Population coupling of V1 and V4 neurons and its relation to local cortical state fluctuations and attention in macaque monkey

Mahyar Doost[1], Jochem van Kempen[1], Michael Boyd[1], Alexander Thiele[1]

[1]Bioscience Institute of Neuroscience, Newcastle University, Framlington Place, Newcastle upon Tyne, UK, NE2 4HH

Corresponding author: Alexander Thiele, alex.thiele@ncl.ac.uk

Disclosures and acknowledgments:

AT was funded by the MRC MR/P013031/1 and BBSRC BB/W006758/1.

The authors have no conflict of interests.

Author contributions: AT designed the study. JvK and MB acquired the data. AT and MD analyzed the data. AT AT and MD wrote the paper, JvK provided feedback on the paper.

Abstract

Neurons couple to various degrees to the activity level of the local neighboring population whereby strongly coupled 'choristers' and weakly coupled 'soloists' have been identified as two extremes of a continuous spectrum. At the same time neuronal populations undergo coordinated On and Off cortical state activity fluctuations, which are locally modulated by attention. The population coupling of soloists and choristers suggests that soloists should show limited alignment with cortical state fluctuations, while choristers should exhibit profound alignment. To test this, we recorded neurons across cortical layers in macaque areas V1 and V4, while animals performed a feature based spatial attention task. As expected, we found a wide range of population coupling strength of neurons. In line with our prediction, coupling of choristers to cortical state changes (ON-OFF transitions) was generally stronger than that of soloists. The strength of population coupling of neurons was similar during spontaneous and stimulus driven activity. Allocation of attention to the receptive field reduced the population coupling strength. Attentional modulation of neurons was positively correlated with population coupling strength. While neurons on averaged retained their coupling strengths across conditions, some neurons can change coupling strength condition dependent, thereby potentially enhancing the coding abilities of cortical circuits.

Significance statement

Cortical neurons can be categorized according to the level with which they couple their activity to the surrounding population, with extremes being referred to as soloists and choristers, respectively. Here we test to what extent this coupling is a constant feature across cortical states, stimulus and attention conditions. We find that at the population level the coupling is modulated by the conditions under which they are investigated. Critically, individual cells can behave as soloists (or choristers) on average while acting like choristers (soloists) during e.g. cortical state On-Off transitions. Thus, neurons can align with the surrounding population under some conditions, and behave independently under different conditions, thereby potentially enhancing the coding abilities of cortical circuits.

Introduction

Global brain states manifest as large-scale spatiotemporal activity pattern (Ferezou *et al.*, 2007; Mohajerani *et al.*, 2013). They evolve on multiple temporal scales, range from minutes to hours during different sleep stages (Lecci *et al.*, 2017), and from seconds to minutes during wakefulness (Harris & Thiele, 2011; Palva & Palva, 2012; McGinley *et al.*, 2015). Brain states affect neuronal responses, neural interactions, and neural information encoding (Poulet & Petersen, 2008; Niell & Stryker, 2010; Harris & Thiele, 2011; McGinley *et al.*, 2015; van Vugt *et al.*, 2018). They are modulated by arousal and link to cognitive function and conscious perception (McGinley *et al.*, 2015; van Vugt *et al.*, 2018). Cortical state fluctuations manifest in single cortical columns by spontaneously transitioning between states of vigorous (On) and faint (Off) spiking periods (Engel *et al.*, 2016; van Kempen *et al.*, 2021; Kharas *et al.*, 2022). Initially, it was assumed that these fluctuations affect activity homogenously throughout the cortex (Harris & Thiele, 2011; Lee & Dan, 2012). However, recent studies show that endogenous fluctuation of cortical states are locally modulated by attention, and the local momentary state (On or Off) predicts detection of visual stimuli and reaction times (Engel *et al.*, 2016; van Kempen *et al.*, 2021).

The alternation of population activity between On-Off spiking periods suggests that neurons behave in unison. However, various studies have shown that neurons have distinct activity pattern and their coupling strength with the neighboring neural population can vary radically. The coupling strength can be quantified as population coupling (PC) (Okun *et al.*, 2015). Strongly coupled 'choristers' and weakly coupled 'soloists' have been identified as two extremes of a continuous spectrum (Bachatene *et al.*, 2015; Okun *et al.*, 2015; Okun *et al.*, 2019; Kharas *et al.*, 2022; Dorman *et al.*, 2023). This raises the question, to what extent soloists exhibit On-Off spiking activity and whether soloists have a more continuous mode of functioning during On-Off fluctuations when compared to choristers. Given the local modulation of cortical state with attention it also raises the question whether soloists and choristers show different levels of attentional modulation. Flexibility in coupling strength might be an important component of increasing coding abilities. Soloists may have the ability to flexibly join multiple different neural ensembles to perform condition dependent encoding. If such flexibility was also permitted to some choristers, it would further expand the encoding potential of the circuit (Miller *et al.*, 2014). To investigate whether condition dependent flexibility of coupling strength is present in soloist and choristers, we recorded from laminar probes across cortical layers simultaneously in macaque area V1 and V4, while monkeys engaged in a feature based spatial attention task. We quantified how soloists and choristers couple to ON-OFF period fluctuations and examined whether PC of soloists and choristers was differently affected by spatial attention. We found that PC was largely a fixed feature across V1 and V4 populations, but coupling strength in some neurons could show condition dependent variation.

Materials and Methods

Subjects and general procedures

Two awake male monkeys (Macaca mulatta, age 10-12 years, weight 8-9.5 kg) were used as subjects in this study. They were housed under conditions described in detail previously (Gray *et al.*, 2016). All procedures were approved by the Newcastle University Animal Welfare Ethical Review Board (AWERB) and the Animal in Sciences

Committee UK, and carried out in accordance with the European Communities Council Directive RL 2010/63/EC, the US National Institutes of Health Guidelines for the Care and Use of Animals for Experimental Procedures, and the UK Animals Scientific Procedures Act.

Surgical procedures

The animals were implanted with a head post and recording chambers over area V1 and V4 under sterile conditions and general anesthesia. Surgical procedures and postoperative care conditions have been described in detail previously (Thiele *et al.*, 2006).

Behavioral paradigm

The behavioral paradigm (**Figure 1**) has been described previously (Ferro *et al.*, 2021; van Kempen *et al.*, 2021), but we copy the description here for ease of access.

Stimulus presentation and behavioral control was managed by Remote Cortex 5.95 (Laboratory of Neuropsychology, National Institute for Mental Health, Bethesda, MD). Stimuli were presented on a cathode ray tube (CRT) monitor at 120 Hz, 1280 × 1024 pixels, at a distance of 54 cm.

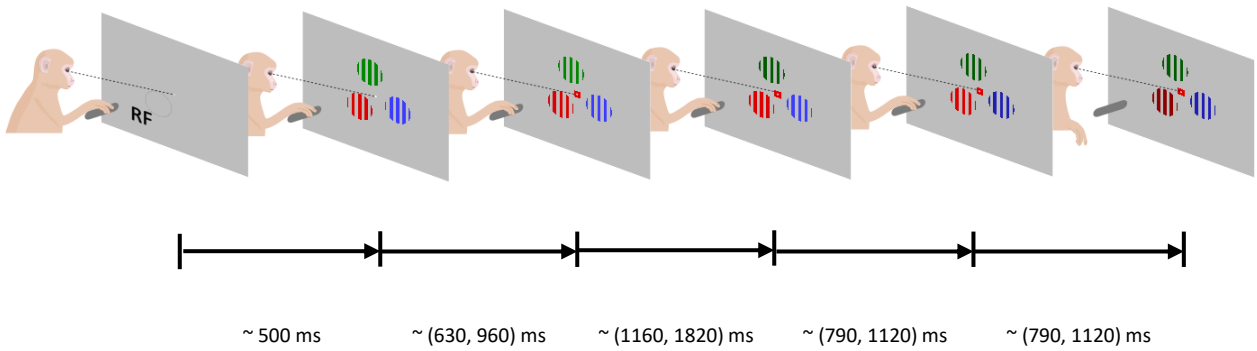

Figure 1: Feature-based spatial attention task. *The monkey held a lever to initiate the trial, whereafter a central fixation spot was turned on. Upon fixation 3 colored gratings appeared, one was centered on the receptive fields (RFs) of the V1 neurons. After a variable delay a cue matching one of the grating colors was presented centrally along with the fixation spot, indicating which grating was behaviorally relevant (target). In pseudorandom order the stimuli decreased in luminance (dimmed). Upon dimming of the target, the monkey had to release the lever to receive a reward.*

During the main task (**Figure 1**), the monkeys initiated a trial by holding a lever and fixating on a central white fixation spot (0.1°) displayed on a grey background (1.41 cd/m^2). After a fixed delay (614, 424 ms for monkey 1 and 2), three colored square wave gratings (for color values see **Table 1**) appeared equidistant from the fixation spot, one was centered on the RF of the V1 neurons under study.

	Red	Green	Blue
Monkey 2 and monkey 1 (n=59)	a. [220 0 0] – 12.8 b. [140 0 0] – 4.2	a. [0 135 0] – 12.9 b. [0 90 0] – 4.6	a. [60 60 255] – 12.2 b. [30 30 180] – 4.6

Monkey 1 (n=1)	b. [170 0 0] –6.7	b. [0 105 0] – 6.4	b. [37 37 210] – 6.6
Monkey 1 (n=1)	b. [175 0 0] –7.2	b. [0 105 0] – 6.4	b. [40 40 220] – 7.7
Monkey 1 (n=8)	b. [180 0 0] –7.7	b. [0 110 0] – 7.3	b. [40 40 220] – 7.7

Table 1. Color values used for the three colored gratings across recording sessions and subjects Methods. Color values are indicated as [RGB] – luminance (cd/m2). a = Undimmed values, b = Dimmed values. For monkey 1, we used a variety of dimmed values across recordings.

The locations of colored gratings were fixed for each recording session but were pseudo-randomly varied across sessions. Stimulus size varied between 2 and 4° diameter, depending on RF eccentricity and size. For most recordings we used drifting gratings but presented one monkey (monkey 2) with stationary gratings during 22 out of 34 recording days. The drifting gratings moved perpendicular to the grating orientation, with the motion direction pseudo-randomly assigned on every trial. After a random delay (618-1131 ms for monkey 1, 618-948 ms for monkey 2; uniformly distributed), a central cue appeared that matched the color of one of the gratings, indicating that this grating would be behaviorally relevant on the current trial. After a variable delay (1162-2133 ms for monkey 1, 1162-1822 ms for monkey 2; uniformly distributed), one of the gratings changed luminance (for color values see Table 1), referred to as dimming. If the cued grating (target) dimmed, the monkey had to release the lever in order to obtain a reward. If, however, a non-cued grating (distractor) dimmed, the monkey had to ignore this and keep hold of the lever until the target dimmed on the second or third dimming event (each after another 792-1331 ms for monkey 1; 792-1164 ms for monkey 2; uniformly distributed).

Data acquisition and analysis

We recorded across cortical layers of visual areas V1 and V4 using 16-contact laminar electrodes (150 µm contact spacing, Atlas silicon probes). We included recordings where the Hidden Markov model analysis (see below) suggested that a 2-phase model best approximated the data. This applied to 34 penetrations from monkey 1 and 35 from monkey 2. Of these, 28 penetrations were performed in V1 and 29 were performed in V4 of monkey 1 with XXX performed simultaneously in V1 and V4. In monkey 2 all 35 included penetrations were from simultaneous recordings in V1 and V4. The electrodes were inserted perpendicular to the cortex on a daily basis.

Raw data were collected using Remote Cortex 5.95 and by Cheetah (Neuralynx) data acquisition interlinked with Remote Cortex 5.95. Neuronal data were acquired with Neuralynx preamplifiers and a Neuralynx Digital Lynx amplifier. Unfiltered data were sampled with 24 bit at 32.7 kHz and stored to disc. Data were replayed offline, sampled with 16-bit and band-pass filtered at 0.5-300 Hz and down sampled to 1 kHz for local field potential (LFP) data, and filtered at 0.6-9 kHz for spike extraction sampled at 32.7 kHz. Eye position and pupil diameter was recorded at 220 Hz (ViewPoint, Arrington Research).

All data analyses were performed using custom written Matlab (the Mathworks) scripts.

Data preprocessing

We corrected for any noise common to all channels via common average referencing, in which the average of all channels of an electrode is subtracted from each individual channel of that electrode.

In order to determine signal-to-noise (SNR) ratios, recording stability, the visual response latency, as well as for the computation of receptive fields (i.e. for preprocessing purposes only), we computed the envelope of MUA (MUAe) by low-pass filtering (<300 Hz, fifth order Butterworth) the rectified 0.6-9 kHz filtered signal. During some recording sessions the electrode moved (e.g. due to movement of the monkey). To account for this, we visually inspected the stability of each recording by investigating the stimulus aligned firing rates, MUAe and their baseline (-500 to -50 ms) energy across all trials and channels. With energy (E) defined as:

$$E = \int_{i}^{t} V(i)^2$$

, where t is the number of time points in the vector (V) representing the single-trial histogram (or MUAe). We selected the largest continuous time window that showed stable activity across all V1 and V4 channels.

Single unit activity (SUA) was extracted by band pass filtering the neural signal between 600 and 9,000 Hz, followed by manual spike sorting using SpikeSort3DSetup (Neuralynx) software. We included SUAs with at least 5Hz firing rate in any time window of interest (see below) and for which we recorded at least 10 trials per attention condition. Multi-unit spiking activity was extracted in the same way, but here the criteria of single unit isolation (less than 0.5% of entries in the first ms bin of the interspike interval) were not fulfilled.

Layer compartments alignment

To identify recording channels located in the gray matter of V1 and V4, a current source density (CSD) analysis was carried out. Here current sinks and sources allowed to determine the relative recording depth, which was then compared to the known cortical anatomy (Schroeder *et al.*, 1998). The CSD profile can be calculated according to the finite difference approximation, taking the inverse of the second spatial derivative of the stimulus-evoked voltage potential φ, defined by:

$$CSD(x) = \frac{\varphi(x+h) - 2\,\varphi(x) + \varphi(x-h)}{h^2}$$

, where x is the depth at which the CSD is calculated and h the electrode spacing (150 µm). We used the iCSD toolbox (Pettersen *et al.*, 2006) to compute the CSD. With this toolbox we used a spline fitting method to interpolate φ smoothly between electrode contacts. We used a diameter of cortical columns of 500 µm (Mountcastle, 1957), and tissue conductance of 0.4 Sm^{-1} (Logothetis *et al.*, 2007).

To aid determination of recording depth, we also computed the signal-to-noise ratio (SNR), the response latencies to stimulus onset for each channel and the receptive field (RF, see below). SNR was computed as:

$$SNR = \frac{Signal - Noise}{\sigma_{noise}}$$

, with $Signal$ defined as the average MUAe amplitude in one of eight 50 ms time windows, from 30 to 80 ms, in 10 ms steps, to 100 to 150 ms after stimulus onset, and $Noise$ as the average MUAe amplitude during the baseline period (-200 to 50 ms) before stimulus onset. SNR in at least one of these eight estimates was required to be higher than 3 for a channel to be included for further analyses.

We computed the response latency to stimulus onset for each channel according to a method previously described (Roelfsema *et al.*, 2007). We fitted the visual response as a combination of an exponentially modified Gaussian and a cumulative Gaussian using a non-linear least-squares fitting procedure (function lsqcurvefit) applied to the average MUAe time course. There are two assumptions implicit in this method. First, the onset latency has a Gaussian distribution across trials and across neurons that contribute to the MUAe, and second, that (part of) the response dissipates exponentially. The visual response y across time t was modelled as:

$$y(t) = d \cdot Exp(\mu\alpha + 0.5\sigma^2\alpha^2 - \alpha t) \cdot G(t, u + \sigma^2\alpha, \sigma) + c \cdot G(t, \mu, \sigma)$$

, where μ is the mean, σ is the standard deviation, α^{-1} is the time constant of the dissipation, $G(t, \mu, \sigma)$ is a cumulative Gaussian, and c and d are the factors scaling the non-dissipating and dissipating modulation of the visual response. The response latency was defined as the time point where $y(t)$ reached 33% of the maximum of the earliest peak of the Gaussian (Roelfsema *et al.*, 2007). The earliest current sink/earliest latency was taken as the presumed thalamic input layer (L4) reference channel. In area V1, channels located 0.25 to 1 mm above the reference channel were assigned to supragranular, channels located within 0.25 mm were assigned to granular, and channels located 0.25 to 0.75 mm below the reference channel were assigned to the infragranular layer compartment. In area V4, channels at distances of 0.1 to 1 mm above the reference channel were labeled as supragranular, channels above or below the reference channel within 0.1 mm were labeled as granular, and channels at distances of 0.1 to 0.75 mm below the reference channel were labeled as being in the infragranular layer compartment. Finally, channels outside the estimated range of gray matter of V1 or V4 were excluded from analysis.

Receptive field mapping

Receptive field (RF) mapping was carried out by a reverse correlation method described in detail previously (Gieselmann & Thiele, 2008). In brief, while monkeys fixated on a fixation point, the RF location and size of neurons in a cortical column was identified performing reverse correlation of spiking activity relative to the presentation of 0.5-2° black squares stimuli (100ms stimulus on with 100ms interstimulus interval) at pseudorandom locations on a 9 × 12 grid (5-25 repetitions for each location) on a grey background.

Offline RFs were determined for each channel via reverse correlation of the MUAe signal (see above for MUAe computation) to stimuli (0.5 – 2 ° black squares). The stimulus-response map was converted to z-scores, after which the RF for each channel was indicated by a contour (thresholded at a z-score of 3) surrounding the peak

activity. RF eccentricity ranged from 3.4 - 7.5° in V1, and from 2.5 to 8.9° in V4, and were located in the lower left quadrant.

Population coupling computation

To calculate the PC of a neuron, we first computed the population rate of neighbouring neurons (those recorded at different electrode channels, i.e. always excluding the activity from the channel the single unit was recorded from). The population rate was calculated by accumulating the single-unit and multi-unit activity of each included channel (electrode contact) at 1ms resolution. The population rate was then smoothed with a Gaussian kernel of half-width $12ms$, followed by subtracting the mean of the smoothed population rate. Then, for every single neuron, its firing rate was calculated within 1ms bins, smoothed with a Gaussian of half width of $12/\sqrt{2}\,ms$ followed by mean subtraction. Finally, we calculated the cross correlation of the mean subtracted smoothed spike train of each neuron with its mean subtracted smoothed population rate. The cross correlation of the smoothed spike train of a neuron with its smoothed neighbouring accumulated spikes was defined by:

$$R_{xy}\,(m) = \begin{cases} \dfrac{1}{\|x_n\|} \displaystyle\sum_{n=0}^{N-m-1} (x_{n+m} - \mu_x)\,(y_n - \mu_y)^* & , \quad m \geq 0 \\ R_{xy}^*(-m) & , \quad m < 0 \end{cases}$$

Here, x_n represents the smoothed spike train of a neuron, y_n represents the smoothed population rate, $\|x_n\|$ represents the norm of x_n (that is, the number of spikes fired), μ_x represents the baseline level of the smoothed spike train of a neuron, μ_y represents the baseline level of the smoothed population rate, m represents the lag of cross correlation, and *N* is the number of samples of x_n or y_n (x_n and y_n have the same number of samples). The output vector of the cross correlation was given by:

$$C\,(m) = R_{xy}\,(m - N)\ , \quad m = 1,\ 2,\ \ldots,\ 2N - 1$$

Here, C represents the output of the cross correlation and m represents the lag of cross correlation. Hence, the population coupling of a neuron at time lag $m = N$ (the time of the lag is 0) was:

$$C\,(N) = R_{xy}\,(0) = \frac{1}{\|x_n\|} \sum_{n=0} (x_{n+N} - \mu_x)\,(y_n - \mu_y)^*$$

We also normalized the strength of population activity by normalizing the cross correlation of the smoothed SUA and smoothed population activity by their covariance. The normalized degree of population coupling of each SUA was given by:

$$C_{norm}\,(N) = R_{xy}\,(0) = \frac{1}{\|x_n\|} \times \frac{\sum_{n=0}(x_{n+N} - \mu_x)\,(y_n - \mu_y)^*}{\sqrt{\sum\,(x_n - \mu_x)^2\,\sum(y_n - \mu_y)^2}}$$

Finally, we calculated the Fisher transform of the correlation coefficient of population coupling ($C_{norm}(N)$) of each SUA as:

$$Z\,(N) = \operatorname{atanh}(\,C_{norm}\,(N)\,)$$

Hidden Markov Model and ON-OFF state estimation

To quantify On-Off dynamics in V1 and V4, we fit a Hidden Markov Model (HMM) to the population activity (MUA) across all laminae (for additional details see van Kempen *et al.*, 2021). For the purpose of HMM fitting we extracted population activity by progressively lowering spike extraction thresholds until approximately 100 Hz spiking activity was detected on each channel between fixation onset and the first dimming event. We fit the HMM both to activity from each individual area, following the procedures described by Engel et al. (2016), as well as to the activity from both areas simultaneously.

Our HMM assumes that spike counts on the recorded channels can be well characterized as a doubly-stochastic process, of which the parameters can be accurately estimated (Rabiner, 1989). In this study, spike counts on each channel are assumed to be produced by a Poisson process with different (constant) mean rates during On or Off phases of the underlying 'hidden' (latent) process s common to all channels that we need to infer (Engel *et al.*, 2016). The mean firing rate on each channel j in phase s is defined by entry λ_j^s in the emission matrix Λ. The transition matrix P gives the probabilities of transitioning between these latent phases. In the transition matrix P, each entry indicates the probability of transitioning between two specific phases. For instance, P_{11} indicates the probability of transitioning from $s = 0$ to $s = 0$ (remaining in the Off phase), whereas P_{12} indicates the probability of transitioning from $s = 0$ to $s = 1$, more formally: $P_{11} = P_{off} = P(s_{t+1} = 0|s_t = 0)$, $P_{12} = 1 - P_{off} = P(s_{t+1} = 1|s_t = 0)$. These probabilities do not depend on time: at any time step t, the probability of transitioning between phases depends only on the value of s at time t (s_t). The latent dynamics estimated by the HMM thus follow a discrete time series in which s_t summarises all information before time t. For each channel, MUA was discretized by determining spike counts in 10 ms bins following each time t, with the probability of observing spike count n on channel j during phase s defined as

$$P(n|s) = \frac{(\lambda_j^s)^n}{n!} e^{-\lambda_j^s}$$

The full description of an HMM is given by the emission matrix Λ, transition matrix P and the probabilities π^0 that indicate the initial values s_0, in which $\pi_i^0 \equiv P(s_0 = i)$. These parameters were estimated using the Expectation Maximization (EM) algorithm (Bishop, 2006), maximizing the probability of observing the data given the model according to the Baum-Welch algorithm (Rabiner, 1989). Because the EM procedure can converge to a local maximum, rather than the global maximum, we repeated the EM procedure ten times with random parameter initializations and chose the model with the highest likelihood. Random values were drawn from Dirichlet distributions for π^0 and P, and from a uniform distribution between zero and twice the channel's mean firing rate for Λ. The EM procedure was terminated if the relative change, computed as $|new - original|/|original|$, in the log-likelihood was smaller than 10^{-3} and the change in the transition and emission matrix was smaller than 10^{-5}, or if it reached the maximum number of iterations (n = 500).

Once the optimal parameters were estimated, we used the Viterbi algorithm to determine the most likely latent trajectory for each individual trial. We applied the HMM separately to each attention condition. For every trial,

we applied the HMM during multiple time periods of the task, during fixation and during the time window from 400 ms after cue onset to 30 ms after the first dimming event.

To determine what number of latent phases best described the data, we fit HMMs with the number of phases ranging from 1 to 8, and used a four-fold cross-validation procedure to compute the leave-one-channel-out cross-validation error for each HMM (Engel *et al.*, 2016). We fit the HMM to a randomly selected subset of 3/4 of the trials and computed the cross-validation error on the remaining 1/4 of trials. This procedure was repeated 4 times using a different 3/4 of trials for training and 1/4 of trials for testing the HMM. We computed the cross-validation error CV_{var} for each channel j across all trials K and time bins T as the difference between the actual and expected spike count according to:

$$CV_{var}[n_j] = \sum_{k=1}^{K} \sum_{t=1}^{T} (n_t^j - \lambda_j^{st})^2$$

We normalized CV_{var} to the error in the 1-phase HMM, averaged across channels, cross-validations and conditions, and determined the difference in CV_{var} with each additional phase in the HMM. For most recordings, and for both V1 and V4, CV_{var} decreased with the addition of a second phase but did not decrease much further with additional phases. This allowed the identification of the elbow (kink) in this error plot as the model with two phases. We included areas/recordings for further analysis that revealed a reduction in cross-validation error of at least 10% with the addition of a second phase but did not decrease by more than 10% with additional phases. For a small subset of recordings, a three or a four-phase model fit the data best [V1: n= 7; V4: n = 1], suggesting that these recordings could contain states with systematically occurring nested fluctuations; these recordings were excluded from further analysis (and are not included in the previously mentioned number of penetrations). In total, we found a reduction of >10% in cross-validation error when fitting a 2-phase versus 1-phase model in 63 V1, and 64 V4 recordings. For these recordings, epoch duration distributions closely followed an exponentially decaying function, consistent with HMM assumptions, indicating that short epoch durations were most prevalent.

While the two-state HMM segments the data into discrete On and Off phases, our results do not depend on the assumption of discrete phases. Previous work showed that the On-Off dynamics can be also modeled with a continuous latent variable, in which case the inferred firing rates showed bimodality and dynamics consistent with that inferred by the HMM (Engel *et al.*, 2016).

Ratio of firing rate of On-Off phases

In each penetration, HMM was fitted to the MUA to extract On and Off phases durations in each trial in two different time windows: 1) Monkey1: 200ms to 500ms after fixation onset, Monkey2: 100ms to 400ms after fixation onset, 2) 400ms after cue onset until 30ms after first dimming onset (identical for both monkeys). Next, we divided the average firing of the On phase by the average firing of the Off phase. This provides the On-Off firing rate ratio for every trial in the two time periods. We averaged the ratios across trials for each attention condition (Attention-in condition and Attention-out condition).

Attentional modulation

The effect of selective attention on neural activity was computed via an attention modulation index ($attMI$), defined as:

$$attMI = \frac{A_{RF} - A_{away}}{A_{RF} + A_{away}}$$

, with A_{RF} as the neural activity when attention was directed towards the RF, and A_{awayt} the activity when attention was directed away from the RF. This index ranges from -1 to 1, with zero indicating no attentional modulation and with positive (negative) values indicating higher (lower) activity when attention was directed towards the RF.

Results

Population coupling

We first describe basic aspects of PC of individual neurons during different epochs of the task, and different attentional conditions. In line with previous reports by Okun and colleagues in 2015 (Okun *et al.*, 2015) we found that PC varied widely between individual neurons (Figure 2B-E).

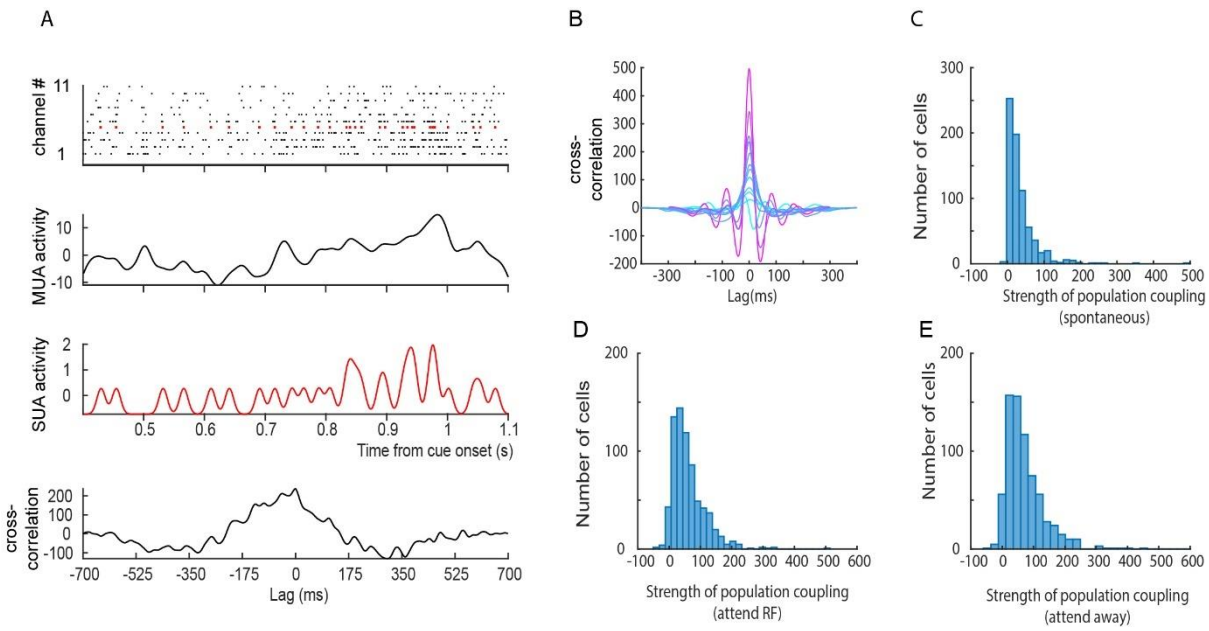

Figure 2: Calculation of population coupling and its distribution for different experimental epochs and conditions. A) Raster plots within a single trial of MUA activity (black) across different channels located in grey matter. Red raster plot shows SUA from a channel, recorded along the MUA activity. Averaged, smoothed and mean subtracted MUA activity across channels is shown below raster plots. Smoothed and mean subtracted SUA activity is shown in red. Bottom plot in A shows cross correlation between MUA activity and SUA activity for this trial. **B)** Range of population coupling shown for selected cells. Couling strength was determined from the height of the cross-correlogram at 0ms time lag. Color coding shows coupling

from least coupled example cells (light blue) to most coupled cells (magenta). **C)** Distribution of coupling strength during spontaneous activity. **D)** Distribution of coupling strength after cue onset during attend RF conditions. **E)** Distribution of coupling strength after cue onset during attend away conditions.

Figure 2 C-E shows that coupling strength shows roughly similar distributions during spontaneous activity periods when compared to activity that occurs after cue onset under attend RF and attend away conditions, but slight differences may be present (analyzed in detail below). To further investigate to what extent coupling strength is a largely fixed feature of a neuron, irrespective of the condition it occurs in, we calculated the correlation between coupling strength during spontaneous activity and after cue onset. To account for the possibility that coupling strength is affected by the firing rate of neurons, we normalized the strength of population coupling of SUAs for each condition, by calculating the correlation coefficient between MUA and SUA (C_{norm} and its Fisher Z-transform Z(N), see Methods). We then computed the correlation of PC (Z(N)) during the spontaneous activity and the activity after cue onset for the different attention conditions. The strength of population coupling of SUAs in two different time periods was highly correlated for both areas (Figure 3A-D), and the correlations were similar and significant for each monkey. This demonstrates that under these conditions the PC is largely a fixed feature of a neuron.

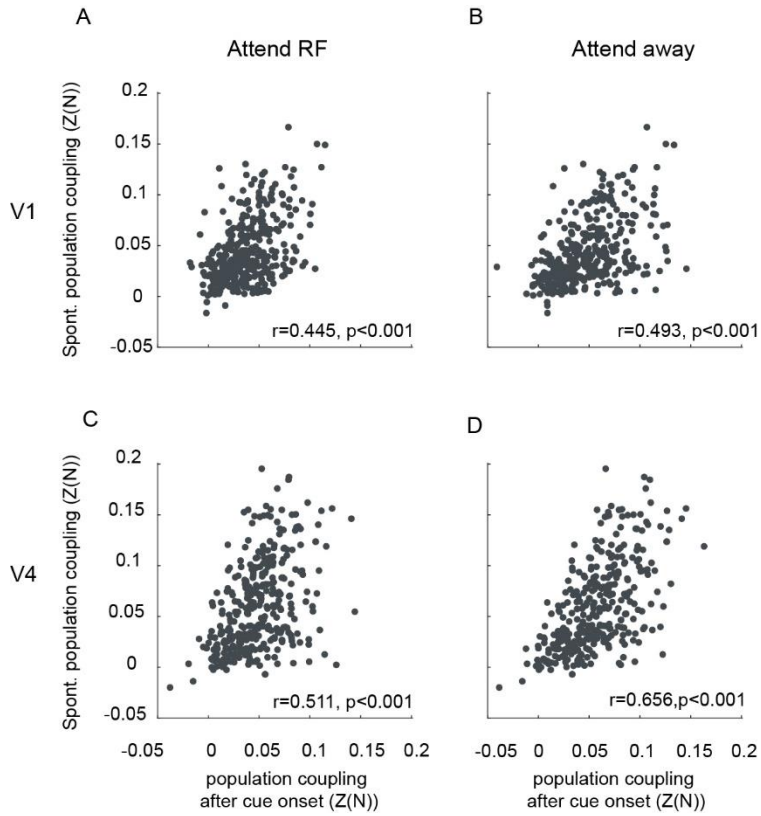

Figure 3: Relation of Population coupling quantified by calculating the Fisher transformed correlation coefficient between single unit and population activity (Z(N)) during spontaneous activity compared to stimulus driven activity. A) Correlation of Population coupling during spontaneous activity and during attend RF stimulus driven activity for V1 neurons (period after cue onset). **B)** Correlation of Population coupling during spontaneous activity and during attend away stimulus driven activity for V1 neurons (period after cue onset). **C)** Correlation of Population coupling during spontaneous activity and during attend RF stimulus driven activity for V4 neurons (period after cue onset). **E)** Correlation of Population coupling during spontaneous

activity and during attend away stimulus driven activity for V4 neurons (period after cue onset). Insets show correlation coefficient between the two measures, and respective p-values.

Despite this pattern, PC may still be modulated by cognitive factors such as attention or cortical state fluctuations, which we explore below.

Effect of spatial attention on population coupling

Previously Okun et al. (Okun *et al.*, 2015) showed that behavioral variables such as motor intention influenced population coupling of V4 neurons. Here we evaluate the effect of spatial attention on the strength of population coupling of neurons. To compare the strength of population coupling of neurons between attend-RF and attend-out conditions, we used the Fisher transformed PC correlation coefficient of a neuron (Z(N), see Methods), and compared these between attention conditions. Figure 4 shows the scatter plot of normalized PCs for attend RF vs attend out conditions for areas V1 and V4, respectively. The majority of data points lie above the diagonal, and consequently PC was significantly stronger during attend out conditions compared to attend Rf conditions in both areas ($p<0.001$, Wilcoxon signed rank test, results were significant for each monkey at $p<0.001$). This demonstrates that allocation of attention to the RF decreased population coupling of neurons.

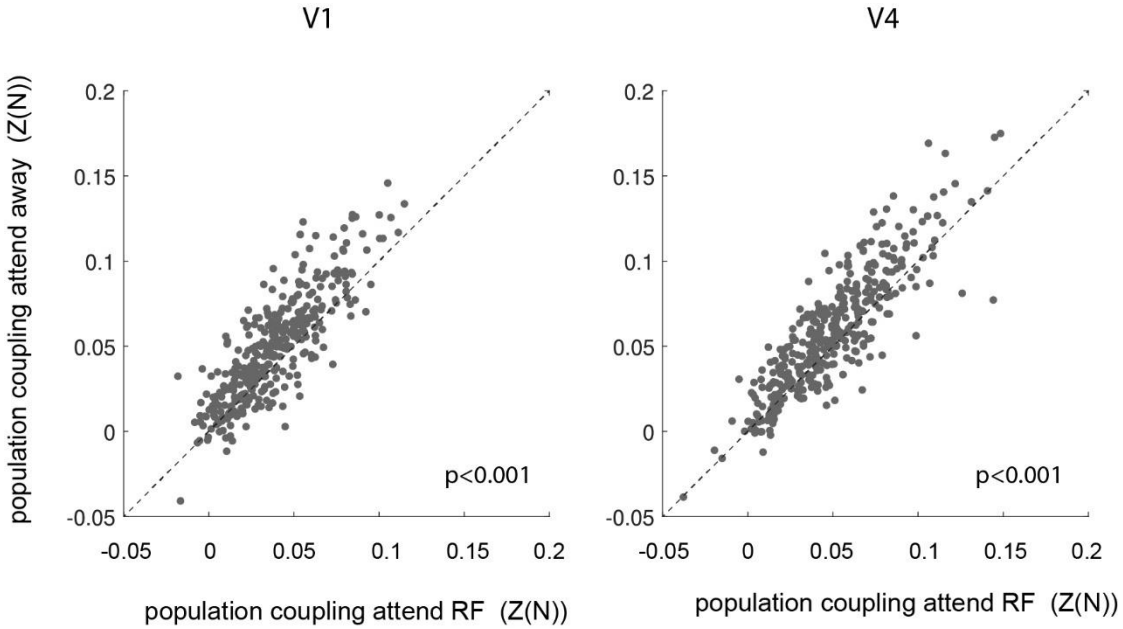

Figure 4: Effect of spatial attention on population coupling of V1 and V4 neurons. Population coupling (Fisher-transformed correlation coefficient Z(N)) during attend RF conditions (x-axis) and attend away conditions (y-axis) during the period after cue-onset for V1 neurons (left) and V4 neurons (right). Insets show p-values of differences in population coupling.

Coupling of soloists and choristers to population state change transitions

Recent studies have demonstrated that attention locally affects brain state. Brain state manifests as On-Off transitions of activity, whereby attention to the receptive field increases On-period durations in V1 and V4, in addition to increasing neural activity during both On- and Off-periods (Engel *et al.*, 2016; van Kempen *et al.*, 2021). Given the broad range of PCs across neurons, we would expect that soloists follow On-Off transitions to a lower degree than choristers. Note, that while this prediction is straightforward, it is not a tautology, as the PC

calculations has been performed without regard of ON-OFF period activity itself, and it has not been tested explicitly before. To examine the coupling of soloists and choristers with cortical state fluctuations during stimulus driven activity, we first calculated the ratio of firing rates during ON to OFF phases for each neuron. This provides a measure how much a neuron changes its activity with ON-OFF transitions. We then calculate the correlation of this ratio with the PC (Z(N)) of the neurons. The results are shown in (**Figure 5**). As predicted, on average cells with lower population coupling, also change their activity rates less during ON-OFF phase transitions during either attention condition, which is evident from the positive correlation between PCs versus ratio of firing rates during ON to OFF phases ($p<0.001$ for both areas and attention conditions). However, the correlation coefficients themselves only explain between 7% to 31% of the variance, and figure 5 demonstrates that some soloists (PC ~0) nevertheless show profound activity differences during ON to OFF phase transitions (ratios of 4 to 6).

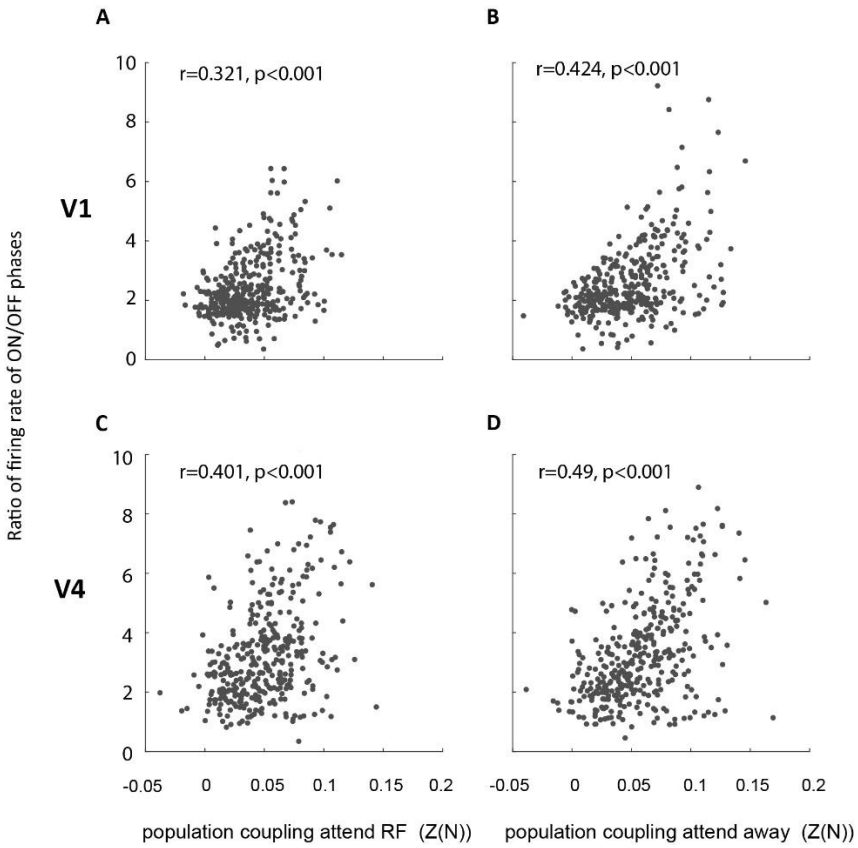

Figure 5: Relation between ON-OFF state fluctuation alignment and population coupling of V1 and V4 neurons. Population coupling (Fisher-transformed correlation coefficient Z(N)) during attend RF conditions and attend away conditions during the period after cue-onset for V1 neurons (left) and V4 neurons (right) compared to the ratio of ON-OFF period firing (y-axis). Insets show correlation coefficients and p-values of this relationship.

Attentional modulation of soloists and choristers

Given that soloists and choristers (on average) differentially align with cortical state On - Off phase transitions, which themselves are affected by spatial attention, we asked whether soloists and choristers are differently affected by spatial attention. To this end we calculated an attentional modulation index for each neuron (Methods) and correlated this with the PC of the neurons (**Figure 6**). A positive correlation would indicate that choristers show stronger attentional modulation, while a negative correlation would indicate the opposite. We find that in both V1 and V4 there is a positive correlation between the attentional modulation index and the PC strength (V1: r= 0158, p=0.003, V4: r= 0199, p<0.001). Thus, there is a tendency for choristers to exhibit stronger attentional modulation when compared to soloists, but overall, this tendency is modest (the correlation explains 2.5-4% of the variance).

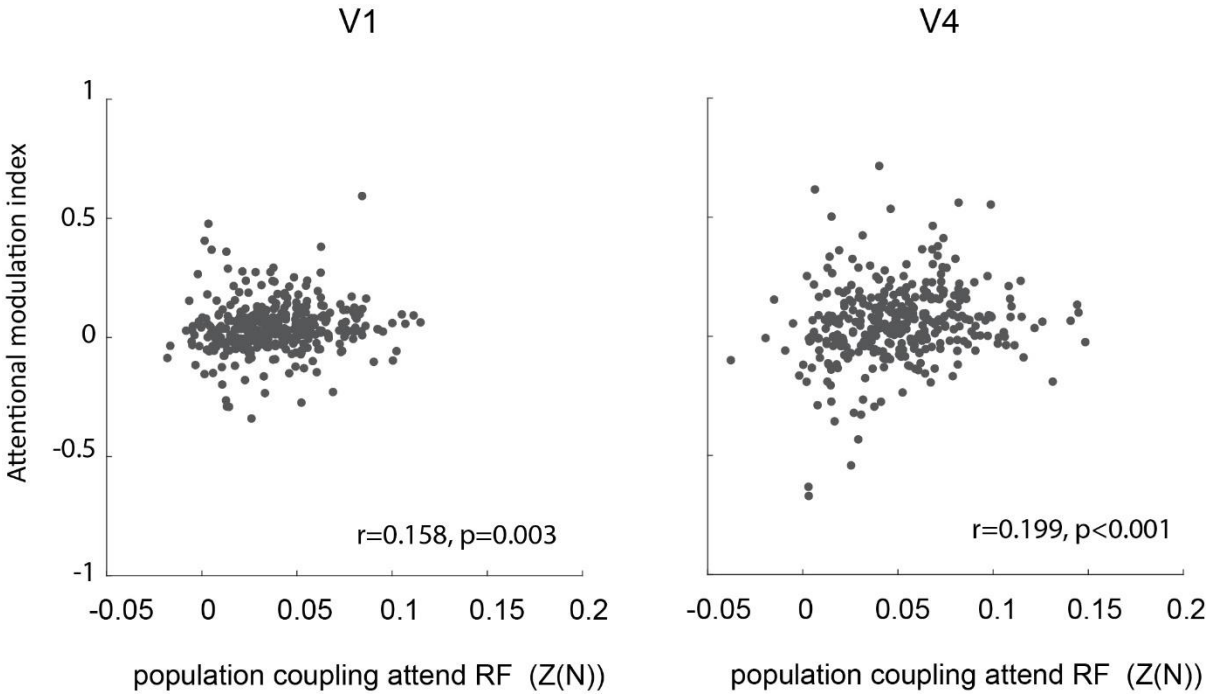

Figure 6: Relation between On-Off state fluctuation alignment and attentional modulation strength of V1 and V4 neurons. Population coupling (Fisher-transformed correlation coefficient Z(N)) during attend RF conditions during the period after cue-onset for V1 neurons (left) and V4 neurons (right) compared to the strength of attentional modulation (y-axis). Insets show correlation coefficients and p-values of this relationship.

Discussion

We show that the strength of population coupling of neurons varied widely in macaque area V1 and V4. In general, the strength of population coupling was similar during spontaneous and stimulus driven activity. Population coupling during attend RF condition was smaller than during attend away condition. Choristers coupled more strongly with cortical state changes (ON-OFF transitions) than soloists and showed stronger attentional modulation. Despite these general tendencies we also found that PC can vary in a condition dependent manner for some neurons.

In line with previous studies we find that PC of neurons forms a continuous spectrum ranging from strongly coupled 'choristers' to weakly coupled 'soloists' (Bachatene *et al.*, 2015; Okun *et al.*, 2015; Okun *et al.*, 2019; Kharas *et al.*, 2022; Dorman *et al.*, 2023). These can be identified in both areas V1 and V4 of the macaque, whereby PC strength was strongly correlated between spontaneous and stimulus driven activity, as well as during different attention conditions. However, PC was generally stronger during attend away conditions when compared to attend RF conditions, i.e. PC is modulated by cognitive factors. This is in line with previous reports, where motor intention affected PC strength of V4 neurons (Okun *et al.*, 2015).

Here we replicate results showing that cortical state fluctuations manifest in single cortical columns by spontaneously transitioning between states of vigorous (On) and faint (Off) spiking periods (Engel *et al.*, 2016; van Kempen *et al.*, 2021; Kharas *et al.*, 2022). These transitions are locally affected by spatial attention, and we found that soloists generally couple less to the On-Off phase activity transitions when compared to choristers, as could have been predicted by the underlying population coupling behavior. However, it is important to note that there are exceptions to the rule, whereby neurons with close to zero PC values still show profound activity changes during On-Off transitions, and neurons with strong PC values, exhibit very limited activity changes during On-Off transitions (see some individual data points in figure 5).

Given the local modulation of cortical state with attention, and the stronger coupling of choristers with On-Off state transitions we predicted that choristers should show stronger attentional modulation when compared to soloists. This prediction turned out to be correct, but the overall alignment was limited, i.e. PC, while a significant predictor of attentional modulation, only explained a limited amount of variance of attentional modulation.

In summary we found that PC of macaque V1 and V4 neurons on average shows limited variation across different stimulus periods, while it is affected by cognitive factors such as spatial attention. This may allow some neurons to flexibly engage in multiple different neural ensembles, which would greatly increase the coding capacity of a given neural population (Miller *et al.*, 2014).

References

Bachatene, L., Bharmauria, V., Cattan, S., Chanauria, N., Rouat, J. & Molotchnikoff, S. (2015) Electrophysiological and firing properties of neurons: Categorizing soloists and choristers in primary visual cortex. *Neuroscience Letters*, **604**, 103-108.

Bishop, C.M. (2006) *Pattern Recognition and Machine Learning* Springer.

Dorman, R., Bos, J.J., Vinck, M.A., Marchesi, P., Fiorilli, J., Lorteije, J.A.M., Reiten, I., Bjaalie, J.G., Okun, M. & Pennartz, C.M.A. (2023) Spike-based coupling between single neurons and populations across rat sensory cortices, perirhinal cortex, and hippocampus. *Cerebral Cortex*, **33**, 8247-8264.

Engel, T.A., Steinmetz, N.A., Gieselmann, M.A., Thiele, A., Moore, T. & Boahen, K. (2016) Selective modulation of cortical state during spatial attention. *Science*, **354**, 1140-1144.

Ferezou, I., Haiss, F., Gentet, L.J., Aronoff, R., Weber, B. & Petersen, C.C. (2007) Spatiotemporal dynamics of cortical sensorimotor integration in behaving mice. *Neuron*, **56**, 907-923.

Ferro, D., van Kempen, J., Boyd, M., Panzeri, S. & Thiele, A. (2021) Directed information exchange between cortical layers in macaque V1 and V4 and its modulation by selective attention. *PNAS*, **118**.

Gieselmann, M.A. & Thiele, A. (2008) Comparison of spatial integration and surround suppression characteristics in spiking activity and the local field potential in macaque V1. *Eur J Neurosci*, **28**, 447-459.

Gray, H., Bertrand, H., Mindus, C., Flecknell, P., Rowe, C. & Thiele, A. (2016) Physiological, Behavioral, and Scientific Impact of Different Fluid Control Protocols in the Rhesus Macaque (Macaca mulatta). *eNeuro*, **3**.

Harris, K.D. & Thiele, A. (2011) Cortical state and attention. *Nat Rev Neurosci*, **12**, 509-523.

Kharas, N., Andrei, A., Debes, S.R. & Dragoi, V. (2022) Brain state limits propagation of neural signals in laminar cortical circuits. *P Natl Acad Sci USA*, **119**.

Lecci, S., Fernandez, L.M., Weber, F.D., Cardis, R., Chatton, J.Y., Born, J. & Luthi, A. (2017) Coordinated infraslow neural and cardiac oscillations mark fragility and offline periods in mammalian sleep. *Sci Adv*, **3**, e1602026.

Lee, S.H. & Dan, Y. (2012) Neuromodulation of brain states. *Neuron*, **76**, 209-222.

Logothetis, N.K., Kayser, C. & Oeltermann, A. (2007) In vivo measurement of cortical impedance spectrum in monkeys: implications for signal propagation. *Neuron*, **55**, 809-823.

McGinley, M.J., Vinck, M., Reimer, J., Batista-Brito, R., Zagha, E., Cadwell, C.R., Tolias, A.S., Cardin, J.A. & McCormick, D.A. (2015) Waking State: Rapid Variations Modulate Neural and Behavioral Responses. *Neuron*, **87**, 1143-1161.

Miller, J.E.K., Ayzenshtat, I., Carrillo-Reid, L. & Yuste, R. (2014) Visual stimuli recruit intrinsically generated cortical ensembles. *P Natl Acad Sci USA*, **111**, E4053-E4061.

Mohajerani, M.H., Chan, A.W., Mohsenvand, M., LeDue, J., Liu, R., McVea, D.A., Boyd, J.D., Wang, Y.T., Reimers, M. & Murphy, T.H. (2013) Spontaneous cortical activity alternates between motifs defined by regional axonal projections. *Nat Neurosci*, **16**, 1426-1435.

Mountcastle, V.B. (1957) Modality and topographic properties of single neurons of cat's somatic sensory cortex. *J Neurophysiol*, **20**, 408-434.

Niell, C.M. & Stryker, M.P. (2010) Modulation of visual responses by behavioral state in mouse visual cortex. *Neuron*, **65**, 472-479.

Okun, M., Steinmetz, N., Cossell, L., Iacaruso, M.F., Ko, H., Bartho, P., Moore, T., Hofer, S.B., Mrsic-Flogel, T.D., Carandini, M. & Harris, K.D. (2015) Diverse coupling of neurons to populations in sensory cortex. *Nature*, **521**, 511-515.

Okun, M., Steinmetz, N.A., Lak, A., Dervinis, M. & Harris, K.D. (2019) Distinct Structure of Cortical Population Activity on Fast and Infraslow Timescales. *Cerebral Cortex*, **29**, 2196-2210.

Palva, J.M. & Palva, S. (2012) Infra-slow fluctuations in electrophysiological recordings, blood-oxygenation-level-dependent signals, and psychophysical time series. *Neuroimage*, **62**, 2201-2211.

Pettersen, K.H., Devor, A., Ulbert, I., Dale, A.M. & Einevoll, G.T. (2006) Current-source density estimation based on inversion of electrostatic forward solution: effects of finite extent of neuronal activity and conductivity discontinuities. *J Neurosci Methods*, **154**, 116-133.

Poulet, J.F. & Petersen, C.C. (2008) Internal brain state regulates membrane potential synchrony in barrel cortex of behaving mice. *Nature*, **454**, 881-885.

Rabiner, L.R. (1989) A Tutorial on Hidden Markov-Models and Selected Applications in Speech Recognition. *P Ieee*, **77**, 257-286.

Roelfsema, P.R., Tolboom, M. & Khayat, P.S. (2007) Different processing phases for features, figures, and selective attention in the primary visual cortex. *Neuron*, **56**, 785-792.

Schroeder, C.E., Mehta, A.D. & Givre, S.J. (1998) A spatiotemporal profile of visual system activation revealed by current source density analysis in the awake macaque. *Cerebral Cortex*, **8**, 575-592.

Thiele, A., Delicato, L.S., Roberts, M.J. & Gieselmann, M.A. (2006) A novel electrode-pipette design for simultaneous recording of extracellular spikes and iontophoretic drug application in awake behaving monkeys. *J Neurosci Methods*, **158**, 207-211.

van Kempen, J., Gieselmann, M.A., Boyd, M., Steinmetz, N.A., Moore, T., Engel, T.A. & Thiele, A. (2021) Top-down coordination of local cortical state during selective attention. *Neuron*.

van Vugt, B., Dagnino, B., Vartak, D., Safaai, H., Panzeri, S., Dehaene, S. & Roelfsema, P.R. (2018) The threshold for conscious report: Signal loss and response bias in visual and frontal cortex. *Science*, **360**, 537-542.