Neural signatures of attention: insights from decoding population activity patterns

Panagiotis Sapountzis[1], Georgia G. Gregoriou[1,2]

[1]*Foundation for Research and Technology Hellas, Institute of Applied and Computational Mathematics, N. Plastira 100, GR70013 Heraklion, Crete Greece,* [2]*University of Crete, Faculty of Medicine, P.O. Box 2208, GR71003, Heraklion, Crete, Greece*

TABLE OF CONTENTS

1. ABSTRACT

Understanding brain function and the computations that individual neurons and neuronal ensembles carry out during cognitive functions is one of the biggest challenges in neuroscientific research. To this end, invasive electrophysiological studies have provided important insights by recording the activity of single neurons in behaving animals. To average out noise, responses are typically averaged across repetitions and across neurons that are usually recorded on different days. However, the brain makes decisions on short time scales based on limited exposure to sensory stimulation by interpreting responses of populations of neurons on a moment to moment basis. Recent studies have employed machine-learning algorithms in attention and other cognitive tasks to decode the information content of distributed activity patterns across neuronal ensembles on a single trial basis. Here, we review results from studies that have used pattern-classification decoding approaches to explore the population representation of cognitive functions. These studies have offered significant insights into population coding mechanisms. Moreover, we discuss how such advances can aid the development of cognitive brain-computer interfaces.

2. INTRODUCTION

Our ability to perceive the world around us is limited by our brain's processing capacity. As a result, we only perceive, remember and respond to a fraction of the visual input that reaches our retinas. Typically, this is the part of the world that we attend to. Attention

serves to prioritize stimuli processing according to their physical salience or their relevance to current behavioral goals. In that sense, attention acts as a selection mechanism to facilitate processing of a subset of sensory information while irrelevant stimuli are filtered out. The ability to flexibly guide our attention to the most relevant stimuli according to the task at hand is a critical part of our cognition and necessary for normal behavior. Indeed, attention is known to be disrupted in several neuropsychiatric disorders, including attention deficit hyperactivity disorder, schizophrenia and Alzheimer's disease. Thus, a thorough understanding of the neural mechanisms of attention is important for two reasons. First, it can provide important insights into the way information is selectively processed in the brain and into the mechanisms that underlie dynamic modulations in functional connectivity in line with current behavioral goals. Second, a detailed description of the neural circuits and computations carried out during attentive behavior is necessary in order to develop more effective treatments and interfaces that can aid people with attention and vision-related disorders.

The deployment of attention to particular locations in space is characterized as spatial attention. Several psychophysical studies have shown that covert shifts of attention (i.e. without movement of the eyes or head) result in improved stimulus detection and faster reaction times toward stimuli at the locus of attention over stimuli located elsewhere (1). However, attention can also be directed to particular features or objects. A typical example is looking for a familiar face in the crowd or for our car in a parking lot. In this case, we rely on the known characteristics of the object we search to guide the search more efficiently. As a result, all stimuli in our visual field that share the searched feature are more likely to attract our attention. For example, if we know that our friend wears a red jacket, all red items in the scene become more likely targets in the search process. This is called feature attention. Psychophysical studies have confirmed an interplay between spatial- and feature-based attention in visual search tasks (2–4).

In this review, we focus on current evidence that attention (mainly spatial attention) can be decoded by measuring population activity in the brain. We will review studies that have decoded population activity to predict behavior in attention and other cognitive tasks. Given that the non-human primate has proved to be an irreplaceable model to understand the complexity of the human brain and particularly cognitive functions, we will mainly focus on studies that have employed behaving macaques and electrophysiological methods, which allow a high enough spatial and temporal resolution to assess decoding principles in the brain. We first provide a brief description of the brain areas known to be involved in attention mechanisms

(section 3). We then review activity measures that are known to be modulated by attention including firing rate responses, neural synchrony, response variability and inter-neuronal correlations (section 4). In section 5, we refer to the advantages of population analysis methods and in section 6, we briefly describe how pattern classification algorithms can be used to decode neural activity. Studies that have used machine-learning algorithms to decode population activity patterns in attention, visual processing and other functions are reviewed in section 7. In section 8, we compare the efficacy of different neural signals in decoding approaches. Finally, in section 9, we discuss the implications of decoding approaches for the development of brain-machine interfaces (BMIs), which aim to restore function in human patients.

3. THE ATTENTION NETWORK

At the neuronal level, significant insights into the neural mechanisms of attention have been obtained from imaging and electrophysiological studies in humans and non-human primates. These studies have implicated a distributed network of brain areas in attention including visual areas in the occipital and temporal lobe as well as higher order brain areas in the prefrontal and parietal lobe (Figure 1) (for reviews see 5, 6, 7). Targeted invasive electrophysiological approaches that employ single-unit and multi-unit extracellular recordings in non-human primates have provided most of our knowledge on the distinct role of different brain areas in attentional mechanisms. These studies have shown that both in the dorsal and ventral visual stream, activity is modulated by spatial and feature-based attention (8–15). Although attentional modulation of neuronal responses has been reported almost in all brain areas with visually responsive neurons (including thalamic nuclei, the superior colliculus, areas V1, V2 etc.), most neurophysiological studies have focused on the effect of attention on neuronal responses in mid- and higher-level visual areas such as area V4, the inferior temporal cortex (IT) and the middle temporal area (area MT) (Figure 1).

Higher order areas in the prefrontal and parietal cortex have been suggested to be the source of these activity modulations in visual areas. Specifically, the prevailing view holds that parietal and prefrontal cortical areas provide top-down biasing signals that modulate processing in posterior visual areas by selectively enhancing activity of neuronal populations that encode the attended feature, object or location at the expense of irrelevant distracters (6, 16). Accordingly, prefrontal and parietal areas are suggested to be critical for the control of visual attention. In line with this view, lesions in these areas produce attentional deficits in both monkeys and humans (17–21). Electrophysiological findings from monkeys have identified two areas within the parietofrontal network
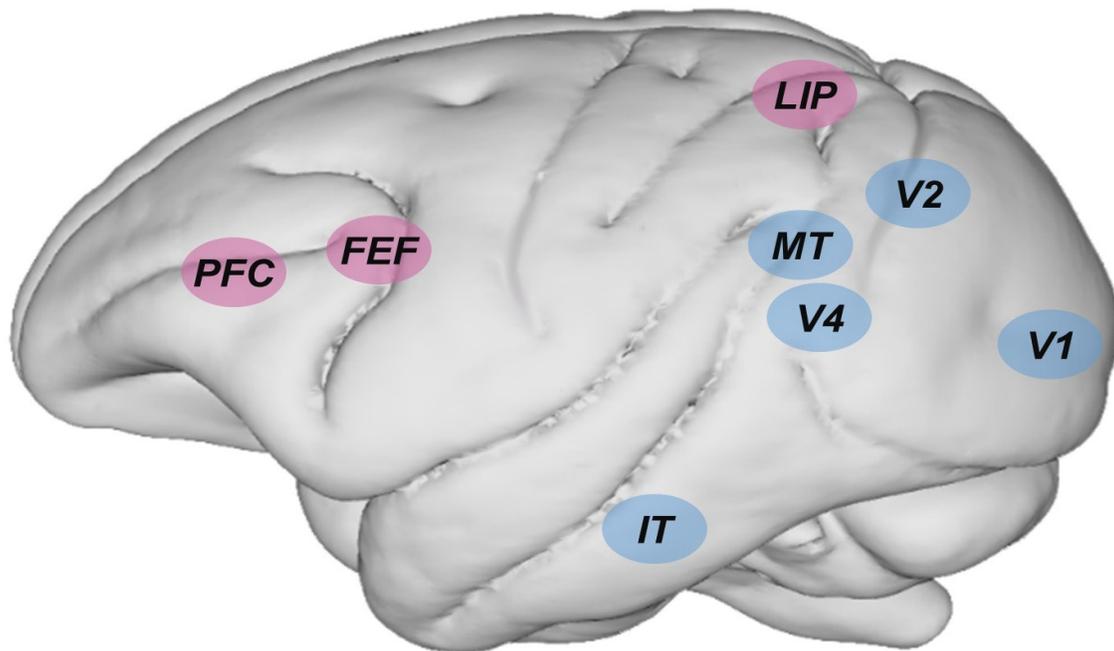
Figure 1. Cortical areas of the attention network. Lateral view of a left hemisphere of a macaque brain. The approximate locations of visual areas V1, V2, V4, MT, and IT are indicated on the brain in blue ellipses. Areas that are considered as attention control areas in the parietal lobe (LIP) and in the prefrontal cortex (PFC) are indicated in pink ellipses. Note that FEF is part of PFC. The brain surface was obtained from the atlas template in (183), through the Scalable Brain Atlas (184).

as potential sources of attention-related signals, the lateral intraparietal area (LIP) in the parietal lobe and the frontal eye fields (FEF) in the prefrontal cortex (PFC) (Figure 1). Both LIP and FEF are thought to integrate information about the physical properties of the stimuli, together with information about expectations and current behavioural goals in order to construct a "saliency" map in which stimuli or spatial locations are represented by a level of activity that reflects their attentional priority (22, 23). It has been hypothesized that this map serves to provide a "top-down" biasing signal, which modulates sensory processing in earlier visual areas so that objects or locations of interest are optimally analyzed and distracting objects are essentially filtered out. Electrophysiological as well as deactivation studies in macaques have corroborated that FEF and LIP are both necessary and sufficient to guide attention. Microstimulation in FEF and LIP (i.e. injection of low amplitude currents into small populations of neurons) biases orienting to a stimulus and improves attentional performance (24, 25). On the other hand, deactivation of the same areas impairs attentional performance (21, 26, 27). Findings from electrophysiological recordings during a covert attention task suggested that whereas LIP has a dominant role in guiding bottom-up, stimulus-driven, exogenous attention, FEF and the nearby PFC have a more prominent role in top-down, goal-directed, endogenous attention establishing a possible distinction between the two areas in attentional functions (28). More recently, an additional distinction between the classical FEF in the anterior bank of the arcuate sulcus

and the cortex anterior to it in the prefrontal lobe was suggested, with FEF being the source of spatial attention signals and the cortex anterior to it serving as the source of feature attention signals (29).

Although the areas and circuits involved in attention have been adequately described, the mechanisms that lead to an improvement in performance with attention are less well understood. It is widely accepted that attentional mechanisms lead to an improvement in the signal to noise ratio. In the next session, we briefly review measures of neural activity that are known to be modulated by attention and can lead to an enhancement of the signal to noise ratio.

4. NEURAL SIGNATURES OF ATTENTION

4.1. Firing rate

In electrophysiological studies, the effect of attention on visual processing is typically measured as an enhancement in the visual response or an increased sensitivity of single neurons to locations or objects of interest at the expense of distracting stimuli (8–15, 30). Specifically, when more than one stimuli exist inside the visual receptive field (RF) of a particular neuron the attended stimulus dominates the neuronal response consistent with the notion that attention serves to filter out irrelevant stimuli. This was initially described as a mechanism that biases the competition among representations of different stimuli in favor of the

attended one (6) and was later formalized as a contrast normalization function (31, 32). This modulation in firing rate with attention is multiplicative and the scaling of the neuronal response depends on the similarity between the preference of the neuron and the attended feature as predicted by the feature similarity gain model (14, 33). It should be noted, that areas higher in the visual hierarchy show more pronounced activity modulation with attention (34). However, modulations of firing rate suggestive of a facilitation of the sensory representation of the attended stimulus have been reported in multiple brain areas at all levels of visual processing including the lateral geniculate nucleus (35), the superior colliculus (36–38), area V1 (10, 39–41), areas V2 and V4 (9–11, 13, 30, 42), the middle temporal (MT) and medial superior temporal (MST) areas (14, 15), IT (8, 29), LIP (28, 43, 44), FEF (45–47) and dorsolateral PFC (29).

4.2. Neural Synchrony

An increase in the firing rate of selected neuronal populations is not the only means through which attention can enhance signal efficacy. More recent studies have shown that attention can also modulate neural synchrony (47–49, 50, for a review see, 51). Given that postsynaptic potentials have limited duration, inputs need to arrive close in time to be summed effectively and lead to the generation of an action potential in the postsynaptic cell. Moreover, rhythmic inputs that arrive at temporal windows during which the postsynaptic cell is more likely to be depolarized are more effective in transmitting information to the next stage. These two conditions that can potentially lead to more effective communication across selected neuronal populations can be implemented through oscillatory synchrony. Thus, synchronization or an enhancement of synchronization of activity in populations of neurons that encode the attended stimulus could render processing of the attended stimulus more effective over the representation of unattended stimuli. Oscillatory synchrony is typically measured using a signal that integrates activity within a small volume of cortex, the local field potential (LFP), by looking at phase locking of action potentials (spikes) to the LFP oscillations.

Accumulating evidence has suggested that indeed, attention leads to an enhancement of local gamma frequency (30–60 Hz) synchronization among neurons that encode the attended location or feature in both humans (52–55) and non-human primates in extrastriate visual areas (34, 42, 48, 50), frontal (e.g. FEF (47)) and parietal (e.g. LIP (49)) areas. This selective enhancement of synchronization in the gamma frequency range with a near zero lag phase locking could ensure synchronous firing of spikes by neurons that encode behaviorally relevant information

and could thus lead to a preferential processing of attended features or locations in downstream areas. Indeed, two studies have provided direct evidence that selective routing of information related to attended stimuli is at least partially reflected in increased gamma band phase locking between neuronal groups in areas V1 and V4 that represent the attended stimulus (56, 57). These together with other studies (47, 54) have also suggested that gamma synchrony with a non-zero phase lag across distant areas could ensure that inputs from one area arrive when the relevant population in the receiving area is at the maximally depolarized state and is, therefore, more excitable. Such a scheme would facilitate selective processing of the relevant inputs across areas (58).

Besides modulations in gamma synchrony, attention can also affect synchrony in lower frequencies. Several studies have shown that local low frequency synchronization in the alpha (8–14 Hz) and beta range (15–25 Hz) is reduced with spatial attention in visual cortices (34, 47, 48, 54, 59–64). Alpha band oscillatory activity, in particular, is enhanced for distracting stimuli and reduced for attended stimuli, in line with the view that alpha band activity is associated with inhibition of processing of external sensory inputs (65–68). Interestingly, studies that have examined the amplitude of oscillatory activity across different cortical layers and its modulation during attention have demonstrated layer-specific patterns for different rhythms, at least in visual cortical areas (34, 69–71).

4.3. Response variability

The response of individual neurons to repeated presentations of the same stimulus varies from trial to trial (72, 73). Spiking variability is typically measured using the Fano factor, defined as the spike count variance across trials normalized by the average spike count. Several factors can affect response variability. For example, the onset of a stimulus inside the neuronal RF leads to a decrease in the variability of neuronal response, an observation that has been reported across many cortical areas (74–76).

Attention related effects on response variability have been less clear. Although, an early V4 study reported a modest, insignificant decrease in the Fano factor with attention (77), a more recent study found a robust attention induced decrease, which was stronger for putative inhibitory interneurons than for putative excitatory pyramidal cells (78). In the FEF, two studies reported no significant modulation with attention (75, 76), whereas more recently a decrease in the Fano factor was reported only for putative pyramidal cells (79). Cell-type specific effects have also been observed in the dorsolateral PFC, where task engagement differentially affects the response variability of putative pyramidal cells (80).

A confound in the use of the Fano factor as an estimate of response variability is its dependence on the mean firing rate. In particular, as the mean firing rate increases, the Fano factor decreases. Recently, a model was introduced, which decomposes variability into a sum of a Poisson variance component and a component arising from slow fluctuations in neuronal excitability (i.e. gain) (81). Following this approach, it was reported that attention induces a reduction in gain variance in both putative pyramidal cells and interneurons in the FEF, with the reduction being more pronounced for putative interneurons (79).

In summary, most studies have found modest attention effects on response variability of single neurons and this effect appears to be cell-type specific. Approaches insensitive to Fano factor biases may provide complementary information. Taking aside these restrictions, the effect of attention on response variability, although small on single neurons, may be maximized when calculated across the population.

4.4. Inter-neuronal correlations

Whether reductions in response variability lead to an increase in the signal to noise ratio, depends on the degree to which the sources of variability are correlated across a population of neurons. If the sources of variability are uncorrelated (e.g. they arise from variations in synaptic transmission and spike generation for each neuron) then, in principle, variability can be averaged out by pooling together a large number of neurons. However, correlated variability (e.g. arising from variability that is shared across a population of neurons due to shared inputs) cannot be averaged out simply by pooling together the responses from many neurons (82).

In the cortex, response variability is typically correlated between neurons. These correlations, usually referred to as noise correlations, are quantified for pairs of neurons as the Pearson correlation of the spike count responses across different trials. Correlation values are typically small and positive (in the order of 0.0.1.-0.2.5. (83)) and tend to become higher for pairs of neurons with overlapping receptive fields (84) and similar tuning properties (85, 86).

Recent studies have shown that correlations are modulated by a variety of sensory, motor and cognitive factors (see 83 for a review). With attention, noise correlations are reduced in V1 (87, 88), (but see 84), V4 (59, 89, 90), MT (88, 91), MST (91) and FEF (92). Interestingly, although small in absolute values, the reduction in noise correlations with attention appears to account for the majority of improvement in signal quality, whereas, attention related increases in firing rate account only for a small proportion (89, 90).

When calculated as a function of window size, noise correlations tend to increase with larger time windows indicating that they reflect predominantly low frequency fluctuations (90). It is, thus, possible that correlations arise from the same mechanisms that cause the low frequency spike-LFP (47, 48) and spike-spike (90) desynchronization with attention. Indeed, according to a recent model, shared excitability signals (i.e. gain) that fluctuate in strength, reduce their fluctuations as a result of attention (93). The reduction in these shared, low-frequency fluctuations accounts for the attention-driven decreases in noise correlations. As a further model prediction, the variance of pooled population activity, as captured by the LFPs, decreases with attention resulting in reduced LFP power in low frequencies, in line with experimental findings (48, see also paragraph 4.2.). Finally, the reduced shared modulation (noise) between neurons leads to low frequency desynchronization, as it is empirically observed with attention (90).

5. ADVANTAGES OF POPULATION ANALYSIS METHODS

To overcome the apparent noisiness in the responses of individual neurons, studies typically average single neuron responses across repetitions (trials) of similar behavioral context. In many cases, responses are subsequently averaged across neurons to obtain population estimates. Such averaging approaches allow for robust estimates of activity and despite their limitations have proven extremely useful in understanding the functional role of single neurons and brain areas. However, with the growing use of multi-electrode arrays that allow simultaneous recordings from a large number of neurons, datasets have (and will) become more diverse, and simple averaging across a heterogeneous population becomes inappropriate. Population analysis methods, such as those discussed in this review, can successfully address the challenges imposed by large and heterogeneous datasets by taking the diversity of neuronal responses into account.

Conventional analyses average responses across trials, yet, we reach decisions on a trial-by-trial basis. To achieve this our brain takes into account responses of populations of neurons, often in short time scales. Although this has been acknowledged for decades, the mechanisms underlying population representations have only recently started to be explored (94). As will be discussed in the following sections, analysis methods that consider the pattern of activity across neuronal ensembles provide insight into the spatial and temporal characteristics of the population code. This is critical in order to understand how representations are constructed in the brain on a moment to moment basis and to address the full complexity of brain function in realistic ways. In

addition, such methods can often predict the behavioral outcome, providing a link between neuronal population activity and behavior.

Population decoding approaches are more powerful in the sense that they are multidimensional. Thus, they can recover information that may not be prominent in single dimensions. For example, population analysis methods have revealed that in higher-order areas the same neuronal population carries information about different task parameters that can be extracted according to behavioral demands (paragraph 7.2.). Trial-averaging approaches may obscure such information especially if it is represented more sparsely in the population. Furthermore, decoding analyses can reveal transient representations at the population level that would be otherwise difficult to observe.

Cognitive processes, such as attention, are accompanied by several physiological changes in the brain (see section 4). These changes may convey information that is relevant to behavior or could simply provide mechanistic support to neuronal processes. Consider the example of a pacemaker cell that generates spikes at a constant frequency in order to synchronize activity in a network. The spiking of this cell does not carry any behaviorally relevant information, however, its activity is necessary for the computation itself. Population decoding analyses, such as those discussed in the following section, can help us understand which physiological modulations convey task-relevant information, and/or whether some of these changes serve computational purposes.

Finally, population decoding techniques have been successfully used in brain-machine interfaces (BMIs) that have been developed to aid patients with motor disabilities. Of particular interest are applications that interpret neuronal signals from higher-order areas to drive cognitive BMIs. We will briefly discuss these applications in section 9.

6. DECODING OF NEURAL ACTIVITY USING PATTERN-CLASSIFICATION ALGORITHMS

Information coding in the brain can be examined using two complementary approaches. The encoding perspective examines how sensory information and behavioral parameters are represented in the brain, whereas, the decoding perspective reads-out the neuronal signals in order to reconstruct the stimulus presented or the internal state of the subject. In Bayesian terms, an encoding model would aim to estimate $p(r|s)$, the conditional probability of obtaining a neuronal response r following presentation of stimulus s, whereas, a decoding model would estimate $p(s|r)$, the probability of stimulus s being present given a response r. The two approaches are tightly coupled

and both contribute to an understanding of coding principles in the brain.

A variety of different algorithms and approaches has been employed to decode neuronal activity. Multivariate pattern-classification methods have been widely used and have offered significant insights into the way large scale representations are formed in the brain. Classification methods work by defining a decision boundary in order to assign inputs to one of a given set of classes (Figure 2). Initially, data are divided into a training and a test set. During training, the classifier learns to discriminate between responses recorded under the different conditions (supervised learning) and defines the decision boundary. The test set is then used to estimate the performance of the classifier, which is typically calculated as the percentage of correct predictions. To avoid overfitting of the data, the generalization of the decision boundary needs to be evaluated on data never experienced by the classifier, thus, the training and test sets have to be drawn from different trials. Overfitting arises when the decision boundary fits idiosyncratic, noise patterns of the training set resulting in poor generalization during testing (Figure 2 C). Cross-validation methods become useful in datasets with a limited number of trials that do not allow the partitioning of data into separate training and test sets (95). Cross-validation methods initially divide data into k groups or folds. Data from k-1 of these groups are assigned to a training set and the remaining to a test set. To reduce variability, the classification performance is evaluated as the mean across such k leave-one-out permutations.

The shape of the decision boundary can be either linear or non-linear and its dimensionality is determined by the number of features (i.e. neurons or signals) included in the analysis. In the linear case, the decision boundary in a two-dimensional feature space would be a line (Figure 2 A). Accordingly, using n features, the decision boundary would be a hyperplane defined in an n-dimensional space (Figure 2 D). Non-linear classifiers create complex, non-planar boundaries that, depending on the data structure, may provide better classification (Figure 2 B), however, they are more prone to overfitting (95) (Figure 2 C). Linear decoders are faster as they require the estimation of fewer parameters, which is advantageous in real time BMI applications. Moreover, they are more appealing given that their estimates can be used to construct plausible biological models (e.g. 96). Specifically, the estimates of a linear decoder are based on a weighted sum of the responses of several neurons. These weights could represent the synaptic weight of each neuron in its connection with a downstream readout neuron. Discussing the type of decoder that readout neurons may use to extract information is beyond the scope of this review. Instead, we consider decoding as an analysis tool that allows us to estimate how
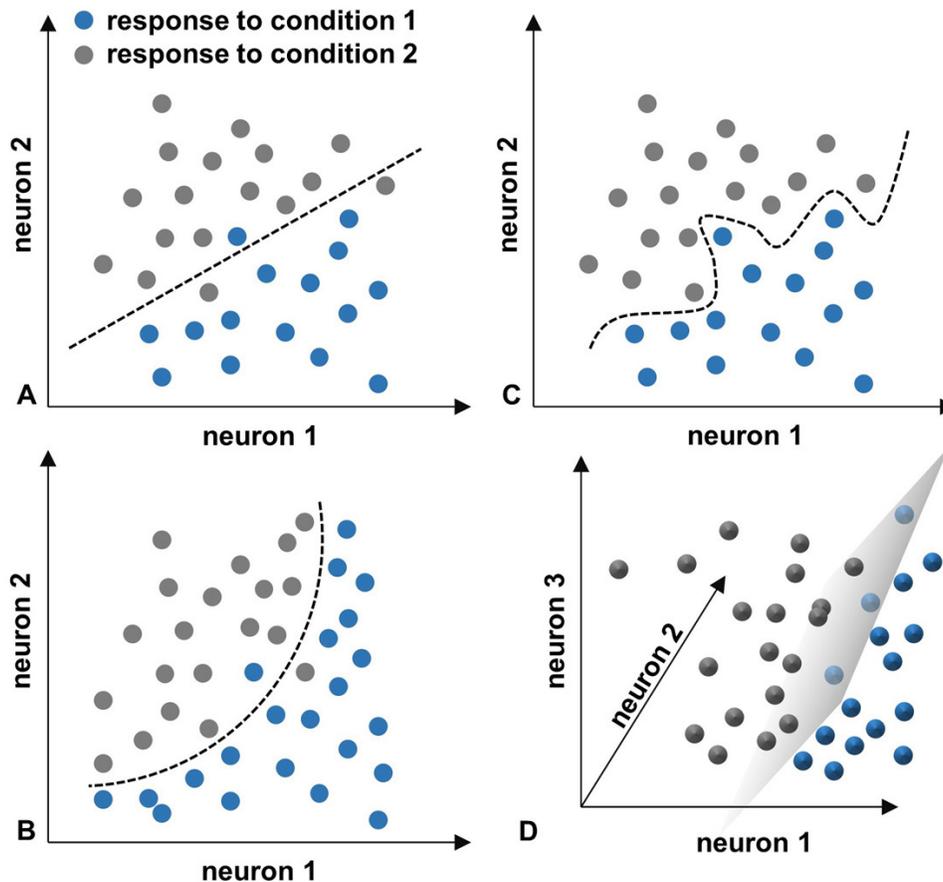
Figure 2. Pattern-classification approaches. A neural response from a single trial can be interpreted as a point in an N-dimensional feature space. Different colors (blue and grey) correspond to responses recorded under two experimental conditions. A. In this hypothetical 2-dimensional feature space a linear decoder is used to classify responses into two-categories. Giving the overlapping response distributions, a univariate, trial-averaging approach cannot separate responses between the two conditions. However, responses can be easily classified when considering the single-trial pattern of activity of two neurons. B. An example where a non-linear decision boundary produces more accurate results than a linear boundary. C. Non-linear classifiers are more prone to overfitting. In this case, the decision boundary fits idiosyncratic, noise patterns and will result in poor generalization. D. In a 3-dimensional feature space the decision boundary is a plane. Accordingly, using *n* features, the decision boundary would be a hyperplane defined in an *n*-dimensional space.

much information can be extracted from a population of neurons, the information carried by different types of signals and examine factors that may limit information.

In order to produce accurate estimates, pattern-classification algorithms, typically, require more samples (trials) than features (neurons or signals). Intuitively, as the dimensionality of the feature space increases, more samples are required in order to increase predictive power. This limitation is known as the "curse of dimensionality" (95). Some classification algorithms are more sensitive to this problem than others. For example, linear discriminant (LD) classifiers require, in practice, more samples than features, whereas, kernel methods such as support vector machines (SVM) are more resilient to this problem (see also 97). When the number of available trials is inadequate, dimensionality reduction methods such as principal component (PCA) or independent component analysis (ICA) can be used

to tackle this issue. In summary, due to the limitation of dimensionality, researchers should be cautious when comparing performance between classifiers trained under different conditions. In order to avoid an underestimation of accuracy, the dimensionality (i.e. number of neurons or signals) and the number of trials need to be equal in the classifiers under comparison.

In addition to classification algorithms, information theoretic approaches, in the Shannon or Fisher sense, have also been used to estimate the information content of neural signals (98). These approaches provide complementary information and both are used to quantify sensory or behavioral information carried in neuronal responses. Information theoretic approaches produce in some cases more accurate estimates than classification approaches (99); however, they are more sensitive to the curse of dimensionality and introduce biases when a limited number of samples is available (100). Thus, pattern-classification algorithms

become more appropriate for the analysis of large neuronal populations.

7. INSIGHTS INTO POPULATION CODING OF COGNITIVE FUNCTIONS USING DECODING APPROACHES

In the following paragraphs, we discuss how cognitive information can be decoded from brain signals such as extracellularly recorded action potentials (spikes) and LFPs. First, we briefly review the current evidence that attention-related information can be decoded from neural signals using machine-learning algorithms. In subsequent paragraphs, we discuss aspects of the neural code that can be best appreciated by using population decoding approaches.

7.1. Population coding of attention

Machine-learning methods have been widely employed to decode movement related parameters from signals recorded in cortical areas involved in movement planning, preparation and execution (101, 102). This bias is justified by the obvious utility of such approaches in the development of BMIs that can restore motor function in human patients. Recent studies, however, have extended the applicability of these methods to explore whether cognitive functions can also be decoded using similar approaches. The results so far have contributed to a better understanding of the way cognition is mediated by neuronal populations and they have opened up new possibilities for the development of cognitive BMIs.

A number of studies have employed machine-learning decoding methods to examine whether attentional variables can be decoded from population activity in areas of the attention network. They have demonstrated that the location of attention can be decoded with very high accuracy from the spiking activity of neuronal populations in extrastriate and frontal areas including the lateral PFC (103, 104), FEF (92, 105–107), MT (108), and V4 (107). These results from single-trial population analyses are in line with the results obtained from univariate approaches. For example, accuracies in FEF are overall higher and reach significance earlier than in V4 (107) in agreement with the notion that the FEF provide top-town feedback to V4 during spatial attention (47). Furthermore, decoding object identity from an IT population, indicated that the main effect of directing attention to an object presented in an array with other objects, was to restore information towards a state in which the object was shown alone (109). This result is in line with previous findings, which have shown that when multiple stimuli co-exist within the RF of a single neuron, attention modulates the response of the neuron so that it responds as if only the attended stimulus is present (8, 13). However, it extends this

initial observation to highlight how activity in the entire population of IT neurons changes to represent the attended stimulus. Thus, it advances our knowledge on how object recognition is implemented in the brain while at the same time it allows the construction of more detailed, biologically plausible models of object recognition.

Important clues into the way distractors are encoded at the population level and how this representation affects the deployment of attention have also been provided by recent studies that used decoding approaches. Early, psychophysical studies had demonstrated that peripheral cues capture attention involuntary and can affect task performance (110). Single-trial population approaches have more recently examined the effect of distractors on decoding the locus of attention. Following a distractor change, a substantial decrease in the information conveyed in spiking activity about the location of attention was found in lateral PFC (103) and in the FEF (92), and a loss of information about object identity was reported in IT (109). A similar interference by the distractor was observed while decoding the locus of attention from LFP signals, particularly for frequencies above 60Hz (104). Effects on spikes and LFPs in the high-gamma range (120–256Hz) were identical. Distractor interference in the mid-gamma (60–120Hz) range was about half as that observed in the high-gamma range. Distractor interference was observed during both correct and error trials, however, in correct trials this effect was transient and rapidly recovered to previous levels. Moreover, both in FEF and in lateral PFC population activity following a distractor change was highly predictive of behavioral outcome (92, 103).

To summarize, decoding approaches that consider the pattern of activity across neuronal ensembles, complement conventional univariate analyses and provide a better understanding of attentional mechanisms at the population level. Given that our perception of the world is based on transient activity patterns of billions of neurons, decoding approaches promise to bridge neuronal population activity and behavior. In the following paragraphs, we discuss how decoding methods can improve our understanding of different aspects of population coding by providing examples from electrophysiology approaches employed in attention or other cognitive tasks.

7.2. Mixed selectivity and adaptive coding

Higher-level structures, such as the parietal, prefrontal and cingulate cortices, have been implicated in several different functions and behaviors and it has been suggested that they control flexible cognitive behavior (e.g. visual attention). A long-standing question is whether

this diversity and flexibility is implemented through different subsets of neurons that are specialized for particular behaviors or variables, or through a single population that participates in different behaviors by flexibly and dynamically adjusting responses within the network. Recent studies have provided insights into the way representations of behavioral variables are constructed at the population level in higher order functions and how these are modified by behavioral context.

In the posterior parietal cortex (PPC) of the rat independent information about decision and multi-modal sensory variables can be reliably decoded from the same neurons indicating that the same population can carry different types of information, which can be extracted according to the animal's needs (111). These results suggest that our perception and our ability to make choices based on multi-modal sensory information depend on extracting information from distributed dynamic patterns of activity in the same neuronal population. Similar patterns of mixed selectivity and adaptive coding at the population level have been demonstrated in monkey PFC. Neurons encoding whether two stimuli were identical or not, carried sufficient information to decode stimulus position (112). In another study, analysis of the pattern of neural activity at the population level in PFC showed that the response pattern to identical stimuli, adapted according to the different behavioral context so that the state of the population reflected a shift in coding depending on the initial instruction (113). In addition, activity of a large proportion of single neurons in the orbitofrontal cortex (OFC) dynamically switched between representations of objects associated with different reward values, as the population representation switched between different choice options, suggesting that the information carried by single cells changes dynamically depending on the network state (114).

The mixed selectivity observed in higher-order areas has a high-dimensional neuronal representation that could provide the basis for the remarkable adaptability of neuronal responses in these structures (115), a prerequisite for complex cognitive behavior (116). Adaptive coding may be the result of activity dependent short-term synaptic plasticity mechanisms (113, 117, 118). Discussing the origins and mechanisms of adaptive coding is beyond the scope of this review. One should, however, note that attentional control is essentially one example of such adaptive coding as it requires facilitation of a subset of inputs and filtering of other inputs in a flexible manner according to the behavioral demands. Thus, studying adaptive coding in higher order brain areas allows us to explore mechanisms that may well underlie attentional behavior.

7.3. Temporal population dynamics

Decoding approaches can also provide insights into the stationarity of the temporal code i.e. whether encoding of information changes over time or whether the representation of the decoded variable remains constant in time. If neuronal populations encode information in a non-stationary manner, the interconnection weights within a network are expected to change in time. An extreme realization of such a dynamic representation would be observed if different neurons encoded information at different time periods. Such fundamental questions remain largely unexplored. One way to examine population dynamics in time is to train a classifier at one time period and test with data from another time period. If the contributions of individual cells within the network remain the same over time, a classifier trained at one time should generalize equally well at other times. On the other hand, if population dynamics change in time, the generalization of a decoder over time will be poor.

Using such a cross-temporal pattern analysis, Crowe *et al* (119) suggested that task-relevant information is encoded dynamically in parietal area 7a of the macaque. They showed that during a spatial cognitive task, the neural representation of space changed dynamically, with distinct neurons being activated sequentially over time, although the value of the spatial variable remained the same. Interestingly, the representation of task-irrelevant information was stationary. Dynamic coding of task relevant information has also been found in IT (120) and in parietal area LIP (121). Importantly, whereas in LIP encoding of spatial attention was found to be non-stationary, the FEF population seemed to encode spatial attention in a more stationary way, and this temporally stable representation was more prominent for the attention selective cells in FEF (121). This finding indicates that although spatial attention signals can be decoded from both LIP and FEF, the nature of the representation of the locus of attention is different in the two areas. Likewise, stationary patterns of PFC activity were observed in the delay period of an association task (113). In contrast, both task-relevant and irrelevant information in the PFC was represented dynamically in a categorization task (120). Therefore, the temporal representation of task-critical information is neither area- nor task-specific. Rather, analysis of the time-resolved pattern of activity across the population in different areas and tasks is necessary to obtain insight into the nature of specific representations at different stages of processing in the brain.

Additional analyses providing a deeper understanding of the potential origin of the observed network dynamics will be most useful in future studies. As mentioned above, the fact that a classifier trained at a given time is unable to decode patterns observed at

another time, could mean that different neurons encode information at different periods in the task. In a match-to-category task, Meyers *et al* (120) removed the most category selective cells in one time bin and trained and tested the classifier in other time bins. If the same cells encoded information over the entire period under study, eliminating the most selective cells in one time bin should have lowered performance at other times. The authors found that this was not the case. Removing the most selective neurons at one time period, reduced accuracy in that period but accuracies in other time periods were unaffected. This finding suggests that different neurons carry task relevant information at different trial periods. Another source of non-stationarity in the population could be the transient selectivity profiles of individual neurons. Indeed, Meyers *et al* (112) described some highly-selective PFC cells that carried task-relevant information in a transient manner. The contribution of such highly-selective cells becomes even more pronounced in populations that rely on a small subset of neurons, as will be discussed in the following section. Thus, although information about the stationarity of the neural code is important in order to understand how representations of behavioral variables are constructed in the brain, additional analyses is required in order to understand the origin of non-stationarity. A more recent study demonstrated that despite the strong dynamic pattern and heterogeneity of individual PFC neurons, the population representation during working memory maintenance was stable across time, in agreement with the notion that working memory relies on persistent activity patterns. Interestingly, it was shown that this stable mnemonic representation coexisted with dynamic coding patterns observed during different task periods (122).

Besides examining population dynamics in the same trial, exploring temporal dynamics in longer timescales that may span several trials can provide important insights into the variability of network dynamics. A recent study showed that, in the mouse PPC, not only the responses of individual neurons varied across trials, but also different sets of neurons could be active in trials of similar behavioral context (123). The transition from one activation pattern to another occurred at a timescale of several seconds. Interestingly, the transition between different activation patterns was not random; activation of a particular set of neurons was predictive of future activation patterns. Furthermore, the initial activation pattern in each trial depended upon the activation patterns in the previous trial. Insightful approaches such as this can help us better understand population dynamics in the brain and assess neuronal variability under a different perspective.

7.4. Sparse and distributed information coding

A fundamental issue relevant to population coding is the number of neurons that are required at a given time to obtain a reliable representation of a stimulus and guide behavior. Decoding analysis can provide insights into the sparseness of information in the population code by calculating decoding performance as a function of the number of contributing neurons. In this context, sparseness refers to the number of neurons in a population that encode information, as opposed to the more commonly used term in relation to spiking activity, which refers to the proportion of neurons that are active at a given time (124). In other words, sparseness, in the context discussed here, does not refer to how often neurons in the population fire but to the information content in spiking activity.

Surprisingly, only a small subset of highly task-selective cells is sufficient to obtain a decoding accuracy almost as high as that of the whole population. Neuronal ensembles in IT and PFC comprising as few as 8 to 16 of the most selective cells contained nearly all of the cognitive information present in the entire population (103, 112, 120). Likewise, a sparse pattern of information coding has been observed with orientation discrimination in V1 (125) or saccade direction in LIP (126) and lateral PFC (127, 128). What are the implications of such a sparse representation for neuronal coding? It is possible that downstream neurons extract information from a small subpopulation of highly informative cells. Such a representation would have the advantage of reducing connectivity and metabolic requirements (129). Another possibility is that highly informative cells represent the output of the computations, whereas, less informative cells may contribute to the actual computations (103).

Although these ideas remain to be tested, decoding studies have reported that even the less selective neurons still contain significant amounts of redundant information. Univariate analyses that average responses across cells would label those neurons as task non-informative. However, taking into account the pattern of activity across the population of those cells reveals that they can carry considerable information (103, 112, 120). This suggests that although a small number of neurons may determine decoding performance, the rest of the population still contains substantial amounts of task-relevant information. To further explore the properties of distributed coding, Rigotti *et al* (130) asked whether task-selectivity is important at all for the neural code in PFC. To examine this, they removed task selectivity from each cell's response by replacing the firing rate responses to a particular task-relevant aspect by the average response to different task-aspects. Interestingly, they found that as task complexity increased within the trial, decoding accuracy increased towards the performance obtained from the intact population, indicating that information is distributed across the population even when it is not present in individual cells.

7.5. Temporal resolution of information code

Decoding approaches allow insights into the timescale of information coding. Two main hypotheses have been proposed about the way information is carried in spiking activity. The rate-coding hypothesis holds that information is carried in the average firing rate within a time window, whereas, the temporal-coding hypothesis states that the precise timing of spikes transmits additional information, beyond that carried by the average firing rate (see 131 for a review). Decoding studies have examined the information carried in different time windows by comparing performance using time windows of different lengths. Using small bins one extracts information primarily from the timing of individual spikes rather than from the average level of activity as it is the case when using longer time windows. Decoding the location of attention was optimal for time windows longer than 80–100ms in the FEF (106) and lateral PFC (103). Similarly, decoding of category information in IT and PFC was optimal for 150ms time windows, with shorter bins resulting in lower accuracies (120). These findings do not necessarily point against a temporal code but, instead, indicate that more information is carried by the average firing rate as measured using longer time windows. Although decoding of spatial attention is optimal for windows longer than 80–100ms, significant information is carried in bins as small as 20ms (103, 106). Another study demonstrated that robust category classification can be achieved from IT neurons using bins as small as 12.5.-50ms (132). Time bins of 12.5.ms that contain on average 0 to 2 spikes decode category information with accuracies higher than 80%, pointing to a sparse spiking representation.

In summary, results so far suggest that cognitive information is optimally decoded in time windows of a few hundreds of milliseconds, which points to rate-coding schemes. Nevertheless, several studies have found that significant information is also carried in finer temporal scales in line with the temporal-coding hypothesis. It should be noted that temporal coding can be assessed using a variety of parameters and the interested reader can refer to recent reviews (e.g. 133).

7.6. The role of correlations

A critical question is how inter-neuronal correlations affect the amount of information carried by a neuronal population (134). Some theoretical studies have suggested that for similarly tuned neurons, correlations limit information (82, 135). However, other studies have noted that under certain assumptions, correlations may actually increase the information carried by populations of neurons (136), especially in heterogeneous populations (137). Recently, it was suggested that not all types of correlations affect information coding (138). Specifically, the authors used a network model to show that information decreases only in the presence of a particular type of correlations, termed differential correlations (correlations proportional to the product of the derivatives of the tuning curves of the neurons in the population). However, the contribution of this type of correlation to overall correlations is likely to be small.

Experimentally, it has been demonstrated that attention, reduces inter-neuronal correlations (see paragraph 4.4.). However, it is not clear, whether reduced correlations alter the amount of information carried by the population, and if so in which direction. The prevailing view is that if correlations do limit information, processes such as attention that reduce correlations become beneficial by increasing the amount of information carried by single neurons and neuronal populations (89, 90). However, a recent theoretical study suggested that correlations within the population may be the result of trial to trial fluctuations in the attentional state and that fluctuations in the strength of attention (gain) do not affect decoding performance (139). If this is the case then the experimentally observed decrease in response variability is not an attentional mechanism that increases the amount of information. Interestingly, the authors showed that trial-to-trial fluctuations in the attentional state during feature attention introduce differential correlations that may actually limit (saturate) information.

Decoding approaches provide a straightforward way to evaluate the effect of correlations on the information that can be extracted from neuronal ensembles. In addition, decoding methods can examine the effect of correlations on large neuronal populations, an approach that can be more informative about the actual role of correlations in the brain than the pairwise Pearson correlations typically calculated between neurons (see paragraph 4.4.). The effect of correlations on population coding can be examined by shuffling trials in the original dataset in order to destroy the correlation structure. Subsequently, the information extracted from the original dataset using the correlations-aware classifier is compared to the information extracted from the shuffled dataset using a correlations-ignorant classifier. To preserve the tuning properties of individual cells, shuffling of trials recorded under the same stimulus condition is carried out.

A number of studies have assessed the role of inter-neuronal correlations in the encoding and decoding of stimulus- and movement-related information. Noise correlations were found to have a positive, albeit small, impact on decoded information from pairs or small populations of neurons (140), but effects were more pronounced for larger ensembles (141). More recently, it was reported that correlations improved the decoding of saccade direction and

eye position from small populations of, on average, seven LIP neurons (142). Examining inter-neuronal correlations within larger populations allows for a more realistic estimate of the actual effect of correlations in the brain. Using multi-electrode arrays, one study reported that V1 correlations significantly increased the decoding accuracy for stimulus orientation (125) (but see 143, 144). Decoding of motion direction also improved when MT correlations were considered (145). Some studies have examined the effect of correlations using LFP signals. Taking temporal correlations into account improved, on average, decoding performance in predicting the saccade goal by 5–10% in PFC (128). In contrast, correlations did not affect decoding of hand-movement direction from four simultaneously recorded LFP signals in the motor cortex (146), whereas in the parietal reach region (PRR) they even reduced decoding accuracy (147). Whether the diversity of results reflects region- and function-specific particularities or is due to methodological differences (e.g. the use of the raw LFP signal (146) or its spectral components (128, 147)) remains to be examined. Overall, in the type of tasks listed above, in most cases, correlations appear to increase decoded information, although, the effect in some cases is modest.

In contrast, correlations appear to be detrimental during attention. Tremblay *et al* (103) recorded simultaneously from 32 electrodes placed in the lateral PFC while monkeys were performing a spatial attention task. The authors used spiking activity to compare classification accuracies between the original and the shuffled dataset. Interestingly, taking correlations into account led to a small (6%) but significant decrease in the accuracy of decoding the locus of attention. Eliminating the correlation structure between similarly tuned neurons and maintaining all other correlations, yielded similar accuracies to the shuffled dataset, indicating that the lower accuracy in the original dataset was mainly caused by the correlations between similarly tuned units. The effect of removing correlations from simultaneously recorded LFP signals during attention was even more pronounced. Eliminating correlations improved decoding accuracy in the mid-gamma frequency range (60–100Hz) by up to 14% but only during the attention epoch (104). Accuracies in other task epochs, at least in the mid-gamma range, were mostly unaffected, suggesting that the effect of correlations on decoding performance was attention specific. These observations are in line with those of an earlier study that decoded the locus of attention from small ensembles (2–4 neurons) in area V1 during a perceptual grouping task (148). Decoding accuracies improved when correlations were ignored but only for ensembles that encoded the same object. Conversely, correlations improved decoding accuracy for ensembles encoding different objects.

At first, these results appear to be in agreement with the notion that attention reduces correlations and as a result the amount of information about the attended object carried by the population is increased (89, 90). This is not directly demonstrated by the decoding studies above, but is instead implied by the fact that eliminating correlations improves the decoding accuracy of attention location. It would be useful in future studies to directly assess the effect of correlations in object information with attention. Previous studies evaluated information about object identity with and without attention (109). Accordingly, future studies could perform a direct estimation of object identity information after taking into account or ignoring correlations. Moreover, given that correlations predominantly reflect low frequency oscillations (see paragraph 4.4.), one would expect that disrupting correlations would predominantly affect the information carried by low frequencies. Notably, the improvement in decoding spatial attention after removing correlations was more evident in the 60–100Hz range (104). Future studies should aim to explain how the disruption in low frequency oscillations may increase the information carried by higher frequencies.

The results presented above highlight the importance of inter-neuronal correlations in information coding. It should be noted that technological advances that allow simultaneous recordings from a large number of sites and neurons have gradually become available only in the last decade. As a result, most studies up to now have used decoding approaches in pseudopopulations of neurons recorded over different sessions. An obvious limitation of this approach is that it ignores the role of inter-neuronal correlations in information coding. Although these studies have contributed significant knowledge on coding principles, the use of larger and more realistic datasets is essential to crack the neural code at the population level.

7.7. Contribution of the LFP signal

While most studies have focused on decoding information from spiking activity, more recent studies have examined the information conveyed by LFP signals. LFPs are the low frequency components of the mean extracellular electric field potentials recorded around the electrode tip and reflect the summed population activity of local synaptic currents within a volume of neural tissue that contains hundreds or thousands of neurons. LFP signals are sensitive to suprathreshold as well as to subthreshold synaptic activity, reflect the balance between excitation and inhibition and carry information about the local network state (149–152). Thus, they are likely to carry complementary information, beyond that measured from spike signals. Moreover, the blood oxygen level-dependent (BOLD) contrast measured with fMRI is better correlated with the LFP signal (152), especially

in the gamma frequency range (153), therefore, LFP decoding results could help interpret aspects of multi-voxel pattern analysis (MVPA) performed with fMRI (154, 155), and relate these results to extracellularly recorded neural signals. Importantly, LFP signals are robust compared to the more noisy spiking signals, and are likely to be useful for decoding purposes, particularly for chronic implants used in BMIs.

LFP oscillatory activity is modulated by attention in a frequency specific manner, as described in paragraph 4.2. Here, we review studies that decoded information from different frequency bands of the LFP signal in an attempt to examine the amount of information carried by oscillatory activity in the brain and the role of different frequency bands in information coding.

Information content in different frequency bands appears to be modality- and area- specific. For example, stimulus related information in the gamma band (60–100Hz) was higher for visual compared to auditory stimuli (156) and information about hand movement direction in the motor cortex was higher in the delta frequency range (less than 4Hz) (157). Cognitive information on the other hand, can be primarily extracted from higher frequency oscillations. Specifically, the location of attention could be reliably decoded from LFP signals in the lateral PFC from frequencies above 60Hz with performance being optimal in the high-gamma range (120–250Hz) (104). Frequencies below 60Hz carried information about stimulus location and saccade direction but no attention related information. Attention has also been reliably decoded from signals obtained through epidural, ECoG electrodes placed over the visual cortex (mainly area V4), with maximal accuracy between 60–80Hz and accuracies significantly above chance for frequencies in the entire gamma range (~30–200Hz) (158). Unpublished results from our lab indicate that decoding of attention in FEF and V4 is optimal for signals in the mid- and high-gamma range (60–140Hz), with accuracies in the low-gamma range (30–60Hz) carrying also substantial information. Performance in other frequency bands is lower but significantly above chance indicating that small amounts of information are also contained in frequencies lower than gamma. Interestingly, in the FEF, but not in V4, we also observed increased accuracy in the theta range (4–8Hz). Decoding attention location from MT signals resulted in somewhat different results. Performance was optimal in the low- and mid-gamma range (30–120Hz) and significantly above chance in lower frequencies (1–30Hz), but was at chance level in the high-gamma range (121–200Hz) (108). Similar to most attention studies, reward value information in the orbitofrontal cortex (OFC) was reliably decoded from frequencies in the high-gamma (70–200Hz) but also in the theta (4–8Hz) range (114). Accuracies obtained

from other frequency bands were significantly lower but mostly above chance. Thus, these first studies indicate that gamma (including mid- and high-gamma frequencies) oscillations carry significant information about the location of attention with lower frequencies also contributing substantial information. Given that only recently studies have started to exploit LFP signals for decoding purposes a number of questions remain open. These include whether particular frequencies are associated with somewhat different aspects of attentional function, whether other aspects of attention besides its spatial focus can be similarly decoded from LFP signals and what the source of some of the discrepancies mentioned above might be. Discrepancies between studies could be due to differences between areas that reflect idiosyncratic characteristics of the local circuits, differences in the recording depth (128), or simply differences arising from the analysis.

A methodological concern relative to the analysis of low frequencies is that time windows into which LFPs are considered should be long enough to include more than one cycles of each frequency of interest. Intuitively, when several cycles are considered, variability is reduced and the signal becomes more informative. Surprisingly, robust decoding information was obtained in the theta range with windows as small as 80ms in the OFC study, and 100ms in our study, with both windows containing less than a theta cycle. Similarly, in another study, information about sensory stimuli could be reliably extracted from low frequencies even when only a fraction of a cycle was considered (156). Although stimulus information did increase for longer windows in all frequency bands as a consequence of increased signal reliability, information in high frequencies (above 50Hz) increased more dramatically when longer windows (a few tens of cycles) were considered. Given these observations, the dependence of cognitive information on the length of the analysis window for the different frequency bands seems to be a factor that should be explored more thoroughly. Future studies that will examine such issues will provide useful insight into the information carried by the LFP signal in different timescales.

It has been previously pointed out that high-gamma activity is strongly correlated with spiking activity as a consequence of the energy in the LFP signal associated with spiking transients and the contribution of small spikes that cannot be filtered out (159). This close relationship between high-gamma oscillations and spiking activity has been observed in decoding studies. The information content of spikes and gamma LFP oscillations (in frequencies above 60Hz) was significantly correlated (104). Furthermore, the authors of this study reported that attention could be reliably decoded by training the classifier with spiking data and testing it with LFP data (frequencies above

60Hz), indicating that the information content in these two signals is very similar. Nevertheless, after applying methods that clean the LFP signal from potentially contaminating spike components recorded on the same electrode (160), decoding accuracy of attention location in the high-gamma range was affected only marginally (104). Although it is possible that remnants of low amplitude spiking activity could be still present in the LFP signal following the cleaning procedure, these results suggest that the high-frequency LFP reflects, at least to a degree, a signal different to the spiking one, which may well carry similar information. In visual and auditory cortices, redundancy between spike and LFP signals was larger for high frequencies, however, it was small in magnitude and accounted for only 10% of information (156). These results are informative for the potential use of LFPs in BMI applications. Moreover, they highlight the importance of using decoding approaches to better understand the role of different oscillatory frequency bands in information processing and the relation of high-frequency oscillations to spiking activity as they allow comparisons on a common scale.

8. COMPARISON BETWEEN DIFFERENT SIGNAL TYPES

A critical question is whether the different types of electrophysiological signals carry different amounts of information. For example, it has been noted that spiking recordings are biased towards larger, pyramidal cells, thus, they mostly represent the output of a cortical area. In contrast, components of the LFP signal arise from synchronized dendrosomatic processing of synaptic inputs (161) and peri-synaptic activity (151) thus, they are considered to mainly reflect the inputs to a cortical region and local processing activity. As a consequence, LFP and spike signals may carry different information. One way to examine the information conveyed by spikes and LFPs is to compare the performance of decoders based on the two signal types. Decoding of object identity in IT was reported to be comparable for spikes and LFPs in the alpha and beta frequency range but was lower for high-frequency LFP oscillations (162). In cognitive tasks, decoding of spatial attention in the lateral PFC was higher for spikes than for LFPs (104). The same pattern was observed while decoding reward value from the OFC (114).

Some studies have suggested that spikes and LFPs may convey information about different behavioral variables. In the parietal reach region (PRR), for example, behavioral state was better predicted by LFPs, whereas, reaching and saccade direction by spikes (163). Behavioral state could also be decoded optimally from LFP frequencies in the 0–20Hz range in LIP, but not from firing rate responses (164). Saccade direction in LIP (164) and PFC (128) was decoded with comparable accuracy from both spike and LFP signals.

Unpublished results from our lab are in line with those of (164) and (128). Classification of attention location in the 60–140Hz range yields comparable results with that from spiking activity in both FEF and V4.

Overall, current findings suggest that in certain cases, spiking and LFP signals convey different amounts of information, and this depends on the task, the cortical area and the decoded variable. However, methodological differences between studies should also be acknowledged. For example, while decoding saccade direction from LFPs, one study used as an input to the classifier the frequency with the highest discriminability in the 30–100Hz range, which was site-specific (164), whereas a second study considered information across all frequencies using a dimensionality reduction approach (128), and a third study concatenated power spectra across different frequency bands and recording sites (163). The first two approaches ensure that the spike and LFP classifiers have the same dimensionality, which is not the case with the third approach. Thus, it is possible that the differences in the performance of the LFP decoder may arise due to theoretical limitations such as the "curse of dimensionality" discussed in section 6.

Another issue, that is of utmost importance for the development of BMI applications, is how stable the decoding is over time for spikes and LFPs. In chronically implanted arrays, signal deteriorates over time due to slight movement of the array within the tissue but also due to the inflammatory response to the foreign object, resulting in gliosis around the electrode tips (165). As a result, the amplitude of action potentials recorded from the electrodes is reduced over time and the signal to noise ratio decreases (166). In principle, the LFP represents activity within a larger volume of brain tissue compared to spikes, thus, it can provide a more reliable signal for decoding purposes in the long term. In line with this, the stability of a fixed classifier, i.e. a classifier that is trained in the first day of recording and is used for testing signals in subsequent sessions, decayed faster for spikes than for LFPs over the period of one month (104). Although these results are promising for the use of LFPs in BMI applications, one long-term study reported that the advantages of LFP decoders in terms of signal stability were modest (167). Nevertheless, LFP signals were more long-lasting and provided reliable decoding when the spike signal completely decayed (168). Technological advances and the development of different recording probes may help to prolong the stability of recording signals that can be used for years long decoding in BMI applications.

In summary, many studies have demonstrated that the LFP signal is a reliable source for decoding behavioral information from neural activity. The degree to which it provides additional information to the spike

signal depends on the task, the cortical area and the decoded variable. In cases where information is carried in the low- and mid- frequency components, the LFP can provide complementary information to that of the spiking signal. In contrast, information in the higher frequencies (above 60Hz) appears to be redundant.

9. IMPLICATIONS FOR THE DEVELOPMENT OF BRAIN-MACHINE INTERFACES (BMIs)

Brain-machine interfaces (BMIs) are applications that allow communication between the brain and the outside world. BMI research has had revolutionary results and applications in the past years with particular emphasis given to the development of interfaces that can drive prosthetic devises (e.g. artificial limbs) to aid patients with motor disabilities. Applications of BMIs include moving a cursor on a screen or controlling a prosthetic arm for reaching and grasping, using brain signals (101, 169–171). Although BMI research has primarily emphasized motor functions, recent studies have highlighted the importance of using signals from areas processing cognitive information to control cognitive BMIs. Cognitive BMIs can be used for at least two types of applications (172). First, for the development of neurofeedback and cognitive control applications for patients with cognitive disorders (e.g. attention deficit hyperactivity disorder) or neurodegenerative diseases (e.g. Parkinson's disease). In such applications, the behavioral information of interest is extracted from brain signals using decoding methods. Patients learn to control their brain signals, through cognitive control, to improve their cognitive function. Second, cognitive BMIs can be used for the development of applications that aid communication of paralyzed patients or patients with neurodegenerative diseases that affect motor function. Devices, such as spelling boards, can be operated by controlling a computer cursor using signals decoded from motor and premotor cortices (173, 174). In paralyzed patients, training the decoder with actual movements is not possible, therefore, users are initially asked to imagine that they are controlling the movement of the cursor and once the decoder is trained, it translates neuronal activity associated with movement intention to control the actual cursor. Another option to operate communication devices is by means of eye movements or movements of facial muscles. However, there are cases that eye tracking cannot be performed accurately (e.g. due to ptosis) or cases that patients lose the ability of muscle control. A number of studies have demonstrated that saccade direction or saccade parameters such as amplitude or timing, can be decoded from the delay period activity, i.e. before the actual saccade occurs, in LIP (126, 164, 175), FEF (176) and lateral PFC (127, 128). A recent study went one step further and developed a BMI that decoded eye movement intentions without the animal performing a saccade (177). Such an application would be advantageous for patients who cannot move their eyes. The studies reviewed here raise the additional possibility to operate devices such as spelling boards with attention signals. The degree to which this can be achieved in practice remains to be explored. An EEG-based spelling board resulted in poor performance when operated with covert attention, however, performance markedly improved when it was operated with overt attention (178). Given that attention and eye movement representations reside in overlapping parietal and prefrontal structures, another possibility would be to use a conjunction of attention and saccade intention signals to design more effective applications.

Besides enhancing our understanding of population coding mechanisms of attention, decoding studies can provide insights that can help determine the optimal design parameters of cognitive and attention based BMIs. The finding that a few highly informative cells are sufficient to obtain a decoding accuracy almost as high as that of the whole population allows for two possible methods for feature extraction. One possibility would be to apply a feature selection method that will identify the most informative cells and use these for decoding. However, although performance is influenced by the most selective neurons, task-relevant information is distributed across the population as discussed in previous sections. Thus, a second possibility would be to use the whole population for decoding, which results in comparable performance (e.g. 120). Although both approaches result in equivalent accuracies, the latter might be preferable as feature extraction procedures require additional computational time. A further limitation of using single unit activity as input to BMIs is that the spike-sorting process used to isolate units requires significant computational time and resources that might be a limiting factor in real time applications. Moreover, in chronically implanted arrays, due to glial scar tissue formation near the electrodes, the amplitude of action potentials decreases over time (166). Furthermore, it may not be possible to record from the same units across days as less than half of the original single units remain stable over a two-week period (179). To overcome these limitations, another option would be to use voltage-thresholded, multiunit activity as input to BMIs. In early recording sessions multiunit performance was found to be comparable, albeit somewhat lower, to that of single units in decoding attention (103) or arm movement direction (180, 181). However, over a recording period of several months, multiunit activity decoders do not lose performance (166) and as signal quality deteriorates, they outperform single-unit decoders (167).

Several studies have demonstrated that robust decoding performance can be obtained from LFP signals as discussed in the previous sections.

Depending on the task and cortical area, LFP classifiers have comparable, and in some cases, better performance than spike decoders. Although, it has been previously hypothesized that LFP signals may be more stable over time, the advantages of LFP signals in decoding stability seem to be modest in the long term (167), however, they provide a reliable decoding source when the spike signal completely decays (168). Taking into account the reduction in signal quality over time, combining spike and LFP features may be beneficial in terms of signal stability and robustness. In terms of decoding performance, combining signals may be beneficial in cases that significant information can be decoded from non-redundant components of the LFP such as low- and mid- frequency oscillations. In contrast, in cases that information is conveyed primarily by the high-frequency components, combining the two signals may not necessarily result in superior performance given the redundancy between the two signals.

In cognitive tasks, optimal decoding performance can be achieved by averaging spiking activity in time windows of a few hundred milliseconds (paragraph 7.5.). Likewise, LFP decoding accuracy improves with longer windows, in all frequency bands, as a result of increased signal reliability (156). Intuitively, longer windows average away more noise and result in more accurate predictions. However, a critical performance factor in online applications is the information transfer rate capacity (ITRC) that measures the rate of information transfer (bits/sec). With longer windows accuracy increases but the ITRC decreases; thus, one needs to determine the optimal accuracy vs. speed ratio for real time decoding that can be efficiently used in BMI applications (182).

Finally, although inter-neuronal correlations in most cases increase the amount of information carried by a population, they become detrimental in decoding the locus of attention from spikes (103) and more so from LFP signals (104). Thus, in attention based BMI applications it might be beneficial to decorrelate signals in order to increase classification performance.

10. CONCLUSION

We have reviewed studies that employed pattern-classification algorithms in the analysis of neural data to understand attention and other cognitive functions. The results show that decoding information approaches can reveal aspects of population coding that remain undetected in single-neuron, trial averaging methods. Decoding information from populations of neurons on a single trial basis allows a deeper understanding of how distributed activity patterns across different neurons can be read out in the brain to form our perception and cognitive behavior

on a moment to moment basis. Moreover, decoding information approaches have quantified the amount of information carried by different signals in the brain (e.g. spikes, LFPs) and have evaluated the utility of these signals in BMI applications. Few studies so far have explored whether information about cognitive variables can be decoded from brain activity and even fewer have addressed whether cognitive signals can be used to help patients with cognitive or motor disabilities. Although these studies have provided important insights into the population representation of cognitive processes, many questions remain to be addressed in the future. Despite the fact that the field of cognitive BMIs is still young, the results so far have pointed to new directions of research and have opened up new possibilities for restoring communication and motor functions in patients.

11. ACKNOWLEDGEMENT

12. REFERENCES

1. M. I. Posner: Orienting of attention. *Q J Exp Psychol*, 32(1), 3–25 (1980)
DOI: 10.1080/00335558008248231

2. K. R. Cave and J. M. Wolfe: Modeling the role of parallel processing in visual search. *Cogn Psychol*, 22(2), 225–71 (1990)
DOI: 10.1016/0010-0285(90)90017-X

3. A. M. Treisman and G. Gelade: A feature-integration theory of attention. *Cogn Psychol*, 12(1), 97–136 (1980)
DOI: 10.1016/0010-0285(80)90005-5

4. J. M. Wolfe: Guided Search 2.0. - a Revised Model of Visual-Search. *Psychonomic Bulletin & Review*, 1(2), 202–238 (1994)
DOI: 10.3758/BF03200774

5. M. Corbetta and G. L. Shulman: Control of goal-directed and stimulus-driven attention in the brain. *Nat Rev Neurosci*, 3(3), 201–15. (2002)
DOI: 10.1038/nrn755

6. R. Desimone and J. Duncan: Neural mechanisms of selective visual attention. *Annu Rev Neurosci*, 18, 193–222 (1995)
DOI:10.1146/annurev.ne.18.030195.001205

7. S. Kastner and L. G. Ungerleider: Mechanisms of visual attention in the human cortex. *Annu Rev Neurosci*, 23, 315–41 (2000)
DOI: 10.1146/annurev.neuro.23.1.315

8.   L. Chelazzi, J. Duncan, E. K. Miller and R. Desimone: Responses of neurons in inferior temporal cortex during memory-guided visual search. *J Neurophysiol*, 80(6), 2918–40. (1998)

9.   L. Chelazzi, E. K. Miller, J. Duncan and R. Desimone: Responses of neurons in macaque area V4 during memory-guided visual search. *Cereb Cortex*, 11(8), 761–72. (2001)
DOI: 10.1093/cercor/11.8.761

10. S. J. Luck, L. Chelazzi, S. A. Hillyard and R. Desimone: Neural mechanisms of spatial selective attention in areas V1, V2, and V4 of macaque visual cortex. *J Neurophysiol*, 77(1), 24–42. (1997)

11. J. Moran and R. Desimone: Selective attention gates visual processing in the extrastriate cortex. *Science*, 229, 782–784 (1985)
DOI: 10.1126/science.4023713

12. B. C. Motter: Focal attention produces spatially selective processing in visual cortical areas V1, V2, and V4 in the presence of competing stimuli. *J Neurophysiol*, 70(3), 909–19 (1993)

13. J. H. Reynolds, L. Chelazzi and R. Desimone: Competitive mechanisms subserve attention in macaque areas V2 and V4. *J Neurosci*, 19(5), 1736–53 (1999)

14. S. Treue and J. C. Martinez Trujillo: Feature-based attention influences motion processing gain in macaque visual cortex. *Nature*, 399(6736), 575–9 (1999)
DOI: 10.1038/21176

15. S. Treue and J. H. Maunsell: Attentional modulation of visual motion processing in cortical areas MT and MST. *Nature*, 382(6591), 539–41 (1996)
DOI: 10.1038/382539a0

16. B. Noudoost, M. H. Chang, N. A. Steinmetz and T. Moore: Top-down control of visual attention. *Curr Opin Neurobiol*, 20(2), 183–90 (2010)
DOI: 10.1016/j.conb.2010.02.003

17. J. Duncan: Disorganization of Behavior after Frontal-Lobe Damage. *Cognitive Neuropsych*, 3(3), 271–290 (1986)
DOI: 10.1080/02643298608253360

18. S. R. Friedman-Hill, L. C. Robertson, R. Desimone and L. G. Ungerleider: Posterior parietal cortex and the filtering of distractors.

Proc Natl Acad Sci U S A, 100(7), 4263–8 (2003)
DOI: 10.1073/pnas.0730772100

19. M. M. Mesulam: A cortical network for directed attention and unilateral neglect. *Ann Neurol*, 10(4), 309–25 (1981)
DOI: 10.1002/ana.410100402

20. A. F. Rossi, N. P. Bichot, R. Desimone and L. G. Ungerleider: Top down attentional deficits in macaques with lesions of lateral prefrontal cortex. *J Neurosci*, 27(42), 11306–14 (2007)
DOI: 10.1523/JNEUROSCI.2939-07.2007

21. C. Wardak, E. Olivier and J. R. Duhamel: A deficit in covert attention after parietal cortex inactivation in the monkey. *Neuron*, 42(3), 501–8 (2004)
DOI: 10.1016/S0896-6273(04)00185-0

22. J. Gottlieb: From thought to action: the parietal cortex as a bridge between perception, action, and cognition. *Neuron*, 53(1), 9–16 (2007)
DOI: 10.1016/j.neuron.2006.12.009

23. K. G. Thompson and N. P. Bichot: A visual salience map in the primate frontal eye field. *Prog Brain Res*, 147, 251–62 (2005)
DOI: 10.1016/s0079-6123(04)47019-8

24. E. B. Cutrell and R. T. Marrocco: Electrical microstimulation of primate posterior parietal cortex initiates orienting and alerting components of covert attention. *Exp Brain Res*, 144(1), 103–13 (2002)
DOI: 10.1007/s00221-002-1032-x

25. T. Moore and M. Fallah: Control of eye movements and spatial attention. *Proc Natl Acad Sci U S A*, 98(3), 1273–6 (2001)
DOI: 10.1073/pnas.98.3.1273

26. I. E. Monosov and K. G. Thompson: Frontal eye field activity enhances object identification during covert visual search. *J Neurophysiol*, 102(6), 3656–72 (2009)
DOI: 10.1152/jn.00750.2009

27. C. Wardak, G. Ibos, J. R. Duhamel and E. Olivier: Contribution of the monkey frontal eye field to covert visual attention. *J Neurosci*, 26(16), 4228–35 (2006)
DOI: 10.1523/JNEUROSCI.3336-05.2006

28. T. J. Buschman and E. K. Miller: Top-down versus bottom-up control of attention in the prefrontal and posterior parietal

cortices. *Science*, 315(5820), 1860–2 (2007)
DOI: 10.1126/science.1138071

29. N. P. Bichot, M. T. Heard, E. M. DeGennaro and R. Desimone: A Source for Feature-Based Attention in the Prefrontal Cortex. *Neuron*, 88(4), 832–44 (2015)
DOI: 10.1016/j.neuron.2015.10.001

30. C. J. McAdams and J. H. Maunsell: Attention to both space and feature modulates neuronal responses in macaque area V4. *J Neurophysiol*, 83(3), 1751–5 (2000)

31. J. H. Reynolds and D. J. Heeger: The normalization model of attention. *Neuron*, 61(2), 168–85 (2009)
DOI: 10.1016/j.neuron.2009.01.002

32. J. Lee and J. H. Maunsell: A normalization model of attentional modulation of single unit responses. *PLoS ONE*, 4(2), e4651 (2009)
DOI: 10.1371/journal.pone.0004651

33. J. C. Martinez-Trujillo and S. Treue: Feature-based attention increases the selectivity of population responses in primate visual cortex. *Curr Biol*, 14(9), 744–51 (2004)
DOI: 10.1016/j.cub.2004.04.028

34. E. A. Buffalo, P. Fries, R. Landman, T. J. Buschman and R. Desimone: Laminar differences in gamma and alpha coherence in the ventral stream. *Proc Natl Acad Sci U S A*, 108(27), 11262–7 (2011)
DOI: 10.1073/pnas.1011284108

35. K. McAlonan, J. Cavanaugh and R. H. Wurtz: Guarding the gateway to cortex with attention in visual thalamus. *Nature*, 456(7220), 391–4 (2008)
DOI: 10.1038/nature07382

36. R. M. McPeek and E. L. Keller: Saccade target selection in the superior colliculus during a visual search task. *J Neurophysiol*, 88(4), 2019–34 (2002)

37. A. Ignashchenkova, P. W. Dicke, T. Haarmeier and P. Thier: Neuron-specific contribution of the superior colliculus to overt and covert shifts of attention. *Nat Neurosci*, 7(1), 56–64 (2004)
DOI: 10.1038/nn1169

38. R. J. Krauzlis, L. P. Lovejoy and A. Zenon: Superior Colliculus and Visual Spatial Attention. *Annu Rev Neurosci* (2013)
DOI:10.1146/annurev-neuro-062012-170249

39. J. L. Herrero, M. J. Roberts, L. S. Delicato, M. A. Gieselmann, P. Dayan and A. Thiele: Acetylcholine contributes through muscarinic receptors to attentional modulation in V1. *Nature*, 454(7208), 1110–4 (2008)
DOI: 10.1038/nature07141

40. C. J. McAdams and R. C. Reid: Attention modulates the responses of simple cells in monkey primary visual cortex. *J Neurosci*, 25(47), 11023–33 (2005)
DOI: 10.1523/JNEUROSCI.2904-05.2005

41. P. S. Khayat, H. Spekreijse and P. R. Roelfsema: Attention lights up new object representations before the old ones fade away. *J Neurosci*, 26(1), 138–42 (2006)
DOI: 10.1523/JNEUROSCI.2784-05.2006

42. N. P. Bichot, A. F. Rossi and R. Desimone: Parallel and serial neural mechanisms for visual search in macaque area V4. *Science*, 308(5721), 529–34 (2005)
DOI: 10.1126/science.1109676

43. J. W. Bisley and M. E. Goldberg: Neuronal activity in the lateral intraparietal area and spatial attention. *Science*, 299(5603), 81–6 (2003)
DOI: 10.1126/science.1077395

44. J. P. Gottlieb, M. Kusunoki and M. E. Goldberg: The representation of visual salience in monkey parietal cortex. *Nature*, 391, 481–484 (1998)
DOI: 10.1038/35135

45. K. G. Thompson, N. P. Bichot and J. D. Schall: Dissociation of visual discrimination from saccade programming in macaque frontal eye field. *J Neurophysiol*, 77(2), 1046–50 (1997)

46. N. P. Bichot and J. D. Schall: Saccade target selection in macaque during feature and conjunction visual search. *Vis Neurosci*, 16(1), 81–9 (1999)
DOI: 10.1017/S0952523899161042

47. G. G. Gregoriou, S. J. Gotts, H. Zhou and R. Desimone: High-frequency, long-range coupling between prefrontal and visual cortex during attention. *Science*, 324(5931), 1207–10 (2009)
DOI: 10.1126/science.1171402

48. P. Fries, J. H. Reynolds, A. E. Rorie and R. Desimone: Modulation of oscillatory neuronal synchronization by selective visual

attention. *Science*, 291(5508), 1560–1563 (2001)
DOI: 10.1126/science.1055465

49. Y. B. Saalmann, I. N. Pigarev and T. R. Vidyasagar: Neural mechanisms of visual attention: how top-down feedback highlights relevant locations. *Science*, 316(5831), 1612–5 (2007)
DOI: 10.1126/science.1139140

50. K. Taylor, S. Mandon, W. A. Freiwald and A. K. Kreiter: Coherent oscillatory activity in monkey area v4 predicts successful allocation of attention. *Cereb Cortex*, 15(9), 1424–37 (2005)
DOI: 10.1093/cercor/bhi023

51. G. G. Gregoriou, S. Paneri and P. Sapountzis: Oscillatory synchrony as a mechanism of attentional processing. *Brain Res*, 1626, 165–82 (2015)
DOI: 10.1016/j.brainres.2015.02.004

52. T. Gruber, M. M. Muller, A. Keil and T. Elbert: Selective visual-spatial attention alters induced gamma band responses in the human EEG. *Clin Neurophysiol*, 110(12), 2074–85 (1999)
DOI: 10.1016/S1388-2457(99)00176-5

53. M. M. Muller and A. Keil: Neuronal synchronization and selective color processing in the human brain. *J Cogn Neurosci*, 16(3), 503–22 (2004)
DOI: 10.1162/089892904322926827

54. M. Siegel, T. H. Donner, R. Oostenveld, P. Fries and A. K. Engel: Neuronal synchronization along the dorsal visual pathway reflects the focus of spatial attention. *Neuron*, 60(4), 709–19 (2008)
DOI: 10.1016/j.neuron.2008.09.010

55. C. Tallon-Baudry, O. Bertrand, M. A. Henaff, J. Isnard and C. Fischer: Attention modulates gamma-band oscillations differently in the human lateral occipital cortex and fusiform gyrus. *Cereb Cortex*, 15(5), 654–62 (2005)
DOI: 10.1093/cercor/bhh167

56. C. A. Bosman, J. M. Schoffelen, N. Brunet, R. Oostenveld, A. M. Bastos, T. Womelsdorf, B. Rubehn, T. Stieglitz, P. De Weerd and P. Fries: Attentional stimulus selection through selective synchronization between monkey visual areas. *Neuron*, 75(5), 875–88 (2012)
DOI: 10.1016/j.neuron.2012.06.037

57. I. Grothe, S. D. Neitzel, S. Mandon and A. K. Kreiter: Switching neuronal inputs by differential modulations of gamma-band phase-coherence. *J Neurosci*, 32(46), 16172–80 (2012)
DOI: 10.1523/JNEUROSCI.0890-12.2012

58. A. M. Bastos, J. Vezoli and P. Fries: Communication through coherence with inter-areal delays. *Curr Opin Neurobiol*, 31, 173–80 (2015)
DOI: 10.1016/j.conb.2014.11.001

59. G. G. Gregoriou, A. F. Rossi, L. G. Ungerleider and R. Desimone: Lesions of prefrontal cortex reduce attentional modulation of neuronal responses and synchrony in V4. *Nat Neurosci*, 17(7), 1003–11 (2014)
DOI: 10.1038/nn.3742

60. A. Bollimunta, J. Mo, C. E. Schroeder and M. Ding: Neuronal mechanisms and attentional modulation of corticothalamic alpha oscillations. *J Neurosci*, 31(13), 4935–43 (2011)
DOI: 10.1523/JNEUROSCI.5580-10.2011

61. P. Fries, T. Womelsdorf, R. Oostenveld and R. Desimone: The effects of visual stimulation and selective visual attention on rhythmic neuronal synchronization in macaque area V4. *J Neurosci*, 28(18), 4823–35 (2008)
DOI: 10.1523/JNEUROSCI.4499-07.2008

62. M. Bauer, R. Oostenveld, M. Peeters and P. Fries: Tactile spatial attention enhances gamma-band activity in somatosensory cortex and reduces low-frequency activity in parieto-occipital areas. *J Neurosci*, 26(2), 490–501 (2006)
DOI: 10.1523/JNEUROSCI.5228-04.2006

63. G. Thut, A. Nietzel, S. A. Brandt and A. Pascual-Leone: Alpha-band electroencephalographic activity over occipital cortex indexes visuospatial attention bias and predicts visual target detection. *J Neurosci*, 26(37), 9494–502 (2006)
DOI: 10.1523/JNEUROSCI.0875-06.2006

64. M. S. Worden, J. J. Foxe, N. Wang and G. V. Simpson: Anticipatory biasing of visuospatial attention indexed by retinotopically specific alpha-band electroencephalography increases over occipital cortex. *J Neurosci*, 20(6), RC63 (2000)

65. B. F. Handel, T. Haarmeier and O. Jensen: Alpha oscillations correlate with the

successful inhibition of unattended stimuli. *J Cogn Neurosci*, 23(9), 2494–502 (2011)
DOI: 10.1162/jocn.2010.21557

66. S. P. Kelly, E. C. Lalor, R. B. Reilly and J. J. Foxe: Increases in alpha oscillatory power reflect an active retinotopic mechanism for distracter suppression during sustained visuospatial attention. *J Neurophysiol*, 95(6), 3844–51 (2006)
DOI: 10.1152/jn.01234.2005

67. S. Palva and J. M. Palva: New vistas for alpha-frequency band oscillations. *Trends Neurosci*, 30(4), 150–8 (2007)
DOI: 10.1016/j.tins.2007.02.001

68. W. Klimesch: Alpha-band oscillations, attention, and controlled access to stored information. *Trends Cogn Sci*, 16(12), 606–17 (2012)
DOI: 10.1016/j.tics.2012.10.007

69. A. Maier, G. K. Adams, C. Aura and D. A. Leopold: Distinct superficial and deep laminar domains of activity in the visual cortex during rest and stimulation. *Front Syst Neurosci*, 4 (2010)
DOI: 10.3389/fnsys.2010.00031

70. D. Xing, C. I. Yeh, S. Burns and R. M. Shapley: Laminar analysis of visually evoked activity in the primary visual cortex. *Proc Natl Acad Sci U S A*, 109(34), 13871–6 (2012)
DOI: 10.1073/pnas.1201478109

71. T. van Kerkoerle, M. W. Self, B. Dagnino, M. A. Gariel-Mathis, J. Poort, C. van der Togt and P. R. Roelfsema: Alpha and gamma oscillations characterize feedback and feedforward processing in monkey visual cortex. *Proc Natl Acad Sci U S A*, 111(40), 14332–41 (2014)
DOI: 10.1073/pnas.1402773111

72. M. N. Shadlen and W. T. Newsome: The variable discharge of cortical neurons: implications for connectivity, computation, and information coding. *J Neurosci*, 18(10), 3870–96 (1998)

73. D. J. Tolhurst, J. A. Movshon and A. F. Dean: The statistical reliability of signals in single neurons in cat and monkey visual cortex. *Vision Res*, 23(8), 775–85 (1983)
DOI: 10.1016/0042-6989(83)90200-6

74. M. M. Churchland, B. M. Yu, J. P. Cunningham, L. P. Sugrue, M. R. Cohen,

G. S. Corrado, W. T. Newsome, A. M. Clark, P. Hosseini, B. B. Scott, D. C. Bradley, M. A. Smith, A. Kohn, J. A. Movshon, K. M. Armstrong, T. Moore, S. W. Chang, L. H. Snyder, S. G. Lisberger, N. J. Priebe, I. M. Finn, D. Ferster, S. I. Ryu, G. Santhanam, M. Sahani and K. V. Shenoy: Stimulus onset quenches neural variability: a widespread cortical phenomenon. *Nat Neurosci*, 13(3), 369–78 (2010)
DOI: 10.1038/nn.2501

75. M. H. Chang, K. M. Armstrong and T. Moore: Dissociation of response variability from firing rate effects in frontal eye field neurons during visual stimulation, working memory, and attention. *J Neurosci*, 32(6), 2204–16 (2012)
DOI: 10.1523/JNEUROSCI.2967-11.2012

76. B. A. Purcell, R. P. Heitz, J. Y. Cohen and J. D. Schall: Response variability of frontal eye field neurons modulates with sensory input and saccade preparation but not visual search salience. *J Neurophysiol*, 108(10), 2737–50 (2012)
DOI: 10.1152/jn.00613.2012

77. C. J. McAdams and J. H. Maunsell: Effects of attention on the reliability of individual neurons in monkey visual cortex. *Neuron*, 23(4), 765–73 (1999)
DOI: 10.1016/S0896-6273(01)80034-9

78. J. F. Mitchell, K. A. Sundberg and J. H. Reynolds: Differential attention-dependent response modulation across cell classes in macaque visual area V4. *Neuron*, 55(1), 131–41 (2007)
DOI: 10.1016/j.neuron.2007.06.018

79. A. Thiele, C. Brandt, M. Dasilva, S. Gotthardt, D. Chicharro, S. Panzeri and C. Distler: Attention Induced Gain Stabilization in Broad and Narrow-Spiking Cells in the Frontal Eye-Field of Macaque Monkeys. *J Neurosci*, 36(29), 7601–12 (2016)
DOI: 10.1523/JNEUROSCI.0872-16.2016

80. C. Hussar and T. Pasternak: Trial-to-trial variability of the prefrontal neurons reveals the nature of their engagement in a motion discrimination task. *Proc Natl Acad Sci U S A*, 107(50), 21842–7 (2010)
DOI: 10.1073/pnas.1009956107

81. R. L. Goris, J. A. Movshon and E. P. Simoncelli: Partitioning neuronal variability. *Nat Neurosci*, 17(6), 858–65 (2014)
DOI: 10.1038/nn.3711

82. E. Zohary, M. N. Shadlen and W. T. Newsome: Correlated neuronal discharge rate and its implications for psychophysical performance. *Nature*, 370(6485), 140–3 (1994)
DOI: 10.1038/370140a0

83. M. R. Cohen and A. Kohn: Measuring and interpreting neuronal correlations. *Nat Neurosci*, 14(7), 811–9 (2011)
DOI: 10.1038/nn.2842

84. J. Poort and P. R. Roelfsema: Noise correlations have little influence on the coding of selective attention in area V1. *Cereb Cortex*, 19(3), 543–53 (2009)
DOI: 10.1093/cercor/bhn103

85. W. Bair, E. Zohary and W. T. Newsome: Correlated firing in macaque visual area MT: time scales and relationship to behavior. *J Neurosci*, 21(5), 1676–97 (2001)

86. M. L. Leavitt, F. Pieper, A. Sachs, R. Joober and J. C. Martinez-Trujillo: Structure of spike count correlations reveals functional interactions between neurons in dorsolateral prefrontal cortex area 8a of behaving primates. *PLoS One*, 8(4), e61503 (2013)
DOI: 10.1371/journal.pone.0061503

87. J. L. Herrero, M. A. Gieselmann, M. Sanayei and A. Thiele: Attention-induced variance and noise correlation reduction in macaque V1 is mediated by NMDA receptors. *Neuron*, 78(4), 729–39 (2013)
DOI: 10.1016/j.neuron.2013.03.029

88. D. A. Ruff and M. R. Cohen: Attention Increases Spike Count Correlations between Visual Cortical Areas. *J Neurosci*, 36(28), 7523–34 (2016)
DOI: 10.1523/JNEUROSCI.0610-16.2016

89. M. R. Cohen and J. H. Maunsell: Attention improves performance primarily by reducing interneuronal correlations. *Nat Neurosci*, 12(12), 1594–600 (2009)
DOI: 10.1038/nn.2439

90. J. F. Mitchell, K. A. Sundberg and J. H. Reynolds: Spatial attention decorrelates intrinsic activity fluctuations in macaque area V4. *Neuron*, 63(6), 879–88 (2009)
DOI: 10.1016/j.neuron.2009.09.013

91. A. Zenon and R. J. Krauzlis: Attention deficits without cortical neuronal deficits. *Nature*, 489(7416), 434-U124 (2012)
DOI: 10.1038/nature11497

92. E. Astrand, C. Wardak, P. Baraduc and S. Ben Hamed: Direct Two-Dimensional Access to the Spatial Location of Covert Attention in Macaque Prefrontal Cortex. *Curr Biol*, 26(13), 1699–704 (2016)
DOI: 10.1016/j.cub.2016.04.054

93. N. C. Rabinowitz, R. L. Goris, M. Cohen and E. P. Simoncelli: Attention stabilizes the shared gain of V4 populations. *Elife*, 4, e08998 (2015)
DOI: 10.7554/eLife.08998

94. R. Yuste: From the neuron doctrine to neural networks. *Nat Rev Neurosci*, 16(8), 487–97 (2015)
DOI: 10.1038/nrn3962

95. R. O. Duda, P. E. Hart and D. G. Stork: Pattern classification. Wiley, New York (2001)

96. M. Jazayeri and J. A. Movshon: Optimal representation of sensory information by neural populations. *Nat Neurosci*, 9(5), 690–6 (2006)
DOI: 10.1038/nn1691

97. E. Astrand, P. Enel, G. Ibos, P. F. Dominey, P. Baraduc and S. Ben Hamed: Comparison of classifiers for decoding sensory and cognitive information from prefrontal neuronal populations. *PLoS One*, 9(1), e86314 (2014)
DOI: 10.1371/journal.pone.0086314

98. R. Quian Quiroga and S. Panzeri: Extracting information from neuronal populations: information theory and decoding approaches. *Nat Rev Neurosci*, 10(3), 173–85 (2009)
DOI: 10.1038/nrn2578

99. I. Kanitscheider, R. Coen-Cagli, A. Kohn and A. Pouget: Measuring Fisher information accurately in correlated neural populations. *PLoS Comput Biol*, 11(6), e1004218 (2015)
DOI: 10.1371/journal.pcbi.1004218

100. S. Panzeri, R. Senatore, M. A. Montemurro and R. S. Petersen: Correcting for the sampling bias problem in spike train information measures. *J Neurophysiol*, 98(3), 1064–72 (2007)
DOI: 10.1152/jn.00559.2007

101. J. M. Carmena, M. A. Lebedev, R. E. Crist, J. E. O'Doherty, D. M. Santucci, D. F. Dimitrov, P. G. Patil, C. S. Henriquez and M. A. Nicolelis: Learning to control a

brain-machine interface for reaching and grasping by primates. *PLoS Biol*, 1(2), E42 (2003)
DOI: 10.1371/journal.pbio.0000042

102. N. Hatsopoulos, J. Joshi and J. G. O'Leary: Decoding continuous and discrete motor behaviors using motor and premotor cortical ensembles. *J Neurophysiol*, 92(2), 1165–74 (2004)
DOI: 10.1152/jn.01245.2003

103. S. Tremblay, F. Pieper, A. Sachs and J. Martinez-Trujillo: Attentional filtering of visual information by neuronal ensembles in the primate lateral prefrontal cortex. *Neuron*, 85(1), 202–15 (2015)
DOI: 10.1016/j.neuron.2014.11.021

104. S. Tremblay, G. Doucet, F. Pieper, A. Sachs and J. Martinez-Trujillo: Single-Trial Decoding of Visual Attention from Local Field Potentials in the Primate Lateral Prefrontal Cortex Is Frequency-Dependent. *J Neurosci*, 35(24), 9038–49 (2015)
DOI: 10.1523/JNEUROSCI.1041-15.2015

105. K. M. Armstrong, M. H. Chang and T. Moore: Selection and maintenance of spatial information by frontal eye field neurons. *J Neurosci*, 29(50), 15621–9 (2009)
DOI: 10.1523/JNEUROSCI.4465-09.2009

106. S. Farbod Kia, E. Astrand, G. Ibos and S. Ben Hamed: Readout of the intrinsic and extrinsic properties of a stimulus from un-experienced neuronal activities: towards cognitive neuroprostheses. *J Physiol Paris*, 105(1–3), 115–22 (2011)
DOI: 10.1016/j.jphysparis.2011.07.015

107. P. Sapountzis, S. Anastasakis, R. Desimone and G. G. Gregoriou: Decoding covert attention from simultaneous recordings in prefrontal and visual cortex. In: *AREADNE Research in Encoding and Decoding of Neural Ensembles*. Ed N. G. Hatsopoulos & J. S. Pezaris. Santorini, Greece (2014)

108. M. Esghaei and M. R. Daliri: Decoding of visual attention from LFP signals of macaque MT. *PLoS One*, 9(6), e100381 (2014)
DOI: 10.1371/journal.pone.0100381

109. Y. Zhang, E. M. Meyers, N. P. Bichot, T. Serre, T. A. Poggio and R. Desimone: Object decoding with attention in inferior temporal cortex. *Proc Natl Acad Sci U S A*, 108(21), 8850–5 (2011)
DOI: 10.1073/pnas.1100999108

110. J. Jonides: Voluntary Versus Automatic Control Over the Mind's Eye's Movement. In: *Attention and Performance IX*. Ed J. B. Long&A. D. Baddely. Lawrence Erlbaum Associates Hillsdale, NJ (1981)

111. D. Raposo, M. T. Kaufman and A. K. Churchland: A category-free neural population supports evolving demands during decision-making. *Nat Neurosci*, 17(12), 1784–92 (2014)
DOI: 10.1038/nn.3865

112. E. M. Meyers, X. L. Qi and C. Constantinidis: Incorporation of new information into prefrontal cortical activity after learning working memory tasks. *P Natl Acad Sci U S A*, 109(12), 4651–4656 (2012)
DOI: 10.1073/pnas.1201022109

113. M. G. Stokes, M. Kusunoki, N. Sigala, H. Nili, D. Gaffan and J. Duncan: Dynamic coding for cognitive control in prefrontal cortex. *Neuron*, 78(2), 364–75 (2013)
DOI: 10.1016/j.neuron.2013.01.039

114. E. L. Rich and J. D. Wallis: Decoding subjective decisions from orbitofrontal cortex. *Nat Neurosci*, 19(7), 973–80 (2016)
DOI: 10.1038/nn.4320

115. J. Duncan: An adaptive coding model of neural function in prefrontal cortex. *Nat Rev Neurosci*, 2(11), 820–9 (2001)
DOI: 10.1038/35097575

116. S. Fusi, E. K. Miller and M. Rigotti: Why neurons mix: high dimensionality for higher cognition. *Curr Opin Neurobiol*, 37, 66–74 (2016)
DOI: 10.1016/j.conb.2016.01.010

117. D. V. Buonomano and W. Maass: State-dependent computations: spatiotemporal processing in cortical networks. *Nat Rev Neurosci*, 10(2), 113–25 (2009)
DOI: 10.1038/nrn2558

118. R. S. Zucker and W. G. Regehr: Short-term synaptic plasticity. *Annu Rev Physiol*, 64, 355–405 (2002)
DOI: 10.1146/annurev.physiol.64.092501.114547

119. D. A. Crowe, B. B. Averbeck and M. V. Chafee: Rapid sequences of population

activity patterns dynamically encode task-critical spatial information in parietal cortex. *J Neurosci*, 30(35), 11640–53 (2010)
DOI: 10.1523/JNEUROSCI.0954-10.2010

120. E. M. Meyers, D. J. Freedman, G. Kreiman, E. K. Miller and T. Poggio: Dynamic population coding of category information in inferior temporal and prefrontal cortex. *J Neurophysiol*, 100(3), 1407–19 (2008)
DOI: 10.1152/jn.90248.2008

121. E. Astrand, G. Ibos, J. R. Duhamel and S. Ben Hamed: Differential dynamics of spatial attention, position, and color coding within the parietofrontal network. *J Neurosci*, 35(7), 3174–89 (2015)
DOI: 10.1523/JNEUROSCI.2370-14.2015

122. J. D. Murray, A. Bernacchia, N. A. Roy, C. Constantinidis, R. Romo and X. J. Wang: Stable population coding for working memory coexists with heterogeneous neural dynamics in prefrontal cortex. *Proc Natl Acad Sci U S A*, 114(2), 394–399 (2017)
DOI: 10.1073/pnas.1619449114

123. A. S. Morcos and C. D. Harvey: History-dependent variability in population dynamics during evidence accumulation in cortex. *Nat Neurosci*, 19(12), 1672–1681 (2016)
DOI: 10.1038/nn.4403

124. B. A. Olshausen and D. J. Field: Sparse coding of sensory inputs. *Curr Opin Neurobiol*, 14(4), 481–487 (2004)
DOI: 10.1016/j.conb.2004.07.007

125. A. B. Graf, A. Kohn, M. Jazayeri and J. A. Movshon: Decoding the activity of neuronal populations in macaque primary visual cortex. *Nat Neurosci*, 14(2), 239–45 (2011)
DOI: 10.1038/nn.2733

126. A. B. Graf and R. A. Andersen: Inferring eye position from populations of lateral intraparietal neurons. *Elife*, 3, e02813 (2014)
DOI: 10.7554/eLife.02813

127. C. B. Boulay, F. Pieper, M. Leavitt, J. Martinez-Trujillo and A. J. Sachs: Single-trial decoding of intended eye movement goals from lateral prefrontal cortex neural ensembles. *J Neurophysiol*, 115(1), 486–99 (2016)
DOI: 10.1152/jn.00788.2015

128. D. A. Markowitz, Y. T. Wong, C. M. Gray and B. Pesaran: Optimizing the decoding of movement goals from local field potentials in macaque cortex. *J Neurosci*, 31(50), 18412–22 (2011)
DOI: 10.1523/JNEUROSCI.4165-11.2011

129. S. B. Laughlin and T. J. Sejnowski: Communication in neuronal networks. *Science*, 301(5641), 1870–4 (2003)
DOI: 10.1126/science.1089662

130. M. Rigotti, O. Barak, M. R. Warden, X. J. Wang, N. D. Daw, E. K. Miller and S. Fusi: The importance of mixed selectivity in complex cognitive tasks. *Nature*, 497(7451), 585–90 (2013)
DOI: 10.1038/nature12160

131. R. C. deCharms and A. Zador: Neural representation and the cortical code. *Annu Rev Neurosci*, 23, 613–47 (2000)
DOI: 10.1146/annurev.neuro.23.1.613

132. C. P. Hung, G. Kreiman, T. Poggio and J. J. DiCarlo: Fast readout of object identity from macaque inferior temporal cortex. *Science*, 310(5749), 863–6 (2005)
DOI: 10.1126/science.1117593

133. S. Panzeri, N. Brunel, N. K. Logothetis and C. Kayser: Sensory neural codes using multiplexed temporal scales. *Trends Neurosci*, 33(3), 111–20 (2010)
DOI: 10.1016/j.tins.2009.12.001

134. A. Kohn, R. Coen-Cagli, I. Kanitscheider and A. Pouget: Correlations and Neuronal Population Information. *Annu Rev Neurosci*, 39, 237–56 (2016)
DOI:10.1146/annurev-neuro-070815-013851

135. K. H. Britten, W. T. Newsome, M. N. Shadlen, S. Celebrini and J. A. Movshon: A relationship between behavioral choice and the visual responses of neurons in macaque MT. *Visual Neuroscience*, 13(1), 87–100 (1996)
DOI: 10.1017/S095252380000715X

136. L. F. Abbott and P. Dayan: The effect of correlated variability on the accuracy of a population code. *Neural Comput*, 11(1), 91–101 (1999)
DOI: 10.1162/089976699300016827

137. A. S. Ecker, P. Berens, A. S. Tolias and M. Bethge: The effect of noise correlations in populations of diversely tuned neurons. *J Neurosci*, 31(40), 14272–83 (2011)
DOI: 10.1523/JNEUROSCI.2539-11.2011

138. R. Moreno-Bote, J. Beck, I. Kanitscheider, X. Pitkow, P. Latham and A. Pouget: Information-limiting correlations. *Nat Neurosci*, 17(10), 1410–7 (2014)
DOI: 10.1038/nn.3807

139. A. S. Ecker, G. H. Denfield, M. Bethge and A. S. Tolias: On the Structure of Neuronal Population Activity under Fluctuations in Attentional State. *J Neurosci*, 36(5), 1775–89 (2016)
DOI: 10.1523/JNEUROSCI.2044-15.2016

140. B. B. Averbeck, P. E. Latham and A. Pouget: Neural correlations, population coding and computation. *Nat Rev Neurosci*, 7(5), 358–66 (2006)
DOI: 10.1038/nrn1888

141. B. B. Averbeck and D. Lee: Effects of noise correlations on information encoding and decoding. *J Neurophysiol*, 95(6), 3633–44 (2006)
DOI: 10.1152/jn.00919.2005

142. A. B. Graf and R. A. Andersen: Predicting oculomotor behaviour from correlated populations of posterior parietal neurons. *Nat Commun*, 6, 6024 (2015)
DOI: 10.1038/ncomms7024

143. P. Berens, A. S. Ecker, R. J. Cotton, W. J. Ma, M. Bethge and A. S. Tolias: A fast and simple population code for orientation in primate V1. *J Neurosci*, 32(31), 10618–26 (2012)
DOI: 10.1523/JNEUROSCI.1335-12.2012

144. D. Nikolic, S. Hausler, W. Singer and W. Maass: Distributed fading memory for stimulus properties in the primary visual cortex. *PLoS Biol*, 7(12), e1000260 (2009)
DOI: 10.1371/journal.pbio.1000260

145. J. S. McDonald, C. W. Clifford, S. S. Solomon, S. C. Chen and S. G. Solomon: Integration and segregation of multiple motion signals by neurons in area MT of primate. *J Neurophysiol*, 111(2), 369–78 (2014)
DOI: 10.1152/jn.00254.2013

146. C. Mehring, J. Rickert, E. Vaadia, S. Cardosa de Oliveira, A. Aertsen and S. Rotter: Inference of hand movements from local field potentials in monkey motor cortex. *Nat Neurosci*, 6(12), 1253–4 (2003)
DOI: 10.1038/nn1158

147. E. J. Hwang and R. A. Andersen: The utility of multichannel local field potentials for brain-machine interfaces. *J Neural Eng*, 10(4), 046005 (2013)
DOI: 10.1088/1741-2560/10/4/046005

148. J. Poort, A. Pooresmaeili and P. R. Roelfsema: Multi-neuron representations of visual attention. In: *Understanding visual population codes: towards a common multivariate framework for cell recording and functional imaging*. Ed N. Kriegeskorte&G. Kreiman. MIT Press, (2012)

149. G. Buzsaki, C. A. Anastassiou and C. Koch: The origin of extracellular fields and currents--EEG, ECoG, LFP and spikes. *Nat Rev Neurosci*, 13(6), 407–20 (2012)
DOI: 10.1038/nrn3241

150. U. Mitzdorf: Current source-density method and application in cat cerebral cortex: investigation of evoked potentials and EEG phenomena. *Physiol Rev*, 65(1), 37–100 (1985)

151. N. K. Logothetis: What we can do and what we cannot do with fMRI. *Nature*, 453(7197), 869–78 (2008)
DOI: 10.1038/nature06976

152. N. K. Logothetis: The underpinnings of the BOLD functional magnetic resonance imaging signal. *J Neurosci*, 23(10), 3963–71 (2003)

153. M. J. Bartolo, M. A. Gieselmann, V. Vuksanovic, D. Hunter, L. Sun, X. Chen, L. S. Delicato and A. Thiele: Stimulus-induced dissociation of neuronal firing rates and local field potential gamma power and its relationship to the resonance blood oxygen level-dependent signal in macaque primary visual cortex. *Eur J Neurosci*, 34(11), 1857–70 (2011)
DOI: 10.1111/j.1460-9568.2011.07877.x

154. P. Sapountzis, D. Schluppeck, R. Bowtell and J. W. Peirce: A comparison of fMRI adaptation and multivariate pattern classification analysis in visual cortex. *Neuroimage*, 49(2), 1632–40 (2010)
DOI: 10.1016/j.neuroimage.2009.09.066

155. J. D. Haynes and G. Rees: Decoding mental states from brain activity in humans. *Nat Rev Neurosci*, 7(7), 523–34 (2006)
DOI: 10.1038/nrn1931

156. A. Belitski, S. Panzeri, C. Magri, N. K. Logothetis and C. Kayser: Sensory information in local field potentials and spikes from visual and auditory cortices:

time scales and frequency bands. *J Comput Neurosci*, 29(3), 533–45 (2010)
DOI: 10.1007/s10827-010-0230-y

157. J. Rickert, S. C. Oliveira, E. Vaadia, A. Aertsen, S. Rotter and C. Mehring: Encoding of movement direction in different frequency ranges of motor cortical local field potentials. *J Neurosci*, 25(39), 8815–24 (2005)
DOI: 10.1523/JNEUROSCI.0816-05.2005

158. D. Rotermund, U. A. Ernst, S. Mandon, K. Taylor, Y. Smiyukha, A. K. Kreiter and K. R. Pawelzik: Toward high performance, weakly invasive brain computer interfaces using selective visual attention. *J Neurosci*, 33(14), 6001–11 (2013)
DOI: 10.1523/JNEUROSCI.4225-12.2013

159. S. Ray and J. H. Maunsell: Different origins of gamma rhythm and high-gamma activity in macaque visual cortex. *PLoS Biol*, 9(4), e1000610 (2011)
DOI: 10.1371/journal.pbio.1000610

160. T. P. Zanos, P. J. Mineault and C. C. Pack: Removal of spurious correlations between spikes and local field potentials. *J Neurophysiol*, 105(1), 474–86 (2011)
DOI: 10.1152/jn.00642.2010

161. U. Mitzdorf: Properties of the evoked potential generators: current source-density analysis of visually evoked potentials in the cat cortex. *Int J Neurosci*, 33(1–2), 33–59 (1987)
DOI: 10.3109/00207458708985928

162. D. A. Kaliukhovich and R. Vogels: Decoding of repeated objects from local field potentials in macaque inferior temporal cortex. *PLoS One*, 8(9), e74665 (2013)
DOI: 10.1371/journal.pone.0074665

163. H. Scherberger, M. R. Jarvis and R. A. Andersen: Cortical local field potential encodes movement intentions in the posterior parietal cortex. *Neuron*, 46(2), 347–54 (2005)
DOI: 10.1016/j.neuron.2005.03.004

164. B. Pesaran, J. S. Pezaris, M. Sahani, P. P. Mitra and R. A. Andersen: Temporal structure in neuronal activity during working memory in macaque parietal cortex. *Nat Neurosci*, 5(8), 805–11 (2002)
DOI: 10.1038/nn890

165. R. Biran, D. C. Martin and P. A. Tresco: Neuronal cell loss accompanies the brain tissue response to chronically implanted silicon microelectrode arrays. *Exp Neurol*, 195(1), 115–26 (2005)
DOI: 10.1016/j.expneurol.2005.04.020

166. C. A. Chestek, V. Gilja, P. Nuyujukian, J. D. Foster, J. M. Fan, M. T. Kaufman, M. M. Churchland, Z. Rivera-Alvidrez, J. P. Cunningham, S. I. Ryu and K. V. Shenoy: Long-term stability of neural prosthetic control signals from silicon cortical arrays in rhesus macaque motor cortex. *J Neural Eng*, 8(4), 045005 (2011)
DOI: 10.1088/1741-2560/8/4/045005

167. J. A. Perge, S. Zhang, W. Q. Malik, M. L. Homer, S. Cash, G. Friehs, E. N. Eskandar, J. P. Donoghue and L. R. Hochberg: Reliability of directional information in unsorted spikes and local field potentials recorded in human motor cortex. *J Neural Eng*, 11(4), 046007 (2014)
DOI: 10.1088/1741-2560/11/4/046007

168. D. Wang, Q. Zhang, Y. Li, Y. Wang, J. Zhu, S. Zhang and X. Zheng: Long-term decoding stability of local field potentials from silicon arrays in primate motor cortex during a 2D center out task. *J Neural Eng*, 11(3), 036009 (2014)
DOI: 10.1088/1741-2560/11/3/036009

169. M. D. Serruya, N. G. Hatsopoulos, L. Paninski, M. R. Fellows and J. P. Donoghue: Instant neural control of a movement signal. *Nature*, 416(6877), 141–2 (2002)
DOI: 10.1038/416141a

170. L. R. Hochberg, D. Bacher, B. Jarosiewicz, N. Y. Masse, J. D. Simeral, J. Vogel, S. Haddadin, J. Liu, S. S. Cash, P. van der Smagt and J. P. Donoghue: Reach and grasp by people with tetraplegia using a neurally controlled robotic arm. *Nature*, 485(7398), 372–5 (2012)
DOI: 10.1038/nature11076

171. L. R. Hochberg, M. D. Serruya, G. M. Friehs, J. A. Mukand, M. Saleh, A. H. Caplan, A. Branner, D. Chen, R. D. Penn and J. P. Donoghue: Neuronal ensemble control of prosthetic devices by a human with tetraplegia. *Nature*, 442(7099), 164–71 (2006)
DOI: 10.1038/nature04970

172. E. Astrand, C. Wardak and S. Ben Hamed: Selective visual attention to drive cognitive brain-machine interfaces: from concepts

to neurofeedback and rehabilitation applications. *Front Syst Neurosci*, 8, 144 (2014)
DOI: 10.3389/fnsys.2014.00144

173. P. Nuyujukian, J. C. Kao, S. I. Ryu and K. V. Shenoy: A Nonhuman Primate Brain–Computer Typing Interface. *Proceedings of the IEEE*, 105(1), 66–72 (2017)
DOI: 10.1109/JPROC.2016.2586967

174. B. Jarosiewicz, A. A. Sarma, D. Bacher, N. Y. Masse, J. D. Simeral, B. Sorice, E. M. Oakley, C. Blabe, C. Pandarinath, V. Gilja, S. S. Cash, E. N. Eskandar, G. Friehs, J. M. Henderson, K. V. Shenoy, J. P. Donoghue and L. R. Hochberg: Virtual typing by people with tetraplegia using a self-calibrating intracortical brain-computer interface. *Sci Transl Med*, 7(313), 313ra179 (2015)
DOI: 10.1126/scitranslmed.aac7328

175. F. Bremmer, A. Kaminiarz, S. Klingenhoefer and J. Churan: Decoding Target Distance and Saccade Amplitude from Population Activity in the Macaque Lateral Intraparietal Area (LIP). *Front Integr Neurosci*, 10, 30 (2016)
DOI: 10.3389/fnint.2016.00030

176. S. Ohmae, T. Takahashi, X. Lu, Y. Nishimori, Y. Kodaka, I. Takashima and S. Kitazawa: Decoding the timing and target locations of saccadic eye movements from neuronal activity in macaque oculomotor areas. *J Neural Eng*, 12(3), 036014 (2015)
DOI: 10.1088/1741-2560/12/3/036014

177. A. B. Graf and R. A. Andersen: Brain-machine interface for eye movements. *Proc Natl Acad Sci U S A*, 111(49), 17630–5 (2014)
DOI: 10.1073/pnas.1419977111

178. M. S. Treder and B. Blankertz: (C)overt attention and visual speller design in an ERP-based brain-computer interface. *Behavioral and Brain Functions*, 6 (2010)
DOI: 10.1186/1744-9081-6-28

179. A. S. Dickey, A. Suminski, Y. Amit and N. G. Hatsopoulos: Single-unit stability using chronically implanted multielectrode arrays. *J Neurophysiol*, 102(2), 1331–9 (2009)
DOI: 10.1152/jn.90920.2008

180. G. W. Fraser, S. M. Chase, A. Whitford and A. B. Schwartz: Control of a brain-computer interface without spike sorting. *J Neural Eng*, 6(5), 055004 (2009)
DOI: 10.1088/1741-2560/6/5/055004

181. S. Todorova, P. Sadtler, A. Batista, S. Chase and V. Ventura: To sort or not to sort: the impact of spike-sorting on neural decoding performance. *J Neural Eng*, 11(5), 056005 (2014)
DOI: 10.1088/1741-2560/11/5/056005

182. G. Santhanam, S. I. Ryu, B. M. Yu, A. Afshar and K. V. Shenoy: A high-performance brain-computer interface. *Nature*, 442(7099), 195–8 (2006)
DOI: 10.1038/nature04968

183. E. Calabrese, A. Badea, C. L. Coe, G. R. Lubach, Y. Shi, M. A. Styner and G. A. Johnson: A diffusion tensor MRI atlas of the postmortem rhesus macaque brain. *Neuroimage*, 117, 408–16 (2015)
DOI: 10.1016/j.neuroimage.2015.05.072

184. R. Bakker, P. Tiesinga and R. Kotter: The Scalable Brain Atlas: Instant Web-Based Access to Public Brain Atlases and Related Content. *Neuroinformatics*, 13(3), 353–66 (2015)
DOI: 10.1007/s12021-014-9258-x

Send correspondence to: Panagiotis Sapountzis, Foundation for Research and Technology Hellas, Institute of Applied and Computational Mathematics, N. Plastira 100, GR 70013, Heraklion Crete, Greece. Tel: 30-2810-394857, Fax: 30-2810-394840, E-mail: pasapoyn@iacm.forth.gr