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) , BBSRC (BB/W006758/1), *and by In2PrimateBrains (European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie Actions Doctoral Networks, Grant Agreement No 956669).*

.

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 average retained their coupling strengths across conditions, some neurons change coupling strength condition dependent, thereby potentially enhancing the coding abilities of cortical circuits.

Introduction

Global brain states manifest as large-scale spatiotemporal activity pattern [1,2]. They evolve on multiple temporal scales, range from minutes to hours during different sleep stages [3], and from seconds to minutes during wakefulness [4-6]. Brain states affect neuronal responses, neural interactions, and neural information encoding [4,6-9]. They underlie arousal and link to cognitive function and conscious perception [6,9]. Cortical state fluctuations manifest in single cortical columns by spontaneously transitioning between states of vigorous (On) and faint (OFF) spiking periods [10-12]. Initially, it was assumed that these fluctuations affect activity homogenously throughout the cortex [4,13]. 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 [10,11].

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) [14]. Strongly coupled 'choristers' and weakly coupled 'soloists' have been identified as two extremes of a continuous spectrum [12,14-17]. 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 [18]. 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 (figure 1). 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 showed condition dependent variation.

Results

Population coupling

Monkeys performed a cued feature based spatial attention task (figure 1, Methods). 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 [14] we found that PC varied widely between individual neurons (Figure 2B-E), and it varied between cortical layers, being generally strongest in supragranular layers (Supplementary Figure S1).

We next investigated whether PC varied for different stimulus conditions. Here we compared PC during spontaneous activity and during stimulus driven activity. Figure 2 C-E shows data pooled across areas. At first glance coupling strength seems similar during spontaneous activity and post cue onset activity, but they differed significantly between areas and periods (details see table 2), which we delineate in the next sections. To further investigate to what extent coupling strength is a largely fixed feature of a neuron, we calculated the correlation between coupling strength during spontaneous activity and after cue onset. We used the normalized population coupling of SUAs for each condition (C_{norm} and its Fisher Z-transform Z(0), see Methods). The strength of population coupling of SUAs during spontaneous and stimulus driven activity 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. As described in Methods, we additionally controlled for the possibly that single-unit activity leaked into neighboring MUA channels, or that neurons would not respond to the stimulus, which both might affect the results reported. Neither was the case as shown in Supplementary Figures S2.

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. [14] showed that behavioral variables such as motor intention influenced population coupling of V4 neurons. Here we evaluate the effect of covert spatial attention on the strength of population coupling of neurons. To compare the strength of population coupling of neurons between attend RF and attend- away conditions, we used the Fisher transformed PC correlation coefficient of a neuron (Z(0), see Methods), and compared these between attention conditions. Figure 4 shows the scatter plot of normalized PCs for attend RF vs attend away 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 (For controls relating to 'leakage', rate differences, and stimulus responsiveness between attention conditions see Supplementary Figures S3). Attentional modulation itself differed between layers in V1, it was weakest in granular layers (Supplementary Figure S4), but no significant differences were found in V4 (Supplementary Figure S4).

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 [10,11]. 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 have 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(0)) of the neurons. The results are shown in (**Figure 5,** see Supplementary Figure S5 for leakage and responsiveness controls, respectively). 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 ~10% to ~24% of the variance, and figure 5 demonstrates that some soloists (PC close to 0) nevertheless show profound activity differences during ON to OFF phase transitions (ratios of 4 to 6).

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,** also see Supplementary Figure S6 for 'leakage' and responsiveness controls). 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: $r= 0.158$, $p=0.003$, V4: $r= 0.199$, $p<0.001$). 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).

Relationship between population coupling and noise correlation

Population coupling should be related to noise correlation measures, as both assess co-variability of neural firing. While the former does so on a larger scale (one neuron's coupling to the local population activity), the latter assesses single neuron pairwise coupling, usually over somewhat longer time scales (often >100 ms integration windows). We did find that attention reduced noise correlations in V1, and in V4 (supplementary figure S7A). This is in line with the effects of attention on population coupling. We also predicted that neurons with strong population coupling (on the chorister spectrum end), would show stronger pairwise coupling (larger pairwise noise correlations). We found this to be true as well (supplementary figure S7B), but overall the effect, while significant, was small, explaining ~5-8% of the variance.

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' [12,14-17]. 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 [14].

Cortical state fluctuations manifest in single cortical columns by spontaneously transitioning between states of vigorous (On) and faint (OFF) spiking periods [10-12]. 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).

How does population coupling relate to more traditional measures such as noise correlations? Noise correlations only measure pairwise coupling, while population coupling is a more large scale measure. We found that noise correlations were reduced by attention in line with numerous previous reports [19-25], and we equally found that noise correlations between pairs of neurons was significantly correlated with their individual population coupling strength. However this correlation, while significant, was relatively small, explaining less 5-8% of the variance. This is not entirely surprising, given that noise correlations between neurons can be positive or negative, whereby the population coupling would be affected by all pairwise combinations (positive and negative), which could thus average to zero even if a single pair shows a large correlation.

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. It suggests that attention induced firing rate changes depend to a very small extent on population coupling and are thus possibly driven by a different mechanism.

Our data were recorded with laminar electrodes, which most likely sampled from single cortical columns, where coupling strength is more profound than between distant columns [26-28]. To what extent our results map to population coupling and its modulation by attention between different columns remains to be determined.

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 [18].

Data availability statement:

The data relevant for the manuscript have been uploaded to https://gin.g-node.org/Mahyar-Doost/md-soloists-and-choristers-data.git. and will be available upon manuscript publication. All relevant analysis scripts have been deposited at https://gitlab.com/nmd174/md-soloists-and-choristers.git, which will be made publicly available upon manuscript publication.

Materials and Methods

Data set

The data set used had previously been recorded, consisting of data that contributed to different publications from our group [11,29,30].

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 [31]. 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 areas V1 and V4 under sterile conditions and general anesthesia. Surgical procedures and postoperative care conditions have been described in detail previously [32].

Behavioral paradigm

The behavioral paradigm (**Figure 1**) has been described previously [11,29], 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.

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.

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 26 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, whereby the average of all channels of an electrode was 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:

$$E = \sum_{t=i}^{j} V[t]$$

, where t is the running index in the vector (V) representing the single-trial histogram (or MUAe), i is the first time point of the trial and j is the last time point of it. 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 1% 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 [33]. 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 [34] 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 [35], and tissue conductance of 0.4 Sm^{-1} [36]. The location of the early sink was identified visually by using interactive CSD plots, whereby 'clicking' on the location of the early sink saved the depth coordinates in relation to all electrode contacts. This in conjunction with the latency analysis (see below) allowed to align depths across sessions, and assign laminar location to a contact.

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 $Noise$ defined 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 [37]. 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 [37]. 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. It assumes a thickness of ~0.5 mm of the granular layer which is in line with published results e.g. [38,39]. 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. This assumes a layer 4 thickness of ~200 um, which estimates it to be thinner than layer 4 of V1 in line with the existing literature V4 layer thickness was estimated based on figures in previous publications e.g. [40,41,42]. While such a thickness may be on the conservative side, it ensures that cells outside the granular layer are not misclassified as granular cells. Since the latter are a smaller sample, misclassification would have larger effects on comparative measures than misclassifying granular cells as supra- or infragranular. For assignment to supra- and infra-granular layers we also used the presence of spiking activity at a given channel. If a channel was >1mm above the granular layer in V1, but showed clear spiking activity we still assigned it to be in supragranular layers. Conversely, if a contact was <1mm above the granular layer, but did not show spiking activity, and contacts above equally did not show spiking activity, all those contacts were considered to be outside the cortex. Similarly, if a layer was >0.75mm below the granular layer, but still showed clear spiking activity (not axon based), we still assigned it to be in the infragranular layers. Based on these approaches, channels outside the estimated range of gray matter of V1 or V4 were excluded from analysis. Note, that layer assignments were made by [29], and for consistency we used these assignments (contact by contact) for the data presented here.

Receptive field mapping

Receptive field (RF) mapping was carried out by a reverse correlation method described in detail previously [43]. 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. As stated previously, positioning of stimuli in the attention task was based on V1 RFs. The rationale was that these were usually quite a bit smaller than V4 RFs, whereby overlap of V1 and V4 RFs would ensure that V4 RFs also get activated , while the opposite might not be true, if stimuli were centered on V4 RFs. In most cases overlap between V1 and V4 RFs was >80% see figure 1C in [11].

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:

$$R_{xy}\,[k] = \frac{1}{\|x[n]\|}\sum_{n=1}^{N}(x[n+k] - \bar{x})\,(y[n] - \bar{y})\,, k = -N+1, -N+2\,, \ldots, N-1$$

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), $\bar{x}$ represents the baseline level of the smoothed spike train of a neuron, $\bar{y}$ represents the baseline level of the smoothed population rate, k 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:

$$C[k] = R_{xy}\,[k]\,, k = -N+1, -N+2\,, \ldots, N-1$$

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

$$C[0] = R_{xy}\,[0] = \frac{1}{\|x[n]\|}\sum_{n=1}^{N}(x[n] - \bar{x})\,(y[n] - \bar{y})$$

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:

$$C_{Norm}[0] = R_{xy}\,[0] = \frac{1}{\|x[n]\|}\,\frac{\sum_{n=1}^{N}(x[n] - \bar{x})\,(y[n] - \bar{y})}{\sqrt{\sum_{n=1}^{N}(x[n] - \bar{x})^2\,\sum_{n-1}^{N}(y[n] - \bar{y})^2}}$$

Finally, we calculated the Fisher transform of the correlation coefficient of population coupling ($C_{Norm}[0]$) of each SUA as:

$$\boldsymbol{Z(0) = atanh(\,C_{Norm}[0]\,)}$$

Leakage controls: Given the electrode contact spacing had an intercontact distance of 150um it could be argued that some single units might contribute to the multi-unit activity one contact apart (we refer to this as 'leakage'), which might in turn confound the results. To account for this possibility we recalculated the strength of coupling for all units, when only contacts with distances of 300um and more were used for estimation of population activity.

Rate matching controls: Attend RF and attend away conditions can result in different firing rates. To ensure that rate differences did not account for changes in population coupling or other measures, we performed rate

matching controls, whereby for each unit we removed trials with the lowest activity (in either condition), until the average activity between the attention condition had less than a 2% activity difference.

Response controls: As a further control only neurons with significant stimulus responses were included in the analysis. Here we compared the pre-stimulus response (-250 to 0 ms) to the stimulus onset response period (50-300ms) (Wilcoxon signed rank test), and only selected neurons where the difference was significant.

Hidden Markov Model and ON-OFF state estimation

ON-OFF state dynamics/periods for this dataset were obtained from the following repositories (G-Node: https://doi.gin.g-node.org/10.12751/g-node.b0mnn2; https://gitlab.com/JvK/cortical-statecoordination). Given that these map 1:1 onto the data set analyzed here, we could determine firing rates of single units during ON as well as during OFF phases, and link their ratio to PC values. For an example of ON-OFF phases in this data set see Figure 1D in [11].

However, to make the process of Hidden Markov Model (HMM) fitting to the data set accessible within the context of this manuscript, we repeat the description of the analysis description here. 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 [11]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 [44]. 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 [10]. 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 [45], maximizing the probability of observing the data given the model according to the Baum-Welch algorithm [44]. 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 [10]. 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:

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

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 [10].

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. The task included 3 different attention locations, whereby only one overlapped with the receptive field of the neurons under study. For attend away conditions we pooled across the two locations that did not overlap with the neurons' RFs, as both were equidistant from the attend RF location.

Noise correlation analysis

We calculated the trial-to-trial spike count correlation (noise correlation) of pairs of neurons in a time window of 500 to 0ms before the first dimming. This was done for attend RF and attend away conditions.

$$r_{ij} = cov\frac{(i,j)}{\sigma_i \sigma_j}$$

Where *cov* represents the covariance of the firing rates of neurons i and j, and σ_i , σ_j their respective standard deviation.

References

1. Ferezou, I., Haiss, F., Gentet, L.J., Aronoff, R., Weber, B., and Petersen, C.C. (2007). Spatiotemporal dynamics of cortical sensorimotor integration in behaving mice. Neuron *56*, 907–923. 10.1016/j.neuron.2007.10.007.
2. Mohajerani, M.H., Chan, A.W., Mohsenvand, M., LeDue, J., Liu, R., McVea, D.A., Boyd, J.D., Wang, Y.T., Reimers, M., and Murphy, T.H. (2013). Spontaneous cortical activity alternates between motifs defined by regional axonal projections. Nat Neurosci *16*, 1426–1435. 10.1038/nn.3499.
3. Lecci, S., Fernandez, L.M., Weber, F.D., Cardis, R., Chatton, J.Y., Born, J., and Luthi, A. (2017). Coordinated infraslow neural and cardiac oscillations mark fragility and offline periods in mammalian sleep. Sci Adv *3*, e1602026. 10.1126/sciadv.1602026.
4. Harris, K.D., and Thiele, A. (2011). Cortical state and attention. Nat Rev Neurosci *12*, 509–523. 10.1038/nrn3084.
5. Palva, J.M., and Palva, S. (2012). Infra-slow fluctuations in electrophysiological recordings, blood-oxygenation-level-dependent signals, and psychophysical time series. Neuroimage *62*, 2201–2211. 10.1016/j.neuroimage.2012.02.060.
6. McGinley, M.J., Vinck, M., Reimer, J., Batista-Brito, R., Zagha, E., Cadwell, C.R., Tolias, A.S., Cardin, J.A., and McCormick, D.A. (2015). Waking State: Rapid Variations Modulate Neural and Behavioral Responses. Neuron *87*, 1143–1161. 10.1016/j.neuron.2015.09.012.
7. Niell, C.M., and Stryker, M.P. (2010). Modulation of visual responses by behavioral state in mouse visual cortex. Neuron *65*, 472–479. 10.1016/j.neuron.2010.01.033.
8. Poulet, J.F., and Petersen, C.C. (2008). Internal brain state regulates membrane potential synchrony in barrel cortex of behaving mice. Nature *454*, 881–885.
9. 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. 10.1126/science.aar7186.
10. Engel, T.A., Steinmetz, N.A., Gieselmann, M.A., Thiele, A., Moore, T., and Boahen, K. (2016). Selective modulation of cortical state during spatial attention. Science *354*, 1140–1144. 10.1126/science.aag1420.
11. van Kempen, J., Gieselmann, M.A., Boyd, M., Steinmetz, N.A., Moore, T., Engel, T.A., and Thiele, A. (2021). Top-down coordination of local cortical state during selective attention. Neuron. 10.1016/j.neuron.2020.12.013.
12. Kharas, N., Andrei, A., Debes, S.R., and Dragoi, V. (2022). Brain state limits propagation of neural signals in laminar cortical circuits. P Natl Acad Sci USA *119*. ARTN e210419211910.1073/pnas.2104192119.
13. Lee, S.H., and Dan, Y. (2012). Neuromodulation of brain states. Neuron *76*, 209–222. 10.1016/j.neuron.2012.09.012.
14. Okun, M., Steinmetz, N., Cossell, L., Iacaruso, M.F., Ko, H., Bartho, P., Moore, T., Hofer, S.B., Mrsic-Flogel, T.D., Carandini, M., and Harris, K.D. (2015). Diverse coupling of neurons to populations in sensory cortex. Nature *521*, 511–515. 10.1038/nature14273.
15. Bachatene, L., Bharmauria, V., Cattan, S., Chanauria, N., Rouat, J., and Molotchnikoff, S. (2015). Electrophysiological and firing properties of neurons: Categorizing soloists and choristers in primary visual cortex. Neuroscience Letters *604*, 103–108. 10.1016/j.neulet.2015.07.049.
16. Dorman, R., Bos, J.J., Vinck, M.A., Marchesi, P., Fiorilli, J., Lorteije, J.A.M., Reiten, I., Bjaalie, J.G., Okun, M., and 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. 10.1093/cercor/bhad111.
17. Okun, M., Steinmetz, N.A., Lak, A., Dervinis, M., and Harris, K.D. (2019). Distinct Structure of Cortical Population Activity on Fast and Infraslow Timescales. Cerebral Cortex *29*, 2196–2210. 10.1093/cercor/bhz023.
18. Miller, J.E.K., Ayzenshtat, I., Carrillo-Reid, L., and Yuste, R. (2014). Visual stimuli recruit intrinsically generated cortical ensembles. P Natl Acad Sci USA *111*, E4053–E4061. 10.1073/pnas.1406077111.
19. Cohen, M.R., and Newsome, W.T. (2008). Context-dependent changes in functional circuitry in visual area MT. Neuron *60*, 162–173. 10.1016/j.neuron.2008.08.007.

20. Cohen, M.R., and Newsome, W.T. (2009). Estimates of the contribution of single neurons to perception depend on timescale and noise correlation. J Neurosci *29*, 6635–6648. 10.1523/JNEUROSCI.5179-08.2009.

21. Herrero, J.L., Gieselmann, M.A., Sanayei, M., and Thiele, A. (2013). Attention-induced variance and noise correlation reduction in macaque V1 is mediated by NMDA receptors. Neuron *78*, 729–739. 10.1016/j.neuron.2013.03.029.
22. Moreno-Bote, R., Beck, J., Kanitscheider, I., Pitkow, X., Latham, P., and Pouget, A. (2014). Information-limiting correlations. Nat Neurosci *17*, 1410–1417. 10.1038/nn.3807.
23. Ruff, D.A., and Cohen, M.R. (2014). Attention can either increase or decrease spike count correlations in visual cortex. Nat Neurosci *17*, 1591–1597. 10.1038/nn.3835.
24. Ruff, D.A., and Cohen, M.R. (2019). Simultaneous multi-area recordings suggest that attention improves performance by reshaping stimulus representations. Nat Neurosci *22*, 1669–1676. 10.1038/s41593-019-0477-1.
25. Cohen, M.R., and Maunsell, J.H. (2011). Using neuronal populations to study the mechanisms underlying spatial and feature attention. Neuron *70*, 1192–1204. 10.1016/j.neuron.2011.04.029.
26. de Oliveira, S.C., Thiele, A., and Hoffmann, K.P. (1997). Synchronization of neuronal activity during stimulus expectation in a direction discrimination task. J Neurosci *17*, 9248–9260.
27. Murphy, J.T., Kwan, H.C., and Wong, Y.C. (1985). Cross correlation studies in primate motor cortex: synaptic interaction and shared input. Can J Neurol Sci *12*, 11–23. 10.1017/s0317167100046527.
28. Smith, M.A., and Kohn, A. (2008). Spatial and temporal scales of neuronal correlation in primary visual cortex. J Neurosci *28*, 12591–12603. 10.1523/JNEUROSCI.2929-08.2008.
29. Ferro, D., van Kempen, J., Boyd, M., Panzeri, S., and Thiele, A. (2021). Directed information exchange between cortical layers in macaque V1 and V4 and its modulation by selective attention. PNAS *118*.
30. Spyropoulos, G., Schneider, M., van Kempen, J., Gieselmann, M.A., Thiele, A., and Vinck, M. (2024). Distinct feedforward and feedback pathways for cell-type specific attention effects. Neuron. 10.1016/j.neuron.2024.04.020.
31. Gray, H., Bertrand, H., Mindus, C., Flecknell, P., Rowe, C., and Thiele, A. (2016). Physiological, Behavioral, and Scientific Impact of Different Fluid Control Protocols in the Rhesus Macaque (Macaca mulatta). eNeuro *3*. 10.1523/ENEURO.0195-16.2016.
32. Thiele, A., Delicato, L.S., Roberts, M.J., and 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.
33. Schroeder, C.E., Mehta, A.D., and 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. DOI 10.1093/cercor/8.7.575.
34. Pettersen, K.H., Devor, A., Ulbert, I., Dale, A.M., and 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. 10.1016/j.jneumeth.2005.12.005.
35. Mountcastle, V.B. (1957). Modality and topographic properties of single neurons of cat's somatic sensory cortex. J Neurophysiol *20*, 408–434. 10.1152/jn.1957.20.4.408.
36. Logothetis, N.K., Kayser, C., and Oeltermann, A. (2007). In vivo measurement of cortical impedance spectrum in monkeys: implications for signal propagation. Neuron *55*, 809–823. 10.1016/j.neuron.2007.07.027.
37. Roelfsema, P.R., Tolboom, M., and Khayat, P.S. (2007). Different processing phases for features, figures, and selective attention in the primary visual cortex. Neuron *56*, 785–792. 10.1016/j.neuron.2007.10.006.
38. Yabuta, N.H., and Callaway, E.M. (1998). Functional streams and local connections of layer 4C neurons in primary visual cortex of the macaque monkey. Journal of Neuroscience *18*, 9489–9499.
39. Zhang L.A., L., P., Callaway, E.M. (2024). High-Resolution Laminar Identification in Macaque Primary Visual Cortex Using Neuropixels Probes. eLife. https://doi.org/10.7554/eLife.97290.2.
40. Pettine, W.W., Steinmetz, N.A., and Moore, T. (2019). Laminar segregation of sensory coding and behavioral readout in macaque V4. P Natl Acad Sci USA *116*, 14749–14754. 10.1073/pnas.1819398116.

41. Günther DM, B.M., Petukhov V, De Oliveira R, Wunderle T, Buchholz CJ, Fries P, Khodosevich K. (2024). Heterogeneity of layer 4 in visual areas of rhesus macaque cortex. bioRxiv. https://doi.org/10.1101/2024.03.11.584345.
42. Stepanyants, A., Hof, P.R., and Chklovskii, D.B. (2002). Geometry and structural plasticity of synaptic connectivity. Neuron *34*, 275–288. Doi 10.1016/S0896-6273(02)00652-9.
43. Gieselmann, M.A., and 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.
44. Rabiner, L.R. (1989). A Tutorial on Hidden Markov-Models and Selected Applications in Speech Recognition. P Ieee *77*, 257–286. Doi 10.1109/5.18626.
45. Bishop, C.M. (2006). Pattern Recognition and Machine Learning (Springer).

Tables

	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.

	Spontaneous activity				Attend RF				Attend away			
	Mean PC	Std	Median	p	Mean PC	Std	Median	p	Mean PC	Std	Median	p
V1	24.75	±23.17	17.04	<0.001	48.22	±39.65	38.64	<0.001	58.52	±48.04	48.13	<0.001
V4	42.91	±50.61	28.62		72.07	±62.65	55.31		81.13	±74.34	63.21	
	Difference between periods (p)			$p_{\text{spont.-attend RF}}$			$p_{\text{spont.-attend away}}$		$p_{\text{attend RF-attend away}}$			
V1	<0.001			<0.001			p < 0.05		<0.001			
V4	<0.001			<0.001			p = 0.3126		<0.001			

Table2. Average, standard deviation and medians of population coupling (PC) in V1 and V4 across different time windows/conditions. P: p-value based on Wilcoxon rank sum test comparing PC between area V1 and V4. Lower rows shows p-values for comparisons across different periods within each area (Kruskal Wallis Anova with rank sum test post-hoc comparison).

Figures

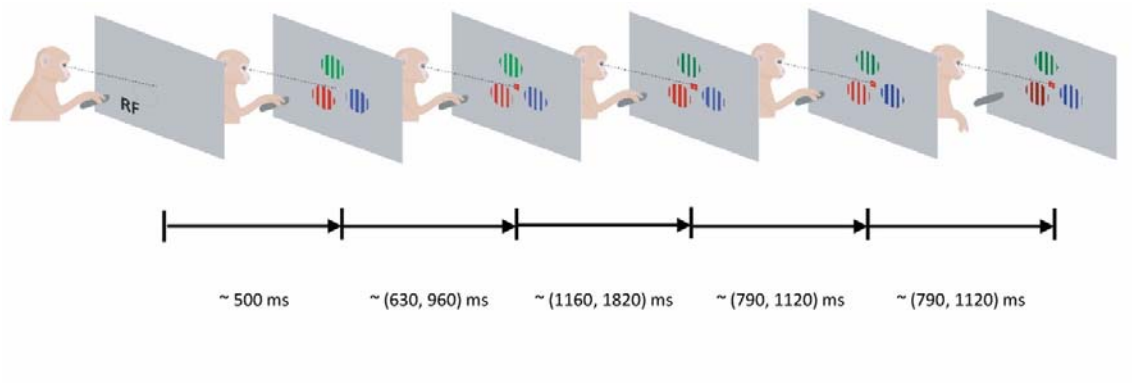

Figure 1: Feature-based spatial attention task. *The monkey held a lever to initiate the trial, whereafter a central fixation spot was turned on. Upon fixation 3 colored gratings appeared, one was centered on the receptive fields (RFs) of the V1 neurons. After a variable delay a cue matching one of the 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.*

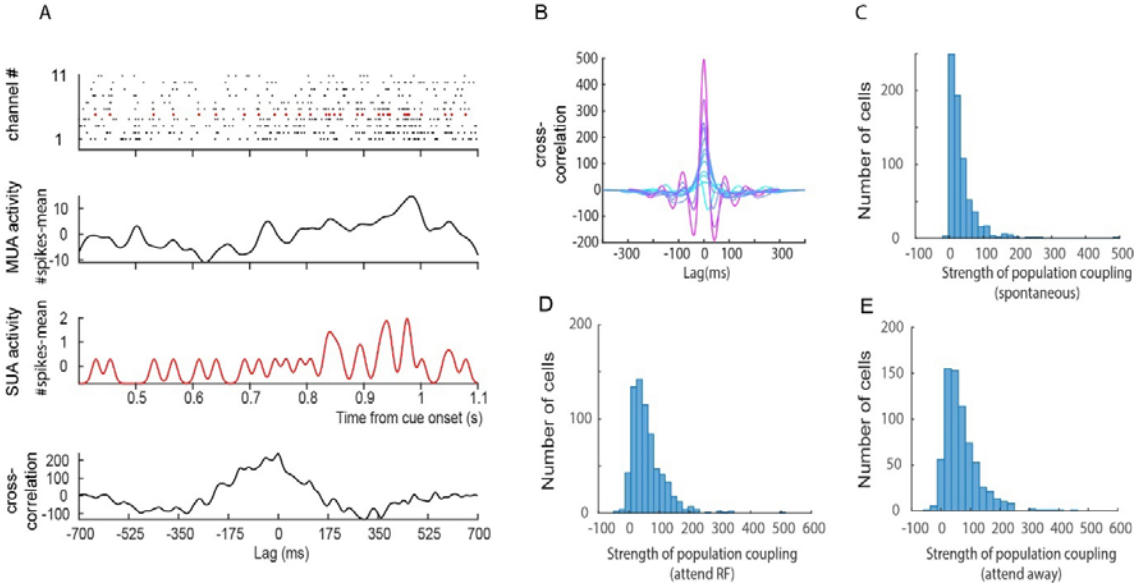

Figure 2: **Calculation of population coupling and its distribution for different experimental epochs and conditions. A)** Raster plots within a single trial of MUA activity (black) across different channels located in grey matter. Red raster plot shows SUA from a channel, recorded along the MUA activity. 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. Coupling 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.

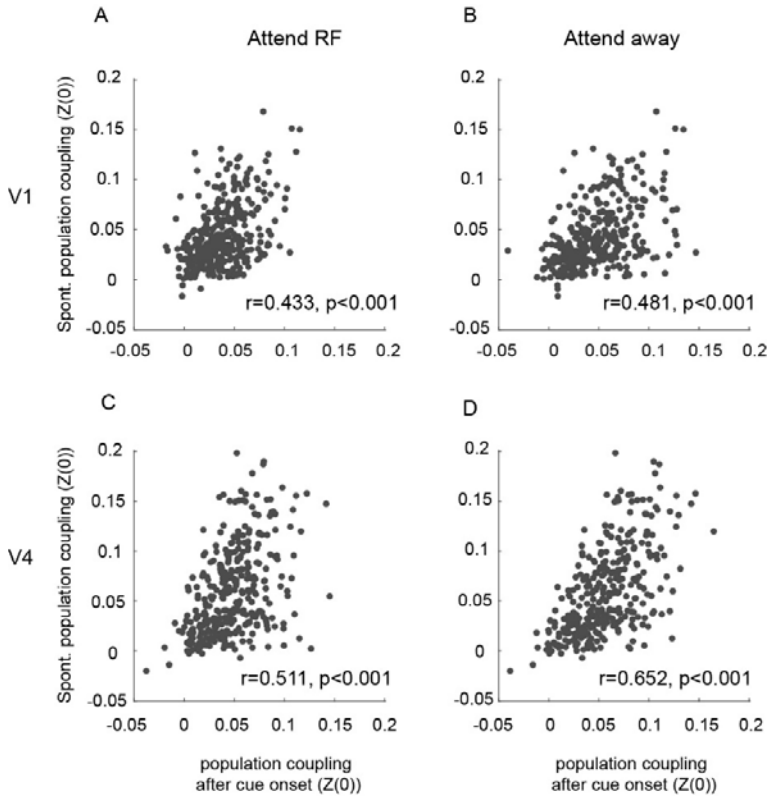

Figure 3: Population coupling during spontaneous activity compared to stimulus driven activity. A) Correlation of Population coupling during spontaneous activity and during attend RF stimulus driven activity for V1 neurons (period after cue onset). **B)** Correlation of Population coupling during spontaneous 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. Population coupling was quantified by calculating the Fisher transformed correlation coefficient between single unit and population activity (Z(0)).

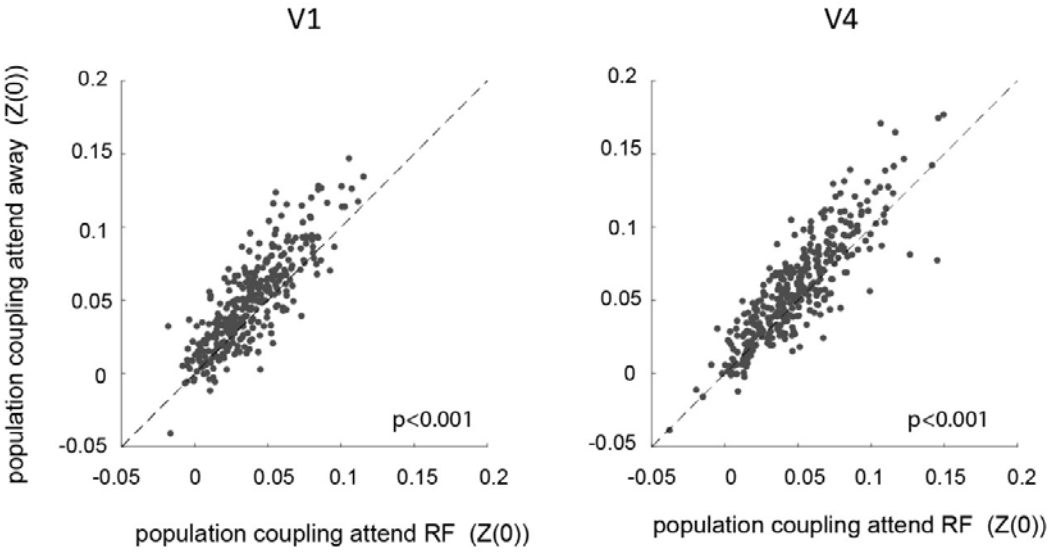

Figure 4: Effect of spatial attention on population coupling of V1 and V4 neurons. Population coupling (Fisher-transformed correlation coefficient (Z(0)) during attend RF conditions (x-axis) and attend away conditions (y-axis) during the period after cue-onset for V1 neurons (left) and V4 neurons (right). Insets show p-values of differences in population coupling.

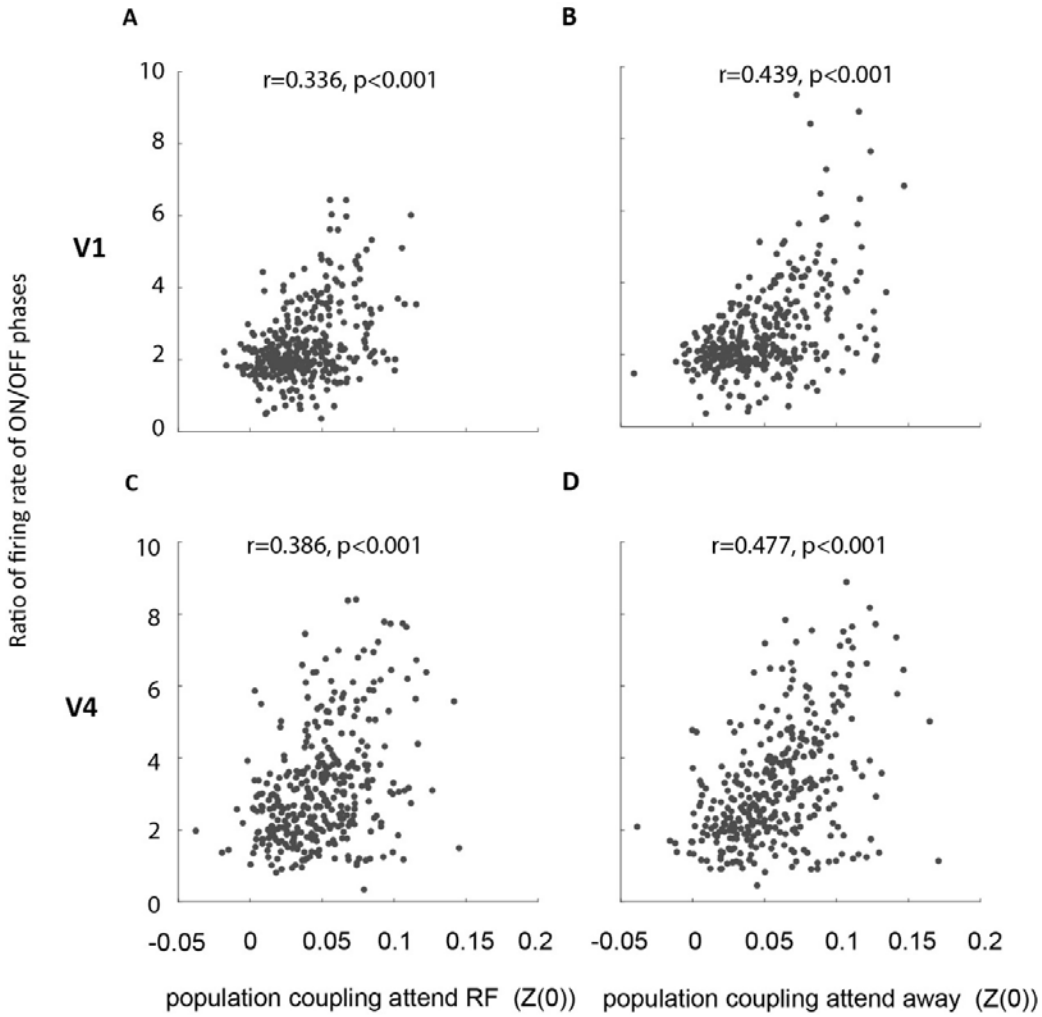

Figure 5: Relation between ON-OFF state fluctuation alignment and population coupling of V1 and V4 neurons. Population coupling (Fisher-transformed correlation coefficient Z(0)) during attend RF conditions and attend away conditions during the period after cue-onset for V1 neurons (left) and V4 neurons (right) compared to the ratio of ON-OFF period firing (y-axis). Insets show correlation coefficients and p-values of this relationship.

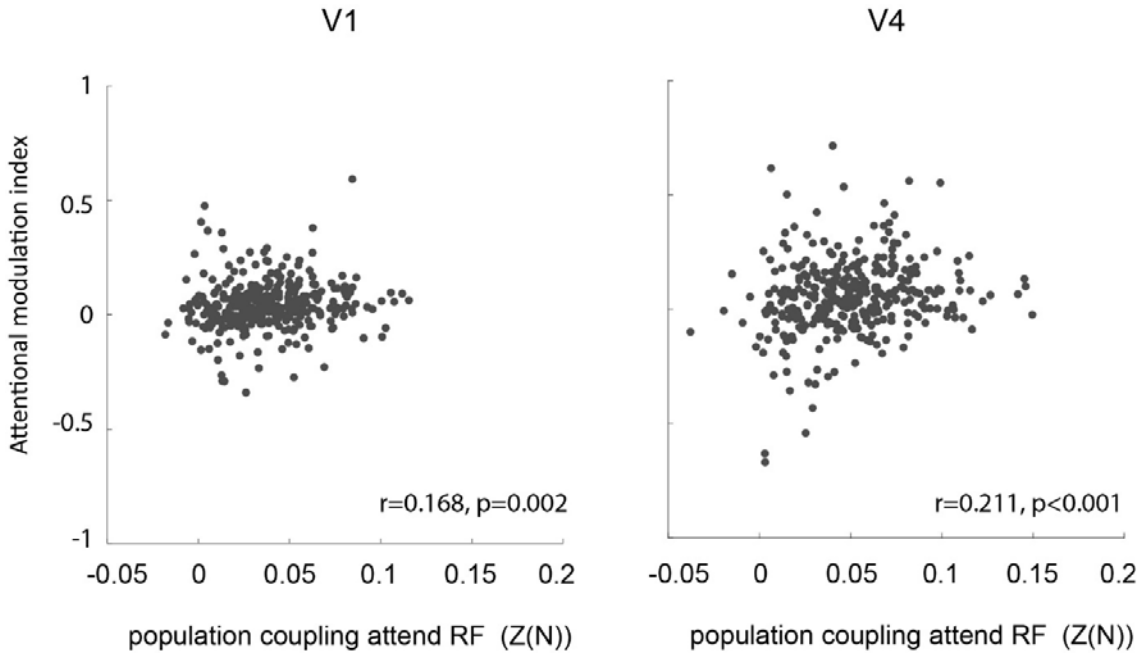

Figure 6: Relation between population coupling of V1 and V4 neurons and their attentional modulation strength. Population coupling (Fisher-transformed correlation coefficient (Z(0)) during attend RF conditions during the period after cue-onset for V1 neurons (left) and V4 neurons (right) compared to the strength of attentional modulation (y-axis). Insets show correlation coefficients and p-values of this relationship.