



Noise correlations and perceptual inference

Mihály Bányai^{a,b} and Gergő Orbán^{a,b}

The presence of stimulus-unrelated covariability in neural responses poses a challenge to models of perception and behavior. Here we review advances in theory and experiments that highlight how noise correlations are implicated in perceptual processes. We focus on the structure of correlations and variations in it to reveal the computations performed by the cortex. We argue that to discover the functional role of correlations in cortical computations, and make predictions about them, we need to harness hierarchical encoding models. We propose that in hierarchical models the source of noise correlations is dual: inference of both task-related and higher-level perceptual variables contribute to their emergence, which is reflected in the task- and stimulus-dependence of the correlation structure.

Addresses

^a Computational Systems Neuroscience Lab, Wigner Research Centre for Physics, Budapest, Hungary

^b Center for Cognitive Computation, Central European University, Budapest, Hungary

Corresponding author: Orbán, Gergő (orban.gergo@wigner.mta.hu)

Current Opinion in Neurobiology 2019, **58**:xx-yy

This review comes from a themed issue on **Computational neuroscience**

Edited by **Máté Lengyel** and **Brent Doiron**

<https://doi.org/10.1016/j.conb.2019.09.002>

0959-4388/© 2019 Elsevier Ltd. All rights reserved.

Introduction

Noise correlation (NC), pairwise correlation between neuronal responses upon repeated presentation of the same stimulus, is a curse for some and a blessing for others. On the one hand, it is considered as a curse because it is a signature of redundancy, which can limit the amount of information that can be encoded in population activity. As a result of covariations, variability in the responses of individual neurons cannot be averaged out by pooling the responses of a population and therefore information about stimuli saturates [1]. On the other hand, NCs can be regarded as a blessing since such correlations can be signatures of the synergies between neurons in a code, where joint activity patterns of neurons carry additional information beyond that carried by

individual activations [2,3]. Here, recurrent interactions of the neural circuitry are expected to shape feedforward information coming from upstream areas. Such duality in the interpretation of this seemingly simple aspect of the neuronal response statistics is fascinating on its own right but recent studies demonstrated that there is more to NCs than the humble name suggests. While the term *noise* correlation suggests that it is related to factors that cannot be traced back to meaningful quantities, systematic and predictable changes in NCs indicate that this form of variability is linked to perceptual and cognitive factors defying the idea of it being merely noise [4*,5,6,7**,8*].

Understanding NCs can potentially answer fundamental questions about the neural code. As a consequence, recent years have seen a surge in analyzing a number of aspects of NCs. First, although the measurement of correlations is only one step beyond the assessment of mean responses on the way to the characterization of the full response statistics, despite the apparent simplicity, multiple studies have been devoted to establishing how reliable measurement of NCs can be achieved [9,10]. Second, NCs provide an insight into the computational processes performed in the neural circuitry, which might be more general than sensory processing. Therefore, the principles underlying the emergence of NCs are important to understand sensory processing and might be relevant in non-sensory areas as well. Accordingly, NCs were explored in a wide range of areas of the central nervous system, from the retina to visual, sensorimotor, and auditory cortices [11–15]. Third, since feedback plays a key role in shaping NCs, network studies have been aiming to explain how network dynamics result in specific patterns in NCs [16–18]. Fourth, theories have been proposed on how correlations are related to information theoretic quantities through the sensitivities of neurons [19–21] or how computations predict the emergence of NCs [22*,23*,24,25]. This point has a particular importance in our view since understanding the role of NCs means to be able to predict changes in them related to changes in experimental conditions. In this review we capitalize on the fact that NCs concern sensory cortices and therefore address a burning question: how are noise correlations related to perception? In particular, we are focusing on the visual cortex since recently important insights were obtained by analyzing visual cortical phenomena. First, we investigate one of the most informative paradigms for studying perceptual processes, the study of discrimination capacities of elementary features. Next, we review how the internal dynamics in sensory cortices contributes to the perception of complex stimuli. Finally, we provide a functional framework that can integrate a wide range of findings on patterns in noise correlations.

Relating neuronal selectivities to elementary features

Two key experimental paradigms central to study low-level perception are discrimination and detection experiments. Using elementary stimulus features to investigate the properties of perception has the appeal that the stimulus can be directly mapped to perceptual performance through an exhaustive parametric exploration of the stimulus space. More importantly, matching the explored stimulus features to the selectivities of neurons provides a chance to relate perceptual performance to neuronal responses. Stimulus selectivities of neurons are identified with receptive field properties, which can be related to a range of simple stimulus features, such as orientation or movement direction, depending on the area of the visual cortex [26,27]. The discrimination task motivates decoding as a computational framework to assess how neuronal responses are related to perceptual performance [28]. In particular, characterizing encoding in neurons with their linear response properties using receptive fields implies the choice of a linear decoder to assess the information carried by neurons about simple stimuli fully characterized by a set of parameters. Decoding performance is affected by the variability present in neuronal responses. Fine discrimination requires that the effects of noise are eliminated, which can be achieved for the variances in individual neuronal responses by pooling the responses of a large number of neurons. However, joint variations introduce redundancy in the neural code, which prevents averaging from completely eliminating the effects of noise. The study of the effect of correlated activity between neurons on decoding performance has a venerable history in theoretical neuroscience [29]. These studies clarified a number of key determinants of decoding performance in the presence of correlations, including the form of correlations that result in a decrease in decoding performance and those that are beneficial for decoding [30]; the way tuning curve properties [31,19,20], or correlations in stimulus characteristics [32] affect decoding. While a decoder can indicate how efficiently stimulus discrimination can be performed, it remains unclear how behavioral performance in a discrimination task is related to the responses of neurons.

The most direct link between performance and neuronal responses is established by the measure of choice probability (CP) [33,34]. CP characterizes how well choices of the animal can be predicted from neural responses and therefore establishes a connection between variations in visual cortical responses and variations in behavior. Early analyses of the hypothesis that variability in neuronal responses results in variations of behavior have revealed that correlated variability among neurons can explain choice-related behavior [35]. This computational study specifically explored the form of NC structure that could affect behavioral output through affecting the readout of information from a population of neurons. Theoretical

considerations have also established a closed form relationship between NCs, choice probability, and the readout weights of downstream neurons reading the output of sensory neurons [36], thus linking neural variability and variations in behavior. A further indication that perceptual performance can only be understood by analyzing the joint aspects of the response statistics of multiple neurons is coming from the analysis of perceptual and neural thresholds. Perceptual threshold measures the signal strength at which behavioral deviates from chance in a discrimination task, while neural threshold measures the signal strength at which the signal can be decoded from responses. Using a single neuron to establish a neural threshold reveals that neural and perceptual thresholds are similar [37], which raises the question why the brain cannot be more effective than a single neuron if it can access the output of large populations of neurons? Choice probabilities and high neural thresholds have been linked by a specific component of NCs, information limiting correlations (ILCs), also termed differential correlations [38^{*}]. ILCs establish the form of correlation that can effectively limit the contribution of individual neurons to decoding stimulus information by linking correlations between pairs of neurons to the properties of their tuning functions [39^{*}]. ILCs characterize the covariation between a pair of neurons that would be expected from changes in the stimulus and is thus detrimental to stimulus identification since covariability induced by stimulus change cannot be distinguished from spontaneously occurring covariation.

ILCs provide an intriguing insight into the neural code: both the correlation of behavioral responses with neural activations and high neural thresholds can be the result of a specific form of NC. This specific form is obtained as the correlation between a pair of neurons that is aligned with the relative rate of change in the mean activities when the stimulus is varied, that is the derivative of tuning curves. An analysis of neural responses in a vestibular heading discrimination task demonstrated that choice probabilities could almost exclusively be traced back to ILCs. In other tasks, however, no such exclusivity was found [40]. Whether this is a result of simplifying assumptions of the original analysis (fine discrimination assumption, invariant response covariance structure, simplistic structure of neural sensitivities), other interfering cognitive processes, or an adaptation that is not specific to the task being performed but more general goals [41] is the subject of future research.

An exciting property of ILCs is revealed by analyzing the ILC structure in different tasks: they are defined through the rate of change of mean responses upon changes in the variable *relevant* for the task. Therefore, switching between different tasks implies a different ILC structure [39^{*}]. As a consequence, if ILCs were an emergent property of task execution, changes in task structure

would imply changes in ILCs and therefore changes in the total NC structure. Indeed, such changes were found in V1 of behaving monkeys [7[•]], which highlights the fact that ILCs are tightly linked to perceptual performance in behavioral tasks. Similar task-dependent modulations of the noise correlation structure were observed in V4 of monkeys while performing a pair of discrimination tasks [8[•]]. In conclusion, behavioral performance in perception of elementary features provides specific indications on the form of correlations that is consistent with the execution of elementary tasks and also with switches between tasks. It remains elusive, however, what principles dictate the emergence of particular patterns in correlations. Furthermore, ILCs constitute only a minor fraction of the total noise correlation structure that is observable in neural populations [42,43[•]], which leaves the question open whether the rest of the structure can be linked to perception (Box 1).

Neural representations related to the perception of complex images

Neural responses are adapted to the statistics of natural stimuli [48] and therefore studying elementary features has a number of limitations. First, tasks such as orientation discrimination are much simpler than those faced by the animal under natural conditions, therefore the full complexity of the neural code (e.g. nonlinear response properties) does not necessarily unfold under such conditions. Mismatch between laboratory tasks and natural tasks to which visual system is adapted has been recently invoked to explain the apparently suboptimal decoding of stimulus-information [43[•],41]. Second, elementary stimuli are not expected to be effective for testing the refined hierarchical machinery dedicated to processing visual stimuli. Linking theories of neural response statistics to perception of complex images can benefit from computational tools that can perform perceptual tasks with an efficiency comparable to the performance of humans. Recent developments in deep neural networks delivered a remarkable performance improvement in machine-classification of images of real-life complexity [49], and could match the performance of humans [50] (see Figure 1a). Interestingly, these artificial neural networks seemed to establish representations matching some characteristics of the neural representations in higher-level vision of primates [51,52]. Noise correlations in strictly feedforward architectures, that are core to these deep neural networks, can emerge but are solely transformed versions of the noise present in lower layers of processing. As a consequence, NCs have no functional role in perception and cannot contribute to shaping behavior. If neuronal representations for perceiving natural images can indeed be efficiently described by feedforward architectures then the functional role of noise correlations is expected to be limited.

Perceptual capacities of deep neural networks have been taken under scrutiny by a number of recent studies [53,54]. Models are characterized not only by the amount of variance they predict in the neural and behavioral responses, but also by how much the patterns of errors in task performance match those of animals or humans. A detailed joint analysis of perceptual performance and neural responses on a core object classification task has revealed deteriorated performance of a deep neural network on a subset of complex synthesized images and this subset was well aligned with the subset of images in which the deep neural network performed poorly on predicting mean responses of inferotemporal cortical neurons [55^{••}]. Such failure of this otherwise effective feedforward architecture both in task performance and in predicting mean neural responses is potentially pointing to the need for recurrent architectures [49,56]. Other approaches that contrast the performance of humans and various models of complex images found differences between humans and feedforward deep neural networks in a task where they were required to learn handwritten characters, but a probabilistic program learner could perform the task similar to humans [57]. These observations indicate that perception of complex images requires a computational architecture that relies on recurrent computations, which incur a rich noise correlation structure. Concurring with this, recent recordings have revealed that noise correlations depend on stimulus identity and stimulus statistics when the animal is presented with complex images [4[•],46^{••}].

Noise correlations in hierarchical probabilistic models

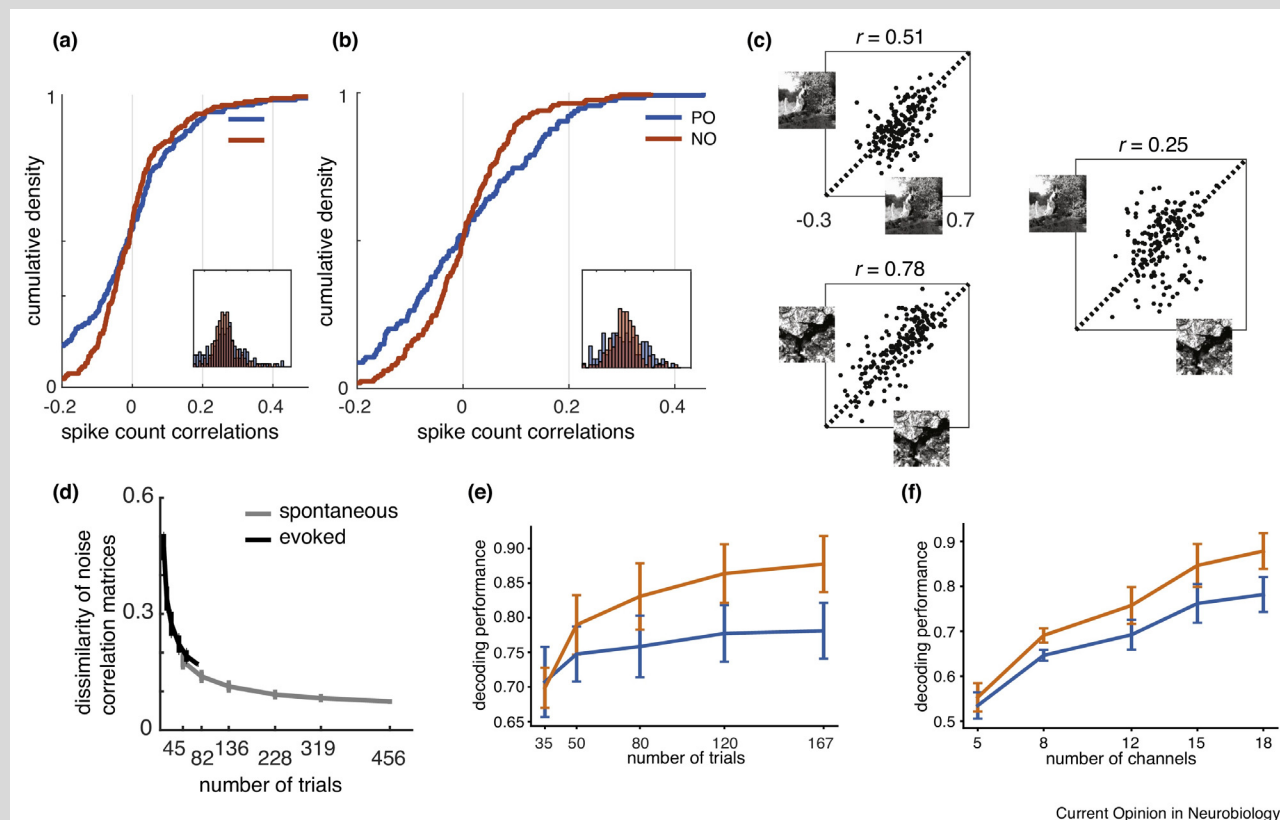
The above evidence demonstrates that NCs have specific structure, such structures can affect perception, and NCs can also undergo systematic changes. The burning question is why the nervous system would invest in producing a code that is inefficient and whether there are computational principles that can reveal the functional role of noise correlations.

We argued that while feedforward architectures are effective in accounting for both perceptual performance and the properties of neural representation, both of these deteriorate for a subclass of images, suggesting the involvement of recurrent connections in cortical computations. Changes in the pattern of NCs [8[•]] motivated a functional approach which argued that performing a task requires inferences at different levels of a computational hierarchy: at the level of the sensory cortex inferences are made about the contribution of visual features to the stimulus, and about task variables at a higher level of processing. Inferences at different levels of the hierarchy are characterized by uncertainty and therefore neurons representing the visual features and task variables need to represent a probability distribution instead of the most probable values. Such probability distributions are

Box 1 Modulation of correlation patterns by the stimulus.

Many studies look at how the population mean of NCs change in response to changes in experimental conditions [44,45]. However, measurements show that often it is not the mean, but other aspects of the NC distribution that are affected by changes in the stimulus. The variance of NCs is modulated by stimulus contrast and orientation. If stimuli are complex, patterns in NCs are also modulated by stimulus identity. The ability to assess such modulations critically depends on the experimental design, as the estimation of the fine structure of correlations is only possible if there was a sufficient number of trials recorded using each stimulus.

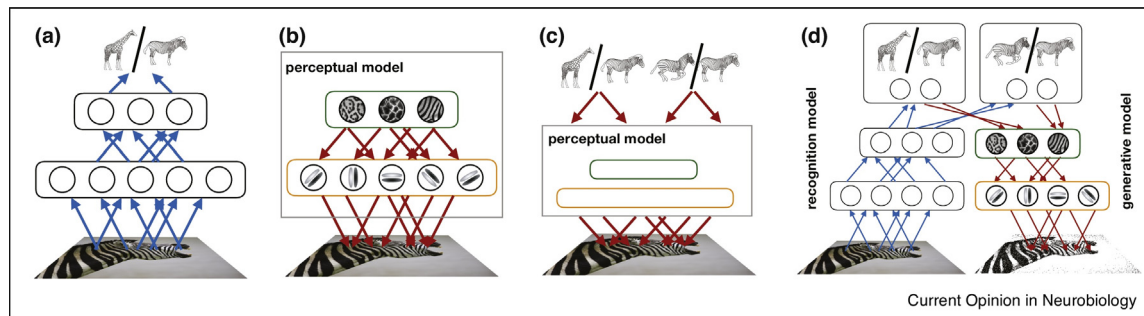
The effect of stimulus-specificity of NCs can be assessed by establishing whether information with respect to stimulus-identity can be decoded more efficiently when relying not only on the mean responses but on the stimulus-dependence of correlations as well. Using mean and correlation structure measured in recorded data (data from [46**]), it can be demonstrated that knowledge of the fine structure of NCs can enhance decoding compared to a case when only a stimulus-independent correlation structure is estimated together with the stimulus-specific mean of the responses (Figure e). More interestingly, such an advantage from stimulus-specific NCs is more moderate at smaller population sizes (Figure f). This insight is surprising because a larger population implies a quadratically increasing number of parameters, which makes the correlation matrix more prone to overfitting. The fact that a decoder based on stimulus-specific correlations becomes more efficient than the one that is agnostic to stimulus-specificity highlights that the advantage of a more intricate correlation structure in high-dimensional data overcomes the increased difficulty of estimating the correlation structure.



Current Opinion in Neurobiology

Stimulus-related effects in the NC distribution beyond the mean and limitations of NC measurability (**a and b**). Distribution of noise correlation across cell-pairs in a population of V1 neurons under different stimulus conditions when monkeys are presented with static grating stimuli. While the distributions (histograms in insets) do not reveal clear differences in the means when comparing responses to high contrast (HC) and low contrast stimuli (LC) (**a**), or preferred orientation (PO) and non-preferred orientation (NO) grating stimuli (**b**), the cumulative distributions (continuous lines) reveal differences in the widths of the distributions (adapted from [47], data from [45]). (**c**) Complex stimulus content modulates the fine structure of NC patterns, as shown by the changes in correlations between pairs of neurons (dots) when NCs are calculated from the responses to two natural image patches (right panel). Full correlation matrices of the population are calculated from 80 presentations of the same image patch. As a control, NCs between pairs of neurons were better aligned when NCs are calculated from non-overlapping sets of trials in response to the same stimulus (left panels). (**d**) Precision of the estimation of noise correlations, measured as the dissimilarity of NC matrices estimated for two sets of trials under the same stimulus condition (evoked: responses to identical images; spontaneous: activity measured without stimulation) strongly depends on the number of trials per stimulus. Panels (**c**) and (**d**) are adapted from [46**]. (**e and f**) Decoding stimulus identity from spiking data as a function of the number of trials used for estimation of decoder parameters (**e**) and the size of the population (**f**). Parameters of the decoders were estimated from training data from three sessions of multiunit activity in response to natural stimuli in V1. One of the decoders estimated both the mean of the responses and a correlation structure for every single stimulus (orange line). The other decoder estimated the mean responses and a correlation structure that was not specific to the stimulus identity (blue line). Difference in decoding performance increases both with the number of simulated trials (**e**) and with the number of simulated channels (**f**).

Figure 1



Hierarchical model structures for vision, featuring perceptual and task variables. **(a)** Deep feedforward networks learn a feature hierarchy from the stimulus that supports the task of object classification. Variables learned by the model can be compared with the activity of neurons while presenting the animal with the same image. Such a perceptual model can capture the mean responses of neurons, but ignores top-down effects and the representation of uncertainty, both key to understanding NCs. **(b)** The model of perception as probabilistic inference posits that the goal of learning is to infer an internal model of the world which describes how observed stimuli are generated. Features in the environment correspond to latent variables in the model, and dependencies between higher and lower-level latents describe the compositional rules for how high-level features of the perceptual hierarchy constrain low-level features. While performing inference in the model, inferences about high-level variables introduce top-down effects at lower levels, resulting in NCs. **(c)** Learning a task corresponds to learning task variables, for example discrimination boundaries between animal species or different potential behaviors of a specific animal, acting on the variables of the perceptual model. Learning the task structure implies that we learn the distribution of possible images that corresponds to different states of task variables. Task variables are inferred in individual trials and this inference introduces top-down effects shaping NC patterns in the perceptual model. **(d)** Inference of feature hierarchies in generative models is possible by augmenting the original model architecture with a feedforward recognition model that infers high-level latent variables from the stimulus, which are fed back to the perceptual hierarchy in a top-down manner. Different tasks may correspond to different sets of high-level latent variables inferred on top of shared recognition and generation hierarchies.

generally multivariate therefore define joint distributions over visual features that are represented by individual neurons or neuron populations. Probabilistic sampling provides a computational framework for the neural representation of probability distributions, such that variability of individual responses and correlations are naturally occurring as part of the representation of uncertainty [58,23[•]]. Such variability in the representation of task variables introduces collective modulations of sensory cortical correlations through top-down influences [22[•],43[•]] (although see [59] for a discussion of temporal dynamics of ILCs). A phenomenological model of the response statistics of V1 neurons also identified global modulations as a potential source of NC patterns which also accounted for a form of stimulus-dependence of the NC structure [6]. A further consequence of the collective modulation by the task variable is the task-dependence of the NC structure, which has been demonstrated in monkey recordings [7^{••}]. Alternatively, such top-down effects can manifest in the fluctuations of the prior expectations which can also introduce correlated variability [60]. Fluctuations in top-down modulations are also consistent with the form of single-cell variability observed along the visual pathway [61] as well as global changes in the amount of shared variability due to attention [62].

Consequently, we can use hierarchical inference as a framework to formulate a normative theory for the presence and structure of NCs: in this framework, correlations

are not detrimental, rather they are signatures of the integration of feedforward and feedback signals and can explain the connection between information limiting correlations, choice probabilities, and perceptual task execution [63,7^{••}]. Crucially, not only the inference of task variables is performed at higher levels of the computational hierarchy but the processing of stimuli also relies on hierarchically organized computational components [64–66] (see Figure 1b). Although an important insight that could be obtained from investigating the inference of task-related variables [24] was that uncertainty at higher levels of the hierarchy can explain the emergence and variations in NCs, feedforward deep neural networks do not attempt to represent uncertainty associated with the inferences at different levels of the hierarchy [49]. Further confirming the role of top-down modulations in perceptual processes is the demonstration of both task-related and perceptual top-down effects in the mean responses of V1 [25,67]. Taken together, similar to the dependence of NCs on task variables, low-level NCs are expected to depend on high-level perceptual inferences, for example about the presence of high-level visual structures. Indeed, such stimulus dependence was found in monkey V1 recordings [5,4[•]]. A more direct confirmation of the effect of high-level inferences on the NC structure at lower levels was obtained by testing the NC structure with natural and synthetic images [46^{••}]. Stimulus-specific NCs were demonstrated for natural stimuli but this specificity was reduced when high-level structure

was removed by synthesizing images which only preserved low-level structure. Such a dependence on stimulus statistics lends support for inference in a hierarchical model of perception since increased stimulus specificity of NCs can be interpreted as increased specificity in high-level inferences. Taken together, hierarchical inference provides a unifying framework for interpreting both task-related and perceptual recurrent interactions (Figure 1c). In hierarchical inference, task variables and perceptual variables jointly establish the NC structure. Such top-down effects implemented by recurrent interactions can predict specific patterns in NCs and have an added benefit that task and stimulus dependence of correlations can also be predicted.

Open questions

Surround suppression as modulation of joint statistics

A motif of cortical computation that is often attributed to recurrent processing is surround suppression. Experiments show that the amount of surround suppression in V1 is dependent on stimulus properties [68] and can be altered by intervention at the higher levels of visual hierarchy [69]. Modeling studies linked these effects to computational motives such as divisive normalization [70,23^{*}], or inference of higher-level latent variables [71,54]. While surround suppression is an effect that clearly alters the joint statistics of neural responses, its study remained practically disjoint from the line of work characterizing correlations. We propose that there are untapped opportunities for gaining a more complete understanding of cortical computations by producing predictions towards surround suppression from models aimed primarily at response correlations and vice versa.

Characterization of joint statistics on low-dimensional manifolds

While the size of neural populations that can be recorded simultaneously is increasing rapidly, we cannot expect the duration of the recordings, and thus the number of recorded trials from behaving animals to increase significantly in the near future. This limitation places severe constraints on the complexity of models that can be evaluated on data. Consequently, if our goal was to describe higher-order statistical structure in neural responses, we would need to restrict the number of parameters in our model. Dimensionality reduction methods [72] transform neural responses to a lower-dimensional manifold preserving as much variability as possible. Latent factors retrieved by dimensionality reduction have been linked to cortical state [73], attentional state [74,75] and task and stimulus variables as well [76,43^{*}], often directly characterizing joint neural response statistics. While such factors can be expected to correspond to latent variables in a perceptual model, such an explicit connection is rarely established.

Signatures of approximate inference in neural correlations

Full Bayesian inference is a computationally intractable problem that cannot be expected to be solved exactly in general in biological systems. Algorithms that approximate Bayesian inference generally fall into either the category of parametric variational or sampling-based approximations, with theories of cortical computations corresponding to both of them [77,58] and proposals attempting to unite the two approaches [78,79]. Approximation by sampling naturally requires the definition of a sampling algorithm that will make predictions about the temporal evolution of neural activations, including joint statistics, or population phenomena such as oscillations [80–82]. The implementation of such algorithms may be facilitated by the architecture of the cortex [16], as architectural constraints are indeed predictive towards joint neural response statistics in themselves [18]. The interpretation of perception as inference in a perceptual model suggests that deviations from the representation of exact posterior distributions are to be regarded as inefficiencies. However, discrepancies between the estimated posterior of a model and corresponding quantities estimated from neural activity and behavior will arise not only because an approximate posterior is being evaluated, but also because the generative model used to describe the subjects' internal model will inevitably make simplifying assumptions and thus only be an approximation to the actual internal model. Notably, most of the generative models investigated in the context of the visual system ignore temporal dynamics, which is certainly an important component of the natural visual environment, and is therefore expected to be represented by the brain.

Efficient implementation of hierarchical probabilistic models

One of the major obstacles one has to overcome when applying hierarchical models to a realistic problem is the combination of high dimensionality and compositional structure in natural stimuli. Most models successfully predicting some aspect of response statistics were either flat [83,23^{*}] or avoided directly fitting parameters to natural stimuli remaining in a low-dimensional regime [22^{*},46^{**}], with some notable exceptions [65]. In general, fitting a high-dimensional, hierarchical probabilistic model to data and performing full Bayesian inference in it remains a challenge. Advances in deep learning with regard to automated model fitting enable not only highly effective discriminative models, but generative models as well. A class of deep generative models with particular relevance to perceptual processing are variational auto-encoders (VAEs) [84,85], already with some applications in neuroscience [86] (see Figure 1d). Although VAEs with a hierarchical structure exist [87], robust fitting of such architectures to naturalistic data is not solved. While VAE-based hierarchical inference solutions come with the appeal of capitalizing on the efficiency of deep

learning infrastructure, alternative approaches, considering cortical architecture specifically, propose inference in hierarchical models based on message passing algorithms [88–90].

Conclusion

Correlations between neural responses are a controversial feature of population response statistics. The rich structure of these correlations and their systematic changes with changes both in stimulus and task pose serious challenges to theories of cortical computations. Hierarchical inference provides a framework under which a disparate set of observations can be unified. Under this framework, correlations are natural consequences of efficient computations rather than detrimental effects that would hamper the processing of stimuli. While the framework has produced successful models, its application to a wider range of phenomena opens up a number of exciting questions.

Conflict of interest statement

Nothing declared.

Acknowledgements

This work was supported by the National Brain Program (NAP-B 2017-1.2.1-NKP-2017-00002, G.O.) and a Human Frontiers Science Program Grant (G.O.).

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- of special interest
- of special interest

1. Zohary E, Shadlen MN, Newsome WT: **Correlated neuronal discharge rate and its implications for psychophysical performance.** *Nature* 1994, **370**:140-143.
2. Tkacik G, Prentice JS, Balasubramanian V, Schneidman E: **Optimal population coding by noisy spiking neurons.** *Proc Natl Acad Sci U S A* 2010, **107**:14419-14424.
3. Kohn A, Coen-Cagli R, Kanitscheider I, Pouget A: **Correlations and neuronal population information.** *Annu Rev Neurosci* 2016, **39**:237-256.
4. Rikhye RV, Sur M: **Spatial correlations in natural scenes**
 - **modulate response reliability in mouse visual cortex.** *J Neurosci* 2015, **35**:14661-14680.

The authors develop a method to control the statistical structure of natural images, yielding a graded reduction of high-level statistics, and study population response statistics in mouse V1. They demonstrate that the statistical structure of the stimulus modulates NCs together with mean responses and signal correlations and link noise correlations with the formation of ensembles.
5. Hofer SB, Ko H, Pichler B, Vogelstein J, Ros H, Zeng H, Lein E, Lesica NA, Mrsic-Flogel TD: **Differential connectivity and response dynamics of excitatory and inhibitory neurons in visual cortex.** *Nat Neurosci* 2011, **14**:1045-1052.
6. Lin I-C, Okun M, Carandini M, Harris KD: **The nature of shared cortical variability.** *Neuron* 2015, **87**:644-656.
7. Bondy AG, Haefner RM, Cumming BG: **Feedback determines the structure of correlated variability in primary visual cortex.** *Nat Neurosci* 2018, **21**:598-606.
8. Ruff DA, Cohen MR: **Attention can either increase or decrease spike count correlations in visual cortex.** *Nat Neurosci* 2014, **17**:E4085-E4094.
9. Cohen MR, Kohn A: **Measuring and interpreting neuronal correlations.** *Nat Neurosci* 2011, **14**:811-819.
10. Schulz DPA, Sahani M, Carandini M: **Five key factors determining pairwise correlations in visual cortex.** *J Neurophysiol* 2015, **114**:1022-1033.
11. Downer JD, Niwa M, Sutter ML: **Task engagement selectively modulates neural correlations in primary auditory cortex.** *J Neurosci* 2015, **35**:7565-7574.
12. Petersen RS, Panzeri S, Diamond ME: **Population coding of stimulus location in rat somatosensory cortex.** *Neuron* 2001, **32**:503-514.
13. Ohiorhenuan IE, Mechler F, Purpura KP, Schmid AM, Hu Q, Victor JD: **Sparse coding and high-order correlations in fine-scale cortical networks.** *Nature* 2010, **466**:617-621.
14. Poulet JFA, Petersen CCH: **Internal brain state regulates membrane potential synchrony in barrel cortex of behaving mice.** *Nature* 2008, **445**:881-885.
15. Franke F, Fiscella M, Sevelev M, Roska B, Hierlemann A, da Silveira RA: **Structures of neural correlation and how they favor coding.** *Neuron* 2016, **89**:409-422.
16. Hennequin G, Ahmadian Y, Rubin DB, Lengyel M, Miller KD: **The dynamical regime of sensory cortex: stable dynamics around a single stimulus-tuned attractor account for patterns of noise variability.** *Neuron* 2018, **98**:846-860.
17. Rosenbaum R, Smith MA, Kohn A, Rubin JE, Doiron B: **The spatial structure of correlated neuronal variability.** *Nat Neurosci* 2016, **19**:107-114.
18. Doiron B, Litwin-Kumar A, Rosenbaum R, Ocker GK, Josić K: **The mechanics of state-dependent neural correlations.** *Nat Neurosci* 2016, **19**:383-393.
19. Shamir M, Sompolinsky H: **Implications of neuronal diversity on population coding.** *Neural Comput* 2006, **18**:1951-1986.
20. 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.
21. Shamir M: **Emerging principles of population coding: in search for the neural code.** *Curr Opin Neurobiol* 2014, **25**:140-148.
22. Haefner RM, Berkes P, Fiser J: **Perceptual decision-making as probabilistic inference by neural sampling.** *Neuron* 2016, **90**:649-660.
23. Orbán G, Berkes P, Fiser J, Lengyel M: **Neural variability and sampling-based probabilistic representations in the visual cortex.** *Neuron* 2016, **92**:530-543.
24. Nienborg H, Roelfsema PR: **Belief states as a framework to explain extra-retinal influences in visual cortex.** *Curr Opin Neurobiol* 2015, **32**:45-52.

25. Gilbert CD, Li W: **Top-down influences on visual processing.** *Nat Rev Neurosci* 2013, **14**:350-363.
 26. Hubel DH, Wiesel TN: **Receptive fields and functional architecture of monkey striate cortex.** *J Physiol* 1968, **195**:215-243.
 27. Dubner RF, Zeki SM: **Response properties and receptive fields of cells in an anatomically defined region of the superior temporal sulcus in the monkey.** *Brain Res* 1971, **35**:528-532.
 28. Parker AJ, Newsome WT: **Sense and the single neuron: probing the physiology of perception.** *Annu Rev Neurosci* 1998, **21**:227-277.
 29. Dayan P, Abbott LF: *Theoretical Neuroscience*. MIT Press; 2005.
 30. Averbeck BB, Latham PE, Pouget A: **Neural correlations, population coding and computation.** *Nat Rev Neurosci* 2006, **7**:358-366.
 31. Seung HS, Sompolinsky H: **Simple models for reading neuronal population codes.** *Proc Natl Acad Sci U S A* 1993, **90**:10749-10753.
 32. Froudarakis E, Berens P, Ecker AS, Cotton RJ, Sinz FH, Yatsenko D, Saggau P, Bethge M, Tolias AS: **Population code in mouse V1 facilitates readout of natural scenes through increased sparseness.** *Nat Neurosci* 2014, **17**:851-857.
 33. Newsome WT, Britten KH, Movshon JA: **Neuronal correlates of a perceptual decision.** *Nature* 1989, **341**:52.
 34. Cumming BG, Nienborg H: **Feedforward and feedback sources of choice probability in neural population responses.** *Curr Opin Neurobiol* 2016, **37**:126-132.
 35. 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.
 36. 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.
 37. Britten KH, Shadlen MN, Newsome WT, Movshon JA: **The analysis of visual motion: a comparison of neuronal and psychophysical performance.** *J Neurosci* 1992, **12**:4745-4765.
 38. Pitkow X, Liu S, Angelaki DE, DeAngelis GC, Pouget A: **How can single sensory neurons predict behavior?** *Neuron* 2015, **87**:411-423.
- This paper links three important concepts: choice probabilities, neural threshold, and information limiting correlations. It proposes that choice probabilities are consequences of the presence of information limiting correlations.
39. Moreno-Bote R, Beck J, Kanitscheider I, Pitkow X, Latham P, Pouget A: **Information-limiting correlations.** *Nat Neurosci* 2014, **17**:1410-1417.
- The study aims to test the hypothesis that NCs limit the information carrying capacity of the population, and that this can be remedied by decorrelating. They show that given a specific task, a small component of NCs is responsible for defining the information limit, the one that causes variation in the population activity along the same direction that the task-relevant changes in the stimulus would.
40. Clery S, Cumming BG, Nienborg H: **Decision-related activity in macaque V2 for fine disparity discrimination is not compatible with optimal linear readout.** *J Neurosci* 2017, **37**:715-725.
 41. Ruff DA, Ni AM, Cohen MR: **Cognition as a window into neuronal population space.** *Annu Rev Neurosci* 2018, **41**:77-97.
 42. Kanitscheider I, Coen-Cagli R, Pouget A: **Origin of information-limiting noise correlations.** *Proc Natl Acad Sci U S A* 2015, **112**:E6973-E6982.
 43. Ni AM, Ruff DA, Alberts JJ, Symmonds J, Cohen MR: **Learning and attention reveal a general relationship between population activity and behavior.** *Science* 2018, **359**:463-465.
- The study demonstrates that learning and attention produces increased performance while exerting similar effects on neural response statistics. They argue that the discrepancy between task-defined and internally represented variables cause the changes in correlations to deviate from the optimum for the task.
44. Kohn A, Smith MA: **Stimulus dependence of neuronal correlation in primary visual cortex of the macaque.** *J Neurosci* 2005, **25**:3661-3673.
 45. Ecker AS, Berens P, Keliris GA, Bethge M, Logothetis NK, Tolias AS: **Decorrelated neuronal firing in cortical microcircuits.** *Science* 2010, **327**:584-587.
 46. Bányai M, Lazar A, Klein L, Klon-Lipok J, Stippinger M, Singer W, Orbán G: **Stimulus complexity shapes response correlations in primary visual cortex.** *Proc Natl Acad Sci U S A* 2019, **116**:2723-2732.
- The study is designed to test the prediction of the hierarchical inference hypothesis that higher-level inferences modulate lower-level population activity statistics. They demonstrate that stimuli with higher-order statistical structure elicit more stimulus-specific NCs than stimuli only containing low-level structure, which is compatible with top-down effects due to the inference of perceptual context.
47. Bányai M, Koman Z, Orbán G: **Population activity statistics dissect subthreshold and spiking variability in V1.** *J Neurophysiol* 2017, **118**:29-46.
 48. Berkes P, Orbán G, Lengyel M, Fiser J: **Spontaneous cortical activity reveals hallmarks of an optimal internal model of the environment.** *Science* 2011, **331**:83-87.
 49. Kriegeskorte N: **Deep Neural Networks: a new framework for modeling biological vision and brain information processing.** *Annu Rev Vis Sci* 2015, **1**:415-446.
 50. Russakovsky O, Deng J, Su H, Krause J, Satheesh S, Ma S, Huang Z, Karpathy A, Khosla A, Bernstein M et al.: **Imagenet large scale visual recognition challenge.** *Int J Comput Vis* 2015, **115**:211-252.
 51. Yamins DLK, Hong H, Cadieu CF, Solomon EA, Seibert D, DiCarlo JJ: **Performance-optimized hierarchical models predict neural responses in higher visual cortex.** *Proc Natl Acad Sci U S A* 2014, **111**:8619-8624.
 52. Güçlü U, van Gerven MAJ: **Deep neural networks reveal a gradient in the complexity of neural representations across the ventral stream.** *J Neurosci* 2015, **35**:10005-10014.
 53. Yamins DLK, DiCarlo JJ: **Eight open questions in the computational modeling of higher sensory cortex.** *Curr Opin Neurobiol* 2016, **37**:114-120.
 54. Berardino A, Laparra V, Ballé J, Simoncelli E: **Eigen-distortions of hierarchical representations.** In *Advances in Neural Information Processing Systems 30*. Edited by Guyon I, Luxburg UV, Bengio S, Wallach H, Fergus R, Vishwanathan S, Garnett R. Curran Associates, Inc.; 2017:3530-3539.
 55. Kar K, Kubilius J, Schmidt K, Issa EB, DiCarlo JJ: **Evidence that recurrent circuits are critical to the ventral stream's execution of core object recognition behavior.** *Nat Neurosci* 2019, **22**:974-983.
- The authors demonstrate that while deep feedforward networks excel in the classification of natural stimuli, there exists a subset of complex images at which performance decreases. Crucially, this set of images is well matching the subset where deep feedforward networks are less efficient in predicting visual cortical activity. The authors argue that this discrepancy is a signature of feedback processing.
56. Rajalingham R, Issa EB, Bashivan P, Kar K, Schmidt K, DiCarlo JJ: **Large-scale, high-resolution comparison of the core visual object recognition behavior of humans, monkeys, and state-of-the-art deep artificial neural networks.** *J Neurosci* 2018, **38**:7255-7269.
 57. Lake BM, Salakhutdinov RR, Tenenbaum JB: **Human-level concept learning through probabilistic program induction.** *Science* 2015, **350**:1332-1338.
 58. Fiser J, Berkes P, Orbán G, Lengyel M: **Statistically optimal perception and learning: from behavior to neural representations.** *Trends Cogn Sci* 2010, **14**:119-130.
 59. Goris RLT, Ziemba CM, Stine GM, Simoncelli EP, Movshon JA: **Dissociation of choice formation and choice-correlated activity in macaque visual cortex.** *J Neurosci* 2017, **37**:5195-5203.

60. Ecker AS, Denfield GH, Bethge M, Tolias AS: **On the structure of neuronal population activity under fluctuations in attentional state.** *J Neurosci* 2016, **36**:1775-1789.
61. Goris RLT, Movshon JA, Simoncelli EP: **Partitioning neuronal variability.** *Nat Neurosci* 2014, **17**:858-865.
62. Rabinowitz NC, Goris RL, Cohen M, Simoncelli EP: **Attention stabilizes the shared gain of V4 populations.** *eLife* 2015, **4**: e08998.
63. Lange RD, Haefner RM: **Characterizing and interpreting the influence of internal variables on sensory activity.** *Curr Opin Neurobiol* 2017, **46**:84-89.
64. Lee TS, Mumford D: **Hierarchical Bayesian inference in the visual cortex.** *J Opt Soc Am A: Opt Image Sci Vis* 2003, **20**:1434-4448.
65. Karklin Y, Lewicki MS: **Emergence of complex cell properties by learning to generalize in natural scenes.** *Nature* 2009, **457**:83-86.
66. Hosoya H, Hyvarinen A: **A hierarchical statistical model of natural images explains tuning properties in V2.** *J Neurosci* 2015, **35**:10412-10428.
67. Lee TS, Nguyen M: **Dynamics of subjective contour formation in the early visual cortex.** *Proc Natl Acad Sci U S A* 2001, **98**:1907-1911.
68. Haider B, Krause MR, Duque A, Yu Y, Touryan J, Mazer JA, McCormick DA: **Synaptic and network mechanisms of sparse and reliable visual cortical activity during nonclassical receptive field stimulation.** *Neuron* 2010, **65**:107-121.
69. Nassi JJ, Lomber SG, Born RT: **Corticocortical feedback contributes to surround suppression in V1 of the alert primate.** *J Neurosci* 2013, **33**:8504-8517.
70. Schwartz O, Simoncelli EP: **Natural signal statistics and sensory gain control.** *Nat Neurosci* 2001, **4**:819-825.
71. Coen-Cagli R, Kohn A, Schwartz O: **Flexible gating of contextual influences in natural vision.** *Nat Neurosci* 2015, **18**:1648-1655.
72. Cunningham JP, Yu BM: **Dimensionality reduction for large-scale neural recordings.** *Nat Neurosci* 2014, **17**:1500-1509.
73. Ecker AS, Berens P, Cotton RJ, Subramanian M, Denfield GH, Cadwell CR, Smirnakis SM, Bethge M, Tolias AS: **State dependence of noise correlations in macaque primary visual cortex.** *Neuron* 2014, **82**:235-248.
74. Cohen MR, Maunsell JHR: **Using neuronal populations to study the mechanisms underlying spatial and feature attention.** *Neuron* 2011, **70**:1192-1204.
75. Denfield GH, Ecker AS, Shinn TJ, Bethge M, Tolias AS: **Attentional fluctuations induce shared variability in macaque primary visual cortex.** *Nat Commun* 2018, **9**:2654.
76. Williams AH, Kim TH, Wang F, Vyas S, Ryu SI, Shenoy KV, Schnitzer M, Kolda TG, Ganguli S: **Unsupervised discovery of demixed, low-dimensional neural dynamics across multiple timescales through tensor component analysis.** *Neuron* 2018, **98**:1099-1115.
77. Ma WJ, Beck JM, Latham PE, Pouget A: **Bayesian inference with probabilistic population codes.** *Nat Neurosci* 2006, **9**:1432-1438.
78. Savin C, Denève S: **Spatio-temporal representations of uncertainty in spiking neural networks.** In *Advances in Neural Information Processing Systems 27*. Edited by Ghahramani Z, Welling M, Cortes C, Lawrence ND, Weinberger KQ. Curran Associates, Inc.; 2014:2024-2032.
79. Shivkumar S, Lange R, Chatteraj A, Haefner R: **A probabilistic population code based on neural samples.** In *Advances in Neural Information Processing Systems 31*. Edited by Bengio S, Wallach H, Larochelle H, Grauman K, Cesa-Bianchi N, Garnett R. Curran Associates, Inc.; 2018:7070-7079.
80. Aitchison L, Lengyel M: **The Hamiltonian brain: efficient probabilistic inference with excitatory-inhibitory neural circuit dynamics.** *PLoS Comput Biol* 2016, **12**:e1005186.
81. Hennequin G, Aitchison L, Lengyel M: **Fast sampling-based inference in balanced neuronal networks.** In *Advances in Neural Information Processing Systems 27*. Edited by Ghahramani Z, Welling M, Cortes C, Lawrence ND, Weinberger KQ. Curran Associates, Inc.; 2014:2240-2248.
82. Echeveste R, Aitchison L, Hennequin G, Lengyel M: **Cortical-like dynamics in recurrent circuits optimized for sampling-based probabilistic inference.** *bioRxiv* 2019 <http://dx.doi.org/10.1101/081661>.
83. Olshausen BA, Field DJ: **Emergence of simple-cell receptive field properties by learning a sparse code for natural images.** *Nature* 1996, **381**:607-609.
84. Dayan P, Hinton GE, Neal RM, Zemel RSS: **The Helmholtz machine.** *Neural Comput* 1995, **7**:889-904.
85. Kingma DP, Welling M: **Auto-encoding Variational Bayes.** *Arxiv.org*; 2013.
86. Barello G, Charles A, Pillow J: **Sparse-coding variational auto-encoders.** *bioRxiv* 2018 <http://dx.doi.org/10.1101/399246>.
87. Sønderby CK, Raiko T, Maaløe L, Sønderby SK, Winther O: **Ladder variational autoencoders.** In *Advances in Neural Information Processing Systems 29*. Edited by Lee DD, Sugiyama M, Luxburg UV, Guyon I, Garnett R. Curran Associates, Inc.; 2016:3738-3746.
88. George D, Hawkins J: **Towards a mathematical theory of cortical micro-circuits.** *PLoS Comput Biol* 2009, **5**:e1000532 <http://dx.doi.org/10.1371/journal.pcbi.1000532>.
89. Bastos AM, Usrey WM, Adams RA, Mangun GR, Fries P, Friston KJ: **Canonical microcircuits for predictive coding.** *Neuron* 2012, **76**:695-711.
90. Pitkow X, Angelaki DE: **Inference in the brain: statistics flowing in redundant population codes.** *Neuron* 2017, **94**:943-953.