Attentional Scanning in Macaque Area V4

A Thesis

submitted to

Indian Institute of Science Education and Research Pune in partial fulfillment of the requirements for the BS-MS Dual Degree Programme

by

Shivangi Patel

Indian Institute of Science Education and Research Pune
Dr. Homi Bhabha Road,
Pashan, Pune 411008, INDIA.

April 2022

Supervisor: Dr. Pascal Fries

Expert: Dr. Assisi Collins

Certificate

This is to certify that this dissertation entitled Attentional Scanning in Macaque Area V4 towards the partial fulfillment of the BS-MS dual degree program at the Indian Institute of Science Education and Research, Pune represents the work carried out by Shivangi Patel at Ernst Strüngmann Institute (ESI) for Neuroscience in Cooperation with Max Planck Society, Frankfurt under the supervision of Dr. Pascal Fries, Director, during the academic year 2021-2022.

Shivangi Patel

Dr. Pascal Fries

Committee:

Supervisor: Dr. Pascal Fries (ESI)

Expert: Dr. Assisi Collins (IISER Pune)

Declaration

I hereby declare that the matter embodied in the report entitled Attentional Scanning in Macaque Area V4 are the results of the work carried out by me at the Fries lab, Ernst Strüngmann Institute for Neuroscience in Cooperation with Max Planck Society, under the supervision of Dr. Pascal Fries, and the same has not been submitted elsewhere for any other degree.

Date: 30/03/2022 **Name:** Shivangi Patel

Table of Content

Abstract

Sensitivity to perception varies on a moment-to-moment basis. It has been shown to be dependent on the phase of spontaneous ongoing oscillations in the low frequency (alpha and theta) bands. When inputs arrive at this optimal phase, the representation of the stimulus is stronger which leads to beneficial processing. Studies have also revealed the existence of traveling waves in the monkey visual cortex and that covert visual attention samples stimuli in a theta-rhythmic way. We hypothesized that attention can scan different locations of an attended object which could lead to beneficial processing of different locations over time. This means the ‘optimal phase’ of neighboring neuronal populations will be systematically shifted in cortical space. In this project, we focused on looking at whether an optimal phase for processing exists by analyzing correlations between the amplitude of evoked response and the prestimulus phase. We observed that, indeed, perception is associated with the prestimulus phase. We also developed and tested a method for phase estimation in the process. Developing this method was crucial in order to be able to estimate the phase reasonably well while dealing with the group delay problem. The role of traveling waves in attentional scanning can be explored in future analysis. It will help us understand the possible mechanisms by which attention scans different attended items (locations or objects or features).

Acknowledgments

I have received a great deal of assistance throughout the writing of this thesis.

I would like to thank my supervisor Dr. Pascal Fries, whose expertise was essential in developing the research question and methodology. I am very grateful for the opportunity to work on such an interesting project. It was a great learning experience and would be invaluable in furthering my research career.

The data I analyzed was recorded by Eleni Psarou (Ph.D. student, Fries lab). I would like to thank her for providing the data, previously done analysis, and figures. I would like to show my gratitude to Elena for her constant guidance, motivation, and patient support.

I would like to acknowledge members of Fries lab for their wonderful collaboration and discussion points. I would like to thank Dr. Assisi Collins for his support. Thank you all for providing me with the tools and assistance that I needed to successfully complete my dissertation.

1. Introduction

Thousands of signals are received by the brain at any moment, but only a few of them are processed. This cognitive ability to give preference to behaviorally relevant stimuli in order to avoid an overload of information processing has been defined as selective attention. It is the selection process that determines which information receives further processing; it can enhance neuronal processing towards behaviorally relevant items (object, location, feature, etc.). It can sample and process multiple items that are presented simultaneously and bias the neuronal processing towards behaviorally relevant ones.

This sampling and selection process is aided by the anatomical hierarchy of the primate visual cortex. The design of the visual system is such that inputs from lower areas of the visual cortex converge to higher visual areas. This means that the higher areas can receive multiple representations, but only some are behaviorally relevant. The competition between these multiple representations is solved by selective attention, where behaviorally relevant stimuli are attended and their representation is processed further. Numerous mathematical models have been developed to explore neuronal mechanisms underlying this selection process which we call selective attention. To simplify, let’s say two neurons with small receptive fields (in the lower visual areas) feed into a neuron in the higher visual area such that the small receptive fields in the lower areas combine to form the receptive field of the neuron in the higher area which can select for complex features (Lund et al., 2003; Salin et al., 1992). The larger receptive fields can contain multiple stimuli, which would have been in different small receptive fields in the lower

visual areas. The effects of directing attention to one of these stimuli can be modeled by increasing the gain of inputs from the neuron in the lower area (that represents the attended stimulus) to the neuron in the higher area (Reynolds and Heeger, 2009; Reynolds et al., 1999; Womelsdorf et al., 2008).

The increase in input gain between lower and higher area neurons can be implemented by neuronal gamma-band (40-90 Hz) synchronization. Ongoing oscillations can modulate the electrical field and excitability of local and long-range neuronal populations; because of this, they might have a role in modifying temporal aspects of information (Buzsáki and Draguhn, 2004; Fries et al., 2007; Sirota et al., 2008). Attention changes the temporal alignment of oscillations to increase synchronization within and between neural ensembles. Attentional effects can be related to ongoing oscillations as well. Studies have shown increased gamma-band synchronization in macaque V4 during attention (Bichot et al., 2005; Fries et al., 2001; Taylor et al., 2005). It is also shown that variation in local gamma-band synchronization is linked to better behavioral performance and shorter reaction time (Womelsdorf et al., 2006).

The gamma-band synchronization makes the connectivity better by timing the arrival of inputs from lower visual areas to higher visual areas such that they arrive at the time of maximal gain or the high-gain phase. This means that the inputs have to arrive at the higher area neuron (refer to figure 1) at a particular phase of the gamma oscillation to benefit from this enhanced gain. It has been shown that the gain of spike responses to visual stimuli is rhythmically modulated by gamma-band activity. This has a direct consequence on the behavior too. A study by Ni et al. in 2016 shows that the gamma phase having maximum firing rate response and shortest behavioral reaction time are the same. It was also found that optogenetically induced gamma rhythmically modulates the gain of spike responses in area 21a of anesthetized cat (Ni et al., 2016) and rodent somatosensory cortex (Cardin et al., 2009; Siegle et al., 2014). Therefore, several studies have shown that gamma rhythmically modulates neuronal gain. Modeling studies have also shown that gamma-band synchronization has a role in gain modulation such that when input arrives at the phase of maximal gain (called the “window of opportunity”), connectivity becomes better or a more excitatory response is elicited (Börgers and Kopell, 2008). It has been shown that behavioral performance improves when gamma synchronization

between lower and higher visual areas occurs at the optimal phase relation (Rohenkohl et al., 2018). In this study, macaques performed selective attention tasks while V1 and V4 were recorded simultaneously. They show that shorter reaction time to an unpredictable stimulus change was associated with an increase in gamma-band synchronization between V1 and V4, and this synchronization occurs at the phase that is optimal for transmission of stimulus to response. The interareal gamma phase relation can predict reaction time which means stimulus transmission is dependent on this phase.

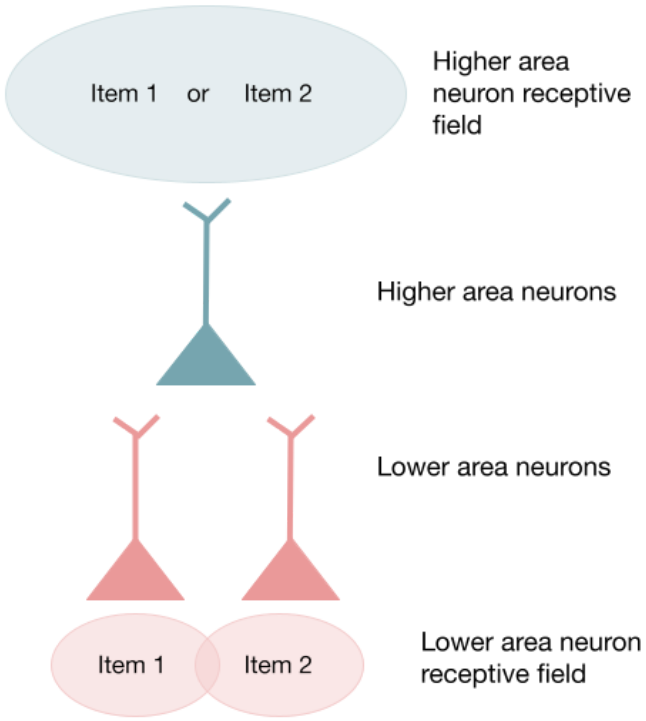

Figure 1. Schematic representation of convergence in anatomical hierarchy in the primate visual system. The small circles at the bottom correspond to the receptive fields of the lower area neurons, and the circle at the top corresponds to the receptive field of higher area neuron. Two neurons in the lower visual area give input to a neuron in the higher visual area such that the receptive field of the higher area neuron covers the receptive fields of lower area neurons.

Several studies have shown phase dependent perceptual fluctuations in the low frequency bands. When a stimulus happens to appear at the optimal oscillatory phase, it receives

beneficial processing leading to a stronger representation. This temporal modulation of information processing could lead to perceptual fluctuations with the phase of spontaneous ongoing oscillations. It was observed that trial-to-trial variability in the detection of stimuli in a signal detection task is linked to the prestimulus phase of ongoing neural oscillations (Busch et al., 2009). This study was done on human subjects, and EEG was recorded while the subject did a detection experiment. The stimulus threshold was such that its detection probability was 50%. Phase distributions of hit and miss trials were evaluated, and they were found to have significant phase concentration at opposite phase angles. A strong difference in phase distribution was observed at 7.1 Hz and 120ms preceding stimulus onset in frontocentral channels. Busch and VanRullen (2010) confirmed that perception is related to the prestimulus phase of ongoing oscillation and found that this relation is observed only when the stimuli are attended. This suggests that the fluctuations in attention are reflected in the ongoing 7 Hz neural oscillation. Effects of attention (like improved perception or detection etc.) are associated with this periodic sampling, and with every 7 Hz cycle, a different attended item (object, location, etc.) is indexed (VanRullen, 2013).

The spotlight of attention moves from one item to the other, and this sampling is said to happen at the frequency of around 7 to 8 Hz (VanRullen, 2016). This mechanism might be useful in the exploration of different items, and many psychophysical studies have reported this behavioral periodicity in visual attention. Attention might operate periodically even when only one item is present (Re et al., 2019). Rhythmic attentional sampling has been found to be at 7 Hz when a single stimulus was presented, around 3-4 Hz when attention was divided over two stimuli, and so on (Holcombe and Chen, 2013). The 3-4 Hz sampling of stimuli could reflect a common sampling process that is divided between the stimuli but operates globally. The same system can sample a single item or divide different sampling cycles across multiple items. This indicates that rhythmic sampling might be a mechanism that resolves interactions among competing stimuli when several behaviorally relevant items are presented.

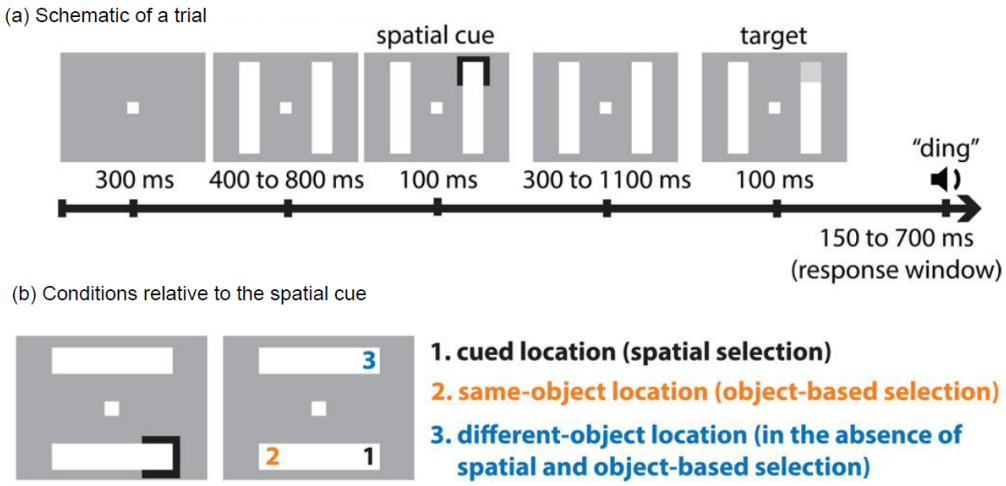

Figure 2. The experimental design of Fiebelkorn et al. (2013) (a) A schematic of a trial in which the target occurs at the cued (spatial cue) location. Behavioral performance was measured at different cue-to-target intervals relative to the spatial cue. (b) The spatial cue and the orientation of the two bars (or objects) determine the conditions. They looked into visual-target detection at a cued location, an uncued location within the same bar, and an uncued location within a different bar. The same- and different-object locations were equidistant from the cued location. (Permission obtained for reusing the figure)

A study by Fiebelkorn et al. (2013) provided more insight into attentional sampling. This study measured behavioral performance at randomly sampled cue-to-target intervals relative to a spatial cue. Human subjects had to maintain central fixation and report a contrast change at one of two bars. In any given trial, two bars were displayed equidistantly from the central fixation such that the four ends of the bar were at equal eccentricities. A spatial cue (brief flash) indicating the behaviorally relevant location occurred at the end of one of the bars (Figure 2). Following a variable time interval from cue presentation, a change in contrast (target) occurred at one of the three locations; – 1) at the cued bar at the cued location (75% of trials). This represented spatial selection condition. 2) at the cued bar but at the uncued location (12.5% of trials). This represented object based selection. 3) at the uncued bar at the location such that same and different object target locations were equidistant from cued location (12.5% of trials). It was seen that the probability of visual-target detection was higher at certain time points (or phases in the behavioral oscillation) than at others. These results are consistent with the attentional sampling at 7-8 Hz. The detection performance fluctuated at 4Hz between the two uncued, and at 8Hz between cued location and the uncued location at the cued bar (with a 30ms

delay). This could suggest that preferential processing spreads from the cued location to the uncued location on the same bar, demonstrating the process of attentional scanning. The attentional scanning could be achieved by traveling waves. It is possible that the flash or spatial cue evokes a traveling wave at the cued location, which travels to the uncued location on the same object. This could explain the preferential processing at the uncued location with a delay of 30 ms from cue onset on the cued edge of the same object (Pascal Fries, personal communication).

Traveling waves could be instrumental in the process of attentional scanning. They could be internally generated or evoked by external stimuli like flashes or saccades. These waves were shown to be triggered by saccadic eye movements in area V4 of the macaque visual cortex (Zanos et al., 2015). The study suggests that the traveling waves can change post-saccadic neuronal processing, which can potentially lead to selectively processing behaviorally relevant items. Another study reporting traveling waves in area MT of marmosets gives evidence towards the role of these waves in attentional scanning. They found that spontaneous traveling waves in area MT modulate sensory processing and gate perception during active vision (Davis et al., 2020).

In summary, studies show that covert visual attention samples stimuli in a theta-rhythmic way (Busch and VanRullen, 2010; Fiebelkorn et al., 2013; Landau and Fries, 2012) and that the detection performance of attended visual stimuli is dependent on the phase of spontaneous ongoing oscillations in the low alpha and theta bands preceding stimulus onset (Busch and VanRullen, 2010; Busch et al., 2009). In other words, there is an alternation between “optimal” and “suboptimal” oscillatory phases during covert attention. When a stimulus happens to appear at the optimal oscillatory phase, it receives a beneficial processing/stronger representation. We hypothesized that attention scans different locations of an attended object, leading to beneficial processing of different locations over time. Recent studies revealed the existence of traveling waves in the visual cortex of monkeys (Davis et al., 2020; Zanos et al., 2015) and humans (Zhang et al., 2018). Attentional scanning could be achieved by traveling waves of oscillatory activity propagating through the visual cortex. Indeed, Davis et al. (2020) showed that spontaneous ongoing traveling waves in marmoset area MT gate visual perception. The main

characteristic of a traveling wave is that its phase varies systematically and approximately monotonically as a function of anatomical space. If an “optimal phase” exists along with a traveling wave in an area with a retinotopic organization as V4, then we would expect that the ‘optimal phase' of neighboring neuronal populations will be systematically shifted in cortical space.

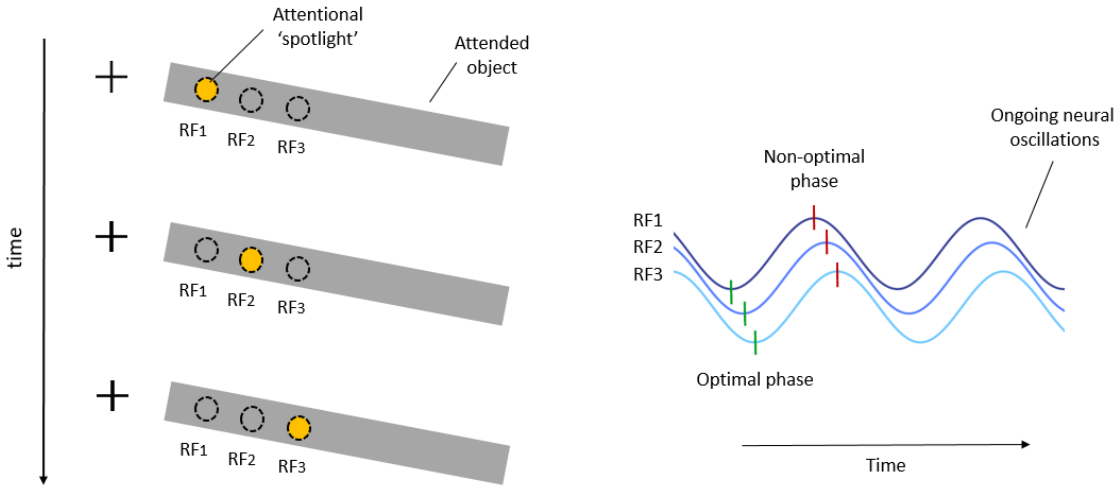

Figure 3. Schematic representation of the attentional-scanning hypothesis. Left: Covert attention as a rhythmic scanning of the visual space. At different time points, the spotlight of attention moves between different spatial locations, leading to beneficial processing of different spatial locations over time. Right: Such an attentional scanning process could be implemented through a traveling wave, which travels in cortical space. The main characteristic of a traveling wave is that its phase varies as a function of anatomical space. Hence, if there is an “optimal” phase and an ongoing traveling wave, we expect this phase to shift in the cortical space as the wave is sweeping the cortex. (Figure provided by Elena Psarou)

This project involves analysis of previously recorded data (by Elena Psarou) from two macaques, chronically implanted with a multichannel Utah array in area V4 while they performed a demanding detection task. This allowed us to simultaneously record the activity of neural populations with neighboring receptive fields while the monkeys attended an object in order to detect a brief target presentation that could randomly appear at different locations within this object. The locations of the target were such that they were covered by the receptive fields of the channels. The stimulus presented was brief (only two monitor update frames) so that we could find the instantaneous phase. If the stimulus were longer, it would be difficult to infer the part of

the cycle at which the stimulus was actually perceived. A difficult target detection task was used so that the monkey would pay attention, and we could have a sensitive measure of perceptual performance fluctuations. Most attention-related work has been done in monkeys because they learn the challenging behavioral tasks needed for attention experiments, and we can monitor their eye movements, which gives information about what they are paying attention to. Attention is studied extensively in the visual system because vision is a primary sensory modality for primates, and they have more neurons devoted to the visual system than any other modality (Felleman and Van Essen, 1991). Recordings were done from area V4 because it is a mid-level processing area in the visual hierarchy, and many studies have shown attentional effects in the area (Luck et al., 1997; Moran and Desimone, 1985). The task design helps us record from neighboring neural populations and inquire about attentional scanning.

The first half of the project revolved around going through literature to get some background, getting familiar with the FieldTrip toolbox, and some initial data analysis. We looked at MUA, LFP, and eye data to establish that behavior is reflected in this data. We compared measures from the neural and eye data in trials with two different behavioral outcomes, namely trials in which the monkey successfully detected the target (hits) with those in which the monkey missed it (misses). We investigated and observed differences in these measures between the two behavioral outcomes. This shows that the difference in detection is reflected in the neural and eye data along with the animal's behavior. The second half of the project revolved around investigating whether the prestimulus LFP phase is systematically predictive of perception. The phase is calculated just prior to the critical point which is the time at which the synaptic inputs arrive at V4. To do so, we needed to determine the phase at different frequencies at the stimulus onset (time at which visual target appears in the task). This turned out to be a complicated task. To find the phase, we needed to calculate the Fourier decomposition of the signal using tapers and then calculate the phase using this Fourier. While calculating the Fourier at the stimulus onset, part of the signal after stimulus onset which contains the ERP (Event Related Potential), is also included. This doesn't give an accurate phase value because the phase will be affected by the ERP, which (as shown in the results section) is dependent on whether a trial is a hit or a miss. To deal with this, we developed and tested a new method based on Ni et al. 2016. In this method, we used autoregressive (AR) extrapolation to get a proxy for the signal after stimulus

onset and then used the signal prior to stimulus onset together with this proxy signal after stimulus onset to calculate the phase at stimulus onset. The details are described in the methods section. After the phase calculation method was established, we tested whether an optimal phase exists by analyzing how the prestimulus phase is related to neural responses reflecting perceptual performance.

It was important to establish the existence of an optimal phase of processing to be able to explore the role of traveling waves in the future. We needed a good and reliable phase estimation method to find the optimal phase. So, the project's primary focus is to establish and test the phase estimation method and discover if an optimal phase for processing exists.

2. Methods

2.1 Experiment

2.1.1 Animals

Two adult monkeys (Macaca mulatta) were implanted with a Utah array in area V4 of the left hemisphere (Figure 4(b)). The array consisted of 64 microelectrodes that could have two different lengths; half of them were 1mm and the other half 0.6 mm long. We analyzed data from one monkey (Monkey K) in this project.

2.1.2 Behavioral Paradigm

Two monkeys were trained to perform a demanding visual detection task that required spatial attention. During the experiments, the monkeys were head-fixed and sat comfortably in a primate chair that was placed in a dark recording booth (faraday cage) that eliminated line noise and attenuated external sounds. All behavioral paradigms were designed and controlled by ARCADE, a stimulus presentation software written in Matlab, and presented on an LCD monitor (Samsung 2233RZ) at a refresh rate of 120 Hz.

Conditions - The visual detection task consisted of three conditions: 1) target, 2) catch, and 3) baseline (Figure 4(a)).

Trials in the ‘target’ and ‘catch’ conditions started with the presentation of a central fixation point that was paired with a placeholder (Figure 4(a)). The placeholder was a gray rectangle, darker than the background gray, indicating the overall area where a target could appear to the monkey. Throughout the trial, the monkey had to maintain central fixation while covertly attending within that area in order to detect and report a brief target presentation. The target presentation lasted only two frames (~17 ms). In 10% of the trials, the target could randomly appear anytime between 1000 to 2000 ms from central fixation onset. In the rest of the trials, the target randomly appeared between 2000 to 3000 ms following fixation onset. The monkey had to saccade towards the memorized target location to report a target detection. When the monkey entered a window of 1 degree around the target location, the target reappeared in order to give visual feedback. The monkey had to maintain his eyes within the target window for 200 ms to receive a reward. Trials in which the monkey failed to detect and report the target presentation were classified as misses and aborted without reward.

The target stimulus consisted of a small Gaussian of gray color. In Monkey K., the target could appear in one of nine predefined locations within the placeholder with eccentricities; 1.7, 2.1, 2.5, 2.9, 3.3, 3.7, 4.1, 4.5, and 4.9 (deg). The target location was pseudo-randomly selected to ensure that all locations were selected equally. In particular, the probability that the target would be presented at one of the possible locations depended on how often this location was chosen in the preceding trials. In a given trial, we counted the number of previous occurrences per location. This vector was padded and smoothed with the Matlab function fastsmooth, with a width of 3 and triangular smooth type (Tom O'Haver (2022), Fast smoothing function). The resulting vector was transformed into a probability distribution over the locations, and this vector of weights was reversed and used for sampling the next location. Finally, the sampling was done using the Matlab function datasample. This resulted in the exclusion of the locations that had the most occurrences and a lower selection chance for the immediately neighboring locations.

In ‘catch’ trials, there was no target presentation, and the monkey was rewarded for maintaining central fixation throughout the trial. Twenty percent of the total trials per session were catch trials

to ensure that the subjects did not randomly saccade towards the placeholder to get a reward but instead reported real target detections.

Five percent of the trials belonged to the ‘baseline’ condition, during which only the central fixation point was presented. During this condition, no spatial attention was required, and the monkey was rewarded at the end of the trial for successful fixation.

Difficulty Levels - The task's difficulty level was manipulated by changing the contrast of the target stimulus to its background (placeholder) throughout the session. During each recording session, we ran an independent staircase procedure per target location in order to maintain a hit rate of ~0.8 within the target condition. To do so, at any given trial, we calculated the detection rate of the subject at the previously presented location, taking into account the behavior of the monkey in the last five ‘target’ trials during which this target location was used. If the hit rate was equal to or higher than 0.8, then the contrast of the target was decreased. If it was smaller than 0.8, its contrast was increased, making the target detection easier.

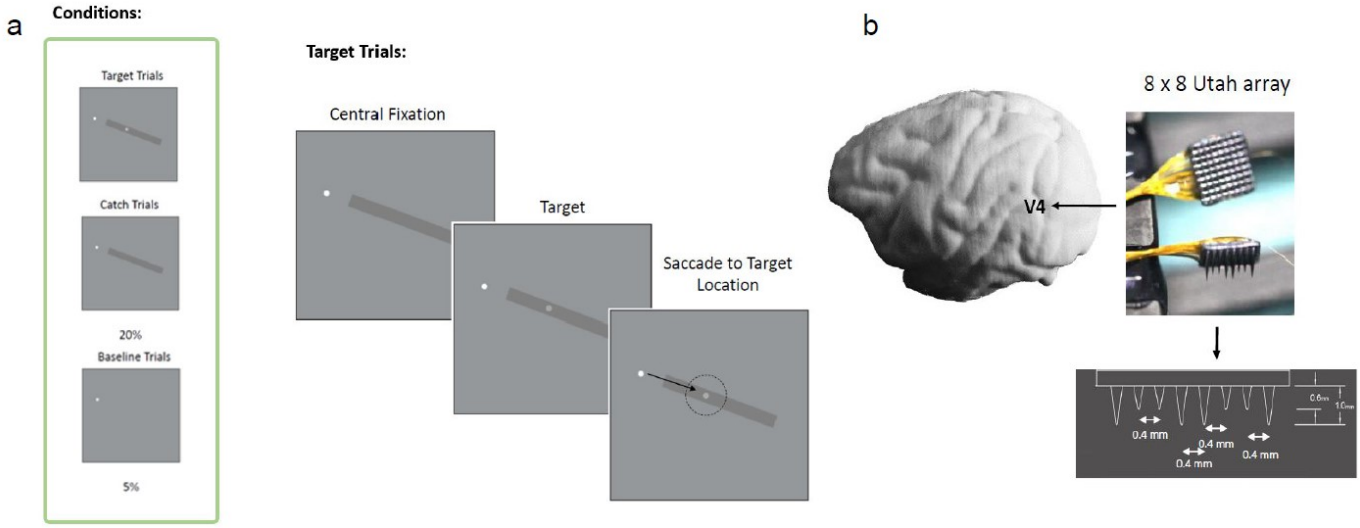

Figure 4. (a) Schematic representation of the behavioral task and conditions. (b) Two monkeys were implanted in left hemisphere V4 with Utah arrays consisting of 64 channels. The inter-electrode distance was 400 μm, and their length was either 0.6 or 1 mm. (Figure provided by Elena Psarou)

2.1.3 Data Acquisition

Electrophysiological recordings were performed using Tucker Davis Technologies (TDT) systems. Data were filtered between 0.35 and 7500 Hz (3 dB filter cutoffs) and digitized at 24.4140625 kHz (TDT PZ2 preamplifier). Eye data (eye movements and pupil size) from both eyes were recorded with an Eyelink 1000 system at a sampling rate of 1000 Hz.

2.2 Analysis methods

2.2.1 MUA and LFP data

LFP and MUA are extracellularly recorded signals from a local network of neurons. The local field potential (LFP) was obtained by low-pass filtering the raw signal at 500 Hz. It is thought to be generated by membrane currents of the neurons in the vicinity of the recording electrode. Multi-unit activity (MUA) represents the average spiking of neuronal populations in the neighborhood of the recording electrode. It was obtained by band-pass filtering the raw recording in the 300-12000 Hz range. FieldTrip (Oostenveld et al., 2011) was used for data analysis. It is an open-source MATLAB software toolbox for MEG, EEG, and iEEG analysis.

Data belonging to the target condition was cut into epochs ranging from -1 to 0.6 sec around target onset. Using the FieldTrip function ft_rejectvisual, we ran an artifact rejection procedure for each recording session. This allowed us to exclude LFP trials that were contaminated with noise (time window: -1 to 0.4 sec from target onset). Channels that showed a high number of artifacts that were not shared with the rest of the grid or very high variance were also excluded from our analysis.

A total of 24 sessions were used in this analysis. Data from each session was z-scored before pooling over sessions. For z-scoring, we concatenated all trials and computed standard deviation (SD) and mean (avg) over all trials of the session.

This SD and mean were used to z-score every single trial as follows:

$$zdata = \frac{(data - avg)}{SD} \quad (1)$$

2.2.1.1 Event Related Potential (ERP)

The visually evoked response in hit and miss trials was computed by averaging trials of each of these behavioral outcomes per channel. This average was normalized by the mean baseline activity (the time period used: -0.4 to 0) across behavioral outcomes (hits and misses).
The difficulty level of the target was varied systematically over trials. To exclude the possibility that a difference in the ERP amplitude between hits and misses could be trivially explained by a discrepancy in the target intensity, we checked the analysis by only including trials from an example target location with the same difficulty level.

2.2.1.2 Time-Frequency analysis

We computed the power values for each frequency bin and each time bin using the mtmconv method of the ft_freqanalysis function of FieldTrip. We used Hann tapers to calculate TFRs (time-frequency representations). Both fixed and variable window length procedures were used. In the fixed window procedure, the frequency resolution is defined according to the length of the time window, which remains constant for all frequencies. In the variable window method, the time window varies with frequency, and usually, the time window gets shorter with an increase in frequency. This approach is advantageous because the temporal smoothing decreases with higher frequencies, but this happens at the expense of frequency smoothing. The results are shown for the variable window method, but similar results were obtained for the fixed window method.

2.2.1.3 Phase Analysis

Phase estimation - We wanted to investigate whether the phase of the ongoing LFP activity just before the time of target-evoked synaptic input (critical point) could predict the amplitude of the target-evoked neuronal responses. Even though we could not estimate the exact time synaptic input arrived in area V4 in each trial, we noticed that the grand average visually-driven ERP responses to the target presentation typically appear after 0.07 sec from target onset. In order to exclude the potential influence of the non-stationary visual responses in the phase estimation, the LFP trials were cut from -1 to 0.03 sec from target onset, excluding the visual response period. In order to be able to compute the phase at critical point or at time points close to it, we modified a previously described method introduced by Ni et al. (2016). In particular, they suggested a method that allowed phase estimation at the edge of a window. To do so, they first removed the time window containing the visual response; then, the trials were downsampled and filtered in different frequency ranges. The filtered signals were fitted with an autoregressive (AR) model (model order: 6) and four cycles per frequency were extrapolated and added at the end of the cut trials. Extending the signal after the critical point allowed spectral estimation without 1) edge artifacts and 2) influence from transient ERP fluctuations. However, filtering can introduce substantial group delays that vary across frequencies. Such delays could move the time of the critical point in a frequency-dependent way that could significantly complicate our analysis. To avoid this problem, we developed a method that is based on the Ni et al (2016) approach but avoids band-pass filtering by using the broadband LFP signal.

Our approach was to estimate the Fourier of the signal using Hann tapers centered at the time of stimulus onset (and at time points before that), and then calculate the phase using this Fourier. The signal with ERP does not give an accurate phase because phase will be affected by the ERP, which is dependent on whether a trial is a hit or a miss (as shown in the results section). To remove the influence of the ERP, we redefined the trials by cutting the data before the ERP occurs and then using an autoregressive model (AR) to extend the data such that the extrapolation does not contain ERP, but has similar spectral properties to the original data.

An autoregressive model predicts how the data will behave in the future based on how they behaved in the past. The length of past data used to predict the future is determined by the model order. To implement this, we used a collection of MATLAB modules called ARfit (Neumaier and Schneider, 2001; Schneider and Neumaier, 2000). The modules in ARfit use a n-variate autoregressive model of order p

$$v_t = w + \sum_{l=1}^{p} A_l v_{t-l} + \varepsilon_t \qquad \varepsilon_t = noise(C) \tag{2}$$

This is a model for stationary time series with vectors v_t such that they have been recorded at equally spaced intervals t. These vectors are of n dimensions in this case (these are different trials). The order p determines the number of past vectors (or time steps) the model takes into account while calculating the future vector. $A_1, A_2, \ldots\ldots, A_p \in \Re^{n \times n}$ are the coefficient matrices of the AR model. $\varepsilon_t = noise(C)$ are uncorrelated random vectors with zero mean and covariance matrix $C \in \Re^{n \times n}$. $w \in \Re^{n}$ is an intercept vector which allows the mean of the time series to be zero.

$$< v_t > \; = (I - A_1 - \ldots\ldots - A_p)^{-1} w \tag{3}$$

Here <·> represents expected value. ARfit has a function called arfit that lets us calculate the coefficient matrices $A_1, A_2, \ldots\ldots, A_p$, intercept vector w and covariance matrix C given the model order. The best order can also be determined using different criteria like SBC (Schwarz's Bayesian Criterion) and FPE (Akaike's Final Prediction Error) given the data and minimum and maximum model order. The ARfit function used is arord. Accuracy of the AR extrapolation can be checked by the arres function which plots the autocorrelation of the residuals. For the model to be adequate, residuals have to be uncorrelated. If 95% of the autocorrelations (for lag > 0) lie within the confidence limits for the autocorrelations of an IID (independent and identically distributed random variables) process of the same length as the time series of residuals then we consider the model order to be accurate.

To extrapolate the data we first needed to find the correct model order and then use it to calculate the coefficient matrices, intercept vector and covariance matrix which are then used to extrapolate the data. The model order was determined by the arord function of the ARfit toolbox. This function found the SBC or BIC (Bayesian information criterion) values for each model order from 1 to 300. SBC is used for model selection among a finite set of models and does so by introducing a penalty term for the number of parameters in the model (Schwarz, 1978). This function was run per channel. In order to get the correct model, the model order with minimum SBC was chosen. Now, for any session we have the optimum model order for each channel. We obtained the median of these orders to get an order for each session. This was repeated throughout the sessions and after getting orders for all the sessions, a median was obtained to get to a common model order for all the sessions. The model order decided upon was 50. Then, we ran several tests in order to assess the quality of the extrapolated signals (Figure 5). First, we plotted the extrapolated signal along with the original signal for one trial to check if the signal is continuous and without any kinks at the point of extrapolation. We also plotted the original and extrapolated signals for trials to check if extrapolations of all the trials have the same variance. And at last we did spectral analysis on the original and extrapolated signals to check if they have similar properties. After all these checks we found that the selected model order was appropriate. We also plotted autocorrelations of the residuals for model order 50 to check the uncorrelatedness of residuals.

After deciding the correct model order, we took the following steps in order to obtain the phase at a given frequency at each trial. After cutting the data at 30ms after stimulus onset, we standardized the data by z-scoring (refer equation (1)) and detrending the data using FieldTrip. An AR model was fitted on the broadband signal of both hits and misses together and the signal was extrapolated for 800 ms using the ARfit toolbox and previously determined model order. We ran this analysis per recording session. Then we estimated the Fourier of the extrapolated signal using the 'mtmconvol' method of the ft_freqanalysis (FieldTrip function) and calculated the angle of the Fourier at the time of stimulus onset (and other time points when needed) to obtain the phase. Fourier decomposition was done for frequencies ranging from 2Hz to 60Hz and frequency dependent Hann taper was used. The length of the taper decreased with an increase

in frequency (length of the taper used was equivalent to 3 cycle length). The taper was used to reduce spectral leakage and control the frequency smoothing.

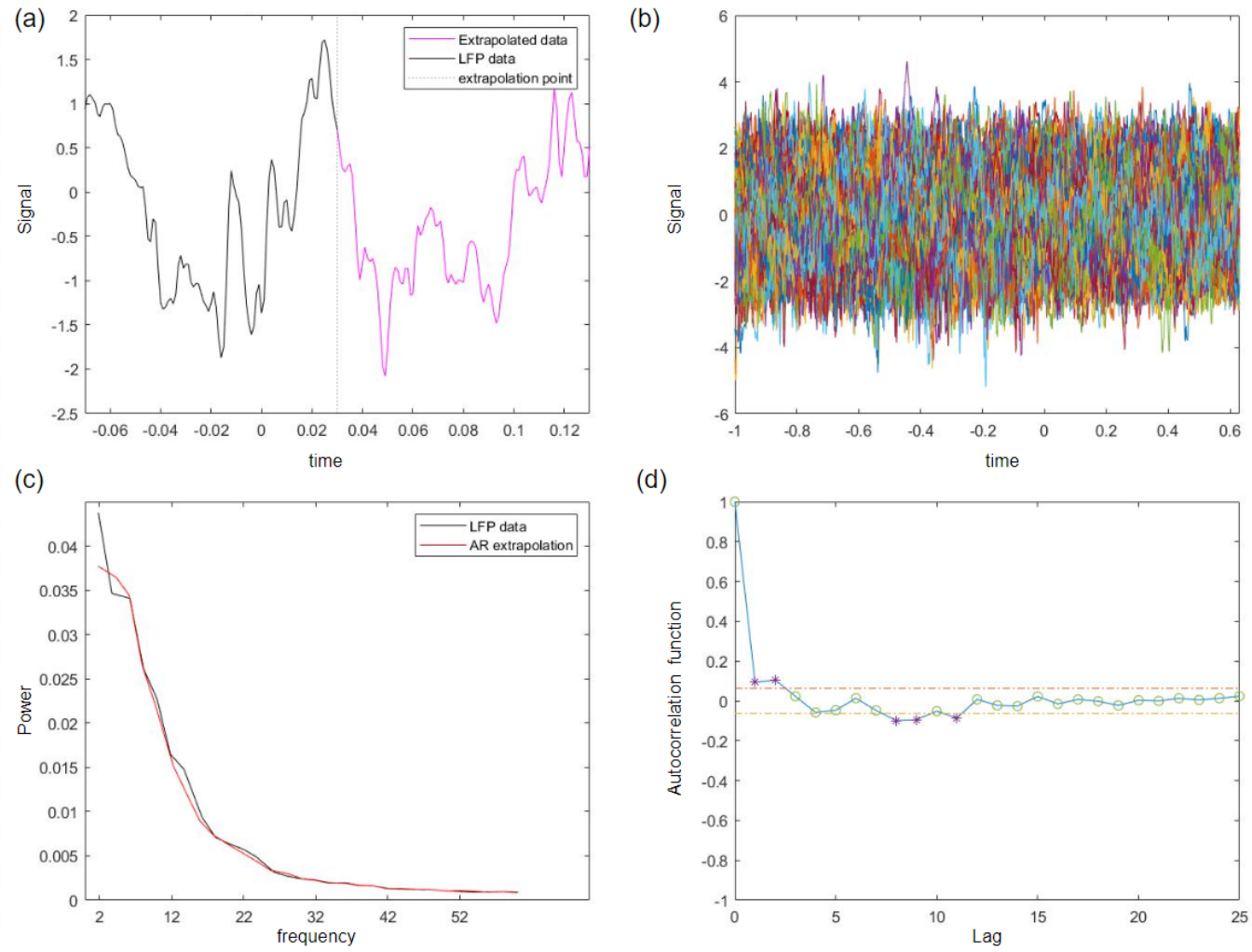

Figure 5. Tests for quality of AR extrapolation. (a) Original signal and extrapolated signal at the point of extrapolation (b) Original and extrapolated signals for all trials to check variance (c) Spectral analysis on the original and extrapolated signals (d) Autocorrelations of the residuals, the green circles show the points where autocorrelation is within bound. Example for one channel is shown here.

When Fourier is calculated using this AR extrapolation, it might not be very accurate because AR extrapolations tend to get noisy as the length of extrapolation increases. To deal with that, we ran 100 iterations of AR extrapolations and then obtained a transform (T) using Fourier of all

these iterations. It is a weighted average of all the Fourier scaled to one multiplied by the mean of length of all the Fourier.

Here eft is Fourier of different iterations

$$phasevec = mean(eft) \,/\, abs(mean(eft)) \quad (4)$$

$$ampvec = mean(abs(eft)) \quad (5)$$

$$T = ampvec. * phasevec \quad (6)$$

To find the final phase at a given frequency at each trial, we obtain the angle of this transform (T)

$$phase = angle(T) \quad (7)$$

Testing phase estimation - In order to test our method, we created synthetic data that allowed us to know the true instantaneous phase at certain frequencies. The simulated data was obtained as follows (by Gregor Moenke, Fries lab).

Phase can be written as a function of angular frequency

$$d\varphi/dt = \omega(t) + \epsilon\xi(t) \quad (8)$$

Here $\xi(t)$ is Gaussian white noise scaled with ϵ
For a time-constant angular frequency, phase can be written as

$$\varphi(t) = \int_0^t \omega + \epsilon\xi(t)\, dt = \omega(t) + \epsilon W(t) \quad (9)$$

Where $W(t)$ is the Wiener Process with diffusion constant ϵ

For discrete time steps $\Delta t,$ phase can be written as

$$\varphi(t_i) = \sum_{0}^{i}(\omega + \xi_i)\Delta t = \omega t_i + W_i \quad (10)$$

This is just a time-discrete diffusion with a linear drift term (angular frequency). This means that we can simulate the exact phase diffusing around the linear (harmonic) phase propagation.

Now to get a $signal\ x(t)$, we needed a 2π periodic waveform function, so we take cosine:

$$x(t_i) = cos[\varphi(t_i)\ ],\ i = 0,\ldots\ldots\ldots, nSamples \quad (11)$$

Note that with this, the time averaged angular frequency still stays ω

$$< \omega + \xi(t) >_t = \omega \quad (12)$$

as the white noise $\xi(t)$ is a zero-mean process. This explains the nicely concentrated spectral peak even with phase diffusion when $\epsilon << \omega$

The data was obtained by combining cosine signals with an AR(1) process or with pink noise

$$signal = acos[\varphi(t_i)\ ]_{\omega 1} + bcos[\varphi(t_i)\ ]_{\omega 2} + cAR(1) \quad (13)$$

or

$$signal = acos[\varphi(t_i)\ ]_{\omega 1} + bcos[\varphi(t_i)\ ]_{\omega 2} + c(pink\ noise) \quad (14)$$

Here $\omega 1$ and $\omega 2$ are different frequencies at which the true phase is known by construction. Here c was varied to get different noise of the AR process and $\xi(t)$ was varied to get different phase diffusion strength. 50 trials were obtained to test the method. The signal was cut at the midpoint and extrapolated, and the taper was centered at the midpoint for Fourier. The method was tested with different noise, phase diffusion and sampling frequency. The differences are discussed in the results and discussion section. Overall the method works pretty well and estimates the phase very close to the true phase and the phase obtained by original data.

ERP amplitude estimation - The ERP amplitude at each trial was calculated by using the LFP signal from 50ms to 150ms after stimulus onset. Trials which have no saccades in this period are used for correlation analysis. During this period, the monkey is not yet allowed to behaviorally respond to the target. Thus, the ERP amplitude can not be attributed to saccade related neuronal responses. If $x_1, x_2 \ldots\ldots\ldots x_{100}$ are the values of the LFP signal within this time interval, then the ERP amplitude was calculated as

$$ERP\ amplitude = \sqrt{\sum_{t=1}^{100} \left(x_t - \bar{x}\right)^2} \quad (15)$$

Here, $\bar{x}$ is the mean of $x_1, x_2 \ldots\ldots\ldots x_{100}$.

Correlation between ERP and phase - The circular-linear correlation between the phase (directional variable) and ERP amplitude (linear variable) was calculated using the circ_corrcl function of the CircStat toolbox (Philipp, 2009). Since phase is a circular measure, the Pearson correlation coefficient cannot be used. Here α is the directional variable, x is the linear variable and ρ_{cl} is the circular-linear correlation (Zar, 1999).

$$\rho_{cl} = \sqrt{\frac{r_{cx}^2 + r_{sx}^2 - 2 r_{cx} r_{sx} r_{cs}}{1 - r_{cs}^2}} \quad (16)$$

$r_{sx} = pc(sin\alpha, x)$, $r_{cx} = pc(cos\alpha, x)$ and $r_{cs} = pc(sin\alpha, cos\alpha)$ where $pc(x, y)$ is the Pearson correlation coefficient.

The phase is obtained for each frequency for a certain time point, so the circular linear correlation value was calculated after pooling relevant trials from all sessions at each frequency. This gives a list of correlation values for each frequency and a statistical analysis is done to

obtain frequencies at which the phase is significantly correlated to ERP amplitude. The statistical analysis is described in section 2.2.3.

2.2.2 Eye data

Microsaccade (MS) Detection - Microsaccades are ballistic eye movements that have small trajectories. We focused our MS analysis during the fixation period (from -1 to 0.1 sec from target onset), before the monkey was allowed to respond (saccade) to the target presentation. .

Two different methods were used for MS detection; 1) the Engbert and Kliegl algorithm and, 2) a convolutional neural network (CNN). Each method is briefly described below.

Engbert and Kliegl algorithm (EK method): Using the Engbert and Kliegl algorithm (Engbert and Kliegl, 2003) according to which MS is defined as outliers in the 2D velocity space. To protect detection from noise, the threshold of this algorithm is based on the median of the velocity-time series. A multiple of the standard deviation of the velocity distribution is used as the detection threshold (6xstandard deviation was used in this analysis).

Convolutional neural network (CNN method): Using a convolutional neural network (CNN) to automatically detect saccades based on manually labeled data (Bellet et al., 2019). The network architecture draws inspiration from U-Net, and it operates on the eye velocity signal. For labeling the MS, a MATLAB app called iLabel was used (the app was developed by Tim Näher, Fries lab). Using these labels, the network was trained (200 trials of a single session, session ‘klecks_20170828_attentional-sampling_1’ was used), and the prediction was done for the rest of the trials and sessions using scripts available at https://github.com/berenslab/uneye. The network segments eye movement recordings into epochs containing MS versus epochs not containing them. The start and the end timing of MS can be extracted from this. We used two different definitions of MS to train the data; they are shown in figure 6.

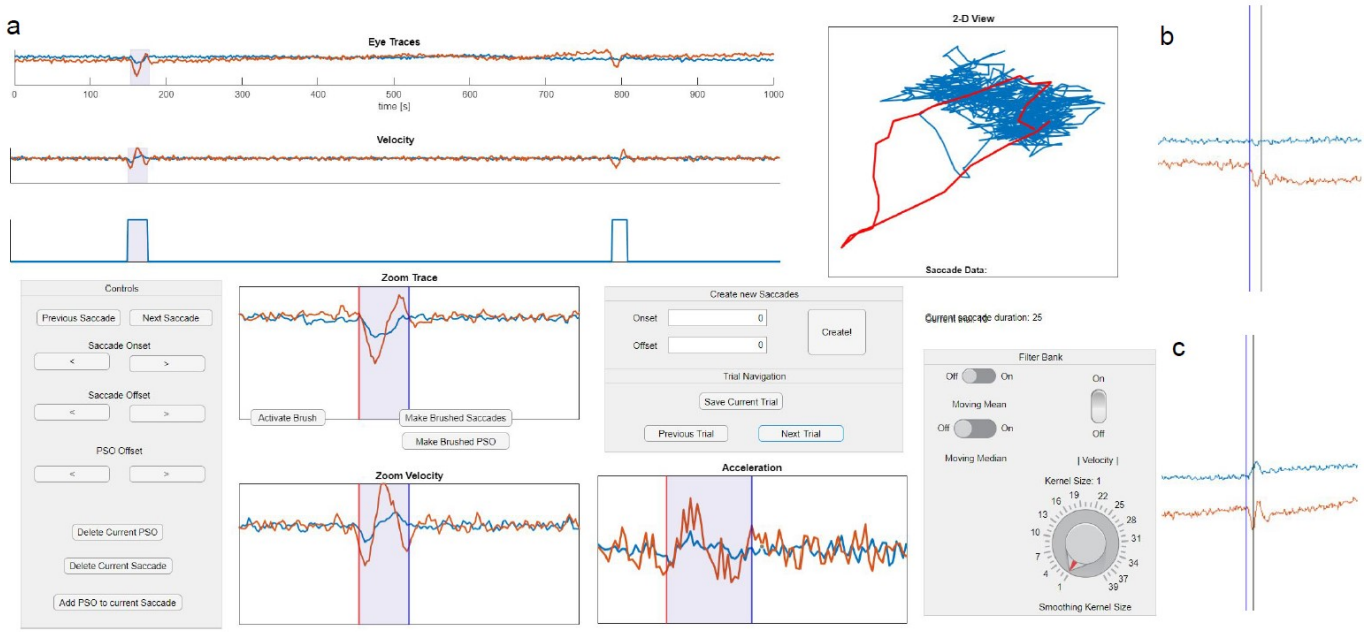

Figure 6. (a) The labeling app ilabel. The orange and blue traces show horizontal and vertical traces. The vertical red and blue lines show the start and end of MS, respectively. (b,c) The blue and black vertical lines show the start and end of MS, respectively. An example detection when the network is trained using (b) Definition 1 of MS, it includes PSOs in the MS end. (c) Definition 2 of MS, does not include PSOs in MS end. This definition is a bit more stringent.

We were interested in comparing the characteristics of MS between hits and misses. To this end, we computed the following microsaccade measures:

- Peak velocity amplitude plot - Microsaccades usually show a fixed relation between peak velocity and amplitude due to their ballistic nature. This relation is usually used to check the validity of detection (Engbert and Kliegl, 2003). Peak velocity is defined as the maximum velocity during the MS, and amplitude is the distance between the start and endpoint of the MS.

- MS distribution and rate - For MS distribution around target onset, we looked at the number of MS in each bin (each bin was 40ms), and to have an equal number of trials for each condition, hit trials were subsampled. For rate calculation, the number of MS at each

time point was calculated for all trials of a given condition, and then this time series was convolved with a gaussian of width 50 ms. To normalize the rate, the time series was divided by the number of trials belonging to each condition and the window size of the gaussian.

- Direction of MS - Two methods were used for the calculation of the direction of MS.

  Method 1: Direction of MS is the angle of vector joining start and endpoint of MS (Engbert and Kliegl, 2003)
  Method 2: Angle at peak velocity is defined as the direction of MS (Xue et al., 2020).

2.2.3 Statistical Analysis

Nonparametric permutation testing was used to assess statistical significance. In this method, we created a distribution from the available data by looking at what the test statistic would be if the null hypothesis were true. For neural and eye data analysis, the distribution was created by iteratively shuffling the condition labels over trials and recomputing the test statistic. The shuffling was done 1000 times. Next, we compared the observed test statistic against a null-hypothesis distribution of test statistic values. If the observed test statistic was not within the ‘boundaries’ of null hypothesis distribution, then the effect is said to be significant, and the null hypothesis is rejected. We corrected for multiple comparisons using pixel-based statistics (Cohen, 2014). In order to do this, we found two extremely valued pixels (for the two-tailed test) from the distribution obtained for each iteration of permutation testing. After running all iterations, we obtained a distribution containing the largest and smallest pixel values. The lower bound is the value corresponding to the 2.5th percentile of the smallest values, and the upper bound is defined as the value corresponding to the 97.5th percentile of the largest values (here $p = 0.05$). Any pixel having a value greater than the upper bound or smaller than the lower bound is considered to be statistically significant.

For ERP analysis, the test statistic is the difference in ERP for the two outcomes. For frequency analysis, the test statistic is the log ratio of power for the two outcomes. For MS rate analysis, the difference between the MS rate for two outcomes was used as the test statistic. To do statistical analysis on pooled data, maximum and minimum pixel values distributions for all the sessions were pooled in and the lower and upper bounds were found using this pooled distribution. This lower and upper bound was applied to the average of all the sessions and the z-scored data was used for averaging.

In the correlation analysis, ERP amplitude is associated with a phase. We shuffled the ERP amplitude values 1000 times and calculated linear circular correlation for each permutation to create the null distribution. By doing so, we obtained 1000 correlation values for each frequency. Similar to the previous analysis, the maximum correlation value for each permutation is obtained to get a distribution (one-sided test). Suppose the test statistic is larger than the value corresponding to the 95th percentile value of this distribution ($p = 0.05$). In that case, the correlation between ERP amplitude and the prestimulus phase is not by chance.

All scripts for the analysis are available at https://github.com/shivangi1399/Attentional-Scanning

3. Results

3.1 MUA and LFP data

3.1.1 Event Related Potential (ERP)

Most channels showed a significant difference between hit and miss conditions for LFP (figure 7) and MUA (figure 8) signals within the first 100 ms of target onset. No saccades occur during this time, which means that the difference is probably not due to saccades.

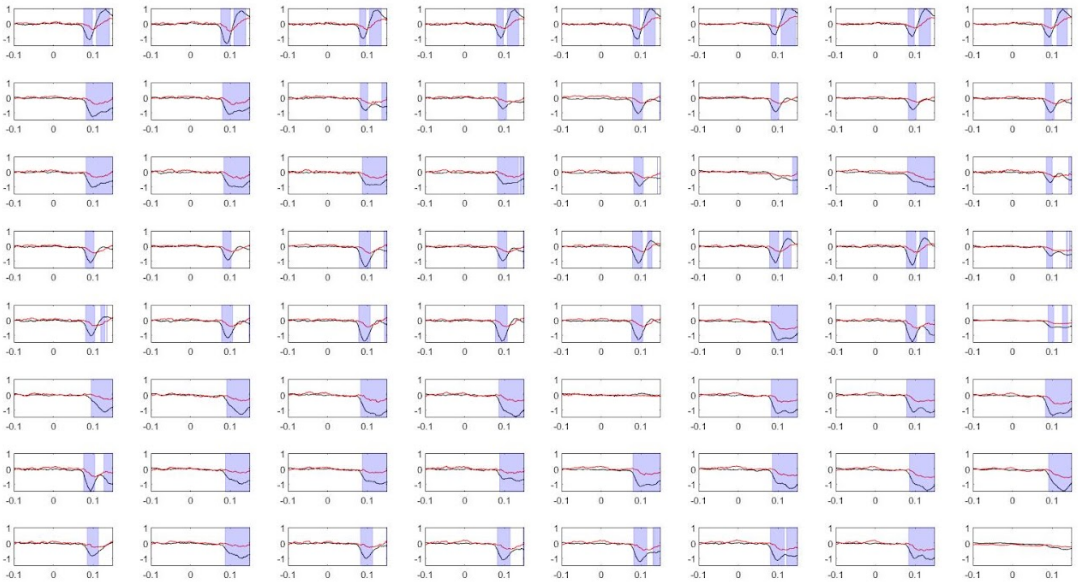

Figure 7. LFP signal for one example session (klecks_20170828_attentional-sampling_1) for each channel. The blue region shows the significance thresholds. The red and black traces are average hits and misses across trials of respective conditions. Target onset is at 0. All trials and channels are included.

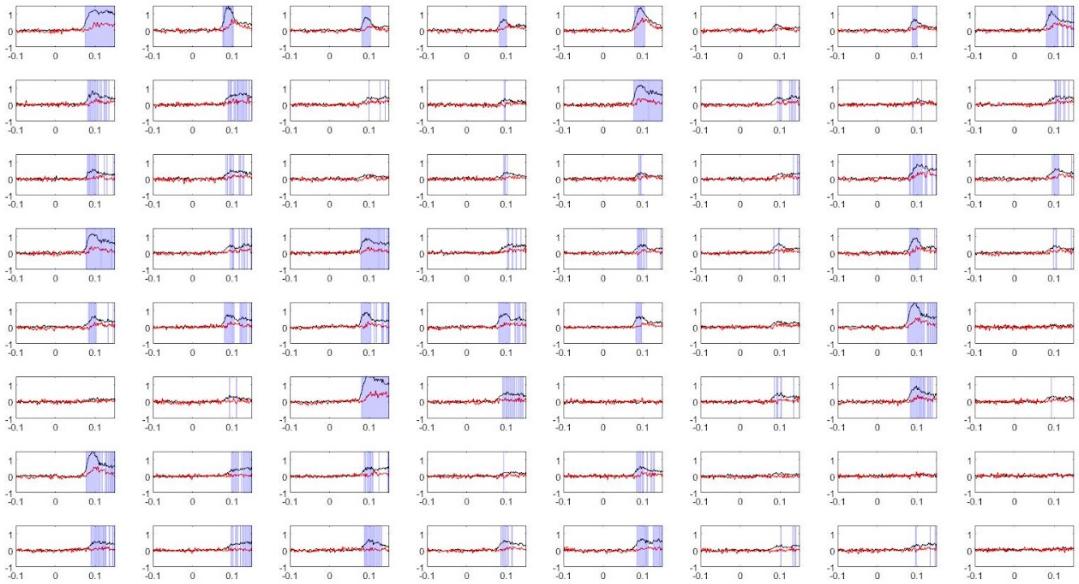

Figure 8. MUA signal for one example session (klecks_20170828_attentional-sampling_1) for each channel. The blue region shows the significance thresholds. The red and black traces are average hits and misses across trials of respective conditions. Target onset is at 0. All trials and channels are included.

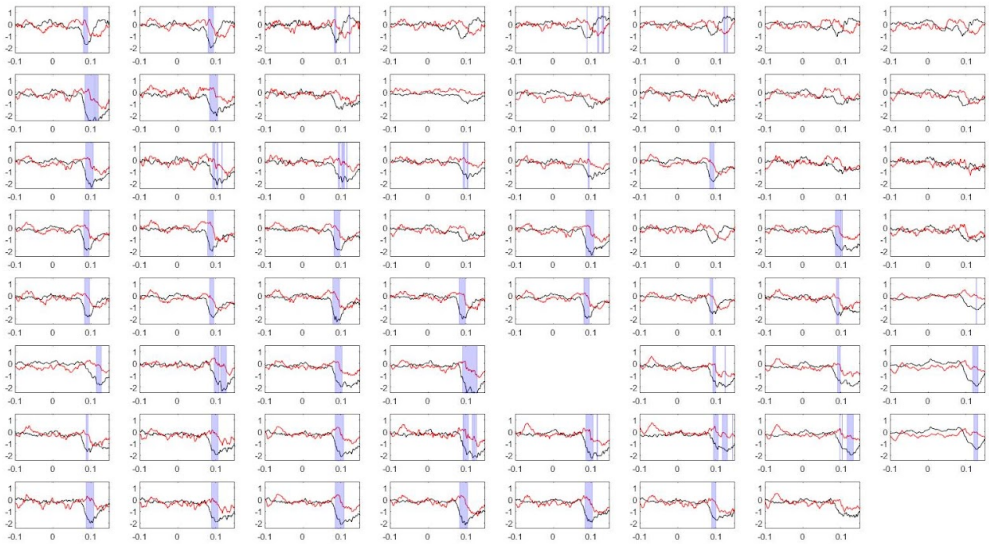

Figure 9. LFP signal for one example session (session klecks_20170822_attentional-sampling_1) for each channel. The trials are for a particular location (103) and difficulty level (115). The blue region shows the significance thresholds. The red and black traces are average hits and misses across trials of respective conditions. Target onset is at 0. The cleaned data is used and trials or channels with high variance or high number of artifacts are excluded.

3.1.2 Time-Frequency analysis

Most channels showed a significant difference between LFP power for hit and miss conditions within the first 100ms of target onset (figure 10 and 11). No saccades occur in this time duration, which means that the difference is probably not due to saccades. As the frequency changes, the window length of time used for power calculation changes. There is a possibility that sometimes the time window is large enough to capture power changes due to an early saccade. The power difference is significant in the 20Hz band of frequency and some higher frequencies like 70 to 100 Hz (probably because of spiking activity).

The difference in detection between hits and misses is reflected in LFP power in the lower gamma band (~ 20 Hz).

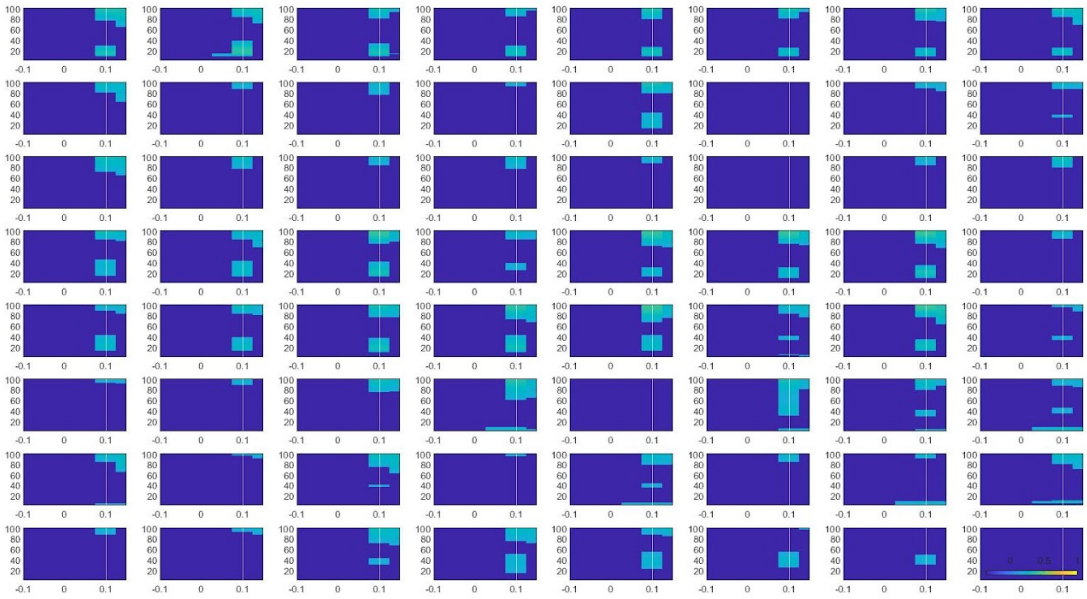

Figure 10. Pooled TFR (log(hit/miss)) for 4 sessions for each channel. Only the pixels crossing significance thresholds are represented here. Target onset is at 0. The white line shows t=0.1 s; after this time, the monkey was allowed to make a saccade.

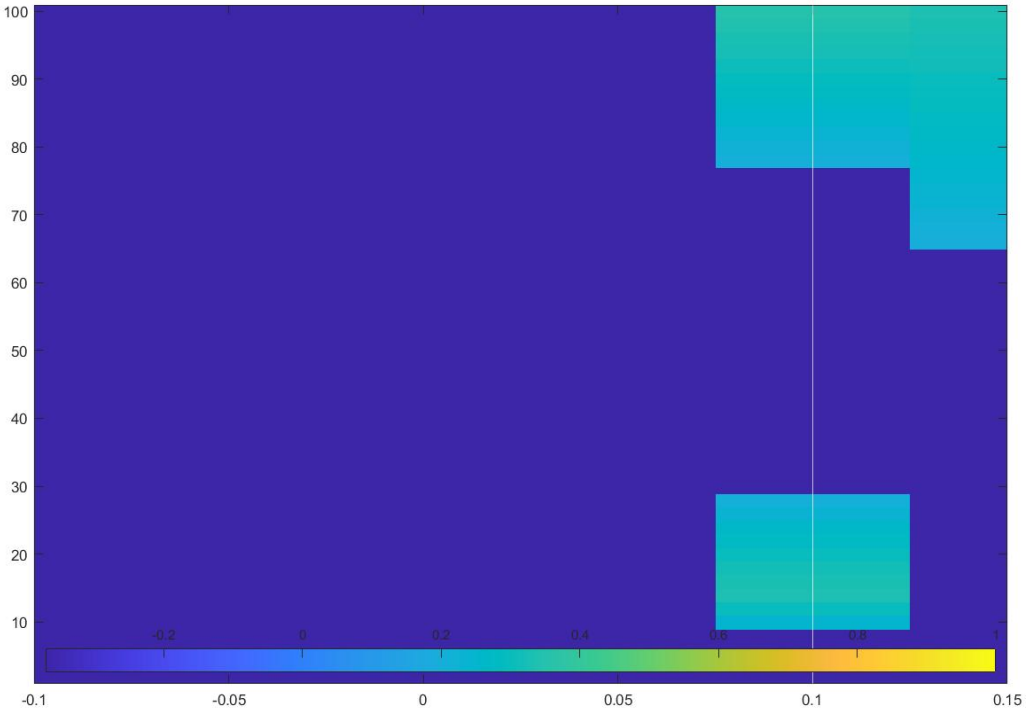

Figure 11. Pooled TFR (log(hit/miss) for 4 sessions, for example channel (ch.1). Only the pixels crossing significance thresholds are represented here. Target onset is at 0. The white line shows t=0.1 s; after this time, the monkey was allowed to make a saccade.

3.1.3 Phase Analysis

Phase estimation - To investigate if the prestimulus phase of the LFP (phase just prior to critical point) correlates with ERP, we need to find the LFP phase at stimulus onset. Phase estimation was done by a method based on Ni et al. 2016 where we used the broadband signal to find the phase. This project assumes that the prestimulus phase of the ongoing oscillatory signal, free of any ERP, determines the perceptual outcome and affects the amplitude of the ERP. To remove the influence of ERP on phase determination, we cut the signal at 30ms after the target onset to be as close to the critical point as possible. We extrapolated the cut LFP signal and then calculated Fourier transform at the point of stimulus onset (and time points before that) as described in the methods section. After establishing this method, we checked the accuracy of phase detection in two ways:

Using real data: We checked the phase estimation method on our data by using the data before the stimulus onset. The trials were 1.6 seconds long (from -1s to 0.6s), and the stimulus onset was at zero seconds. So we could use the data from -1s to 0s as it doesn't have ERP. To test the method, we decided to find the phase at -0.5s (midpoint of the cut signal). For this, we cut data at -0.480 seconds (20ms after the point of phase determination so that a larger part of the original signal be included in the phase determination), extrapolated it, and calculated Fourier transform at -0.5 seconds. To compare the predicted phase to the phase using the original signal, we cut the data at zero and found the phase at -0.5s. The method seemed to work well in this case. The phase determination is better for higher frequencies in this case because the data was cut 20 ms after the point of phase determination, which means a bigger part of the real signal is used in Fourier of higher frequency. Signals used in the lower frequencies will be noisier because AR extrapolation has more noise as it is extended more.

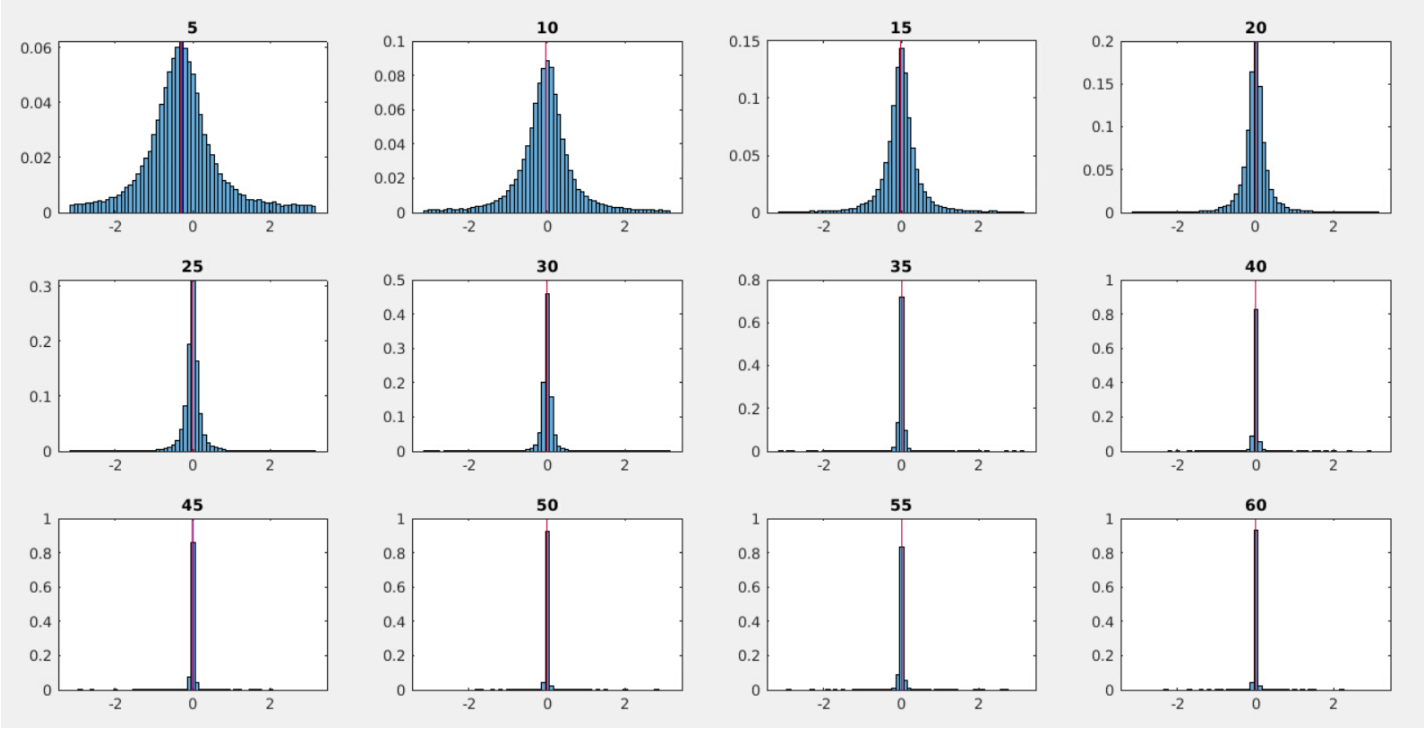

Figure 12. The histogram for the difference between AR phase and phase using the original signal for all trials and all channels in an example session. Different subplots show different frequencies in Hz. x axis shows the difference in radians and y axis shows the proportion of trials having that difference. The red and blue lines represent mean and median respectively.

Using synthetic data: As mentioned in the methods section, the phase estimation method was also tested using synthetic data with various phase diffusion strength, noise and sampling frequency. Different signals were generated and AR and Fourier analysis was run on each of them. The signal was constructed as mentioned in the methods section in equations (13) and (14).

In case 1, the signal was constructed using the AR(1) process and the two frequencies used were 10 Hz and 35 Hz. a = 1, b = 0.75 and c = 0.2 (refer equation (13)). The phase diffusion strength is 0.0025 and sampling frequency is 200Hz. The model order came out to be 44. The estimation in this case looks quite good because the difference between true phase and predicted phase concentrates around zero (Figure 13(a) and 14(a)). This shows that the phase estimation method works well.

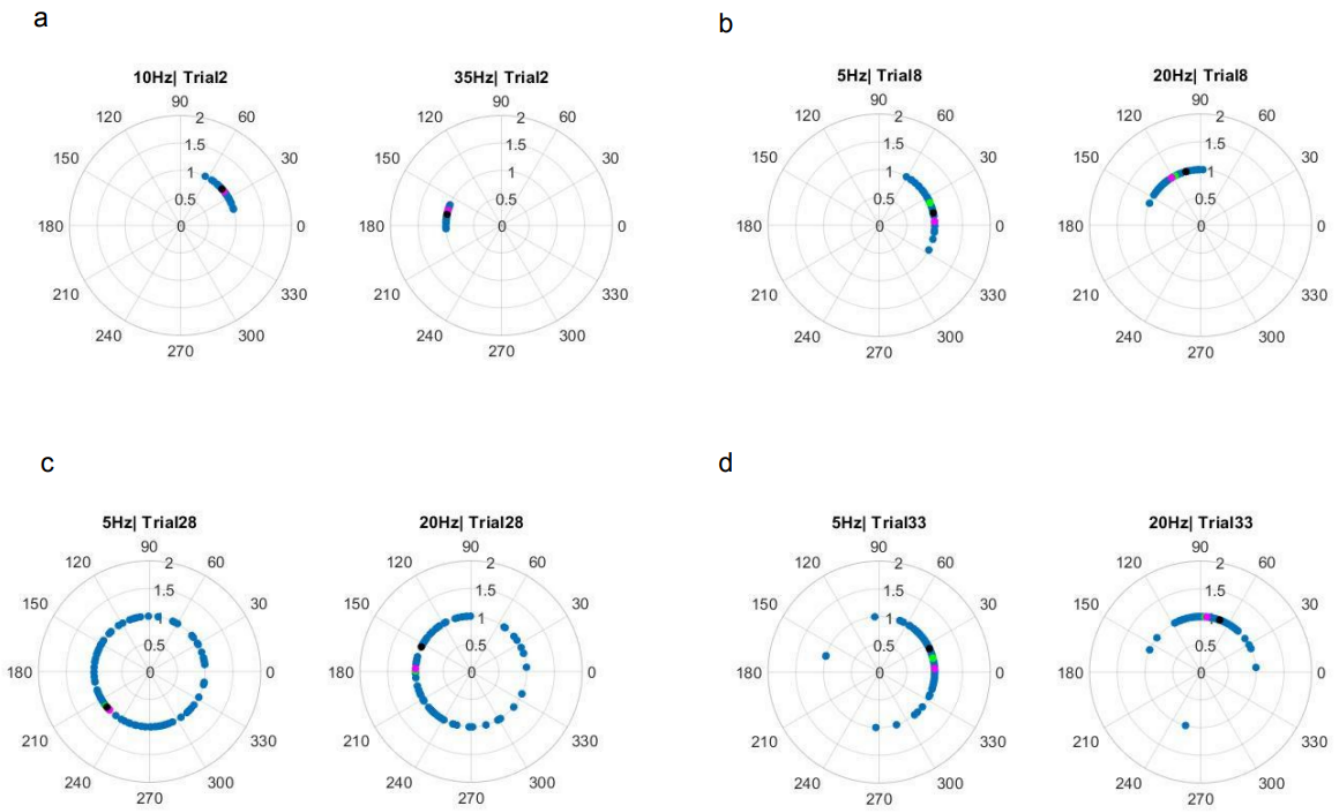

Figure 13. Example trial where blue dots are of all the phases obtained in different iterations, Green is the predicted phase (using AR signal), magenta is the true phase and black is the phase using the original signal (we can use this since it doesn't have any ERP). (a) Case 1 (b) Case 2 (c) Case 3 (d) Case 4.

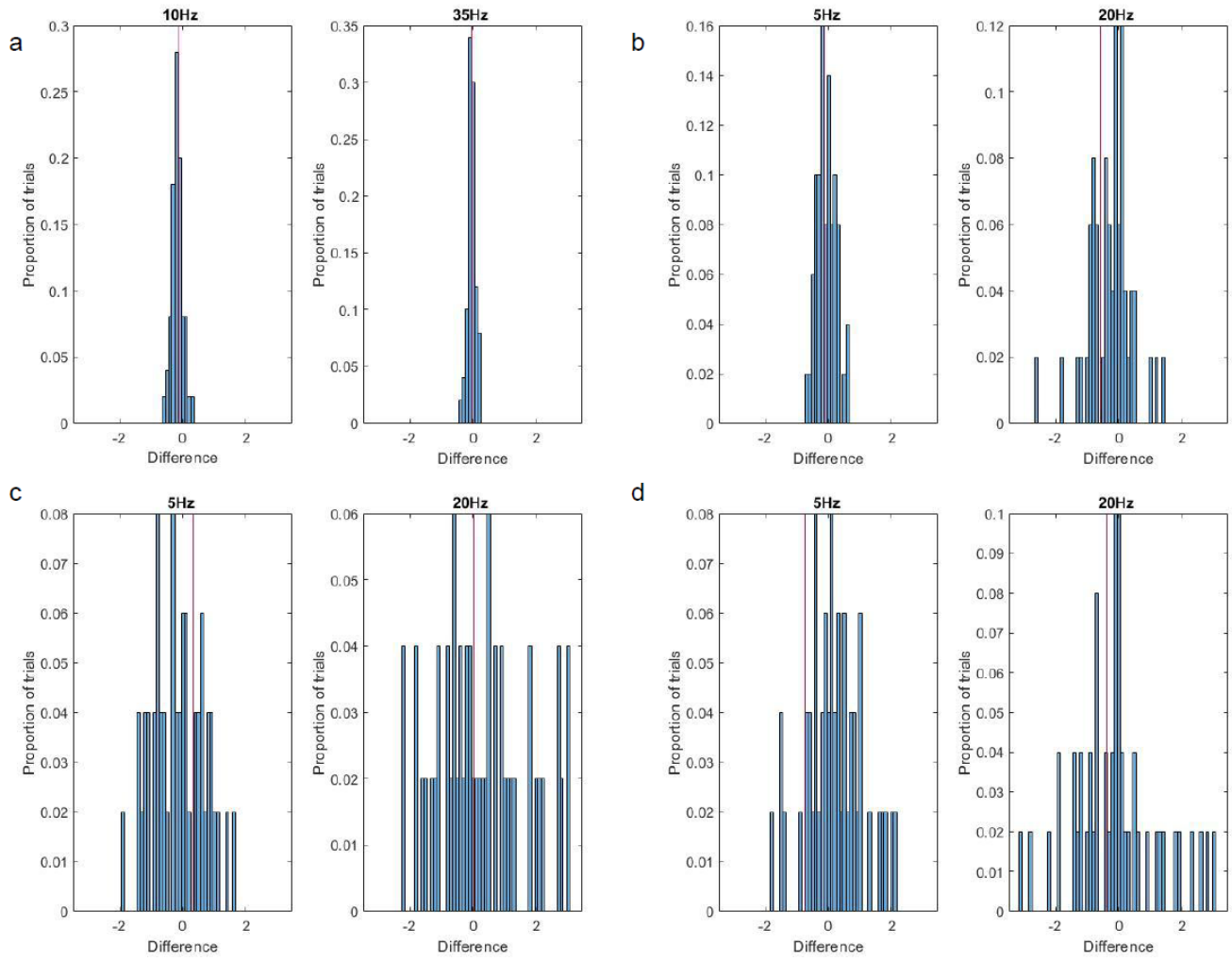

Figure 14. The histograms show the difference between the predicted phase and the true phase for all the trials for the two frequencies (a) Case 1 (b) Case 2 (c) Case 3 (d) Case 4. The labels are different frequencies for which the true phase was known. The red and blue lines represent mean and median respectively.

In the next three cases, we tried the method on a signal with 1000Hz sampling frequency so that the signal is as similar to the original signal as possible. We changed noise and phase diffusion strength to test how accuracy of phase determination changes with these different parameters. To increase noise, the parameter c was increased. Pink noise was used instead of the AR(1) process so the model order is comparable to a real signal. Refer equation (14) for the last three cases. Since we are more interested in phase relations at lower frequencies, we looked at 5Hz and 20Hz frequency signals in the next cases to check accuracy of phase estimation at these

frequencies. Here parameters a and b correspond to 5Hz and 20Hz signals respectively. In all the cases, the signal was cut and extrapolated at the midpoint of the signal, and Fourier transform was calculated at the midpoint of the signal.

In case 2, strong phase diffusion, weak background noise is used: a = 1, b = 0.5 and c = 0.5. The phase diffusion strength is 0.1. The model order came out to be 70. In case 3, weak phase diffusion, strong background noise is used: a = 1, b = 0.5 and c = 2. The phase diffusion strength is 0.01. The model order came out to be 11. In case 4, both strong phase diffusion and strong background noise is used: a = 1, b = 0.5 and c = 2. The phase diffusion strength is 0.1. The model order came out to be 12. Model order came out to be low in the last two cases because of strong pink noise but the signal is still comparable to real signal.

As shown in figure 14, the difference between the true phase and predicted phase is concentrated around zero for case 1, which means phase determination was good. When we compare case 2 to case 3 and 4, the spread around zero increases in both cases. So, as noise increases and phase diffusion strength decreases, the accuracy of phase determination decreases. And it looks like background noise has more influence on phase determination than phase diffusion strength. Further statistical analysis needs to be done to confirm that. Real data has strong noise and low phase diffusion strength, which makes phase determination even harder, but the tests show that our phase calculation method works reasonably well.

After the phase determination method was established and tested by using synthetic data, we looked at the correlation between ERP amplitude and phase.

Correlation between ERP and phase - If perception changes with ongoing LFP oscillations, then the prestimulus phase of LFP should be correlated with post-stimulus ERP on a trial-by-trial basis. A circular-linear correlation measure is used as described in the methods section. This analysis was done for a particular location (location 161) and all the difficulty levels pooled together. Trials for both the behavioral outcomes - hits and misses were pooled in for the analysis.

This whole analysis was done at four different time points namely t = 0 seconds (stimulus onset), t = - 0.05 seconds, t = -0.1 seconds and t = -0.15 seconds. Correlation analysis at stimulus onset tells us about prestimulus phase (phase just prior to critical point) correlation with post stimulus ERP amplitude. Correlation analysis at time points before stimulus onset was done to check how this correlation changes with time.

We observed that prestimulus single-trial phases were correlated with post-stimulus neural responses for a few channels (Figure 15 and 16). The analysis for the four prestimulus time points is shown in figure 16, circular-linear correlation is significant for lower frequencies (6 to 10Hz) for a few channels in the last three time points leading up to stimulus onset. Significant correlation is also seen in some higher frequencies (~ 50 Hz) for one channel for t = 0 seconds and t = -0.05 seconds.

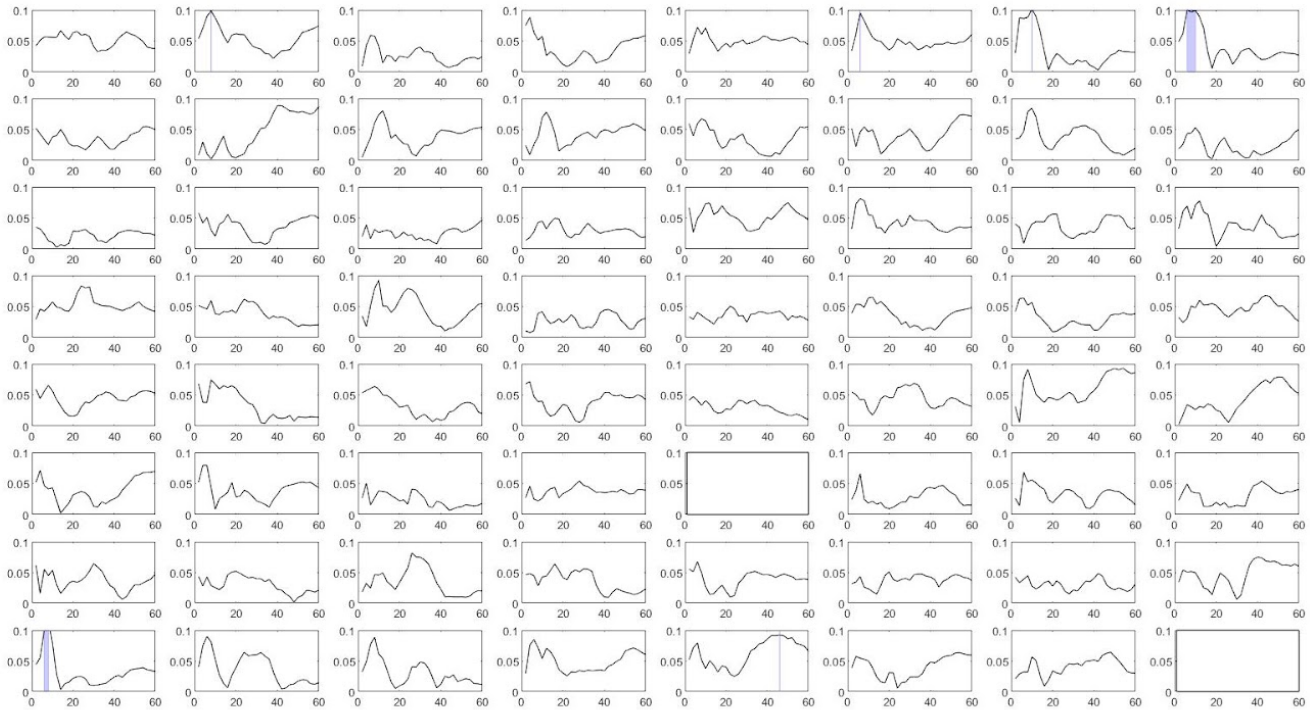

Figure 15. Circular-linear correlation vs frequency for all channels att = 0 (or stimulus onset). Channels 45 and 64 are blank because they did not have good signal quality (selection of trials is described in methods section). Blue area represents the frequencies which had significant correlations when we looked at significant correlations across frequencies. All difficulty levels are used. 70 iterations of AR were used in this plot.

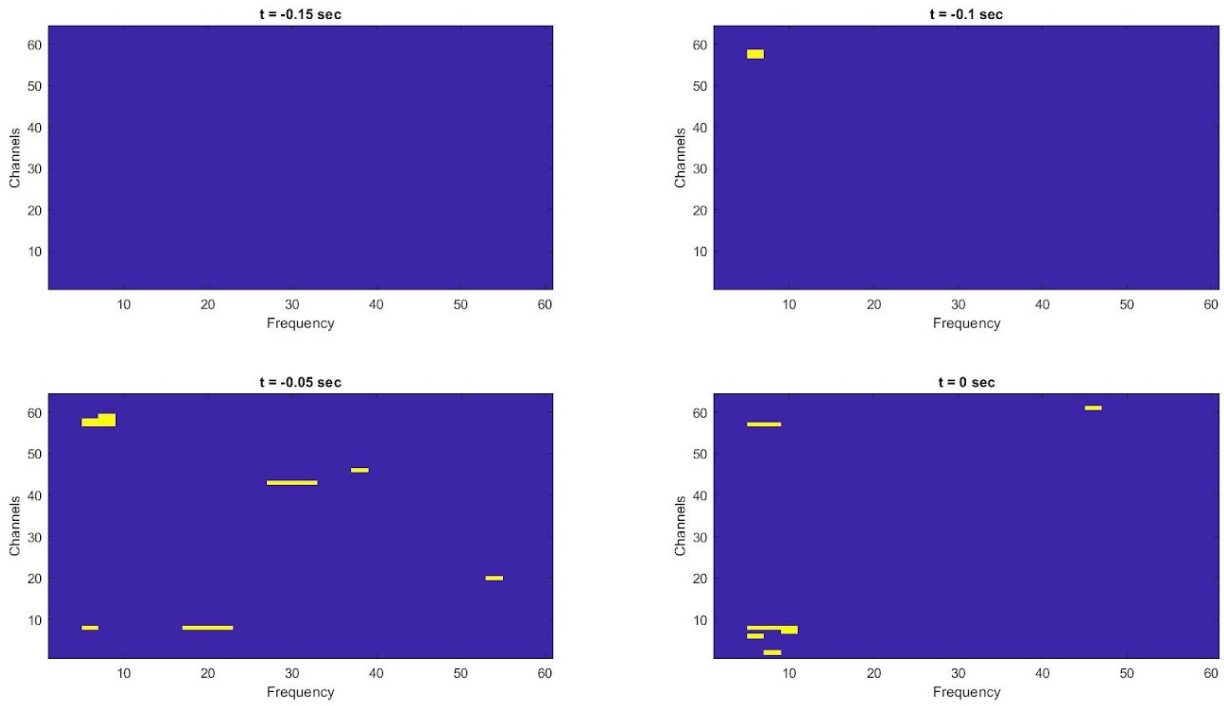

Figure 16. Each subplot represents different prestimulus time points. In each plot, the yellow areas represent the frequencies and channels which had significant circular-linear correlation between phase at the corresponding time point and post-stimulus ERP amplitude. All difficulty levels are used. 70 iterations of AR were used in this plot.

We also did the correlation analysis by selecting trials of only intermediate difficulty levels (115, 116, 117 and 118) because the extreme difficulty levels had very few trials and different signals compared to intermediate difficulty levels. This might be decreasing the observed significant correlation. Indeed, we see that the number of channels having significant correlation increases when only intermediate difficulty level trials are used (Figure 17) compared to when all the difficulty levels were used (Figure 15). Till now we have been using both hit and miss trials but there might be a lot of reasons to miss a trial apart from non-optimal phase, so we wanted to check if the correlation becomes better when only hits trials are considered. Looking at only hit trials can be helpful because in this case we might see a clearer correlation between ERP amplitude and phase. We do observe slightly more channels showing significant correlation when only hit trials are used (Figure 18). In all these cases, we observe that the circular-linear correlation is significant for lower frequencies (6 to 10 Hz).

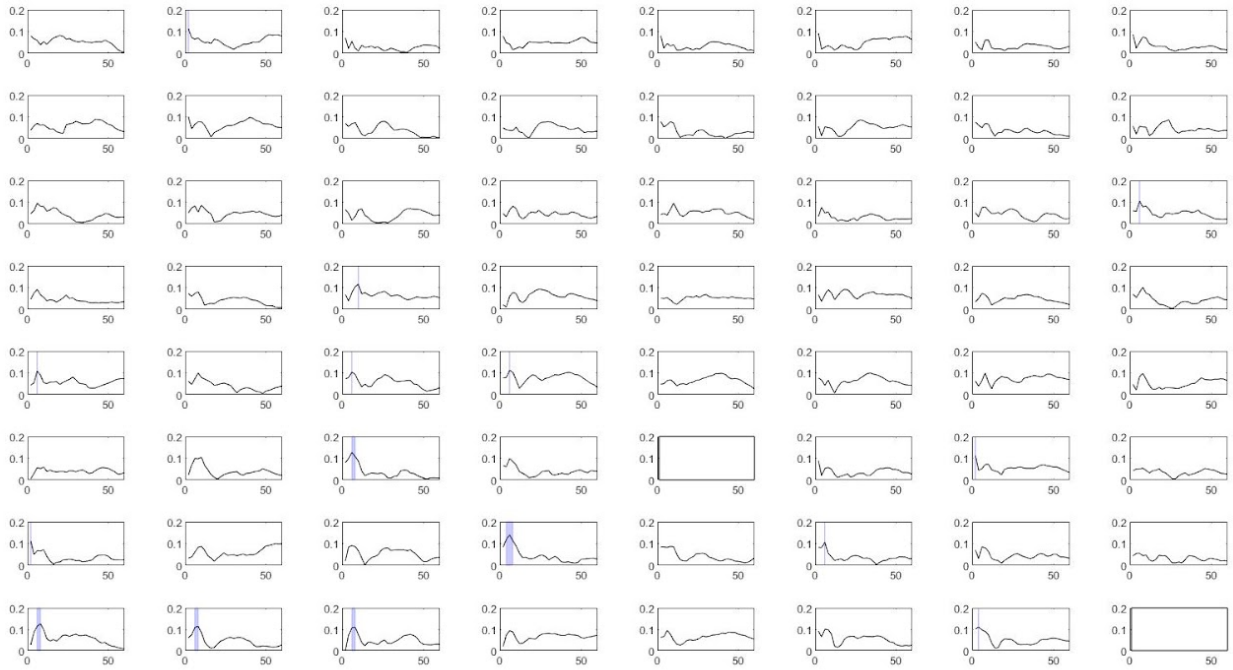

Figure 17. Circular-linear correlation vs frequency for all channels for t = 0 (or stimulus onset). Channels 45 and 64 are blank because they did not have good trials (selection of trials is described in methods section). Blue area represents the frequencies which had significant correlations when we looked at significant correlations across frequencies. This analysis is for trials with intermediate difficulty levels, and for both hits and miss trials.

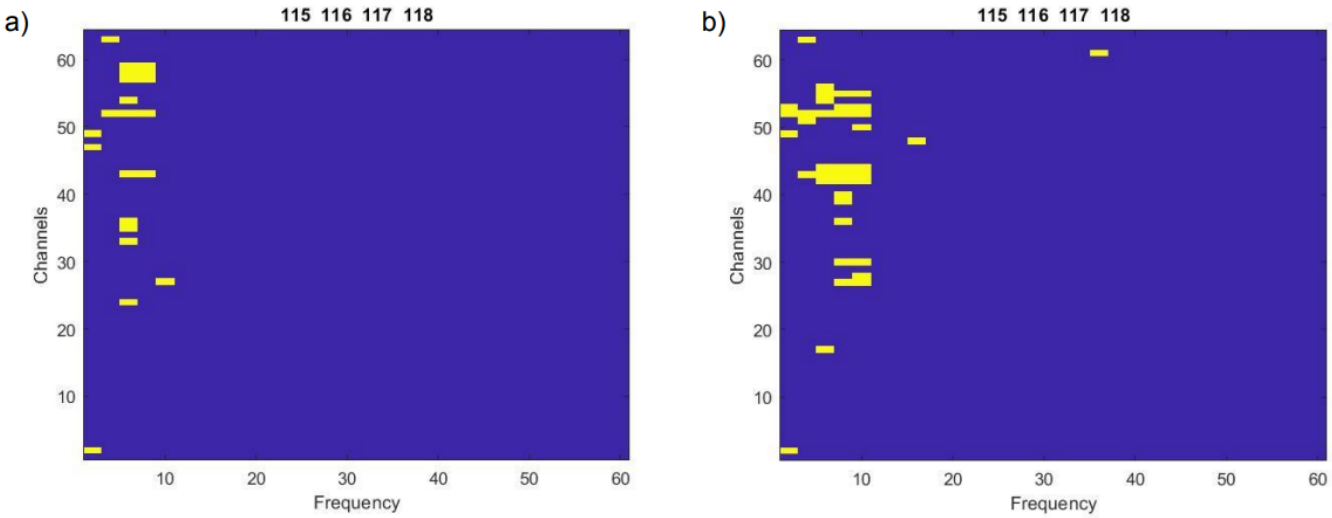

Figure 18. In each plot, the yellow areas represent the frequencies and channels which had significant circular-linear correlation (at stimulus onset) when multiple comparison was done across frequency for each channel. This analysis is for trials with intermediate difficulty levels. a) Correlation for both hits and miss trials. b) Correlation for only hit trials.

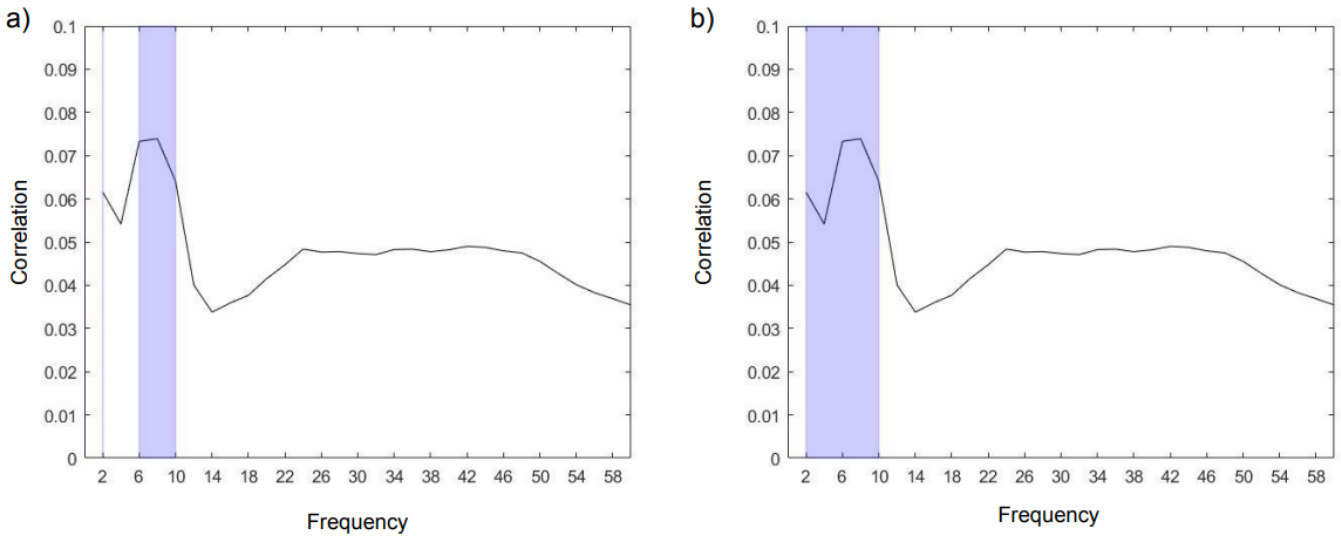

Figure 19. Average circular-linear correlation vs frequency for all channels (62 channels) for t = 0 (or stimulus onset). Blue area represents the frequencies which had significant correlations when we looked at significant correlations across frequencies. This analysis is for trials with intermediate difficulty levels. a) Correlation for both hits and miss trials. b) Correlation for only hit trials.

We also looked at the average correlation for channels for intermediate difficulty levels and observed that significant correlation in 6-10Hz band for both hit and miss trials and in 2-10Hz band for only hit trials (Figure 19). This shows that prestimulus phase in the 6-10Hz frequency band is correlated to post-stimulus ERP amplitude.

3.2 Eye data

Microsaccade

Engbert and Kliegl algorithm (EK method): Some measures of MS are discussed here:

- Peak velocity amplitude plot: As expected, there is a fixed relation between peak velocity and amplitude for all sessions. Two example sessions are shown in figure 21.

- Distribution and rate - To check if there are more MSs around target onset in misses compared to hits, we looked at MS rate and found that right after target onset, there is a significant difference between hits and misses MS rate (figure 20). Similar results were observed for MS distribution. There are more MS after target onset in the miss condition. It is probably because MS around target onset does not allow good detection of the target, and as a result, the monkey fails to perform the task correctly (miss trial).

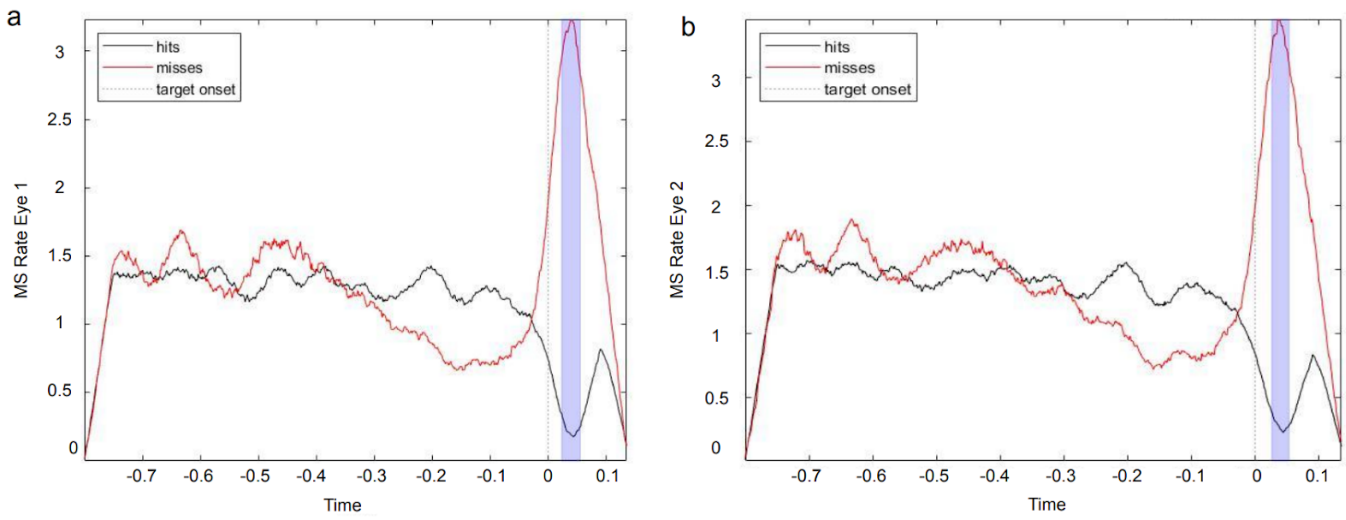

Figure 20. (a,b) Average MS rate difference and statistical thresholds for all sessions pooled together.

- Direction of MS - We wanted to check if more microsaccades are in the direction of the target compared to any other direction because this might benefit target detection. This effect has been observed in some human psychophysical studies (Engbert and Kliegl, 2003; Hafed and Clark, 2002; Rolfs et al., 2005). MS is found to be in the general direction (quadrant) of the target, but it's still quite far from the exact target position. This could be because microsaccades are biased towards the attended location and not the exact target location (Xue et al., 2020). Figure 21 (b and c) shows the direction of MS in hits and misses. Proper statistical analysis needs to be done in order to verify whether there is a significant difference.

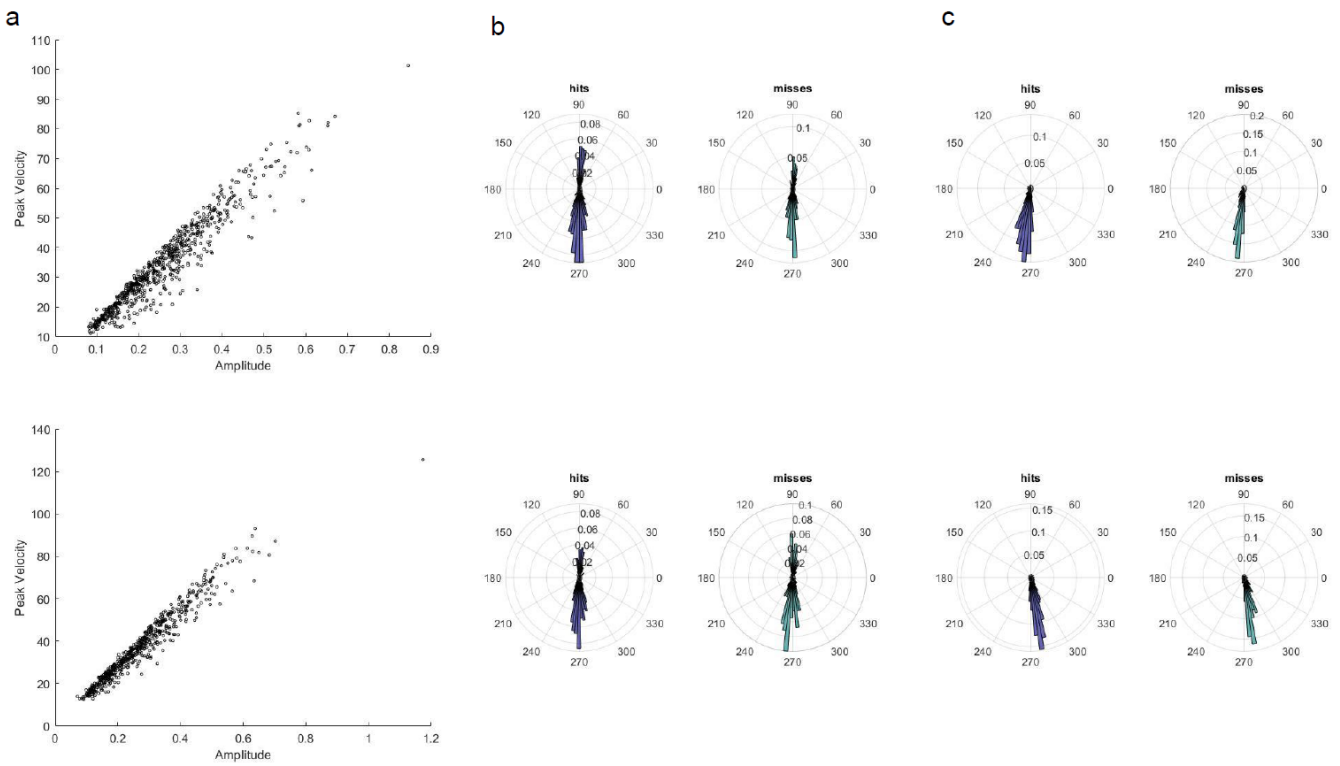

Figure 21. These plots are for the EK method. (a) Peak velocity-amplitude plots for sessions 3 (row 1) and 4 (row 2)). (b) Direction plots according to method 1 (Direction of MS is the angle of vector joining start and endpoint of MS) for sessions 3 and 4. (c) Direction plots according to method 2 (Angle at peak velocity is defined as the direction of MS) for session 3 and 4. The differences between these two methods are due to different endpoints of MS. Session 3 is klecks_20170808_attentional-sampling_1 and session 4 is klecks_20170810_attentional-sampling_1.

Convolutional neural network (CNN method): The MS rate was similar to the EK method using this method. There is a difference between the distribution of the direction of MS and the peak velocity-amplitude plot depending on the definition of MS used. In this method, the amplitude peak velocity plot has more spread, probably due to the absence of strong thresholding, which was present in the EK method.

- Definition 1 - This is a less stringent way of marking MS end, so the end is usually defined as the point where the eye finally settles to, after making a MS. Because of this difference in endpoints, the direction plot looks quite different for the two definitions. The

peak velocity-amplitude plot is also concentrated towards smaller amplitudes in definition 1. The vector joining start and endpoint will be quite small as the eye tends to go back near the fixation point (or window).

- Definition 2 - This definition is similar to the one used in detection by the EK method, and the distribution obtained is similar to the one obtained in the EK method. Since it is a stricter way of marking the MS end and does not include PSOs, the end of MS is defined as the point to which the big saccade was made (similar to the EK method MS endpoint). Here, the endpoint will be quite far from the start point because of which there is a fixed relation between peak velocity and amplitude.

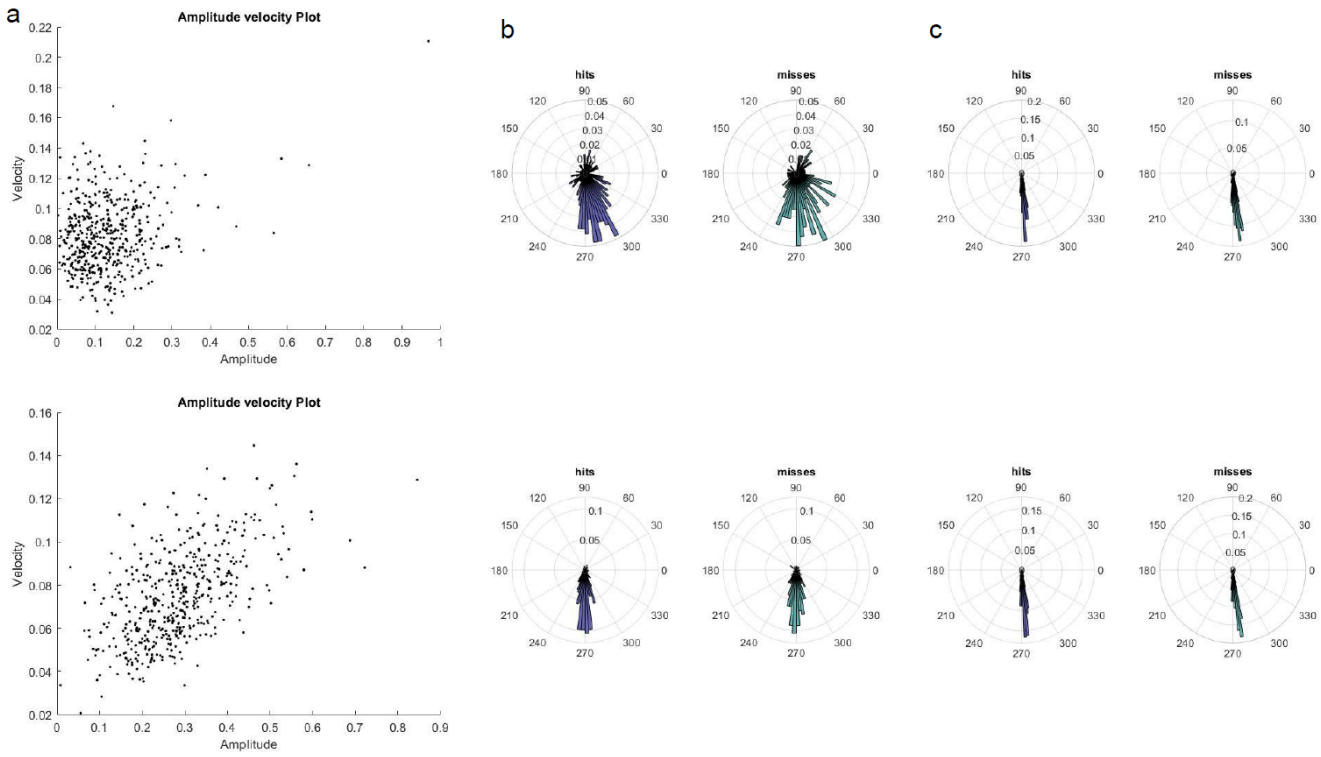

Figure 22. These plots are for CNN method (a) Peak velocity-amplitude plots for definitions 1 (row 1) and 2 (row 2). (b) Direction plots according to method 1 (Direction of MS is the angle of vector joining start and endpoint of MS) for definition 1 and 2. (c) Direction plots according to method 2 (Angle at peak velocity is defined as the direction of MS) for definition 1 and 2. All these plots are from session 1 (klecks_20170804_attentional-sampling_1).

4. Discussion

Attention scans different locations of an attended object using traveling waves according to the attentional scanning hypothesis. This can lead to beneficial processing of different locations over time. To explore the role of traveling waves in attentional scanning, we had to establish that the perception when the stimuli is attended fluctuates with the prestimulus phase of the spontaneous ongoing oscillations. This project revolves around showing that perception is dependent on the phase just prior to the time at which synaptic inputs arrive at V4.

The oscillations in the local field potential reflect the changes in the membrane potentials, which represent the variation in excitability of the neuronal ensemble (Bishop, 1932; Buzsáki and Draguhn, 2004; Rajkai et al., 2008; Schroeder and Lakatos, 2009). There are certain points in the oscillation where the neural excitability is high, which can be referred to as the high-gain phase or the phase of maximal gain. The phase of ongoing oscillation can affect the processing of the stimulus because when the input arrives at the postsynaptic neuron or neural ensemble during this phase of maximal gain, it receives beneficial processing. This phase can affect how well the stimulus is processed or even perceived. This means that perception is dependent on the phase of ongoing oscillation. Phase can affect other dependent variables like perception, behavioral response, neural response, etc.

In our task, the behavior of the monkey - whether the target is detected correctly (hits) or not (miss) reflects perception. We first established that this difference in the two behavioral

outcomes (or difference in perception) is reflected in the visually evoked responses (both in MUA and LFP) and in the LFP power in the lower gamma band. The difference in ERP signal shows there is a neural basis for the detection difference between hits and misses. The difference in ERP for different conditions has been shown previously using EEG (Busch and VanRullen, 2010) and MUA data for different regions including V4 (van Vugt et al., 2018). We also found a significant difference in microsaccade rate between hits and misses, which means the difference in detection is also reflected in MS rate (Figure 20). To exclude the possibility that the difference in neuronal responses was due to microsaccades, the ERP difference analysis was repeated by Elena only using trials which did not have microsaccades. The difference in ERP persisted in trials where there are no microsaccades, this confirms that the difference in ERP is not due to MS but to difference in perception.

Next, we established the existence of an optimal phase by looking at the link between neural response (which reflects perception) and the prestimulus phase. As shown in section 3.1.1, perception and ERP amplitude are related. There is a significant difference between the hit and miss ERP within the first 100 ms of target onset. Here the single trial behavioral outcome or monkey's response - hit or miss is a proxy for perception. If an optimal phase for information processing exists, then the prestimulus phase would predict perception. Since ERP amplitude and perception are related, then post-stimulus ERP amplitude and prestimulus phase should also be related. In this project, we tried to find if an optimal phase for processing exists and whether it is linked to neural response or ERP (which is linked to perception). To do so we looked at whether the trial-to-trial variability in post-stimulus ERP in a target detection task is linked to the prestimulus phase of ongoing neural oscillations. We did this analysis by calculating the circular-linear correlation between prestimulus phase and ERP amplitude, and observed significant correlations for a few channels in the frequency range of 6 to 10 Hz. This analysis was done for four different prestimulus time points and significant correlations were observed till about 100ms before stimulus onset. The correlation becomes better when only intermediate difficulty levels are considered. Apart from this, on taking average correlation for channels, we observed significant correlation in the 6 to 10Hz range. A human EEG study by Busch and VanRullen in 2010 also showed that the detection performance (which represents perception) fluctuated with the prestimulus phase of oscillations in the theta frequency band (~7 Hz) for the

attended stimuli. Our study used the neural response instead of performance of the animal which is a continuous measure of perception and showed that the perception is dependent on the prestimulus phase.

For the correlation analysis, trials with one location and all (or multiple) difficulty levels were pooled. The significant correlations map might get better after doing a regression analysis to control for difference in signal due to different difficulty levels. The statistical analysis was also done across frequencies for each time point, in the future we would like to do statistical analysis across time and frequencies to obtain time and frequency at which correlation is significant. Statistical analysis across frequency and channels also needs to be done to check which part of the array and which depths show significant correlations. The correlation analysis could be done for different locations which might give us more information about transition of the effect in time and space.

The correlation between ERP amplitude and prestimulus phase shows that perception is linked to the prestimulus phase which indicates that the optimal phase of processing might exist. This can also be checked by looking at whether hit trials are associated with the prestimulus phase; this analysis will be done in the future. Next, we can look at the role of traveling waves in attentional scanning. If the traveling waves are used to scan different object locations, then the optimal phase of neighboring neuronal populations will be systematically shifted in cortical space because of the retinotopic organization of V4. This project was a step towards exploring the attentional scanning hypothesis which can help us understand possible mechanisms using which attention can scan different items.

5. Conclusions and Further Plans

We showed that the difference in detection is reflected in the visually evoked responses (both in MUA and LFP) and in the LFP power in the lower gamma band. Miss trials show a higher MS rate just after target onset. These results are based on the analysis of all sessions. We also showed that the phase of the ongoing LFP activity before the target onset is systematically correlated to ERP amplitude which is related to perception. This points towards the existence of an optimal phase. Next we will look at whether this phase is systematically shifted between neighboring recording sites. If that is the case, it would suggest the existence of a traveling wave which would help us explore the role of traveling waves in attentional scanning.

We saw differences in the rate and direction of microsaccades around the time of target onset and this suggests that microsaccades might have a role to play in how well the target is detected. To explore this further we will look at the phase relation of MS with ongoing oscillation. We also plan on analyzing pupil size differences around target onset for different conditions.

References

Bellet, M.E., Bellet, J., Nienborg, H., Hafed, Z.M., and Berens, P. (2019). Human-level saccade detection performance using deep neural networks. J. Neurophysiol. *121*, 646–661.

Bichot, N.P., Rossi, A.F., and Desimone, R. (2005). Parallel and serial neural mechanisms for visual search in macaque area V4. Science *308*, 529–534.

Bishop, Geo.H. (1932). Cyclic changes in excitability of the optic pathway of the rabbit. Am. J. Physiol.-Leg. Content *103*, 213–224.

Börgers, C., and Kopell, N.J. (2008). Gamma oscillations and stimulus selection. Neural Comput. *20*, 383–414.

Busch, N.A., and VanRullen, R. (2010). Spontaneous EEG oscillations reveal periodic sampling of visual attention. Proc. Natl. Acad. Sci. U. S. A. *107*, 16048–16053.

Busch, N.A., Dubois, J., and VanRullen, R. (2009). The phase of ongoing EEG oscillations predicts visual perception. J. Neurosci. Off. J. Soc. Neurosci. *29*, 7869–7876.

Buzsáki, G., and Draguhn, A. (2004). Neuronal oscillations in cortical networks. Science *304*, 1926–1929.

Cardin, J.A., Carlén, M., Meletis, K., Knoblich, U., Zhang, F., Deisseroth, K., Tsai, L.-H., and Moore, C.I. (2009). Driving fast-spiking cells induces gamma rhythm and controls sensory responses. Nature *459*, 663–667.

Cohen, M.X. (2014). Analyzing Neural Time Series Data: Theory and Practice (Cambridge, MA, USA: MIT Press).

Davis, Z.W., Muller, L., Trujillo, J.-M., Sejnowski, T., and Reynolds, J.H. (2020). Spontaneous Traveling Cortical Waves Gate Perception in Behaving Primates. Nature *587*, 432–436.

Engbert, R., and Kliegl, R. (2003). Microsaccades uncover the orientation of covert attention. Vision Res. *43*, 1035–1045.

Felleman, D.J., and Van Essen, D.C. (1991). Distributed hierarchical processing in the primate cerebral cortex. Cereb. Cortex N. Y. N 1991 *1*, 1–47.

Fiebelkorn, I.C., Saalmann, Y.B., and Kastner, S. (2013). Rhythmic sampling within and between objects despite sustained attention at a cued location. Curr. Biol. CB *23*, 2553–2558.

Fries, P., Reynolds, J.H., Rorie, A.E., and Desimone, R. (2001). Modulation of oscillatory neuronal synchronization by selective visual attention. Science *291*, 1560–1563.

Fries, P., Nikolić, D., and Singer, W. (2007). The gamma cycle. Trends Neurosci. *30*, 309–316.

Hafed, Z.M., and Clark, J.J. (2002). Microsaccades as an overt measure of covert attention shifts. Vision

Res. *42*, 2533–2545.
Holcombe, A.O., and Chen, W.-Y. (2013). Splitting attention reduces temporal resolution from 7 Hz for tracking one object to <3 Hz when tracking three. J. Vis. *13*, 12.
Landau, A.N., and Fries, P. (2012). Attention samples stimuli rhythmically. Curr. Biol. CB *22*, 1000–1004.
Luck, S.J., Chelazzi, L., Hillyard, S.A., and Desimone, R. (1997). Neural mechanisms of spatial selective attention in areas V1, V2, and V4 of macaque visual cortex. J. Neurophysiol. *77*, 24–42.
Lund, J., Angelucci, A., and Bressloff, P. (2003). Anatomical Substrates for Functional Columns in Macaque Monkey Primary Visual Cortex. Cereb. Cortex N. Y. N 1991 *13*, 15–24.
Moran, J., and Desimone, R. (1985). Selective attention gates visual processing in the extrastriate cortex. Science *229*, 782–784.
Neumaier, A., and Schneider, T. (2001). Schneider, T.: Estimation of parameters and eigenmodes of multivariate autoregressive models. ACM Trans. Math. Softw. 27(1), 27-57. ACM Trans. Math. Softw. *27*, 27–57.
Ni, J., Wunderle, T., Lewis, C.M., Desimone, R., Diester, I., and Fries, P. (2016). Gamma-Rhythmic Gain Modulation. Neuron *92*, 240–251.
Oostenveld, R., Fries, P., Maris, E., and Schoffelen, J.-M. (2011). FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. Comput. Intell. Neurosci. *2011*, 156869.
Philipp, B. (2009). CircStat: a MATLAB toolbox for circular statistics. J. Stat. Softw. *31*.
Rajkai, C., Lakatos, P., Chen, C.-M., Pincze, Z., Karmos, G., and Schroeder, C.E. (2008). Transient cortical excitation at the onset of visual fixation. Cereb. Cortex *18*, 200–209.
Re, D., Inbar, M., Richter, C.G., and Landau, A.N. (2019). Feature-Based Attention Samples Stimuli Rhythmically. Curr. Biol. CB *29*, 693-699.e4.
Reynolds, J.H., and Heeger, D.J. (2009). The normalization model of attention. Neuron *61*, 168–185.
Reynolds, J.H., Chelazzi, L., and Desimone, R. (1999). Competitive mechanisms subserve attention in macaque areas V2 and V4. J. Neurosci. Off. J. Soc. Neurosci. *19*, 1736–1753.
Rohenkohl, G., Bosman, C.A., and Fries, P. (2018). Gamma Synchronization between V1 and V4 Improves Behavioral Performance. Neuron *100*, 953-963.e3.
Rolfs, M., Engbert, R., and Kliegl, R. (2005). Crossmodal coupling of oculomotor control and spatial attention in vision and audition. Exp. Brain Res. *166*, 427–439.
Salin, P.A., Girard, P., Kennedy, H., and Bullier, J. (1992). Visuotopic organization of corticocortical connections in the visual system of the cat. J. Comp. Neurol. *320*, 415–434.
Schneider, T., and Neumaier, A. (2000). Algorithm 808: ARfit - a MATLAB package for the estimation of parameters and eigenmodes of multivariate autoregressive models. ACM Trans. Math. Softw.

Schroeder, C.E., and Lakatos, P. (2009). Low-frequency neuronal oscillations as instruments of sensory selection. Trends Neurosci. *32*, 9–18.
Schwarz, G. (1978). Estimating the Dimension of a Model. Ann. Stat. *6*, 461–464.
Siegle, J.H., Pritchett, D.L., and Moore, C.I. (2014). Gamma-range synchronization of fast-spiking interneurons can enhance detection of tactile stimuli. Nat. Neurosci. *17*, 1371–1379.
Sirota, A., Montgomery, S., Fujisawa, S., Isomura, Y., Zugaro, M., and Buzsáki, G. (2008). Entrainment of neocortical neurons and gamma oscillations by the hippocampal theta rhythm. Neuron *60*, 683–697.
Taylor, K., Mandon, S., Freiwald, W.A., and Kreiter, A.K. (2005). Coherent oscillatory activity in monkey area v4 predicts successful allocation of attention. Cereb. Cortex N. Y. N 1991 *15*, 1424–1437.
VanRullen, R. (2013). Visual Attention: A Rhythmic Process? Curr. Biol. *23*, R1110–R1112.
VanRullen, R. (2016). Perceptual Cycles. Trends Cogn. Sci. *20*, 723–735.
van Vugt, B., Dagnino, B., Vartak, D., Safaai, H., Panzeri, S., Dehaene, S., and Roelfsema, P.R. (2018). The threshold for conscious report: Signal loss and response bias in visual and frontal cortex. Science *360*, 537–542.
Womelsdorf, T., Fries, P., Mitra, P.P., and Desimone, R. (2006). Gamma-band synchronization in visual cortex predicts speed of change detection. Nature *439*, 733–736.
Womelsdorf, T., Anton-Erxleben, K., and Treue, S. (2008). Receptive Field Shift and Shrinkage in Macaque Middle Temporal Area through Attentional Gain Modulation. J. Neurosci. *28*, 8934–8944.
Xue, C., Calapai, A., Krumbiegel, J., and Treue, S. (2020). Sustained spatial attention accounts for the direction bias of human microsaccades. Sci. Rep. *10*, 20604.
Zanos, T.P., Mineault, P.J., Nasiotis, K.T., Guitton, D., and Pack, C.C. (2015). A sensorimotor role for traveling waves in primate visual cortex. Neuron *85*, 615–627.
Zar, J.H. (1999). Biostatistical Analysis (Prentice Hall).
Zhang, H., Watrous, A.J., Patel, A., and Jacobs, J. (2018). Theta and Alpha Oscillations Are Traveling Waves in the Human Neocortex. Neuron *98*, 1269-1281.e4.