Neurons with similar tuning curves (panel **a**) typically have positive signal correlations, because when s increases, the mean responses of both neurons tend to increase, or decrease, together. Conversely, neurons with dissimilar tuning curves typically have negative signal correlations. Unless stated otherwise, ‘correlated’ in this article means ‘noise correlated’.

In panels **c** and **d** we illustrate the response of a population of neurons. The x-axis corresponds to the preferred orientation of the neuron, the response of which is plotted on the y-axis. Each dot corresponds to the firing rate of one neuron in this example trial, and the purple curve shows the average response of each neuron in the population. Although the neurons in both panels **c** and **d** exhibit noise fluctuations, there is a difference in the structure of those fluctuations: on individual trials the responses of nearby neurons in panel **c** are uncorrelated (fluctuating up and down independently), whereas in panel **d** they are correlated (tending to fluctuate up and down together). Note that nearby neurons in panel **d** are positively correlated (as in panel **b**, left) whereas those that are far apart are negatively correlated (as in panel **b**, right).



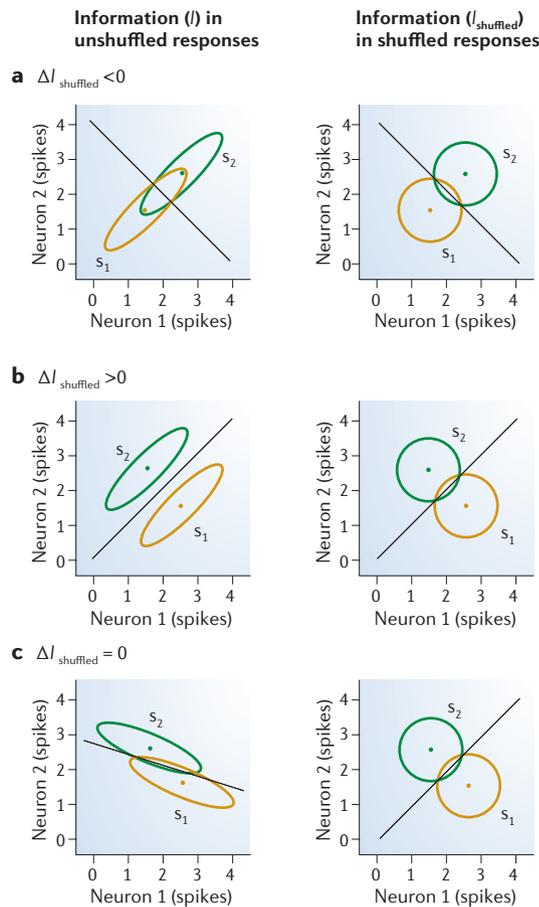


Figure 1 | Effects of correlations on information encoding. In all three cases, we show the response distributions for two neurons that respond to two different stimuli. The panels on the left show the unshuffled responses, those on the right show the shuffled responses. Each ellipse (which appears as a circle in the uncorrelated plots) indicates the 95% confidence interval for the responses. Each diagonal line shows the optimal decision boundary — that is, responses falling above the line are classified as stimulus 2 and responses below the line are classified as stimulus 1. The x-axis is the response of neuron 1, the y-axis the response of neuron 2. **a** | A larger fraction of the ellipses lie on the ‘wrong’ side of the decision boundary for the true, correlated responses than for the independent responses, so $\Delta I_{\text{shuffled}} < 0$. **b** | A smaller fraction of the ellipses lie on the wrong side of the decision boundary for the correlated responses, so $\Delta I_{\text{shuffled}} > 0$. **c** | The same fraction of the ellipses lies on the wrong side of the decision boundary for both the correlated and independent responses, so $\Delta I_{\text{shuffled}} = 0$. I_{shuffled} , uncorrelated information; $\Delta I_{\text{shuffled}} = I - I_{\text{shuffled}}$

The larger the overlap of the ellipses, the more mistakes are made during decoding, and the less information is contained in the neural code. Therefore, these plots allow us to see, graphically, how correlations affect the information in the neuronal responses.

An important point about correlations is that the interaction between the signal correlations (which in this case correspond to the relative positions of the mean responses, see BOX 1 for definitions) and the noise

correlations controls whether correlations increase or decrease information. To illustrate this, in FIG. 1a we have constructed responses such that the signal and noise correlations are both positive. This leads to larger overlap between the ellipses for the correlated than for the uncorrelated responses, which makes the correlated responses harder to decode. The correlated responses carry less information, so $\Delta I_{\text{shuffled}} < 0$. In FIG. 1b, on the other hand, the signal is negatively correlated whereas the noise is positively correlated. Here, there is less overlap in the correlated than the uncorrelated responses, which makes the correlated responses easier to decode. In this figure, then, the correlated responses carry more information, and $\Delta I_{\text{shuffled}} > 0$. Importantly, there is also an intermediate regime (FIG. 1c) in which I and I_{shuffled} are the same ($\Delta I_{\text{shuffled}} = 0$). So, the presence of correlations does not guarantee an effect on the amount of information encoded. The decrease in information when the signal and noise are both positively correlated (or both negatively correlated) and the increase when the signal and noise have opposite correlations is a general feature of information coding in pairs of neurons, and has been observed by a number of authors^{18–20}.

These examples illustrate two important points. First, if we know only the individual responses of each neuron in a pair, and not their correlations, we do not know how much information they encode. Second, just because neuronal responses are correlated does not necessarily mean that they contain more (or less) information. This is important, as it has been suggested that correlations between neurons provide an extra channel of information^{14,21}.

In all of the examples shown in FIG. 1, the correlations are the same for both stimuli, meaning the ellipses in each panel have the same size and orientation. However, it is possible for the correlations to depend on the stimulus, in which case the ellipses would have different sizes or orientations. Such correlations are often referred to as stimulus-modulated correlations²², and they affect information encoding in the same way as the examples discussed above: if the correlations increase the overlap, the information goes down, whereas if the correlations decrease the overlap then the information goes up. In extreme cases, it is even possible for neurons to have identical mean responses to a pair of stimuli, but different correlations (for example, one ellipse at +45° and the other at –45°). Although beyond the scope of this review, the effect of these stimulus-modulated correlations, which are just beginning to be investigated²³, can be large.

What is actually observed in the brain? Do correlations increase or decrease the amount of available information? Various empirical studies have measured $\Delta I_{\text{shuffled}}$ in pairs of neurons and have found that it is small in the rat barrel cortex²⁴, and macaque V1 (REFS 25,26), prefrontal²⁷ and somatosensory cortices²⁸. The results of these studies have also shown that $\Delta I_{\text{shuffled}}$ can be either positive or negative, which means that in real neurons — not just in theory — noise correlations can either increase or decrease the amount of information encoded in pairs of simultaneously recorded neurons. Overall, however, the observed effects have been quite small²⁹.

These results have direct bearing on (and are at odds with) the binding-by-synchrony hypothesis. This hypothesis, originally put forward by Milner³⁰ and von der Malsburg³¹, and championed by Singer and colleagues^{32–34}, states that noise correlations (more specifically, synchronous spikes) could solve the binding problem³⁵ by signalling whether different features in a visual scene belong to the same object. Specifically, they suggested that the number of synchronous spikes across a pair of neurons depends on whether the pair represents the same or different objects. If this hypothesis were true, it would imply that $\Delta I_{\text{shuffled}}$ would be large and positive, at least for some pairs, because shuffling data removes synchronous spikes. To test this directly, Golledge *et al.*²⁵ calculated $\Delta I_{\text{shuffled}}$ (I_{cor} in their study) using an experimental paradigm similar to that used by Singer and colleagues. They found that shuffling the data eliminated little information about whether two features in a visual scene belonged to the same object, a finding that argues against the binding-by-synchrony hypothesis.

These empirical studies suggest that *in vivo* correlations have little impact on the amount of information in pairs of neurons. Whether this holds for large populations, however, is unknown. In fact, as pointed out by Zohary and colleagues³⁶, small effects of correlations in pairs can have large effects in populations. But, as with the two neuron example given above, the effect can be either positive or negative. To illustrate this, consider a population of neurons with bell-shaped tuning curves in which neurons with similar tuning curves are more strongly correlated than neurons with dissimilar tuning curves. As, in this example, neurons with similar tuning curves show positive signal correlations, we expect, on the basis of our two neuron, two stimulus example above, that positive noise correlations will lead to a reduction in information and negative correlations to an increase. This is exactly what is found. Specifically, as the number of neurons increases, I_{shuffled} (FIG. 2a, correlation coefficient (c) = 0) becomes much larger than I when noise correlations are positive (FIG. 2a, $c = 0.01$ or $c = 0.1$) and much smaller when they are negative (FIG. 2a, $c = -0.005$). Interestingly, however, these effects are small for pairs of neurons, and only become pronounced at the population level. FIGURE 2b shows how I_{shuffled} compares with I as the number of neurons increases. For a model in which the maximum correlations are 0.1, the difference between I_{shuffled} and I is minimal (<1%) for a pair of neurons ($n = 2$). However, as the size of the population grows to only a few thousand neurons, correlations begin to have a large effect on the encoded information, reducing it by a factor of almost 25 relative to I_{shuffled} . Although it is not yet clear whether this model accurately reflects the effects of noise correlation in the brain, it provides us with an important lesson: small, perhaps undetectable, correlations in pairs of neurons can have a large effect at the population level. Therefore, it may be typical for I_{shuffled} and I to be very different. This, in turn, implies that studies^{14,37–44} in which I_{shuffled} is used as a surrogate for the true information, I , should be treated with caution.

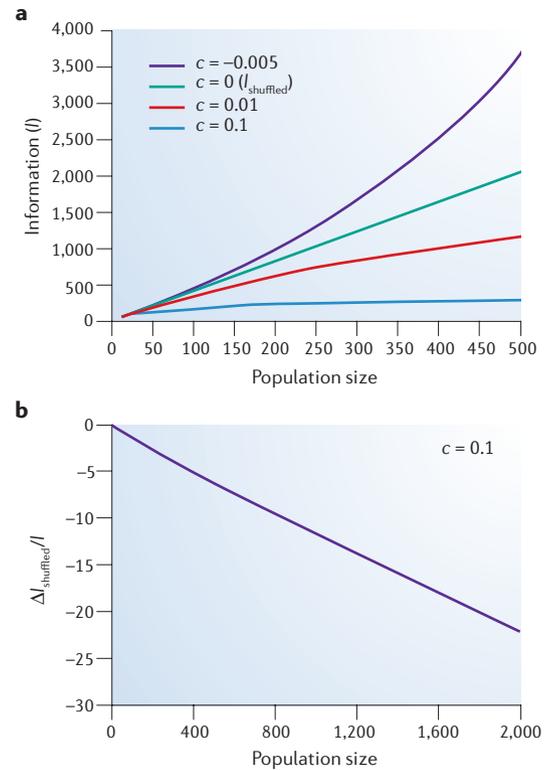


Figure 2 | Information, I , and $\Delta I_{\text{shuffled}}$ versus population size. a | Information, I , versus population size, for different correlation coefficients, c . Positive correlations ($c = 0.01$ or $c = 0.1$) decrease information with respect to the uncorrelated ($c = 0$) case. Furthermore, for positive correlations, information saturates as the number of neurons increases. **b** | $\Delta I_{\text{shuffled}}/I$ versus population size. An important feature of this plot is that correlations have large effects at the population level even though $\Delta I_{\text{shuffled}}/I$ is small for individual neuronal pairs. I_{shuffled} , uncorrelated information; $\Delta I_{\text{shuffled}} = I - I_{\text{shuffled}}$. Encoding model in panels **a** and **b** was taken from REF. 45 and the information measure was Fisher.

A corollary of these results is that noise correlations can cause the amount of information in a population of neurons to saturate as the number of neurons approaches infinity^{36,45–47} (FIG. 2a). One of the first studies to address this question empirically suggested that the pattern of noise correlations observed in the medial temporal visual area (MT) was such that information would saturate³⁶. This was subsequently challenged by theoretical studies^{46,47} that pointed out that the correlations measured in MT do not necessarily imply that the information would saturate as the number of neurons increased.

Although the question of whether or not a particular correlational structure will cause the information to saturate is interesting from a theoretical perspective, it may not be so relevant to networks in the brain. This is because the nervous system can extract only a finite amount of information about sensory stimuli, and, in subsequent stages of processing, the amount of information cannot exceed the amount extracted by, for example, the retina or the cochlea. Therefore, as the number of neurons increases, the correlations must be such that information

No-sharpening model

A model in which the orientation tuning curves of cortical cells are solely the result of the converging afferents from the LGN, without further sharpening in the cortex.

Sharpening model

A model in which the LGN afferents provide broad tuning curves to orientation that are sharpened in the cortex through lateral interactions.

Box 2 | Assuming independence when decoding

What do we mean by ‘ignoring correlations when decoding’? Consider the following situation: a machine generates a number, x , which we would like to know. Unfortunately, every time we query the machine, the sample it produces is corrupted by independent, zero mean noise. To reduce the noise, we collect 10,000 samples. As the samples are independent, the best estimate of x is a weighted sum, with each sample weighted by $1/10,000$.

Imagine now that the machine gets lazy, and only the first two samples are independent; the other 9,998 are the same as the second. In this case, the optimal strategy is to weight the first sample by $1/2$ and the other 9,999 by a set of weights that adds up to $1/2$. If, however, we decide not to measure the correlations, and assume instead that the samples are independent, we would assign a weight of $1/10,000$ to all samples. This is of course suboptimal, as the first sample should be weighted by $1/2$, not $1/10,000$. The difference in performance of the optimal strategy (weight of $1/2$ on the first sample) versus the suboptimal strategy (weights of $1/10,000$ for all samples) is what ΔI_{diag} measures.

But why should we settle for the suboptimal strategy? The answer is that the suboptimal strategy is simple: the weights are determined by the number of samples, which is easy to compute. For the optimal strategy, on the other hand, it is necessary to measure the correlations. In this particular example, the correlations are so extreme that we would immediately notice that the last 9,999 examples are perfectly correlated. In general, however, measuring correlations is hard, and requires large amounts of data. Therefore, when choosing a strategy, there is a trade-off between performance and how much time and data one is willing to spend measuring correlations.

Neurons face the same situation: they compute some function of the variables encoded in their inputs, and to perform this computation optimally they must know the correlations in the $\sim 10,000$ inputs that they receive⁷¹. If they ignore the correlations, they may — or may not — pay a price in the form of suboptimal computations.

saturates. As such, the question is not whether information saturates in the nervous system — it does — it's how quickly it saturates as the number of neurons increases, and whether it saturates at a level well below the amount of information available in the input. These remain open experimental and theoretical questions⁴⁵.

Pitfalls of using I_{shuffled} in place of I

As we have seen, the value of I_{shuffled} compared with I quantifies the impact of correlations on information in population codes. However, this is not the only use of I_{shuffled} . This measure is also commonly used as a surrogate for the true information¹⁴, primarily because estimating the true information in a large neuronal population would require simultaneous recordings of all of the neurons, whereas I_{shuffled} requires only single cell recordings, as well as fewer trials. Similarly, correlations are often ignored in theoretical and computational work, as they can be difficult to model^{37–43}. Instead, information is often estimated under the assumption of independent noise, which is an estimate of I_{shuffled} rather than I .

Unfortunately, using I_{shuffled} instead of the true information can be very misleading because, as discussed in the previous section, I_{shuffled} is not guaranteed to provide a good estimate of I . Orientation selectivity provides a good example of the problem that can arise. Two types of model have been proposed to explain the emergence of orientation selectivity in V1. One is a no-sharpening model in which the tuning to orientation is due to the convergence of lateral geniculate nucleus (LGN) afferents onto cortical neurons (this is essentially the model that was proposed by Hubel and Wiesel⁴⁸). The other is a sharpening model in which the LGN afferents produce only weak tuning, which is subsequently sharpened by lateral connections in the cortex. It is possible to build these models in such a way that they produce identical tuning curves. Can we conclude from this that they contain the same amount of information about orientation?

If we were to use I_{shuffled} as our estimate of information, we would answer ‘yes’. For instance, if we assume that the

noise is independent and Poisson in both models, identical tuning curves imply identical information. How about the true information, I ? To compute the true information, we need to know the correlations. Seriès *et al.*⁴⁹ have simulated these two models in a regime in which the tuning curves and the variability were matched on average. They then estimated the true information, I , and found that, across many architectures, the no-sharpening models always contained more information than the sharpening models, despite identical tuning curves. The difference in information is the result of using different architectures, which lead to different neuronal dynamics and, therefore, different correlations. This point is lost if only I_{shuffled} is measured.

Similar problems often emerge in other models. For example, one approach to modelling the neural basis of attention is to simulate a network of analogue neurons and modify the strength of the lateral connections to see if this increases information⁴⁴. If the information is computed under the assumption of independent Poisson noise, these simulations only reveal whether I_{shuffled} increases. Unfortunately, as we have shown above, without knowing the correlations, the true information might have either increased or decreased.

A common theme in these examples is that the noise correlations are not independent of the architecture of the network. If the architecture changes, so will the correlations. Assuming independence before and after the change is not a valid approximation, and can therefore lead to the wrong conclusions.

The decoding perspective: ΔI_{diag}

Above, we asked how correlations affect the total amount of information in a population code. Our ultimate interest, however, is in how the brain computes with population codes, so what we really want to know is how correlations affect computations. This, however, requires us to specify a computation, and to also specify how it is to be performed. To avoid such details, and also to derive a measure that is computation and

implementation independent, we ask instead about decoding, and, in particular, whether downstream neurons have to know about correlations to extract all the available information. We focus on this question because its answer places bounds on computations. Specifically, if ignoring correlations means a decoder loses, for example, half the information in a population code, then a computation that ignores correlations will be similarly impaired. This does not mean that decoding is a perfect proxy for computing; the effect of correlations on decoding will always depend, at least to some degree, on the computation being performed. However, there is one fact that we can be sure of: if all the information in a population can be extracted without any knowledge of the correlations, then, formally, any computation can perform optimally without knowledge of the correlations.

To investigate the role of correlations in decoding, then, we can measure the difference between the information in a population code, I , and the information, denoted I_{diag} , that would be extracted by a decoder optimized on the shuffled data but applied to the original correlated data (BOX 2). We refer to this difference as ΔI_{diag} ($\Delta I_{\text{diag}} = I - I_{\text{diag}}$), although it has been given different names depending on the details of how it is measured. The name ΔI_{diag} is often used when working with Fisher information^{13,50}, whereas ΔI (REFS 51,52) and $I_{\text{cor-dep}}$ (REF. 22) have been used for Shannon information⁵³ (both ΔI (REF. 52) and $I_{\text{cor-dep}}$ (REF. 22), which are identical, are upper bounds on the cost of using a decoder optimized on shuffled data; see REF. 52 for details). For this discussion the details of the information measure are not important.

Although the encoding perspective (discussed above) and the decoding perspective are related¹³, they are not as tightly coupled as might be expected. For example, $\Delta I_{\text{shuffled}}$ can be non-zero — even very far from zero — when correlations have no effect on decoding ($\Delta I_{\text{diag}} = 0$). The opposite is also possible: $\Delta I_{\text{shuffled}}$ can be zero when correlations both exist and have a large effect on decoding ($\Delta I_{\text{diag}} > 0$)^{13,54}. To understand this intuitively, let us investigate how correlations affect decoding for our two neuron, two stimuli example (FIG. 3). In general, a decoder is just a decision boundary, and in FIG. 3, in which we have only two stimuli and the correlational structure is fairly simple, the decision boundary is a line. Examining the panels in FIG. 3a, we see that the decision boundaries are the same whether they are estimated on shuffled (left column) or correlated (right column) responses. However, in the example shown in FIG. 3b, using a decision boundary based on shuffled responses (black line) can lead to a strongly suboptimal decoding algorithm, as it would produce wrong answers much more often than the optimal decision boundary (red line in FIG. 3b). (Although $\Delta I_{\text{diag}} = 0$ in FIG. 3a, for technical, but potentially important, reasons, correlations can be crucial for decoding in this case; in fact, $\Delta I \neq 0$. A discussion of this issue is beyond the scope of this review, but see REFS 52,54 for details.)

Importantly, although ΔI_{diag} is zero in FIG. 3a, the correlations clearly affect the amount of information encoded ($\Delta I_{\text{shuffled}} < 0$, as can be seen in FIG. 1a). Conversely, in the

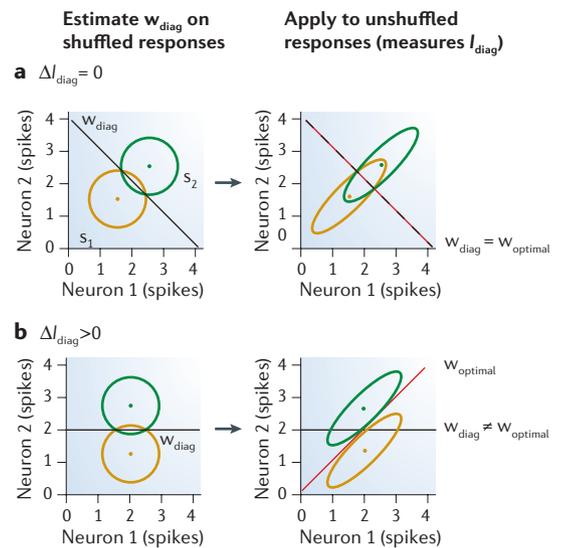


Figure 3 | Effects of correlations on information decoding. The panels on the left show the shuffled responses, those on the right show the unshuffled responses. Each ellipse (which appears as a circle in the uncorrelated plots) indicates the 95% confidence interval for the responses. Each diagonal line shows the optimal decision boundary — that is, responses falling above the line are classified as stimulus 2 and responses below the line are classified as stimulus 1. The x-axis is the response of neuron 1, the y-axis the response of neuron 2. The panels on the left show the decoding boundary (black line) constructed using the uncorrelated responses (green and yellow circles). The panels on the right show the decoding boundary (red line) constructed using the correlated responses (green and yellow ellipses). This is the optimal decoding boundary. The ‘independent’ decoding boundary is included on this panel for easy comparison. **a** | The two decoding boundaries (indicated by a dashed red and black line) are identical, so the fraction of trials decoded correctly is the same whether or not the decoding algorithm was constructed using the correlated responses, and $\Delta I_{\text{diag}} = 0$. **b** | The two decoding boundaries are different, so fewer trials are decoded correctly using the decoding algorithm constructed from the correlated responses, and $\Delta I_{\text{diag}} > 0$. I , information; I_{diag} , information that would be extracted by a decoder optimized on the shuffled data but applied to the original correlated data; ΔI_{diag} , $I - I_{\text{diag}}$; w_{diag} , decoding boundary estimated on shuffled data.

example in FIG. 3b, ΔI_{diag} is greater than zero, even though the effect of correlations on encoding is rather small ($\Delta I_{\text{shuffled}}$ is close to zero, and could be made exactly zero by adjusting the angle of the ellipses).

So how much information is lost when neural responses measured in the brain are decoded using algorithms that ignore correlations? To our knowledge, the first researchers to address this question were Dan *et al.*⁵⁵, who asked whether or not synchronous spikes in the LGN carried additional information. They found that pairs of synchronous spikes did carry extra information: for their most correlated pairs, 20–40% more information was available from a decoder that took synchronous spikes into account than a decoder that did not.

Fisher information
Measures the variance of an optimal estimator.

Shannon information
Measures how much one’s uncertainty about the stimuli decreases after receiving responses.

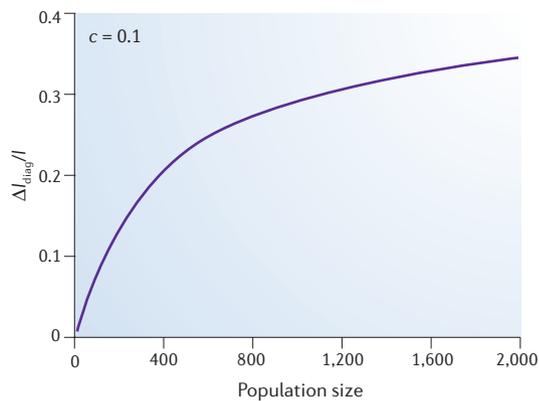


Figure 4 | $\Delta I_{\text{diag}}/I$ versus population size. As was the case in FIG. 2b, correlations can have a small effect when decoding pairs of neurons, but a large effect when decoding populations. *c*, correlation coefficient; *I*, information; I_{diag} , information that would be extracted by a decoder optimized on the shuffled data but applied to the original correlated data; ΔI_{diag} , $I - I_{\text{diag}}$. Encoding model taken from REF. 45.

Almost all subsequent studies found that the maximum value of ΔI_{diag} across many pairs of neurons was small, of the order of 10% of the total information. This has been shown in the mouse retina⁵¹, rat barrel cortex²⁴, and the supplementary motor area^{13,56}, V1

(REF. 25) and motor cortex⁵⁷ of the macaque. So, almost all of the empirical data suggest that little additional information is available in the noise correlations between neurons.

Do the small values of I_{diag} that have been observed experimentally extrapolate to populations? Amari and colleagues⁵⁸ were the first to study this question theoretically. They looked at several correlational structures and tuning curves, and in most cases found that ΔI_{diag} was small compared with the total information in the population. These results, however, should not be taken to imply that ΔI_{diag} is always small for populations. In FIG. 4 we plot ΔI_{diag} as a function of population size. As was the case with $\Delta I_{\text{shuffled}}$, the effect of correlations on decoding increases for larger populations.

This analysis tells us that the effect of correlations on decoding strategies can be anything from no effect at all to a large effect. In some sense the observation that correlations can be present and large and ΔI_{diag} can be small or even zero is the most surprising. This has the important ramification that, when studying population codes, one has to go beyond simply showing that noise correlations exist: their effect on decoding spike trains must be directly measured.

Conclusions

During the past decade it has become increasingly clear that if we want to understand population coding we need to understand neuronal noise. This is not just because noise makes population coding probabilistic, but is also because correlated noise has such a broad range of effects. First, correlations can either increase or decrease the amount of information encoded by a population of neurons. Importantly, decreases can have especially severe effects in large populations, as many correlational structures cause information to saturate as the number of neurons becomes large^{36,45,46}. Such correlations, if they occur in the brain, would place fundamental constraints on the precision with which variables can be represented^{36,45–47}. Second, correlations might or might not affect computational strategies of networks of neurons. A decoder that can extract all the information from a population of independent neurons may extract little when the neurons are correlated, or it may extract the vast majority^{54,58}. Incidentally, other measures of information coding in populations, including synergy, have been put forward, but they do not directly address the questions we are considering here (BOX 3).

These two aspects of correlations — how they affect the amount of information encoded by a population and how they affect decoding of the information from that population — can be quantified by the measures $\Delta I_{\text{shuffled}}$ and ΔI_{diag} , respectively. The first of these, $\Delta I_{\text{shuffled}}$, is the difference between the information in a population code when correlations are present and the information in a population code when the correlations are removed. This measure can be either greater than or less than zero^{46,47}. The second, ΔI_{diag} , which is more subtle, measures the difference between the amount of information that could be extracted from

Box 3 | Other measures of the impact of correlations

An information-theoretic measure that has been applied to pairs of neurons is redundancy^{20,22,27,61–63}. This quantity is the sum of the information from individual cells minus the total information,

$$\Delta I_{\text{redundancy}} = \sum_i I_i - I \quad (1)$$

where I_i is the Shannon information from neuron *i* and *I* is the total information^{24,27,61,63–65}. The negative of $\Delta I_{\text{redundancy}}$ is known as $\Delta I_{\text{synergy}}$, and neural codes with positive $\Delta I_{\text{synergy}}$ (negative $\Delta I_{\text{redundancy}}$) are referred to as synergistic.

The quantity $\Delta I_{\text{redundancy}}$ is often taken to be a measure of the extent to which neurons transmit independent messages, an interpretation based on the observation that if neurons do transmit independent messages, then $\Delta I_{\text{redundancy}}$ is zero. However, this interpretation is somewhat problematic, because the converse is not true: $\Delta I_{\text{redundancy}}$ can be zero even when neurons do not transmit independent messages^{51,52}. Therefore, despite the fact that $\Delta I_{\text{redundancy}}$ has been extensively used^{24,27,61–65}, its significance for population coding is not clear (for a more detailed discussion, see REFS 51, 52).

More recently, $\Delta I_{\text{redundancy}}$ has also been interpreted^{63,65} as a measure of how well neurons adhere to the famous redundancy reduction hypothesis of Attneave, Barlow and others^{56–59}. However, this interpretation is due to a rather unfortunate duplication of names; in fact $\Delta I_{\text{redundancy}}$ is not the same as the redundancy referred to in this hypothesis. For Barlow, as for Shannon⁵³, redundancy is defined to be $1 - H/H_{\text{max}}$, where H_{max} is the maximum entropy of a discrete distribution (subject to constraints) and *H* is the observed entropy. This definition has also been extended to continuous distributions⁷⁰, for which redundancy is $1 - I/I_{\text{max}}$, where I_{max} is the channel capacity. The redundancy given in equation 1, however, corresponds to neither of these definitions, nor does its normalized version, $\Delta I_{\text{redundancy}}/I$. Measuring $\Delta I_{\text{redundancy}}$ therefore sheds little, if any, light on the redundancy reduction hypothesis.

Studies that have estimated synergy or redundancy have in general, but not always⁶³, found that pairs of neurons can be either redundant or synergistic^{24,27,61,63–65}, whereas larger populations are almost always redundant. The latter result is not surprising: populations typically use many neurons to code for a small number of variables, so the marginal contribution of any one neuron to the information becomes small as the population size increases⁶³.

a population by a decoder with full knowledge of the correlations, and the amount that could be extracted by a decoder with no knowledge of the correlations. Because correlations are not removed from the responses when computing ΔI_{diag} , this quantity is very different from $\Delta I_{\text{shuffled}}$. In particular, unlike $\Delta I_{\text{shuffled}}$ it can never be negative, because no decoder can extract more information than one with full knowledge of the correlations. So, if ΔI_{diag} is zero, correlations are not important for decoding, and if ΔI_{diag} is positive, they are. In the latter case, the ratio $\Delta I_{\text{diag}}/I$ quantifies just how important correlations are⁵².

A somewhat counterintuitive result that has emerged from quantitative studies of $\Delta I_{\text{shuffled}}$ and ΔI_{diag} is that the two are not necessarily related: $\Delta I_{\text{shuffled}}$ can be either positive or negative when ΔI_{diag} is zero, and $\Delta I_{\text{shuffled}}$ can be zero when ΔI_{diag} is positive^{52,54}. So, these two quantities answer different questions, and the use of both of them together can provide deeper insight into population codes.

Both $\Delta I_{\text{shuffled}}$ and ΔI_{diag} have usually been found to be <10% for pairs of neurons. Therefore, it would seem that correlations are not important, and, in particular, that correlations caused by synchronous spikes — the type of correlations implicated in the binding by synchrony hypothesis — do not carry much extra information. However, whether correlations are important for populations, which is the relevant question for the brain, remains an open question, because even small correlations can have a significant effect in large populations^{23,36} (FIGS 2b,4).

In summary, theoretical studies have greatly increased our understanding of the effects of correlations on the amount of information encoded in a population^{36,45–47}, and we are even beginning to understand how to build networks that could extract a large fraction of the information from a population in cases in which correlations are important²³. The somewhat optimistic nature of both of these statements should, however, be tempered by two observations. First, essentially all these studies assumed Gaussian noise, and it is not clear how well they generalize to the non-Gaussian noise found in the brain. Second, experimentally quantifying the role of correlations in large populations has proved extremely difficult: we have good measurements only for pairs of neurons, and, as discussed above, results for pairs do not give much indication of what is going on at the population level. It is therefore crucial that we develop methods that can be used to study large populations experimentally. Because of data limitations it is not possible to directly compute information^{59,60}, so we are left with two options. One is to assess the role of correlations by decoding spike trains using algorithms that do and do not take correlations into account, and comparing their performance. The other is to develop better models of how noise is correlated in populations, and to carry out theoretical computations based on those noise models. By applying both methods, we should ultimately understand the role of correlated noise in the brain, and, in particular, how the brain carries out computations efficiently in the presence of this noise.

- Lee, C., Rohrer, W. H. & Sparks, D. L. Population coding of saccadic eye movements by neurons in the superior colliculus. *Nature* **332**, 357–360 (1988).
- Sparks, D. L., Holland, R. & Guthrie, B. L. Size and distribution of movement fields in the monkey superior colliculus. *Brain Res.* **113**, 21–34 (1976).
- Georgopoulos, A. P., Schwartz, A. B. & Kettner, R. E. Neuronal population coding of movement direction. *Science* **233**, 1416–1419 (1986).
- 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).
- Pouget, A., Dayan, P. & Zemel, R. Information processing with population codes. *Nature Rev. Neurosci.* **1**, 125–132 (2000).
- Seung, H. S. & Sompolinsky, H. Simple models for reading neuronal population codes. *Proc. Natl Acad. Sci. USA* **90**, 10749–10753 (1993).
- Salinas, E. & Abbott, L. F. Vector reconstruction from firing rates. *J. Comput. Neurosci.* **1**, 89–107 (1994).
- Deneve, S., Latham, P. E. & Pouget, A. Reading population codes: a neural implementation of ideal observers. *Nature Neurosci.* **2**, 740–745 (1999).
- McAdams, C. J. & Maunsell, J. H. Effects of attention on the reliability of individual neurons in monkey visual cortex. *Neuron* **23**, 765–773 (1999).
- Schoups, A., Vogels, R., Qian, N. & Orban, G. Practising orientation identification improves orientation coding in V1 neurons. *Nature* **412**, 549–553 (2001).
- Yang, T. & Maunsell, J. H. The effect of perceptual learning on neuronal responses in monkey visual area V4. *J. Neurosci.* **24**, 1617–1626 (2004).
- Ghose, G. M., Yang, T. & Maunsell, J. H. Physiological correlates of perceptual learning in monkey V1 and V2. *J. Neurophysiol.* **87**, 1867–1888 (2002).
- Averbeck, B. B. & Lee, D. Effects of noise correlations on information encoding and decoding. *J. Neurophysiol.* (in the press).
Combines a theoretical and empirical examination of the way in which studies of information encoding and decoding are related, as well as investigating the role of stimulus-modulated correlations.
- Hung, C. P., Kreiman, G., Poggio, T. & DiCarlo, J. J. Fast readout of object identity from macaque inferior temporal cortex. *Science* **310**, 863–866 (2005).
- Rolls, E. T., Treves, A. & Tovee, M. J. The representational capacity of the distributed encoding of information provided by populations of neurons in primate temporal visual cortex. *Exp. Brain Res.* **114**, 149–162 (1997).
- Gochin, P. M., Colombo, M., Dorfman, G. A., Gerstein, G. L. & Gross, C. G. Neural ensemble coding in inferior temporal cortex. *J. Neurophysiol.* **71**, 2325–2337 (1994).
- Georgopoulos, A. P. & Massey, J. T. Cognitive spatial-motor processes. 2. Information transmitted by the direction of two-dimensional arm movements and by neuronal populations in primate motor cortex and area 5. *Exp. Brain Res.* **69**, 315–326 (1988).
- Oram, M. W., Foldiak, P., Perrett, D. I. & Sengpiel, F. The 'Ideal Homunculus': decoding neural population signals. *Trends Neurosci.* **21**, 259–265 (1998).
- Johnson, K. O. Sensory discrimination: decision process. *J. Neurophysiol.* **43**, 1771–1792 (1980).
- Panzeri, S., Schultz, S. R., Treves, A. & Rolls, E. T. Correlations and the encoding of information in the nervous system. *Proc. R. Soc. Lond. B* **266**, 1001–1012 (1999).
Although somewhat technical, this was one of the first studies to clearly define a set of measures that can be used to assess the role of correlations in information coding. The basic approach presented in this manuscript was further elaborated in reference 22.
- Engel, A. K., Konig, P. & Singer, W. Direct physiological evidence for scene segmentation by temporal coding. *Proc. Natl Acad. Sci. USA* **88**, 9136–9140 (1991).
- Pola, G., Thiele, A., Hoffmann, K. P. & Panzeri, S. An exact method to quantify the information transmitted by different mechanisms of correlational coding. *Network* **14**, 35–60 (2003).
- Shamir, M. & Sompolinsky, H. Nonlinear population codes. *Neural Comput.* **16**, 1105–1136 (2004).
- Petersen, R. S., Panzeri, S. & Diamond, M. E. Population coding of stimulus location in rat somatosensory cortex. *Neuron* **32**, 503–514 (2001).
- Golledge, H. D. *et al.* Correlations, feature-binding and population coding in primary visual cortex. *Neuroreport* **14**, 1045–1050 (2003).
- Panzeri, S., Golledge, H. D., Zheng, F., Tové, M. J. & Young, M. P. Objective assessment of the functional role of spike train correlations using information measures. *Vis. Cogn.* **8**, 531–547 (2001).
- Averbeck, B. B., Crowe, D. A., Chafee, M. V. & Georgopoulos, A. P. Neural activity in prefrontal cortex during copying geometrical shapes. II. Decoding shape segments from neural ensembles. *Exp. Brain Res.* **150**, 142–153 (2003).
- Romo, R., Hernandez, A., Zainos, A. & Salinas, E. Correlated neuronal discharges that increase coding efficiency during perceptual discrimination. *Neuron* **38**, 649–657 (2003).
- Averbeck, B. B. & Lee, D. Coding and transmission of information by neural ensembles. *Trends Neurosci.* **27**, 225–230 (2004).
- Milner, P. M. A model for visual shape recognition. *Psychol. Rev.* **81**, 521–535 (1974).
- von der Malsburg, C. The correlation theory of brain function. *Internal Report, Dept Neurobiology, MPI for Biophysical Chemistry* (1981).
- Singer, W. & Gray, C. M. Visual feature integration and the temporal correlation hypothesis. *Annu. Rev. Neurosci.* **18**, 555–586 (1995).
- Schnitzer, M. J. & Meister, M. Multineuronal firing patterns in the signal from eye to brain. *Neuron* **37**, 499–511 (2003).
- Vaadia, E. *et al.* Dynamics of neuronal interactions in monkey cortex in relation to behavioural events. *Nature* **373**, 515–518 (1995).
- Roskies, A. L. The binding problem. *Neuron* **24**, 7–9 (1999).
- Zohary, E., Shadlen, M. N. & Newsome, W. T. Correlated neuronal discharge rate and its

- implications for psychophysical performance. *Nature* **370**, 140–143 (1994).
The first study to show, in the context of neural coding, that small correlations between neurons can have a large effect on the ability of a population of neurons to encode information. The main conclusion of this manuscript was that the correlations in MT cause information to saturate as the population reaches ~100 neurons. Whether or not this is correct remains an open experimental question (see also references 45 and 46).
37. Olshausen, B. A. & Field, D. J. Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature* **381**, 607–609 (1996).
 38. Simoncelli, E. P. & Olshausen, B. A. Natural image statistics and neural representation. *Annu. Rev. Neurosci.* **24**, 1193–1216 (2001).
 39. Hyvarinen, A. & Hoyer, P. O. A two-layer sparse coding model learns simple and complex cell receptive fields and topography from natural images. *Vision Res.* **41**, 2413–2423 (2001).
 40. Hyvarinen, A., Karhunen, J. & Oja, E. *Independent Component Analysis* (John Wiley and Sons, New York, 2001).
 41. Bell, A. J. & Sejnowski, T. J. An information-maximization approach to blind separation and blind deconvolution. *Neural Comput.* **7**, 1129–1159 (1995).
 42. Nadal, J. P. & Parga, N. Nonlinear neurons in the low-noise limit: a factorial code maximizes information transfer. *Network* **5**, 565–581 (1994).
 43. Gold, J. I. & Shadlen, M. N. Neural computations that underlie decisions about sensory stimuli. *Trends Cogn. Sci.* **5**, 10–16 (2001).
 44. Lee, D. K., Itti, L., Koch, C. & Braun, J. Attention activates winner-take-all competition among visual filters. *Nature Neurosci.* **2**, 375–381 (1999).
 45. Sompolinsky, H., Yoon, H., Kang, K. & Shamir, M. Population coding in neuronal systems with correlated noise. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* **64**, 051904 (2001).
 46. Abbott, L. F. & Dayan, P. The effect of correlated variability on the accuracy of a population code. *Neural Comput.* **11**, 91–101 (1999).
One of the most influential theoretical studies of the effect of noise correlations on information encoding.
 47. Wilke, S. D. & Eurich, C. W. Representational accuracy of stochastic neural populations. *Neural Comput.* **14**, 155–189 (2002).
 48. Hubel, D. H. & Wiesel, T. N. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J. Physiol. (Lond.)* **160**, 106–154 (1962).
 49. Serriès, P., Latham, P. E. & Pouget, A. Tuning curve sharpening for orientation selectivity: coding efficiency and the impact of correlations. *Nature Neurosci.* **7**, 1129–1135 (2004).
One of the first papers to show that manipulations that increase the information in single cells (through, for example, sharpening tuning curves) can, because the manipulation modifies correlations, reduce the information in the population.
 50. Casella, G. & Berger, R. L. *Statistical Inference* (Duxbury Press, Belmont, California, 1990).
 51. Nirenberg, S., Carcieri, S. M., Jacobs, A. L. & Latham, P. E. Retinal ganglion cells act largely as independent encoders. *Nature* **411**, 698–701 (2001).
 52. Latham, P. E. & Nirenberg, S. Synergy, redundancy, and independence in population codes, revisited. *J. Neurosci.* **25**, 5195–5206 (2005).
 53. Shannon, C. E. & Weaver, W. *The Mathematical Theory of Communication* (Univ. Illinois Press, Urbana Champagne, Illinois, 1949).
 54. Nirenberg, S. & Latham, P. E. Decoding neuronal spike trains: how important are correlations? *Proc. Natl Acad. Sci. USA* **100**, 7348–7353 (2003).
 55. Dan, Y., Alonso, J. M., Usrey, W. M. & Reid, R. C. Coding of visual information by precisely correlated spikes in the lateral geniculate nucleus. *Nature Neurosci.* **1**, 501–507 (1998).
 56. Averbeck, B. B. & Lee, D. Neural noise and movement-related codes in the macaque supplementary motor area. *J. Neurosci.* **23**, 7630–7641 (2003).
 57. Oram, M. W., Hatsopoulos, N. G., Richmond, B. J. & Donoghue, J. P. Excess synchrony in motor cortical neurons provides redundant direction information with that from coarse temporal measures. *J. Neurophysiol.* **86**, 1700–1716 (2001).
 58. Wu, S., Nakahara, H. & Amari, S. Population coding with correlation and an unfaithful model. *Neural Comput.* **13**, 775–797 (2001).
The first study to theoretically investigate the effects of ignoring correlations when decoding a large population of neurons. As such it is the decoding complement to reference 46.
 59. Strong, S. P., Koberle, R., de Ruyter van Steveninck, R. R. & Bialek, W. Entropy and information in neural spike trains. *Phys. Rev. Lett.* **80**, 197–200 (1998).
 60. Treves, A. & Panzeri, S. The upward bias in measures of information derived from limited data samples. *Neural Comput.* **7**, 399–407 (1995).
 61. Gawne, T. J. & Richmond, B. J. How independent are the messages carried by adjacent inferior temporal cortical neurons? *J. Neurosci.* **13**, 2758–2771 (1993).
 62. Schneidman, E., Bialek, W. & Berry, M. J. Synergy, redundancy, and independence in population codes. *J. Neurosci.* **23**, 11539–11553 (2003).
 63. Narayanan, N. S., Kimchi, E. Y. & Laubach, M. Redundancy and synergy of neuronal ensembles in motor cortex. *J. Neurosci.* **25**, 4207–4216 (2005).
 64. Gawne, T. J., Kjaer, T. W., Hertz, J. A. & Richmond, B. J. Adjacent visual cortical complex cells share about 20% of their stimulus-related information. *Cereb. Cortex* **6**, 482–489 (1996).
 65. Puchalla, J. L., Schneidman, E., Harris, R. A. & Berry, M. J. Redundancy in the population code of the retina. *Neuron* **46**, 493–504 (2005).
 66. Attneave, F. Informational aspects of visual perception. *Psychol. Rev.* **61**, 183–193 (1954).
 67. Barlow, H. B. in *Current Problems in Animal Behaviour* (eds Thorpe, W. H. & Zangwill, O. L.) 351–360 (Cambridge Univ. Press, Cambridge, 1961).
 68. Srinivasan, M. V., Laughlin, S. B. & Dubs, A. Predictive coding: a fresh view of inhibition in the retina. *Proc. R. Soc. Lond. B* **216**, 427–459 (1982).
 69. Barlow, H. Redundancy reduction revisited. *Network* **12**, 241–253 (2001).
 70. Atick, J. J. & Redlich, A. N. Towards a theory of early visual processing. *Neural Comput.* **2**, 308–320 (1990).
 71. Braitenberg, V. & Schüz, A. *Anatomy of the Cortex* (Springer, Berlin, 1991).

Acknowledgements

P.E.L. was supported by the Gatsby Charitable Foundation, London, UK, and a grant from the National Institute of Mental Health, National Institutes of Health, USA. A.P. was supported by grants from the National Science Foundation. B.B.A. was supported by a grant from the National Institutes of Health.

Competing interests statement

The authors declare no competing financial interests.