Published in final edited form as:
Curr Opin Neurobiol. 2016 April ; 37: 126–132. doi:10.1016/j.conb.2016.01.009.

Feedforward and feedback sources of choice probability in neural population responses

Bruce G. Cumming[1] and **Hendrikje Nienborg**[2]

[1]National Eye Institute, National Institutes of Health, Bethesda, MD, USA

[2]Werner Reichardt Centre for Integrative Neuroscience, University of Tuebingen, 72076 Tuebingen, Germany

Abstract

How the processing of signals carried by sensory neurons supports perceptual decisions is a longstanding question in neuroscience. The ability to record neuronal activity in awake animals while they perform psychophysical tasks near threshold has been a key advance in studying these questions. Trial-to-trial correlations between the activity of sensory neurons and the decisions reported by animals ("choice probabilities"), even when measured across repeated presentations of an identical stimulus provide insights into this problem. But understanding the sources of such co-variability between sensory neurons and behavior has proven more difficult than it initially appeared. Below, we discuss our current understanding of what gives rise to these correlations.

Our perceptual experience of the outside world depends upon the signals delivered to the brain by spiking activity of sensory neurons. How the processing of these inputs allows us to make decisions about the world is a long-standing puzzle in neuroscience . The ability to record neuronal activity in awake animals while they perform psychophysical tasks near threshold has been a key advance in studying these questions. One observation early on was that there are trial-to-trial correlations between the activity of sensory neurons and the decisions reported by animals, even when measured across repeated presentations of an identical stimulus . Thus the activity of sensory neurons contains some information about an animals' upcoming decision, in addition to information about the physical stimulus. The question of what this finding may or may not tell us about how the activity of sensory neurons is linked to perceptual decisions, has engaged experimental and theoretical neuroscientists for many years. Below, we discuss our current understanding of what gives rise to these correlations. In the interest of focus, we will restrict the review of experimental work to studies in the macaque monkey but similar observations have increasingly been made in other species.

Corresponding authors: hendrikje.nienborg@cin.uni-tuebingen.de and bgc@lsr.nei.nih.gov.

Feed-forward interpretations and the role of noise-correlations

The first study to show a correlation systematically predicting choice , studied neurons in primate MT while monkeys performed a direction discrimination task with moving random dot patterns. They quantified the correlation with a non-parametric measure called “Choice Probability” (CP), which has been widely used since. This calculates the probability that a random pick from the measured spike count distribution associated with “preferred” choices is greater than a random pick from the distribution for “null” choices. This is equal to the proportion of responses that an ideal observer would predict choice correctly, given the spike count from two neurons. Britten et al reported a mean CP of 0.55, indicating a modest, but systematic correlation between spike count and choice. Similar observations have been reported in a number of sensory cortical areas [–], and for a variety of tasks [{Bo –].

In a task with two possible responses, uncorrelated activity would produce a CP of 0.5, so a mean value of 0.55 seems small at first sight. But from one perspective it is puzzlingly large. Suppose that CP arises because stochastic fluctuations in the activity of sensory neurons determine the animals’ choices near threshold. If the random fluctuations in firing rate were independent for all the neurons involved, and there are a large number of neurons (more than 100), then the contribution of any one neuron to the choice would be much smaller than this . This study simulated a simple pooling framework, in which a binary decision was based on the summed activity of model neurons in two pools, each contributing evidence towards one of the choices. Shadlen et al pointed out that the paradox can be resolved if the fluctuations are not independent, but instead are positively correlated. Indeed, such “noise correlations”, i.e. correlations between pairs of sensory neurons that cannot be explained by changes in the stimulus, are commonly observed in cortical neurons (see for a review). A crucial requirement for the emergence of CP in this model is that these noise correlations have a particular structure, with high correlations between neurons that belong to the same pool, and weaker correlations between neurons in opposite pools [,]. Again this is compatible with the observed properties of noise correlations, which tend to be stronger in neurons with similar response properties (e.g. [–]). Thus, observed CP are quantitatively compatible with a simple model in which the summed activity of sensory neurons produces psychophysical choices.

Recently, derived an analytical expression for the magnitude of CP in pooling models in which the decision is based on a linear weighting of sensory neurons. It helps to understand many simulation results, including those described by , and highlights the dependence of a neuron’s CP on it’s noise-correlation with the rest of the population.

The importance of noise correlations for neuronal population codes

We describe above how noise correlations play a central role in explaining CP. They also significantly influence how reliably a population of neurons encodes information about the sensory stimulus [,]. Some simple cases can be understood intuitively. If noise is independent in each neuron, then averaging the responses of many neurons will reduce the impact of the noise. If in a large pool of neurons the pairwise correlations are uniform and positive, this implies that a single common input drives the correlated fluctuations (Fig. 1). No matter how many neurons are pooled, the fluctuations driven by this common input will remain, and so the information available about the stimulus saturates as a function of the number of neurons.

While intuitions derived from these simple cases are helpful, they can be misleading. An important result , (Fig. 2A, B) leads to a very different picture: if the members of a given pool have some variation in their response functions (e.g. differences in amplitude) an optimal linear decoder can in general recover information that does not saturate with pool size. Figure 2A illustrates a simple case. Notice that the optimal decoder has weights that depend upon knowledge of noise correlations. A “correlation-blind” decoder will in general be suboptimal.

An important recent study discovered the conditions under which correlations lead to a limit in the growth of information with pool size even for an optimal linear decoder . For a single pair of neurons, the important parameter is the product of their sensitivity (the slopes of the tuning curves of the two neurons) to stimulus changes. If the noise fluctuations in a large population lie along a direction defined by the sensitivity of all the neurons, they will be information-limiting. One way to think about this result is as the pattern of correlations that would be observed if there were no neuronal noise, but if the experimenter changed the visual stimulus. This implies that the pattern of correlations would mimic changes in the stimulus (Fig. 2C). A pattern of correlations that cannot be distinguished from a change in the stimulus obviously produces uncertainty about the stimulus. If the correlations in a population contain a significant component of this sort, they place an upper limit on the information available. Moreno-Bote et al go on to show that it is difficult to detect the presence of such information limiting correlations simply by measuring pairwise correlations. It is necessary to study the activity of entire populations to reveal this feature.

Evidence for feed-forward models and optimal linear readout?

Thus the relationship between information, correlations, and choice probability is complicated, and at first sight it looks like progress will be difficult with the current tools. But a recent study provided a powerful way to examine the relationship under a set of assumptions. First, the psychophysical behavior is at threshold. Second, the brain uses an optimal linear decoder of a sensory population. This implies, third, that the decision process is noise free. For such a scheme, in the case where information-limiting correlations are present, there is a simple relationship between the neurometric threshold (a measure of the discrimination performance of individual neurons based on signal detection theory) and the CP for a given neuron. It predicts the relationship between CP and the ratio (neurometric threshold)/(psychometric threshold) for neurons in that sensory population. Pitkow at al. then analysed CP data from a study of neurons in area MST while animals reported the direction of heading in a virtual reality display, and found that they were compatible with the predicted relationship.

Neurons in area MSTd also show CP for the direction heading task in complete darkness, where the only available information is from vestibular afferents. This produces an opportunity to explore the contribution of peripheral afferents to CP, since for the vestibular system these are fairly simple to interpret. found significant CP in neurons recorded in the vestibular nucleus, which at first sight seems to support a feedforward explanation. However, there are feedback projections from the cortex to the vestibular nucleus, so these could equally play a role in generating CP (see next section). Whatever the origin, this group provided the strongest empirical test that noise correlations are indeed central role in producing CP. Unilateral labyrinthectomy significantly altered the structure of noise correlations in the vestibular nucleus , and this produced changes in CP as predicted by theory .

More recently, the same group also reported the activity of the otolith afferents , which pose an interesting challenge to the linear read-out framework. Similar to cortical neurons, the threshold performance of single neurons was only slightly poorer than that of the animal. Given the large number of afferent neurons (ca. 6,000;), this suggests that there must be information limiting correlations. Despite this, the observed CP did not deviate significantly from 0.5 – that is there was no evidence for a correlation between neuronal activity and choice for the otolith afferents. The lack of significant CPs can be readily explained in a feedback framework (see below), as there are no feedback projections to the otolith. But to explain this result in a linear framework as e.g. used by , seems to require an extremely sub-optimal read-out , contrary to what has been found in the cortex. (Similarly, a recent comparison of the sensitivity for color between the signals in the retina and behavior or cortical signals in V1 also suggested sub-optimal read-out of the primary afferents .) One possibility is that nonlinear transformations that are applied downstream make it impossible to recover all of the information available to a linear decoder of the afferents.

Such nonlinearites may play an important role in sensing translation, since information in the otolith afferents cannot differentiate translation from changes in pose relative to gravity. The ambiguity is resolved by nonlinear combination of afferent signals in the vestibular nucleus . However, it is important to note that in the context of the particular task used by Yu et al (2015) it was not necessary to perform the disambiguation because a linear decoder applied to the otolith afferents is sufficient for the task. For future work it is therefore important to extend the linear framework successfully applied to the cortex to include nonlinear downstream computation. It is possible that many of the principles we have learned from linear decoders will remain broadly applicable. On the other hand, if nonlinear operations downstream of a given sensory population can substantially change these relationships, it is essential we understand the differences.

Feed-back interpretations of choice probability and the origin of noise-correlations

The theoretical insights above [,,] seem entirely feedforward – noise in the sensory representation simply propagates all the way to the animal's choice. However, the central role played by structured noise correlations in these accounts make other explanations equally possible. Indeed, a number of observations[–], including the discrepancy in time course between CPs and a metric quantifying the weights the animal gives to the visual stimulus ('psychophysical kernel amplitude') as a function of time , have challenged purely feedforward explanations. Importantly, a positive noise correlation between a pair of neurons only indicates that they receive some common input that is not derived from the stimulus. A frequent interpretation is that this reflects noise on shared afferent inputs (entirely feedforward), but any other process that generates a common input to a group of neurons could in principle explain the results (Fig. 1). These could include “top-down” or feedback phenomena, such as effects of attention. Note that we here use a broad definition of feedback (top-down) signals, as reflecting signals arising anywhere but the ascending sensory processing chain preceding the recorded neuron. The question as to whether CPs reflect only feed-forward or also feedback phenomena therefore becomes a question about the origin of noise-correlations with the required structure.

It is useful to consider one simple way in which a well-recognized top-down process could give rise to these structured correlations. In area MT, it is well established that feature-based attention can alter the firing rates of neurons in a way that depends on their preferred directions . If an animal is instructed to attend to leftward motion, the activity of neurons with preferred directions near this direction is increased on average, while the activity of rightward preferring neurons is reduced. The opposite pattern is produced if animals are instructed to attend to rightward motion. Now consider the possibility that during a left-right discrimination task the animal engages this same mechanism, sometimes attending to leftward motion, sometimes attending to rightward motion. This fluctuation in feature-based attention will produce exactly the pattern of noise correlations required to produce CP, with high correlations between pairs that both prefer leftward (or rightward) motion (because they are affected in the same way by changes in feature attention), and low correlations between neurons with opposite preferred directions (which are affected differently by feature attention).

One prediction of this account of CP is that changes in task instruction should give rise to changes in noise correlations – when doing a left-right discrimination, fluctuations in feature based attention to left and right will not much alter correlations involving neurons with preferred directions that are up or down. Cohen and Newsome (2008) recorded from pairs of MT neurons while animals performed direction discrimination. The animals performed two versions of the task, along different axes, but crucially used the same zero-signal stimulus in both tasks. Thus any changes in noise-correlation measured during presentation of that stimulus must reflect an effect of task instruction. Cohen and Newsome did indeed find changes with task instruction that were compatible with the scheme laid out above.

Because fluctuations in feature based attention give rise to structured correlations, they will also generate CPs, if the sensory population is read out with a linear decoder. This situation is almost indistinguishable from the standard feedforward model – the only difference is the proposed source of noise correlations. This illustrates how the question of what causes CP is reduced to the question of what gives rise to structured noise correlations (Fig. 1).

Recent studies have found that noise correlations can be influenced by a number of brain states such as attention (e.g. – task engagement , task difficulty, learning , or anesthesia . Most of these studies observed uniform changes in noise correlation, not changes in the structure of the noise correlations, and would therefore not influence CPs. (But see also who observed changes in the structure of noise correlations with spatial attention for a discrimination task.) Statistical modeling , has explained noise correlations in anesthetized and alert animals not performing a task with slow stochastic fluctuations in a global response gain of the neuronal population. A recent elegant extension of these approaches could explain the changes in correlation with spatial attention with gain fluctuations that affected the neuronal population in each hemisphere independently. The gain fluctuations were strongest in task relevant neurons, while task irrelevant neurons were barely affected. This suggests that these fluctuating gain modulations are not noise but rather “meaningful intrinsic signals” . Global gain modulations could reflect fluctuations in alertness, possibly mediated by neuromodulators , while the gain modulations restricted to one hemisphere could reflect fluctuations in spatial attention as observed behaviorally . Note that while these particular fluctuations with spatial attention did not give rise to changes in the structure of noise correlations and hence would not affect CPs, they may be analogous to the fluctuations in feature attention suggested by Cohen and Newsome (2008) , which would impact CPs. The observation of a multitude of such systematic gain fluctuations raises the question about their computational role in the brain.

A number of studies have proposed useful functions for this kind of feedback, (e.g. ,) but one framework in which they play a particularly principled role is probabilistic inference. This combines prior knowledge about the world with incoming sensory information to infer the most likely source of inputs and has long been proposed to underlie perception (e.g. –). Indeed, a recently proposed neural implementation of probabilistic inference using neural sampling , could explain a number of experimental observations on noise correlations, their structure, and CPs. In this framework , cortical sensory neurons (e.g. neurons in MT) are influenced by feedforward input from the sensory periphery (e.g. the retina) and by top-down influences reflecting prior information about the likely structure of the sensory inputs. In the model this prior information reflects the subject’s beliefs about the sensory inputs. In a psychophysical task (e.g. up vs down direction discrimination), a subject typically knows what the discriminanda are (e.g. upward motion and downward motion). The top-down influences on MT neurons therefore will reflect this knowledge. If the belief about the most likely stimulus fluctuates from trial to trial (on some trials the belief is more upward motion, on other trials more downward motion), this will introduce noise-correlations. Importantly, the structure of the noise correlations will reflect the subject’s knowledge of the task. It will change with the task, exactly as observed by Cohen and Newsome (2008), and more recently by in an orientation discrimination task. If the belief relies on information from the preceding choice, this could also explain recently observed effects of preceding choices on the activity of visual neurons in a discrimination task . The model also predicts that neurons with higher selectivity for the task show stronger noise correlations (similar to the observation by). Its strengths are testable predictions about the structure of noise correlations depending on the task and offering a computational role for feedback.

Enormous progress has been made in understanding how sensory neurons support perceptual decisions. Computational models that apply a linear decoder to neurons in a given sensory representation have tremendously advanced our understanding of the nature of how sensory neurons support perceptual decisions. They not only describe CPs, but also constrain the relationship between neuronal signals and psychophysical performance. An important step for the future is to extend this framework to include nonlinear computations downstream of a given sensory population. Neural variability, long regarded as noise, is increasingly viewed as partly reflecting the effect of previously unexplained or ignored signals. Understanding their origin and functional role are important challenges not only for understanding CP, but also the representation of all types of information in the brain.

Acknowledgments

This work was supported by the Intramural Research Program of the National Institutes of Health–National Eye Institute and a Starting Independent Researcher grant to H.N. from the European Research Council (project acronym: NEUROOPTOGEN).

References and papers to highlight

1. Parker AJ, Newsome WT. Sense and the single neuron: probing the physiology of perception. Annu Rev Neurosci. 1998; 21:227–277. [PubMed: 9530497]
2. Logothetis NK, Schall JD. Neuronal correlates of subjective visual perception. Science. 1989; 245:761–763. [PubMed: 2772635]
3. Britten KH, Newsome WT, Shadlen MN, Celebrini S, Movshon JA. A relationship between behavioral choice and the visual responses of neurons in macaque MT. Vis Neurosci. 1996; 13:87–100. [PubMed: 8730992]
4. Celebrini S, Newsome WT. Neuronal and psychophysical sensitivity to motion signals in extrastriate area MST of the macaque monkey. J Neurosci. 1994; 14:4109–4124. [PubMed: 8027765]
5. Uka T, Tanabe S, Watanabe M, Fujita I. Neural correlates of fine depth discrimination in monkey inferior temporal cortex. J Neurosci. 2005; 25:10796–10802. [PubMed: 16291953]
6. Nienborg H, Cumming BG. Decision-related activity in sensory neurons may depend on the columnar architecture of cerebral cortex. J Neurosci. 2014; 34:3579–3585. [PubMed: 24599457]
7. Nienborg H, Cumming BG. Macaque V2 neurons, but not V1 neurons, show choice-related activity. J Neurosci. 2006; 26:9567–9578. [PubMed: 16971541]
8. Shiozaki HM, Tanabe S, Doi T, Fujita I. Neural activity in cortical area V4 underlies fine disparity discrimination. J Neurosci. 2012; 32:3830–3841. [PubMed: 22423103]
9. Bosking WH, Maunsell JH. Effects of stimulus direction on the correlation between behavior and single units in area MT during a motion detection task. J Neurosci. 2011; 31:8230–8238. [PubMed: 21632944]
10. Chen A, Deangelis GC, Angelaki DE. Functional specializations of the ventral intraparietal area for multisensory heading discrimination. J Neurosci. 2013; 33:3567–3581. [PubMed: 23426684]
11. Cook EP, Maunsell JH. Dynamics of neuronal responses in macaque MT and VIP during motion detection. Nat Neurosci. 2002; 5:985–994. [PubMed: 12244324]
12. Dodd JV, Krug K, Cumming BG, Parker AJ. Perceptually bistable three-dimensional figures evoke high choice probabilities in cortical area MT. J Neurosci. 2001; 21:4809–4821. [PubMed: 11425908]
13. Grunewald A, Bradley DC, Andersen RA. Neural correlates of structure-from-motion perception in macaque V1 and MT. J Neurosci. 2002; 22:6195–6207. [PubMed: 12122078]
14. Liu J, Newsome WT. Correlation between speed perception and neural activity in the middle temporal visual area. J Neurosci. 2005; 25:711–722. [PubMed: 15659609]
15. Nienborg H, Cumming BG. Psychophysically measured task strategy for disparity discrimination is reflected in V2 neurons. Nat Neurosci. 2007; 10:1608–1614. [PubMed: 17965712]
16. Palmer C, Cheng SY, Seidemann E. Linking Neuronal and Behavioral Performance in a Reaction-Time Visual Detection Task. J Neurosci. 2007; 27:8122–8137. [PubMed: 17652603]
17. Purushothaman G, Bradley DC. Neural population code for fine perceptual decisions in area MT. Nat Neurosci. 2005; 8:99–106. [PubMed: 15608633]
18. Roelfsema PR, Spekreijse H. The representation of erroneously perceived stimuli in the primary visual cortex. Neuron. 2001; 31:853–863. [PubMed: 11567622]

19. Romo R, Salinas E. Flutter discrimination: neural codes, perception, memory and decision making. Nat Rev Neurosci. 2003; 4:203–218. [PubMed: 12612633]

20. Smolyanskaya A, Haefner RM, Lomber SG, Born RT. A Modality-Specific Feedforward Component of Choice-Related Activity in MT. Neuron. 2015; 87:208–219. [PubMed: 26139374]

21. Uka T, DeAngelis GC. Contribution of area MT to stereoscopic depth perception: choice-related response modulations reflect task strategy. Neuron. 2004; 42:297–310. [PubMed: 15091344]

22. Shadlen MN, Britten KH, Newsome WT, Movshon JA. A computational analysis of the relationship between neuronal and behavioral responses to visual motion. J Neurosci. 1996; 16:1486–1510. [PubMed: 8778300]

23. Cohen MR, Kohn A. Measuring and interpreting neuronal correlations. Nat Neurosci. 2011; 14:811–819. [PubMed: 21709677]

24. Nienborg H, Cumming B. Correlations between the activity of sensory neurons and behavior: how much do they tell us about a neuron's causality? Curr Opin Neurobiol. 2010; 20:376–381. [PubMed: 20545019]

25. Nienborg H, Cohen MR, Cumming BG. Decision-related activity in sensory neurons: correlations among neurons and with behavior. Annu Rev Neurosci. 2012; 35:463–483. [PubMed: 22483043]

26. Zohary E, Shadlen MN, Newsome WT. Correlated neuronal discharge rate and its implications for psychophysical performance. Nature. 1994; 370:140–143. [PubMed: 8022482]

27. Kohn A, Smith MA. Stimulus dependence of neuronal correlation in primary visual cortex of the macaque. J Neurosci. 2005; 25:3661–3673. [PubMed: 15814797]

28. Ecker AS, et al. Decorrelated neuronal firing in cortical microcircuits. Science. 2010; 327:584–587. [PubMed: 20110506]

29. Haefner RM, Gerwinn S, Macke JH, Bethge M. Inferring decoding strategies from choice probabilities in the presence of correlated variability. Nat Neurosci. 2013; 16:235–242. [PubMed: 23313912]

30. Abbott LF, Dayan P. The effect of correlated variability on the accuracy of a population code. Neural Comput. 1999; 11:91–101. [PubMed: 9950724]

31*. Shamir M, Sompolinsky H. Implications of neuronal diversity on population coding. Neural Comput. 2006; 18:1951–1986. In a population of neurons with identical tuning curves, positive noise correlations limit the information available about stimulus changes as the pool size increases. This paper was the first to demonstrate that with heterogeneous tuning curves (e.g. variation in amplitude or tuning width) in general information grows indefinitely with pool size. [PubMed: 16771659]

32. Ecker AS, Berens P, Tolias AS, Bethge M. The effect of noise correlations in populations of diversely tuned neurons. J Neurosci. 2011; 31:14272–14283. [PubMed: 21976512]

33*. Moreno-Bote R, et al. Information-limiting correlations. Nat Neurosci. 2014; 17:1410–1417. This paper identified the circumstances under which noise correlations limit the information available about a stimulus even from an infinite population of neurons with heterogenous tuning curves. They also showed that measurements of pairwise correlations alone are not sufficient to see if these conditions are satisfied. [PubMed: 25195105]

34. Pitkow X, Liu S, Angelaki DE, DeAngelis GC, Pouget A. How Can Single Sensory Neurons Predict Behavior? Neuron. 2015; 87:411–423. [PubMed: 26182422]

35. Liu S, Gu Y, DeAngelis GC, Angelaki DE. Choice-related activity and correlated noise in subcortical vestibular neurons. Nat Neurosci. 2013; 16:89–97. [PubMed: 23178975]

36**. Yu, X-j; Dickman, JD.; DeAngelis, GC.; Angelaki, DE. Neuronal thresholds and choice-related activity of otolith afferent fibers during heading perception. Proceedings of the National Academy of Sciences. 2015; 112:6467–6472. This was the first study to measure CP systematically in peripheral afferents, using recordings from Otolith afferents while animals were moved on a platform in complete darkness. Although cortical neurons show significant CP in this task, the Otolith afferents did not. The performance of individual afferents implies that either there are noise limiting correlations (which should produce CP), or that a simple decoding of the afferent neurons would produce much better performance than the animals.

37. GACEK RR, RASMUSSEN GL. Fiber analysis of the statoacoustic nerve of guinea pig, cat, and monkey. Anat Rec. 1961; 139:455–463. [PubMed: 13946037]

38. Beck JM, Ma WJ, Pitkow X, Latham PE, Pouget A. Not noisy, just wrong: the role of suboptimal inference in behavioral variability. Neuron. 2012; 74:30–39. [PubMed: 22500627]

39. Hass CA, Angueyra JM, Lindbloom-Brown Z, Rieke F, Horwitz GD. Chromatic detection from cone photoreceptors to V1 neurons to behavior in rhesus monkeys. J Vis. 2015; 15:1. [PubMed: 26523737]

40. Angelaki DE, Shaikh AG, Green AM, Dickman JD. Neurons compute internal models of the physical laws of motion. Nature. 2004; 430:560–564. [PubMed: 15282606]

41*. Nienborg H, Cumming BG. Decision-related activity in sensory neurons reflects more than a neuron's causal effect. Nature. 2009; 459:89–92. This study used simultaneous measurements of time-resolved perceptual strategy and CP, and tuning curves separated by choice. Together these findings were difficult to reconcile in a purely feed-forward account of CP, arguing for a top-down contribution to CP. [PubMed: 19270683]

42. Wimmer K, et al. Sensory integration dynamics in a hierarchical network explains choice probabilities in cortical area MT. Nat Commun. 2015; 6:6177. [PubMed: 25649611]

43. Siegel M, Buschman TJ, Miller EK. Cortical information flow during flexible sensorimotor decisions. Science. 2015; 348:1352–1355. [PubMed: 26089513]

44. Treue S, Martinez Trujillo JC. Feature-based attention influences motion processing gain in macaque visual cortex. Nature. 1999; 399:575–579. [PubMed: 10376597]

45**. Cohen MR, Newsome WT. Context-dependent changes in functional circuitry in visual area MT. Neuron. 2008; 60:162–173. This is the first study to provide compelling empirical evidence for the view that noise correlations do not just reflect noise resulting from sensory processing but also depend on cognitive factors, such as task instruction. [PubMed: 18940596]

46. Cohen MR, Maunsell JH. Attention improves performance primarily by reducing interneuronal correlations. Nat Neurosci. 2009; 12:1594–1600. [PubMed: 19915566]

47. Cohen MR, Maunsell JH. Using neuronal populations to study the mechanisms underlying spatial and feature attention. Neuron. 2011; 70:1192–1204. [PubMed: 21689604]

48. Mitchell JF, Sundberg KA, Reynolds JH. Spatial attention decorrelates intrinsic activity fluctuations in macaque area V4. Neuron. 2009; 63:879–888. [PubMed: 19778515]

49. Downer JD, Niwa M, Sutter ML. Task engagement selectively modulates neural correlations in primary auditory cortex. J Neurosci. 2015; 35:7565–7574. [PubMed: 25972181]

50. Ruff DA, Cohen MR. Global cognitive factors modulate correlated response variability between V4 neurons. J Neurosci. 2014; 34:16408–16416. [PubMed: 25471578]

51. Gu Y, et al. Perceptual learning reduces interneuronal correlations in macaque visual cortex. Neuron. 2011; 71:750–761. [PubMed: 21867889]

52. Ecker AS, et al. State dependence of noise correlations in macaque primary visual cortex. Neuron. 2014; 82:235–248. [PubMed: 24698278]

53. Ruff DA, Cohen MR. Attention can either increase or decrease spike count correlations in visual cortex. Nat Neurosci. 2014; 17:1591–1597. [PubMed: 25306550]

54. Goris RLT, Movshon JA, Simoncelli EP. Partitioning neuronal variability. Nature neuroscience. 2014; 17:858–865. [PubMed: 24777419]

55*. Rabinowitz NC, Goris RL, Cohen M, Simoncelli E. Attention stabilizes the shared gain of V4 populations. Elife. 2015; 4 This study analyzed population recordings from V4 in a spatial attention task. It finds that the observed changes in noise correlation can be explained by four separate time-varying gain modulators that are shared between neurons in a population. Two of these modulate the responses in only one hemisphere, providing a very simple explanation for observed changes with spatial attention.

56. Harris KD, Thiele A. Cortical state and attention. Nat Rev Neurosci. 2011; 12:509–523. [PubMed: 21829219]

57. Cohen MR, Maunsell JH. A neuronal population measure of attention predicts behavioral performance on individual trials. J Neurosci. 2010; 30:15241–15253. [PubMed: 21068329]

58. Engel TA, Chaisangmongkon W, Freedman DJ, Wang X-J. Choice-correlated activity fluctuations underlie learning of neuronal category representation. Nature communications. 2015; 6

59. Gregory RL. Perceptions as hypotheses. Philos Trans R Soc Lond B Biol Sci. 1980; 290:181–197. [PubMed: 6106237]

60. Helmholtz, vHLF. Handbuch der physiologischen Optik. Voss; Leipzig: 1867.
61. Lee TS, Mumford D. Hierarchical Bayesian inference in the visual cortex. J Opt Soc Am A Opt Image Sci Vis. 2003; 20:1434–1448. [PubMed: 12868647]
62. Ma WJ, Beck JM, Latham PE, Pouget A. Bayesian inference with probabilistic population codes. Nat Neurosci. 2006; 9:1432–1438. [PubMed: 17057707]
63. Gold JI, Shadlen MN. The neural basis of decision making. Annu Rev Neurosci. 2007; 30:535–574. [PubMed: 17600525]
64. Pouget A, Beck JM, Ma WJ, Latham PE. Probabilistic brains: knowns and unknowns. Nat Neurosci. 2013; 16:1170–1178. [PubMed: 23955561]
65. Nienborg H, Roelfsema PR. Belief states as a framework to explain extra-retinal influences in visual cortex. Current opinion in neurobiology. 2015; 32:45–52. [PubMed: 25463564]
66. Haefner RM, Berkes P, Fiser J. Perceptual decision-making as probabilistic inference by neural sampling. 2014 arXiv:1409.0257.
67. Buesing L, Bill J, Nessler B, Maass W. Neural dynamics as sampling: a model for stochastic computation in recurrent networks of spiking neurons. PLoS Comput Biol. 2011; 7:e1002211. [PubMed: 22096452]
68. Fiser J, Berkes P, Orban G, Lengyel M. Statistically optimal perception and learning: from behavior to neural representations. Trends Cogn Sci. 2010; 14:119–130. [PubMed: 20153683]
69. Bondy, AG.; Cumming, BG. Top down signals infuence the distribution of noise correlations amongst sensory neurons. Society for Neuroscience Annual Meeting; 2013. Abstract Nr. 311.02
70. Nienborg, H.; Macke, JH. Using sequential dependencies in neural activity and behavior to dissect choice related activity in V2. Society for Neuroscience Annual Meeting; 2014. Abstract: 11628
71. Renart A, Machens CK. Variability in neural activity and behavior. Curr Opin Neurobiol. 2014; 25:211–220. [PubMed: 24632334]

Highlights

- Decision-related activity (choice-probabilities, CP) in sensory neurons is widely found
- Correlated variability in sensory neurons limits information in some cases
- The structure of correlated variability in sensory neurons influences CPs
- CPs and correlated variability likely have feed-forward and feed-back sources

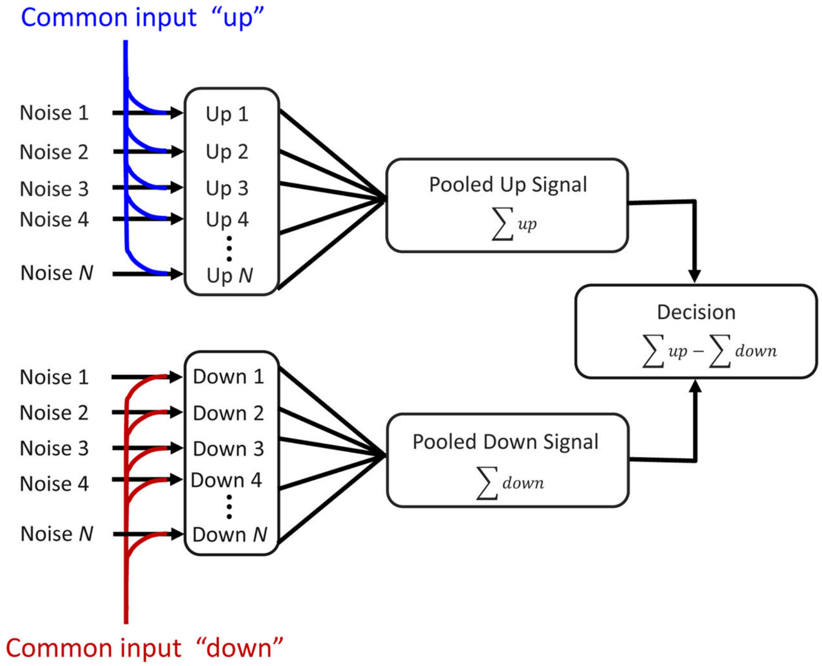

Figure 1.

A simple scheme that produces the structured correlations required for CP in pooling models. The response of each neuron on any trial is the sum of two terms – a noise term that is independent for each neuron, and a common input that is the same for each neuron within a pool. The common inputs produce uniform positive noise correlations between pairs of neurons within a pool, but no correlation between pools. As N becomes large the pooled signals become dominated by the common input terms, since the independent noise terms tend to cancel. Thus the choices of the model (whether pooled up > pooled down) are largely determined by the common input terms. These therefore also determine CP in the model neurons. This property of linear pooling models remains true regardless of what gives rise to the common input terms. Whether these reflect noise in afferent neurons, or feedback from higher areas makes no difference.

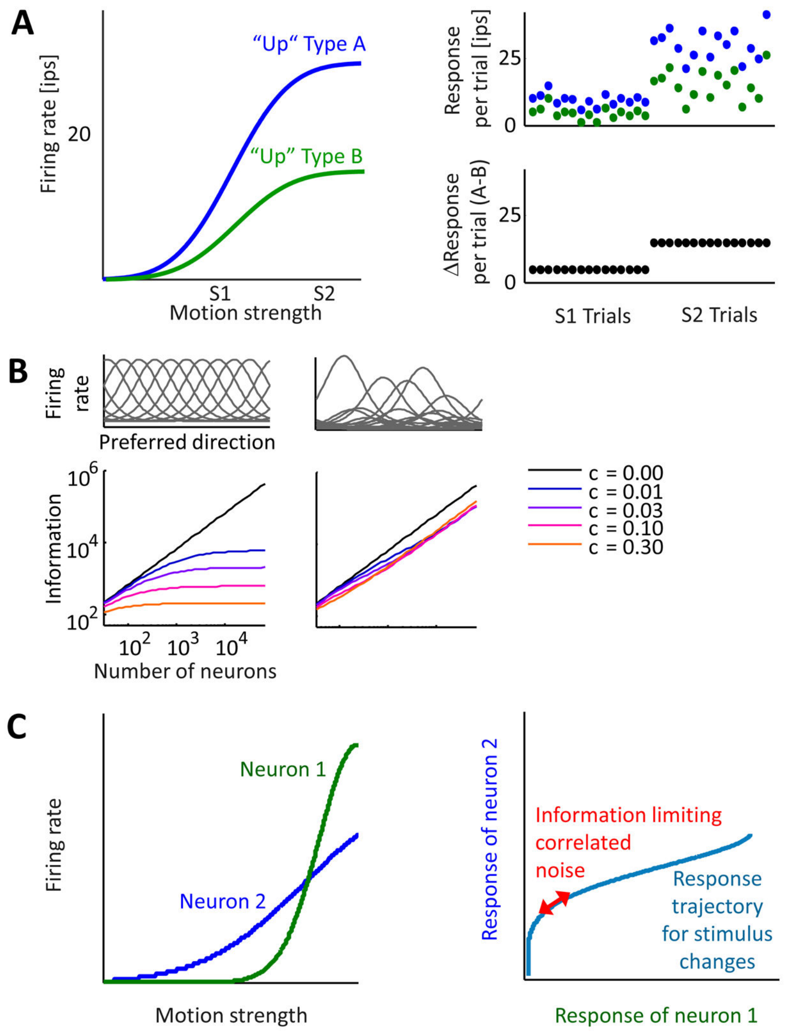

Figure 2.

Positive noise correlations need not limit the information available in a single pool, provided there is heterogenous tuning. **A** Consider a population of upward preferring neurons composed of two types: one with stronger rate modulation (type A) and the other weaker (type B). A common input to the pool (both type A and type B) generates uniform positive correlations between all pairs (see Fig. 1). The right panel illustrates across different trials of two stimulus strengths (S1 and S2), showing a high correlation between the trial-by-trial responses of the two neurons (blue and green circles, respectively). However, taking the difference of the responses between neuron A and B (bottom panel, black circles) removes this correlated noise, without losing the information about the stimulus. For illustration, a high correlation between the two neurons was used. Note that if the trial responses here represented the sum over a large population of type A neurons and type B neurons, the correlation would indeed approach 1 as pool size increases. This is because mean is dominated by the common input. The tuning heterogeneity means this subtraction produces a signal whose mean still depends on stimulus strength, so that here the ratio signal/noise increases with the number of neurons, with no upper bound. This is illustrated in panel **B**, where Fisher information is plotted for increasing homogeneous neuronal populations (left) and heterogeneous neuronal populations (right). The two top panels schematically depict the tuning curves in the respective populations (modified, with permission, after Ecker et al. 2011). **C** The situation when even an optimal linear decoder cannot remove the effect of correlated noise is illustrated with only two neurons. Left panel: the tuning curves of two neurons to the task relevant stimulus are shown. Right panel: The blue line depicts the response trajectory that corresponds to changes of the stimulus along the task-relevant dimension. If the correlated noise affects the population response along this trajectory, it is information limiting, since it cannot be differentiated from changes in the stimulus.