

Neuronal Population Coding

Dharmesh Tailor
University of Edinburgh

August 2017

1 Neural coding

The neuron or nerve cell makes up half of all cells in the human brain; the human brain being the least understood of all organs in the human body. Neurons are distinguished from other cells by two properties: excitability and synaptic communication. A voltage gradient is maintained across neuronal membranes in which the interior of the cell has a negative voltage compared with the exterior. During a period of excitability, the membrane potential rises and falls in quick succession (voltage spike) resulting in an action potential. Neurons form connections with other neurons by structures called synapses. This allows for the propagation of the action potential which may in turn cause other neurons to spike. The spikes are not uniform, varying in duration, amplitude and shape. Despite this, they are considered all-or-none events and for purposes of modelling, we consider all behaviour to be represented by the sequence of spikes (spike train).

An unresolved question is how the brain encodes stimuli in the external world. Experimental data supports the hypothesis that stimuli are encoded in the spike trains of neurons. This assumption underpins research in neural coding which involves “measuring and characterising how stimulus attributes, such as light or sound intensity, or motor actions, such as the direction of an arm movement, are represented by action potentials” [Dayan and Abbott, 2001]. A major breakthrough came with the discovery of the orientation selectivity of neurons in the primary visual cortex (V1) of the cat [Hubel and Wiesel, 1962] and monkey [Hubel and Wiesel, 1968]. The experiments involved moving a light-dark edge such as a slit or bar (stimulus) at varying orientations across the neuron’s receptive field. For neurons in the visual cortex, the receptive field is the portion of the visual space for which stimulation influences the response of the neuron. It was observed that the number of spikes depends on the orientation of the stimulus (Figure 1a).

This motivates the assumption that the response of a neuron during presentation of a stimulus can be solely characterised by the number of spikes (spike count). Whether or not the precise timing of spikes encodes information regarding the stimulus is an active research area, known as the temporal coding

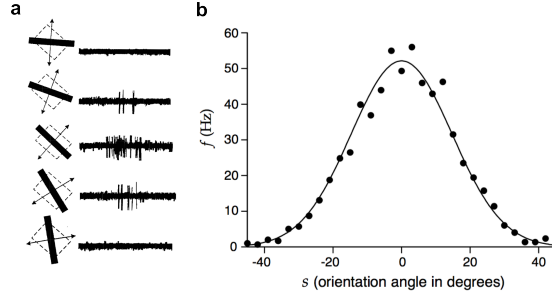


Figure 1: Hubel and Wiesel plot [Dayan and Abbott, 2001]

hypothesis. In a typical experiment, the stimulus is presented a number of times at a fixed orientation. The spike count is averaged over these trials and then divided by the duration of stimulus presentation giving the *average firing rate*, measured in spikes per second or hertz (Hz). This is repeated for different orientations of the stimulus resulting in the data points shown in Figure 1b. The points are then fitted by a Gaussian function,

$$f(s) = r_{max} \exp \left[- \frac{(s - s_{pref})^2}{2\omega^2} \right], \quad (1)$$

where s is the orientation of the stimulus, r_{max} is the maximum average firing rate attained, s_{pref} is the stimulus orientation corresponding to the maximum rate (the *preferred stimulus* of the neuron) and ω determines the width of the Gaussian curve (*tuning width*). The fitted curve is an example of a neural response *tuning curve* which characterises the response of a neuron as a function of the stimulus attribute (black line in Figure 1b). Other commonly used functions for tuning curves include the cosine and sigmoid function. The stimulus attribute of interest (e.g. orientation) is often referred to as the stimulus and we will use this convention thereafter in this note.

By recording the responses of neurons in a contiguous area of the primary visual cortex, Hubel and Wiesel observed that V1 neurons were arranged according to their stimulus preference [Hubel and Wiesel, 1962]. For a given neuron, the preferred stimuli of neighbouring neurons only differed by a small amount. Following onset of a stimulus, the activity of multiple neurons would increase but with different firing rates. There appeared to be a set of neurons sensitive to the particular attribute. This led to the development of the population coding hypothesis which proposes that stimulus attributes are represented by the set of responses from the neural population (population response). Furthermore, experimental data indicated a coarse code for stimulus selectivity. A V1 neuron responded maximally for a narrow band of stimuli but its response decreased gradually as the distance from the preferred stimulus increased. This corresponds to a broad tuning curve in which a significant portion of the stimulus space gives rise to a non-zero firing rate.

Evidence for population codes exists beyond the primary visual cortex. Neurons in the medial temporal visual area (MT) of the macaque monkey are selective to stimulus direction and speed [Maunsell and Van Essen, 1983]. This highlights an important characteristic of population codes which is that they can encode multiple stimulus attributes simultaneously. As with area V1, nearby neurons demonstrated preference for similar stimuli. Population codes also have other uses including encoding motor commands and sensory input. For example, M1 neurons from the primary motor cortex of the macaque monkey encode the direction of arm movement [Georgopoulos et al., 1986]. Cerebral interneurons in crickets encode the direction of air current signalling the movement of nearby predators [Theunissen and Miller, 1991]. There is evidence that pyramidal neurons in the hippocampus of rats encode body location [Wilson and McNaughton, 1993]. Although spike timing may provide additional information, simple decoders have been shown to extract the location from population responses to a reasonable degree of accuracy.

Population coding theory makes the assumption that the map from stimulus to average response is a deterministic process. Thus, we expect the dots in Figure 1b to exactly follow the tuning curve as the number of trials is increased. Responses from single trials are expected to vary about their average value i.e. modelled using a probability distribution. This trial-to-trial variability is considered *noise* and restricts the reliability of single neural responses for tasks such as stimulus discrimination and estimation. It is suggested that population codes overcome this uncertainty by pooling multiple responses.

Noise in neuronal responses is due to apparent randomness in the biophysical processes underpinning spike firing. However, other factors which are not strictly noise also give rise to trial-to-trial variability. These include: different initial states of neural circuitry at start of each trial; response reflecting internal states of the animal; variable levels of arousal and attention [Faisal et al., 2008]. For our purposes these are included in our definition of noise.

2 From population codes to psychophysics

The existence of broadly tuned neurons that are selective to a stimulus attribute or intended motor action leads us to ask whether the neuronal activity of these populations is the cause of perception or behaviour on an organism level. Experiments first sought to determine whether the responses of single neurons could represent sensory stimuli or motor actions. One such experiment involved a macaque monkey moving its arm towards a target whilst electrophysiological recordings of neurons in its motor cortex were being made [Georgopoulos et al., 1986]. In repeated trials, the offset of the final location of the monkey’s arm from the target was considerably smaller than the tuning width of motor neurons. This indicated that single neural responses lacked the precision to represent fine motor actions.

The experiment then moved on to exploring whether the population response could predict the direction of arm movement. The direction was represented by

a vector relative to the location of the animal and the targets were positioned at equally spaced intervals in 3-D space. Assuming the population response encodes the direction, the problem was restated as decoding the direction from the population response. A method known as *population vector* decoding was developed in which the estimated direction was computed from the sum of the neurons' preferred directions weighted by their responses. Figure 2 illustrates this for the simplified case of direction in two dimensions. After a number of trials, the actual direction of arm movement or *movement vector* fell within a 95% confidence cone centred on the decoded direction or *population vector*. The similarity in the two directions demonstrated that the behavioural response of an animal could be explained by the activity of neuronal populations from the relevant brain area.

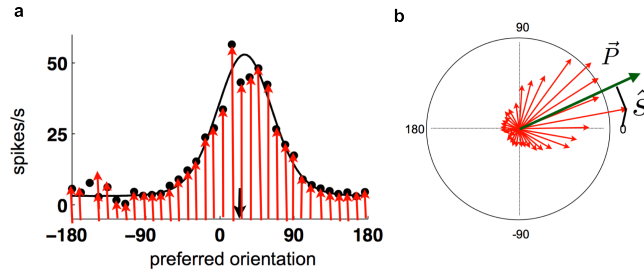


Figure 2: Population vector decoding [Seriès, 2017]

This raises an important question as to whether neural mechanisms which give rise to motor action implement population vector decoding. Population vector decoding is regarded as being neurobiologically plausible as it simply evaluates a linear sum of the responses. Furthermore, there is experimental evidence for this decoding method being implemented in the brain, in the generation of saccadic eye movements in macaque monkeys [Lee et al., 1988]. However, it is worth noting that the performance of population vector decoding depends on the involvement of a large number of neurons. Also, *heterogeneity* in the population such as varying amplitudes of the tuning curves and non-uniformity in the distribution of preferred directions has a considerable effect on the decoding accuracy [Salinas and Abbott, 1994].

The perceptual system in humans has been extensively studied in psychophysics. One such investigation involves quantifying the limits of visual discrimination. A typical experiment involves presenting two stimuli successively (e.g. bars) which differ in orientation. Through repeated trials, the smallest difference that can be detected for a given human subject is determined. Previous results indicate that an orientation difference of 0.42 degrees is discriminated 75% of the time [Westheimer et al., 1976]. This is an example of a psychophysical threshold or *just-noticeable difference* which quantifies the difference at which two stimuli can be reliably discerned for a given level of confidence. Similarly to the behavioural experiment discussed above, the response of single neurons can be compared to these thresholds. The just-noticeable dif-

ference in orientation has been found to be considerably smaller than the tuning width of neurons in the visual cortex. This also holds true for other visual stimulus attributes such as spatial frequency and speed [Orban, 2012]. Thus, the empirical evidence indicates that single neuron responses lack the sensitivity to encode sensory variables with precision equivalent to human performance.

We move on to looking at whether the population response of broadly tuned neurons can encode sensory variables such as stimulus orientation with precision observed in perception. To do so, an approach known as *ideal observer analysis* is followed in which the performance of an ideal observer, that is able to decode the population response optimally, is analysed [Geisler, 2003]. This ideal performance is constrained by the amount of information in the population response about the stimulus. In other words, performance is limited by the accuracy with which the population response encodes the stimulus. It is theoretically impossible for any other decoder to exceed the performance. A result from statistical estimation theory, the Cramér-Rao bound, quantifies this upper limit on performance [Cover and Thomas, 2012]. For encoded stimulus s , we have the following lower bound on the average squared estimation error of any estimator \hat{s} (subject to constraints),

$$\mathbb{E}[(\hat{s} - s)^2] \geq \frac{1}{I_F(s)}, \quad (2)$$

where \mathbb{E} is the expectation operator denoting an average over trials and $I_F(s)$ is the Fisher information. The bound stated above is only guaranteed to hold for unbiased estimators, that is estimators for which there is no systematic difference between the estimate averaged over many trials and the true stimulus value. Biased estimators may exist for which the estimation error is smaller than the Cramér-Rao bound. Despite this, the nature of certain stimulus attributes such as orientation mean that over- and underestimation of the stimulus are equally likely. This symmetry results in the correct estimate being obtained on average.

The Fisher information is computed from $p(\mathbf{r}|s)$ which is the probability distribution characterising the population response conditioned on the stimulus. Whilst the form of the tuning curve can be determined after just several trials for each stimulus in the stimulus space, the nature of trial-to-trial variability is harder to determine. A larger amount of experimental data is required and characterising the variability is further complicated by the presence of correlations in the response. In order to simplify the calculation of Fisher information, a distribution from the exponential family is typically used such as a Gaussian distribution. This is motivated by the central limit theorem which states that the combined effect of a large number of stochastic processes tends to a Gaussian distribution. Furthermore, individual neuronal responses are assumed to be independent. Experimental data indicates that there exists a considerable amount of heterogeneity in the responses of neurons, for instance tuning widths and amplitudes are not uniform across a population. Despite this, homogeneous populations are often modelled in which tuning curves are considered translations of each other and preferred stimuli are uniformly distributed across

the stimulus space (Figure 3a). This allows for the ideal performance to be evaluated in terms of single neuron response properties such as the number of neurons, amplitude, tuning width and noise variance [Paradiso, 1988].

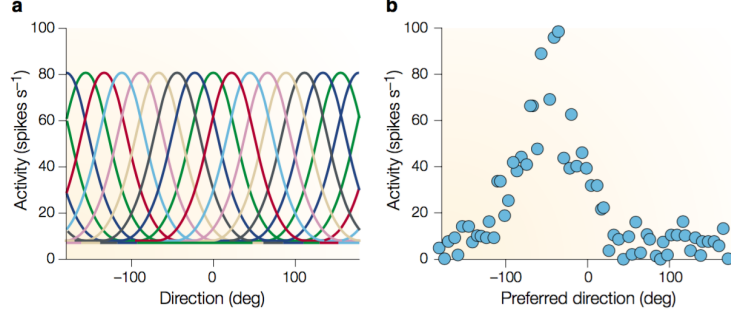


Figure 3: Modelling population response [Pouget and Latham, 2007]

From all of the response properties considered, the number of neurons (i.e. population size) was found to be most critical for performance. The black line in Figure 4 shows the relationship between estimation error and population size (with other response properties fixed). Taking the estimation error to be a measure of discriminability, the population size required to give a minimum estimation error matching the psychophysical threshold of 0.42 degrees was calculated to be 2000 neurons [Paradiso, 1988]. Compared to the number of neurons in the mammalian cortex, this population size is relatively insignificant. One might expect psychophysical performance not to agree with the performance of an ideal observer. However, the consistency suggests the neural mechanism used in the brain may be close to optimal (i.e. efficient) at extracting sensory variables (Figure 4). Thus, it is possible to discriminate orientation with precision matching human perception from the population response of broadly tuned neurons.

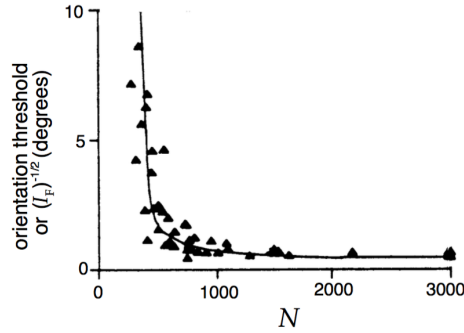


Figure 4: Psychophysical and theoretical threshold [Paradiso, 1988]

3 Fisher information

Fisher information limits the performance of an unbiased decoder by quantifying the accuracy to which a stimulus is encoded in a population response. It is evaluated from $p(\mathbf{r}|s)$ for a particular stimulus s but is not a function of the response since the functional form is averaged over trials. In this form, $p(\mathbf{r}|s)$ (or more specifically its logarithm) is known as the likelihood function. The general form of Fisher information is:

$$I_F(s) = \mathbb{E} \left[- \frac{\partial^2 \log p(\mathbf{r}|s)}{\partial s^2} \right], \quad (3)$$

which characterises the expected curvature of the log-likelihood at some stimulus s [Dayan and Abbott, 2001]. It can also be interpreted as a measure of the *sharpness* of the log-likelihood [Kay, 1993]. A large value for the Fisher information corresponds to a ‘sharp’ likelihood in which responses typical for stimulus s are less likely to occur for slightly different stimuli. In other words, the typical response gives a good indication of the stimulus.

Now, we consider a simple encoding model in which the population of neurons are modelled by uniformly distributed, Gaussian tuning curves with independent, Poisson noise. The Fisher information simplifies to:

$$I_F(s) = \sum_{i=1}^N \frac{f'_i(s)^2}{f_i(s)}, \quad (4)$$

where $f_i(s)$ is the tuning curve for neuron i . We see that the contribution of each neuron to the overall Fisher information is directly proportional to the square of the tuning curve gradient and inversely proportional to the average firing rate. It is worth noting that the direct proportionality component is also present if Gaussian noise is used instead. Since the gradient of the Gaussian tuning curve is zero at its peak, neurons with the largest response contribute little to the Fisher information. In contrast, neurons with an *intermediate* response have a large contribution due to the tuning curve being highly sloped. Intuitively this makes sense since a large slope means a small change in the stimulus causes a large change in the average response Figure 5 illustrates this for a single neuron.

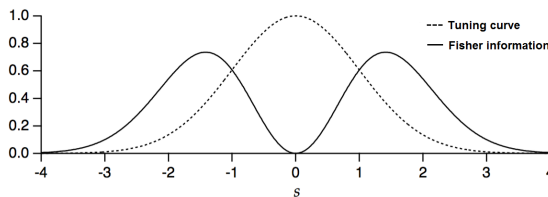


Figure 5: Fisher information [Dayan and Abbott, 2001]

4 Decoder performance

The Fisher information is a useful measure of performance however it does not guarantee that the performance is attainable. An alternative strategy is to analyse the performance of a specific decoder. This performance can be compared to the psychophysical threshold and the decoder can be rejected if it cannot account for the precision of perceptual accuracy. A decoder is typically evaluated by computer simulations which involve *sampling* a large number of population responses from the conditional probability distribution $p(\mathbf{r}|s)$. This distribution, also known as the *encoding model*, makes assumptions on the nature of neuronal responses. For a given population response, the decoder computes an estimate. The estimates will vary trial to trial even when the stimulus is kept constant due to response variability. After a large number of trials, the set of estimates is used to evaluate the performance of the decoder. A decoder is sought that has zero bias and minimal variance.

The decoder used depends on the type of study. Theoretical studies seek to analyse the statistical properties of an encoding model. The decoder used is optimal in the sense that it exploits full knowledge of the response distribution. This is to minimise the loss of information introduced at the decoding step. On the other hand, studies that seek to extract the value of a stimulus attribute from electrophysiological recordings usually do not have access to the complete distribution or make assumptions on its form. Methods such as population vector decoding are commonly used which only utilise the preferred stimuli of neurons. Other methods may use the form of tuning curves such as template matching. There may also be a preference to use linear estimators (due to a preference for neurobiologically plausible decoders) as opposed to decoders that involve nonlinear statistical calculations such as the maximum likelihood decoder. Despite this, it is worth noting that linear decoders are not efficient (i.e. do not saturate the Cramér-Rao bound) under normal conditions [Snippe, 1996].

5 Noise correlations

Trial-to-trial fluctuations from the average response (noise) of cortical neurons have been observed to be positively correlated with other neurons in its region. That is, on a given trial, if the response of a neuron is greater than its average response then the response of another neuron will tend to be above its average. It has been hypothesised these *noise correlations* are a consequence of the hierarchical nature of information processing that is ubiquitous in the brain. The noise correlations observed are weak e.g. average correlation coefficient of 0.12 across pairs of neurons in visual area MT of macaque monkeys [Zohary et al., 1994]. It is thought that the advantage of population coding is its ability to overcome the variability in single neuron responses. Thus, it is natural to ask the question as to whether weak, often undetectable, correlations in pairs of neurons can affect the accuracy with which an encoded stimuli is extracted from a population.

For directionally selective neurons in visual area MT of macaque monkeys, a group of neurons with preferred stimuli that differed by less than 90 degrees had a larger average correlation coefficient than a group with preferred stimuli whose separation was greater than 90 degrees [Zohary et al., 1994]. This is an example of noise correlations with *functional* dependence i.e. a pair of neurons are more strongly correlated if their tuning is similar. This relationship has also been observed in other areas including the motor cortex [Lee et al., 1998] and retinal ganglion cells [Mastronarde, 1983]. Spatial dependence is another form of noise correlation dependency that is found in the brain i.e. the greater the physical distance between two neurons, the weaker their correlation. It has been observed that neurons with similar tuning show uncorrelated noise if the neurons belong to separate cortical columns [Lee et al., 1998]. Determining the precise structure of noise correlations is difficult due to the lack of experimental data. There are many correlation structures which are consistent with the data. As a result, any single model may not reflect the true nature of noise correlations in the brain. It is believed this will improve as recording techniques continue to become more sophisticated.

We focus our attention to models of functional dependence. Theoretical studies typically involve the following two correlation structures: the limited range correlation model [Snippe and Koenderink, 1992] and stimulus-dependent correlations [Josić et al., 2009]. Limited range correlation models the noise correlation as a smoothly decaying function (e.g. exponential function) with increasing separation in preferred stimuli. Stimulus-dependent correlation models the covariance of two neurons as the product of their average responses. This corresponds to a larger correlation for neurons with overlapping tuning curves; clearly the area of overlap increases with increased similarity in tuning. Early work on modelling noise correlations also analysed the effect of uniform correlations [Abbott and Dayan, 1999]. However, uniform correlations are a special case of limited range correlations where the correlation length approaches infinity.

We previously saw that for a population code with independent responses, the mean squared estimation error decays as $\frac{1}{N}$ where N is the population size (line in Figure 4). This corresponds to a positive, linear relationship between Fisher information and population size (dashed line in Figure 6b) i.e. information continues to increase as the number of neurons increases. It is worth noting that the *isotropy* of the system used in modelling ensures that the Fisher information is identical across all stimulus values [Sompolinsky et al., 2001]. Isotropy is guaranteed by the use of homogeneous, bell-shaped tuning curves that are broadly tuned and uniformly distributed. Theoretical studies have shown that a limited range correlation model with weak, positive correlations results in the saturation of information as population size approaches infinity (Figure 6b) [Zohary et al., 1994, Sompolinsky et al., 2001]. A further constraint is that the correlations need to be broad, that is a neuron is correlated with a substantial fraction of the population, typically close in value to the tuning width. This indicates that noise correlations can limit the precision to which stimuli are encoded. Previously, any level of accuracy could be attained since the number of

neurons used by the nervous system in stimulus extraction is unknown. Despite this, it has been shown recently that heterogeneity and stimulus-dependent correlations provide a way to overcome the limit on precision caused by limited range correlations [Ecker et al., 2011, Josić et al., 2009].

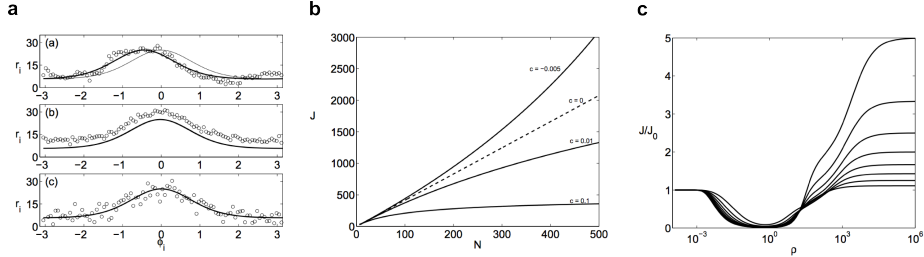


Figure 6: Limited range noise correlations [Sompolinsky et al., 2001]

An interesting property of the limited range correlation model is the non-monotonic relationship between Fisher information and correlation length (Figure 6c). For extremely small values of the correlation length (i.e. approximately less than 10^{-3}), information is identical to that of the independent model. As the correlation length increases to near the value of the tuning width, information decreases to 0. Consequently, as the correlation length further increases, information increases to beyond that of the independent model although it eventually saturates. The amount of information at which it saturates depends on the correlation coefficient used. Despite this, the relationship is independent of the particular coefficient used. The subplots in Figure 6a give intuition as to the cause of this behaviour. The middle subplot shows the response of a model with uniform correlations which is equivalent (for illustration purposes) to limited range correlations with a large correlation length. We observe that the deviation from the average population response is mostly in the amplitude and consistent across all neurons. Thus, it is expected that the accuracy of the population code will increase. This can be seen if we consider a template matching decoder in which the average population responses for different stimuli correspond to the templates. Decoding would only be affected by lateral shifts. The top subplot shows a model in which the correlation length is approximately equal to the tuning width. Trial-to-trial responses appear to shift laterally increasing the possibility of the response being matched by a template generated from a stimulus different to the true value. Hence, we expect the accuracy of the population code to decrease. We have explained the behaviour of Fisher information, which is computed by averaging over all possible responses, from just the responses of single trials. It is important to note that a large correlation coefficient was used in Figure 6a. Despite the coefficient measured experimentally being much smaller, we recall from Figure 6c that the behaviour is independent of the coefficient used.

From this point onwards, we will refer generally to the use of limited range correlations as noise correlations. Regardless of the nature of noise corre-

lations, it is argued that the independent model is not biologically realistic [Averbeck et al., 2006]. The amount of information extracted by sensory receptors (i.e. 1st layer in the information processing hierarchy) is clearly finite. Therefore, in subsequent layers, the amount of information cannot exceed this amount and is expected to saturate for increasing numbers of neurons. We should note that the results in Figure 6 used Gaussian statistics to model the noise. In addition, the variance of single neuron responses was assumed to be identical across neurons (i.e. uniform variance). There is no guarantee that these results hold for other noise distributions. In particular, a Poisson distribution is often used for independent populations, especially when neuronal responses are represented by integer spike counts. The difficulty in generalising the Poisson model to multivariate, correlated populations means a comparison to previous studies is lacking. Most studies also restrict their investigation of noise correlations to the impact on Fisher information. Whilst Fisher information quantifies performance independent of the decoding mechanism, the bound is limited to unbiased estimators. Whether results differ significantly for biased estimation remains to be seen. Lastly, and perhaps most importantly, is the question of whether stimulus extraction in the brain utilises correlation structure in the decoding process. The observation of reduced correlations following attention or learning suggests not since this also coincides with improved psychophysical performance [Cohen and Maunsell, 2009, Gu et al., 2011]. Theoretical studies have also looked at the cost of ignoring correlations with so-called *unfaithful* decoders [Wu et al., 2001].

References

- [Abbott and Dayan, 1999] Abbott, L. and Dayan, P. (1999). The effect of correlated variability on the accuracy of a population code. *Neural computation*, 11(1):91–101.
- [Averbeck et al., 2006] Averbeck, B. B., Latham, P. E., and Pouget, A. (2006). Neural correlations, population coding and computation. *Nature reviews. Neuroscience*, 7(5):358.
- [Cohen and Maunsell, 2009] Cohen, M. R. and Maunsell, J. H. (2009). Attention improves performance primarily by reducing interneuronal correlations. *Nature neuroscience*, 12(12):1594–1600.
- [Cover and Thomas, 2012] Cover, T. M. and Thomas, J. A. (2012). *Elements of information theory*. John Wiley & Sons.
- [Dayan and Abbott, 2001] Dayan, P. and Abbott, L. F. (2001). *Theoretical neuroscience*, volume 806. Cambridge, MA: MIT Press.
- [Ecker et al., 2011] Ecker, A. S., Berens, P., Tolias, A. S., and Bethge, M. (2011). The effect of noise correlations in populations of diversely tuned neurons. *Journal of Neuroscience*, 31(40):14272–14283.

- [Faisal et al., 2008] Faisal, A. A., Selen, L. P., and Wolpert, D. M. (2008). Noise in the nervous system. *Nature reviews. Neuroscience*, 9(4):292.
- [Geisler, 2003] Geisler, W. S. (2003). Ideal observer analysis. *The visual neurosciences*, 10(7):12–12.
- [Georgopoulos et al., 1986] Georgopoulos, A. P., Schwartz, A. B., and Kettner, R. E. (1986). Neuronal population coding of movement direction. *Science*, pages 1416–1419.
- [Gu et al., 2011] Gu, Y., Liu, S., Fetsch, C. R., Yang, Y., Fok, S., Sunkara, A., DeAngelis, G. C., and Angelaki, D. E. (2011). Perceptual learning reduces interneuronal correlations in macaque visual cortex. *Neuron*, 71(4):750–761.
- [Hubel and Wiesel, 1962] Hubel, D. H. and Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat’s visual cortex. *The Journal of physiology*, 160(1):106–154.
- [Hubel and Wiesel, 1968] Hubel, D. H. and Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. *The Journal of physiology*, 195(1):215–243.
- [Josić et al., 2009] Josić, K., Shea-Brown, E., Doiron, B., and de la Rocha, J. (2009). Stimulus-dependent correlations and population codes. *Neural computation*, 21(10):2774–2804.
- [Kay, 1993] Kay, S. M. (1993). *Fundamentals of statistical signal processing*. Prentice Hall PTR.
- [Lee et al., 1988] Lee, C., Rohrer, W. H., and Sparks, D. L. (1988). Population coding of saccadic eye movements by neurons in the superior colliculus. *Nature*, 332(6162):357–360.
- [Lee et al., 1998] Lee, D., Port, N. L., Kruse, W., and Georgopoulos, A. P. (1998). Variability and correlated noise in the discharge of neurons in motor and parietal areas of the primate cortex. *Journal of Neuroscience*, 18(3):1161–1170.
- [Mastronarde, 1983] Mastronarde, D. N. (1983). Correlated firing of cat retinal ganglion cells. i. spontaneously active inputs to x-and y-cells. *Journal of Neurophysiology*, 49(2):303–324.
- [Maunsell and Van Essen, 1983] Maunsell, J. H. and Van Essen, D. C. (1983). Functional properties of neurons in middle temporal visual area of the macaque monkey. ii. binocular interactions and sensitivity to binocular disparity. *Journal of neurophysiology*, 49(5):1148–1167.
- [Orban, 2012] Orban, G. A. (2012). *Neuronal operations in the visual cortex*, volume 11. Springer Science & Business Media.

- [Paradiso, 1988] Paradiso, M. (1988). A theory for the use of visual orientation information which exploits the columnar structure of striate cortex. *Biological cybernetics*, 58(1):35–49.
- [Pouget and Latham, 2007] Pouget, A. and Latham, P. E. (2007). Population codes. *Bayesian Brain: Probabilistic Approaches to Neural Coding*, page 115.
- [Salinas and Abbott, 1994] Salinas, E. and Abbott, L. (1994). Vector reconstruction from firing rates. *Journal of computational neuroscience*, 1(1):89–107.
- [Seriès, 2017] Seriès, P. (2017). Lecture slides for MSc course Computational Cognitive Neuroscience: Decoding. <http://homepages.inf.ed.ac.uk/pseries/CCN17/CCN-wk3-1.pdf>. Accessed: Aug 17, 2017.
- [Snippe, 1996] Snippe, H. P. (1996). Parameter extraction from population codes: A critical assessment. *Neural Computation*, 8(3):511–529.
- [Snippe and Koenderink, 1992] Snippe, H. P. and Koenderink, J. J. (1992). Information in channel-coded systems: correlated receivers. *Biological cybernetics*, 67(2):183–190.
- [Sompolinsky et al., 2001] Sompolinsky, H., Yoon, H., Kang, K., and Shamir, M. (2001). Population coding in neuronal systems with correlated noise. *Physical Review E*, 64(5):051904.
- [Theunissen and Miller, 1991] Theunissen, F. E. and Miller, J. P. (1991). Representation of sensory information in the cricket cercal sensory system. ii. information theoretic calculation of system accuracy and optimal tuning-curve widths of four primary interneurons. *Journal of neurophysiology*, 66(5):1690–1703.
- [Westheimer et al., 1976] Westheimer, G., Shimamura, K., and McKee, S. P. (1976). Interference with line-orientation sensitivity. *JOSA*, 66(4):332–338.
- [Wilson and McNaughton, 1993] Wilson, M. A. and McNaughton, B. L. (1993). Dynamics of the hippocampal ensemble code for space. *Science*, 261(5124):1055–1059.
- [Wu et al., 2001] Wu, S., Nakahara, H., and Amari, S.-I. (2001). Population coding with correlation and an unfaithful model. *Neural Computation*, 13(4):775–797.
- [Zohary et al., 1994] Zohary, E., Shadlen, M. N., and Newsome, W. T. (1994). Correlated neuronal discharge rate and its implications for psychophysical performance. *Nature*, 370(6485):140–143.