Experimental Brain Research
https://doi.org/10.1007/s00221-021-06166-w

RESEARCH ARTICLE

Dynamics of coherent activity between cortical areas defines a two-stage process of top-down attention

E. Levichkina[1,2,3] · M. Kermani[1] · Y. B. Saalmann[1,4] · T. R. Vidyasagar[1,3]

Received: 23 December 2020 / Accepted: 23 June 2021

Abstract

Analysing a visual scene requires the brain to briefly keep in memory potentially relevant items of that scene and then direct attention to their locations for detailed processing. To reveal the neuronal basis of the underlying working memory and top-down attention processes, we trained macaques to match two patterns presented with a delay between them. As the above processes are likely to require communication between brain regions, and the parietal cortex is known to be involved in spatial attention, we simultaneously recorded neuronal activities from the interconnected parietal and middle temporal areas. We found that mnemonic information about features of the first pattern was retained in coherent oscillating activity between the two areas in high-frequency bands, followed by coherent activity in lower frequency bands mediating top-down attention on the relevant spatial location. Oscillations maintaining featural information also modulated activity of the cells of the parietal cortex that mediate attention. This could potentially enable transfer of information to organize top-down signals necessary for selective attention. Our results provide evidence in support of a two-stage model of visual attention where the first stage involves creating a saliency map representing a visual scene and at the second stage attentional feedback is provided to cortical areas involved in detailed analysis of the attended parts of a scene.

Keywords Visual attention · Parietal cortex · Mid-temporal cortex · Oscillations · Primates

Introduction

Cognitive functions such as attention depend heavily upon communication between brain regions. Synchronized oscillations of membrane potentials of groups of neurons in different areas have been postulated as a basic mechanism for the interaction between brain regions, a concept often referred to as 'communication through coherence', or CTC (Fries 2005; Buzsáki et al. 2013), and illustrated in Fig. 1a).

Communicated by Melvyn A. Goodale.

✉ T. R. Vidyasagar
trv@unimelb.edu.au

1 Department of Optometry and Vision Sciences, The University of Melbourne, Parkville, VIC 3010, Australia

2 Institute for Information Transmission Problems, Russian Academy of Sciences, Moscow 127051, Russia

3 ARC Centre of Excellence in Integrative Brain Function, University of Melbourne Node, Melbourne, Australia

4 Department of Psychology, University of Wisconsin – Madison, Madison, WI 53706, USA

If groups of neurons from two different areas have oscillations of similar frequencies it can help in transmitting signals from one area to the other by improving the chances to reach the threshold for initiating action potentials. CTC has been demonstrated in macaques between parietal and an early visual area, namely the middle temporal area (V5/MT) (Saalmann et al. 2007), prefrontal and parietal cortices (Buschman and Miller 2007) and between prefrontal cortex and visual area V4 (Gregoriou et al. 2009) in tasks requiring top-down attention. It has also been suggested (Wolfe 1994; Vidyasagar 1999; Bullier 2001) that attention-demanding tasks such as finding an object in a cluttered scene or matching a stimulus to an item in working memory may involve a two-stage process. At the first stage the most relevant spatial locations are selected based on coarse featural information from the visual scene, and at the second stage a spotlight of attention focuses attention on one selected location at a time for detailed visual processing and to enable the object of interest to be identified.

Recent findings (Levichkina et al. 2017) obtained in monkeys performing a Delayed Match to Sample Task (DMS, Fig. 1b), namely matching visual stimuli presented with a

Published online: 09 July 2021

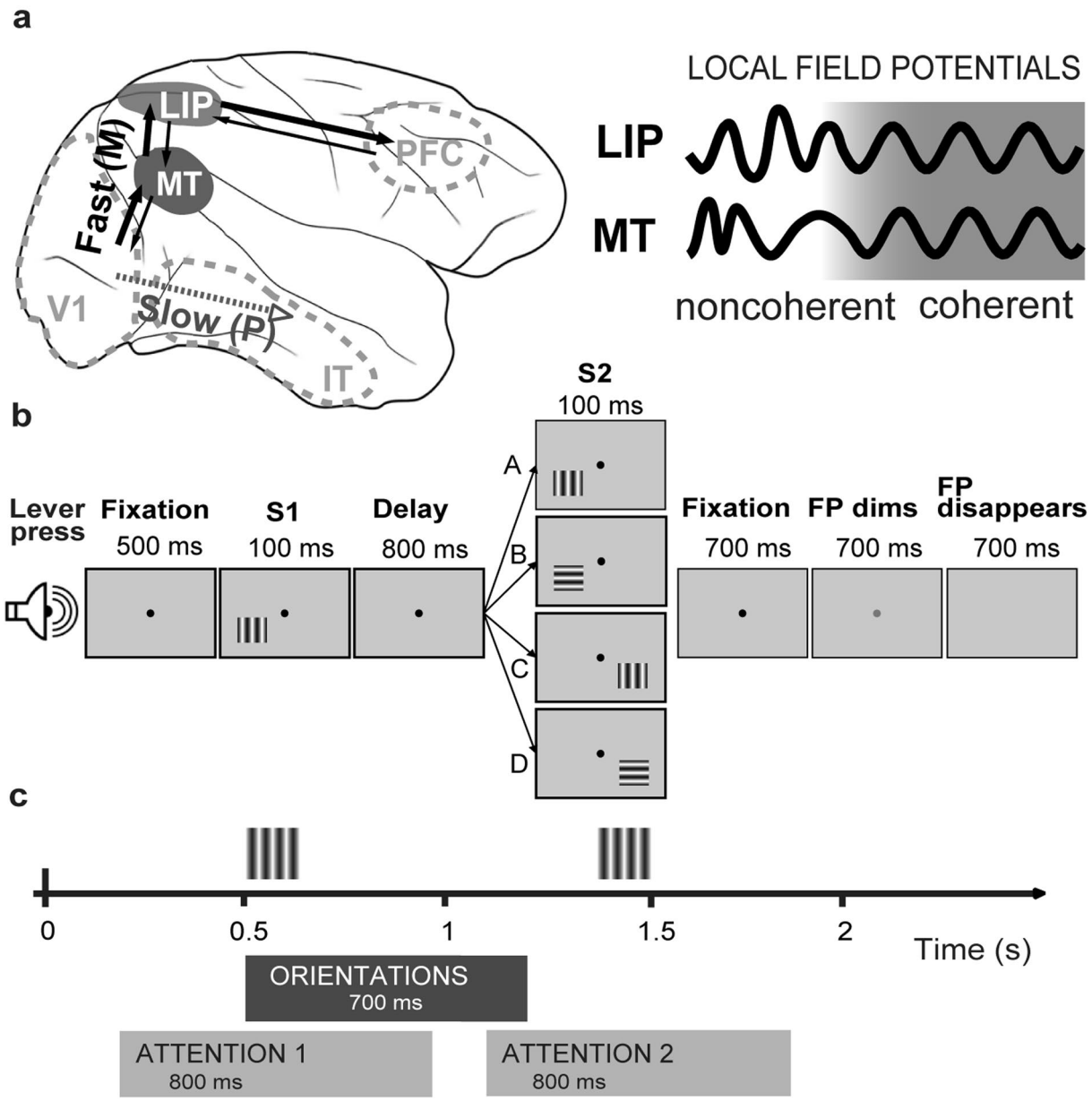

Fig. 1 Recording areas in the monkey brain and the experiment paradigm. **a** Areas MT and LIP within fast dorsal visual pathway are shown on the left, and two LFP signals on the right, showing initial lack and subsequent presence of synchronized (coherent) oscillations. Fast dorsal pathway is also shown here as providing spatial information in order to facilitate detailed recognition of the objects by the slower ventral pathway. **b** Schematic depiction of the DMS task. The task involved matching two gratings (S1 and S2) presented briefly for 100 ms each with an interval of 800 ms between them (Delay). The monkey presses a lever when ready for a trial and fixates on the central point (FP) during the trial. Gaze fixation was controlled by an infrared oculometer and the trial was aborted if gaze deviated by more than 1º. The monkey was required to report whether each trial was a match (both gratings appearing at the same location and having the same orientation, case A) or a non-match (whenever the locations or the orientations or both were different—cases B, C, and D). For match trials, the monkey had to release the lever when the FP dimmed, and for the non-match trails when it disappeared. **c** Time course of the match trial shown with the intervals used for coherence comparisons. Feature-related coherence bands were analysed by comparing coherences evoked by stimuli of 2 orthogonal orientations in 700 ms long period after S1 onset (ORIENTATIONS), whereas attention-related coherence bands were obtained by comparing 800 ms intervals around S1 and S2 (ATTENTION1 and ATTENTION2)

delay between them, suggest that the lateral intraparietal area (LIP) of the macaque's posterior parietal cortex may possess the neural substrate for such a two-stage process. This is particularly plausible given the presence of feature selective neurons in LIP, e.g., cells with selective responses to differently orientated lines (Sereno and Maunsell 1998;

Toth and Assad 2002; Ibos and Freedman 2016; Levichkina et al. 2017). We have also recently shown that LIP contains two groups of neurons having distinctly different responses to the stimuli in our DMS task, which occur during different parts of the delay period between the stimulus pair (Levichkina et al. 2017). One group of cells are feature selective and show higher activity in an early part of the delay period between the two stimuli in each trial, but not any attention-related elevation around the second stimulus of the pair (Attentional Enhancement negative, or AE- cells). The second group of cells (AE+) shows poor feature selectivity, but exhibits enhanced responses before and during the second stimulus when attention is attracted to that location by the first stimulus. Area MT provides the major afferent input to LIP (Baizer et al. 1991), which is necessary for the feature selectivity expressed by LIP. In turn, area MT receives top-down attention signals from LIP which increases responses of MT cells (Saalmann et al. 2007; Herrington and Assad 2010) during the presentation of the second stimulus, but MT cells do not demonstrate ramping of the response during the delay period, essentially demonstrating AE- behaviour. Figure 2 shows examples of LIP cells of both AE+ and AE- types, and a typical AE- MT cell. See Levichkina et al. (2017) for a detailed description of cell responses under different conditions.

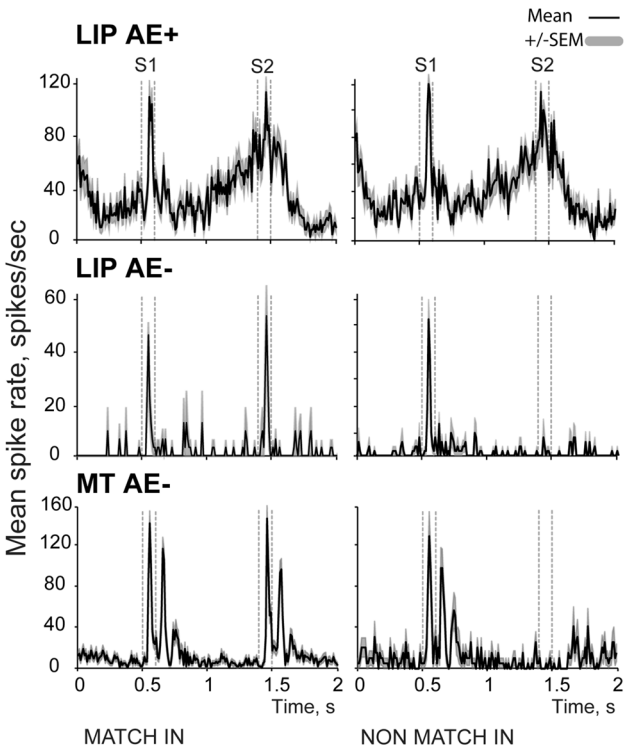

Fig. 2 Types of cell responses in LIP and MT areas. Peristimulus time histograms for LIP and MT cells with standard error, with top row showing responses of an AE+cell from LIP area, middle row an AE–cell of LIP and bottom row AE- cell of area MT. 'Match' trials with preferred orientation for both stimuli (S1 and S2) are shown in the left column. The right column shows responses to 'Non-Match' trials with preferred orientation for S1 stimulus and non-preferred orientation for S2. Stimulus presentation times are indicated by dashed vertical lines and denoted as S1 and S2 on the top of the figure. For detailed description of the cells' behaviours and trial types refer to Levichkina et al. (2017)

However, since spiking activity of feature-selective cells in MT usually declines within 200 ms from stimulus offset (Saalmann et al. 2007; Mendoza-Halliday et al. 2014), transfer of the featural information by neuronal spiking ostensibly occurs within a short interval. Therefore in the DMS task, featural information that would drive the LIP AE+cells for eliciting top-down attention during the later part of the delay period is stored in some way other than as a simple increase in spike rate in both areas during the early part of the delay, after the stimulus-evoked afferent signals from MT ceases. This question of the nature and site of the transition from coding feedforward feature information to coding feedback attention signals remains open. Here, we investigated whether synchronization of the oscillatory activities between MT and LIP could serve as the crucial transition mechanism. We also tested whether this transition is accompanied by a change in the coherence frequency, consistent with suggestions that CTC mediating interareal communications may use different oscillating frequency bands in the bottom-up and top-down pathways (Buzsáki and Draguhn 2004; Buzsáki 2006).

Materials and methods

Animal care and behavioural training

Data was collected from two male macaque monkeys (*Macaca nemestrina)*. The study was conducted as per the guidelines of the National Health and Medical Research Council Australian Code of Practice for the Care and Use of Animals for Scientific Purposes and approved by the University of Melbourne Animal Experimentation Ethics Committee. Monkeys were housed together and had *ad lib* access to water. They were trained to come voluntarily to the training chair and perform a visual delayed match-to-sample task (DMS, Fig. 1b). For detailed description of husbandry, surgical and training procedures, please refer to Supplementary Online Material of Saalmann et al. (2007).

The task included presentation of two sinusoidal gratings ($8^0 \times 8^0$, 30% contrast with a mean luminance of 15 cd/m^2) with an interstimulus interval (ISI) between them. The monkeys matched two gratings with respect to both the orientation and location of the gratings. The orientation preference of neurons at the recording site was first assessed by hand held stimuli, then verified with gratings, before each set

of recordings; and an orientation close to that preferred for multiunit activity in both MT and LIP, and the orientation orthogonal to the preferred, were used for the recordings. Up to 5 different locations in the visual hemifield around the fixation point and the two orientations were presented in pseudorandom succession in such a way that approximately 50% of the trials were 'Match' trials, where the second grating appeared at the same location as the first and the orientations of both gratings were also the same. The pseudorandomisation was also mildly biased for the first stimulus (S1) to fall on the receptive field location of the MT and LIP recording sites to ensure collection of a sufficient amount of data from the superimposed receptive field sites in the two cortical areas.

The monkey initiated each trial by pressing a lever. This led to the black fixation spot (FP, 0.1° diameter) being presented at the centre of a uniform grey screen (of luminance 15 cd/m^2). The monkey was required to maintain fixation during the whole trial, but could break it between trials. If an eye movement of more than 1°, monitored using an infrared oculometer (Dr. Bouis) was detected, the trial was aborted.

S1 was presented for 100 ms duration, 500 ms after the start of fixation. After a delay period, the second grating (S2) was also presented for 100 ms duration. The delay was 800 ms in 29 recordings, 900 ms in 1 recording, 1000 ms in 1 recording and 500 ms in 5 recordings. The monkey had to keep fixating for an additional 700 ms after S2 offset. After that, the FP was dimmed for 700 ms before it disappeared. In Match trials (when locations and stimulus orientations of S1 and S2 were the same) the monkey was required to release the lever within the 700 ms dimming interval to obtain fruit juice reward. In all Non-Match trials, the lever had to be released only after disappearance of the FP. Correct responses were rewarded with fruit juice and incorrect responses were followed by prolongation of the ISI by 1–3 s. The window allowed for response was from 200 to 650 ms after either the start of the dimming period (in the case of match trials) or after FP disappeared (in the case of non-match trials).

Electrophysiology

A low invasive "halo" surgical technique was used for implantation (Pigarev et al. 1997, 2009; Saalmann et al. 2007). Structural MRI was performed prior to the start of experiments to guide craniotomies (2.5 mm diameter) necessary to record from LIP and MT areas.

We recorded neuronal activities from areas LIP and MT using tungsten or platinum–iridium microelectrodes (FHC, ME USA). The recorded signal was filtered either at 1–4000 or 10–4000 Hz and sampled at 10,000 Hz by the Cambridge Electronic Design (CED) Micro1402 data collection system. It was further filtered using inbuilt filters of CED's Spike 2 programme to obtain separate spike and local field potential (LFP) data. A band-pass filter was applied in the 300–4000 Hz range for spikes and a low-pass filter up to 250 Hz for LFPs. Single or multiunit activity was acquired with the WaveClus spike sorting toolbox running on Matlab (Quiroga et al. 2004) with amplitude threshold set to be above 3 standard deviations from baseline noise of the band-pass filtered data.

36 paired LIP-MT recordings with overlapping receptive fields and matched preferred stimulus orientations were used for the present analysis.

Data analysis and statistics

General approach to coherence analysis

Coherence analysis allows one to measure synchronization between two signals, which in our case, was the degree of synchronization between the neuronal activities recorded simultaneously from areas MT and LIP. Since neuronal activity can potentially occupy a large range of frequencies, it is necessary to define an approach to select specific frequency bands for further analysis. The usual way to analyse neuronal activity in the frequency domain largely depends on assumptions regarding putative functions of the oscillations in different frequency bands. To overcome such semi-arbitrary character of choosing bands of interest and the questionable applicability of widely defined bands to activities of different cells with different soma sizes, axonal arbours and synaptic characteristics, we compared coherences by applying Aversen's technique to multiple comparisons across frequencies, implemented as "two group_test_coherence" function in Matlab-based Chronux toolbox (http://chronux.org/; Mitra and Bokil 2008; for details, Bokil et al. 2007). This approach permits splitting the analysis into two parts. First one addresses the relevance of a particular oscillation to the behavioural task and the second step involves analysing the data selectively to observe the dynamics of any change of coherence specifically within the task-relevant frequency bands. At both steps, we applied multitaper spectral estimations implemented in Chronux toolbox.

Detailed description of the coherence comparison can be found in (Bokil et al. 2007). Here we describe only the details relevant to the application of this method to our data.

The coherence is calculated as:

$$C(f) = S12(f) \Big/ \sqrt{S11(f)S22(f)},$$

where S(f) is the spectrum, 1 and 2 refer to the neural activity in MT and LIP respectively. Coherence was calculated involving 3 orthogonal Slepian taper functions and a time bandwidth product of 2 (Mitra and Pesaran 1999;

Jarvis and Mitra 2001; Bokil et al. 2007). For data length N and frequency bandwidth W, the first $K = 2NW - 1$ Slepian sequences are optimally concentrated in the frequency range $[-W, W]$. Therefore the minimal frequency range for estimating significance of the coherence differences between 2 conditions is 2 W = (K + 1)/N.

The method of comparison of coherences is based on the fact that the Fisher *transformed coherences,* $\tanh^{-1}(C(f))$, has a Gaussian distribution and therefore differences of transformed coherences for two populations are distributed as $\Delta z(f) \sim N(0, 1)$ when the two population coherences are equal.

Statistical estimation of the coherence difference utilizes jackknifed variance based on computing *m = N(trials) K(tapers)* Fourier transform pairs, leaving out one taper and one trial in turn. To summarize, coherence differences ($\Delta z(f)$) are considered significant if they exceed jackknife-based error bars consecutively for at least 2 W frequencies. We report here only those differences that exceeded these criteria.

Figure 3 illustrates oscillatory behaviour of LFPs in areas MT and LIP in one example Match trial taken from a recording session where multiunit (MU) activity in response to S1 and S2 were significantly above the baseline and significant coherence between the two sites was observed. Note that high gamma oscillation was evoked by the first stimulus and lasted over approximately half of the delay period in both MT and LIP and thereafter replaced by lower frequency oscillations. We applied coherence analysis to this specific pattern of activities.

Feature-related coherence differences

The two conditions that were compared to estimate feature-related activity were presentations of two orthogonal orientations of the grating, each presented as the first of the pair in all trials where stimulus S1 fell on the receptive fields of the neurons in the two sites.

Since the first grating could appear in any of five different locations, attention could not have been already focussed on any one location at the time of S1. Furthermore, the monkey had to code in memory both the location and the orientation of the S1 grating to be able to match them with those of S2. Therefore, since the location of S1 (being on the RF) does not change in the subset of trials we analysed, any difference in neural activity between the two orientations in the period from the onset of S1 and extending into the early part of the delay period can be assumed to be orientation-related activity. Our previous results also indicate that feature-related neuronal activity in LIP can indeed be retained for some time during delay period but ceases towards the end of it (Saalmann et al. 2007; Levichkina et al. 2017). Therefore for the recording sessions where delays were equal to or more than 800 ms ($N = 31$), we decided to choose a 700 ms interval from S1 onset for the first step of our analysis, as it included S1 presentation and a reasonably long interval after that (600 ms), but not the later part of the delay period (last 200 ms). Coherence for all trials, where one orientation was presented as S1, were compared to the coherence for all trials with the orthogonal orientation as S1, using the above described statistical technique (Bokil et al. 2007). Thus for the 700 ms interval used for feature-related activity estimation, the 2 W frequency bin was equal to 5.71 Hz. Therefore, coherence differences exceeding jackknife-based error bars for more than 5.71 Hz were taken as feature-relevant. The 2 W value was adjusted according to the same principle for the recordings with 500 ms delay period ($N = 5$) where the

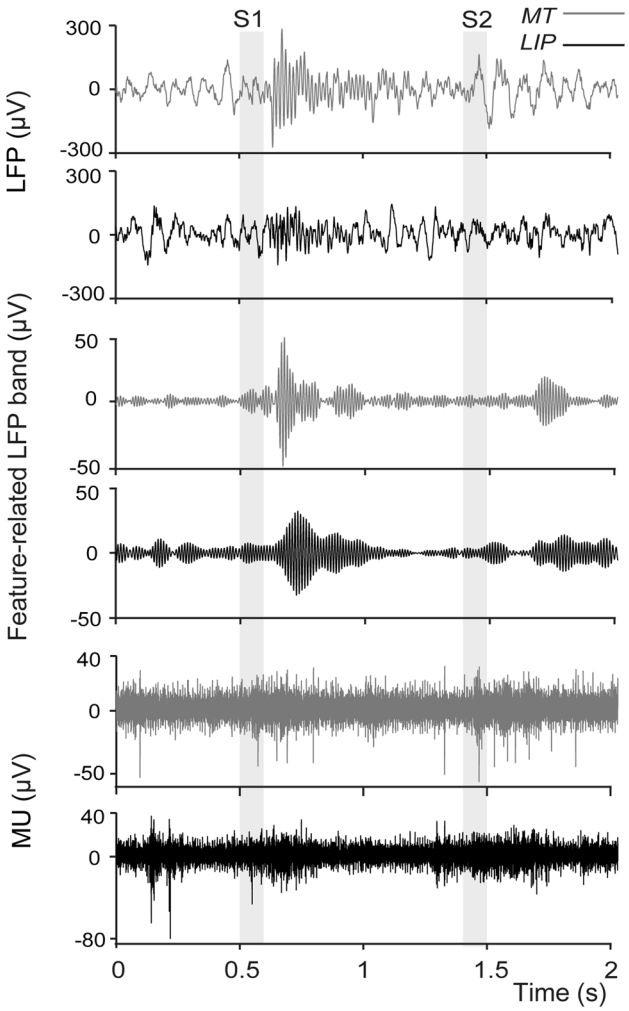

Fig. 3 Example of neuronal activity recorded in single match trials. Raw LFPs recorded from MT and LIP areas (upper panel), band-pass filtered LFP where all but the coherent gamma frequencies are filtered out (middle panel), and the corresponding MU responses (bottom panel). Vertical grey intervals correspond to presentations of S1 and S2

interval of the analysis was reduced to 500 ms starting from S1 onset.

The second step of the analysis involved constructing the dynamics of the feature-related coherence. Coherence was calculated in 300 ms intervals in 10 ms steps along the trial. Since direct estimation of orientation preference in the frequency domain is problematic, we used as preferred the orientation that produced more widespread changes in terms of frequency bandwidths that were deemed relevant in step 1. To visualize the dynamics of coherence, coherences for all relevant frequency bands were averaged within each recording site for that orientation. For visualization purposes only Match trials are used because responses to different S2 gratings as in non-match trials can be dramatically different (Saalmann et al. 2007; Levichkina et al. 2017). Latency of the coherence maximum within the interval of interest (700 ms or 500 ms after S1 onset) was also compared to the corresponding latencies of the multiunit responses in MT and LIP in the same interval using *Wilcoxon Rank-Sum test.*

Attention-related coherence differences

In our DMS task, covert attention is required to match location and orientation of the second stimulus to the first one. Attention to the S1 location is known to cause ramping of the spike rates in the late delay period and in location-matched trials, also enhanced response to S2 (Saalmann et al. 2007). To estimate attentional modulation we compared activity around S1 to that around S2 in Match trials. As described before, attentional modulation starts before S2 onset in the late part of the delay period (Saalmann et al. 2007; Levichkina et al. 2017). Thus, we chose two 800 ms windows, one around S1 and another around S2, from 300 ms before stimulus onset to 400 ms after stimulus offset. For the chosen 0.8 s interval used for attention-related activity estimation, the minimal 2 W frequency range for significant coherence differences was 5 Hz. As short delay periods are known to show an attentional blink in LIP responses (Maloney et al. 2013), we included in this analysis only those recordings where delay period was set to 800, 900 or 1000 ms (31 recordings).

We considered increased coherence during S2 compared to S1 in particular frequency bands as a sign of positive attentional modulation that sustains attention and lower coherence values as negative attentional modulation that suppresses neural activity related to unattended items.

The next step of the analysis aimed to illustrate the temporal dynamics in the frequency bands identified as attention-related as above. For this, we used the same parameters as we did for demonstrating the dynamics of feature-related coherence differences, namely 300 ms windows in 10 ms steps along the trial. We analysed positive S2 > S1 (attentional enhancement) and negative S2 < S1 (suppression) coherences separately.

To visualize coherence dynamics, coherences for all significant frequency bands were averaged within each recording site.

MU responses in MT and LIP areas in relation to feature-related coherence

Since feature-related activity occurs in the high frequency range where the LFP can be potentially affected by leakage of spiking-related frequencies (Ray and Maunsell 2011; Scheffer-Teixeira et al. 2013, Waldert et el. 2013), it was necessary to test if any observed LFP coherence at higher gamma frequencies can be simply explained by the higher spike rate during the response to the visual stimulus. Spike removal from the LFP signal may not be particularly effective in the case of extracellular recordings made with relatively low impedance electrodes. Therefore we compared latencies of the response peak of MU to the latencies of the peak LFP coherence, since the spike leakage can be expected to be maximal for the maximal spike rate.

For every recording site, we identified the presence of the response to S1 by comparing mean background spike rate in 300 ms before the S1 onset to the mean spike rate within 300 ms after S1 onset using Wilcoxon Signed Rank Test. Recording sites with significant S1 response to any of the two orientations ($p < 0.05$) were used for further analysis. In order to investigate if the coherence maximum corresponds to the MU maximum, latency of the S1 response maximum was estimated using the same averaging window and step parameters as those used for analysis of the coherence dynamics, namely 300 ms window duration in 10 ms steps. Latencies of MT and LIP MU response maxima were compared to the latency of the LFP coherence maxima for trials with significant feature-related coherence, using the Wilcoxon Rank-Sum test.

Another possibility to consider is that the activity of other cells that do not respond to the stimulus may be active in the delay period and contribute to the coherence seen in the delay period. This synchronized activity is best observed in the multiunit response rather than in single unit recordings, especially in the case of higher oscillation frequencies, because an individual cell that is in phase with an oscillation may not fire action potentials with every cycle (Maloney et al. 2013; Luczak et al. 2015). Therefore we studied spike-spike coherence dynamics between multiunit spiking activities in MT and LIP areas for feature-related processes in order to compare it to the LFP-LFP coherence, applying again the same parameters as for the LFP analysis described earlier. The spike-spike coherence difference between the two orientations reached significance only in two of the recording pairs. However it is to be admitted that

spike-spike coherence is generally weak as action potentials do not accompany each cycle of LFP oscillation, resulting in difficulties in identifying spike-spike coherence, especially in the high frequency range.

Spike-triggered average (STA) of feature-related oscillations with multiunit activity as the trigger

The LFP was band-pass filtered in those frequency bands where we observed significant feature-related coherence difference, using bidirectional Butterworth filter to achieve zero phase-shift. The filter had just three poles to avoid ripples when band-pass filtering using relatively narrow frequency bands. Instantaneous amplitude of the oscillation – the envelope of a filtered signal – was obtained by calculating the Hilbert transform of the detrended and filtered signal and using its complex modulus (magnitude) as shown in Fig. 4a. The resulting signal provides information regarding the amplitude of the oscillation in question, which allows testing whether the cell spikes are related to the modulation amplitude of that oscillation.

In each of the relevant frequency bands, STA was calculated for both the oscillation and its envelope using multiunit spikes as triggers and z-scored the oscillation/envelope signals. Z-scoring is done by subtracting the mean and dividing the signal by its standard deviation in order to normalize all signals to one scale. For every recording site, we compared the peak-to-trough amplitude of the averaged signal wave triggered by multiunit spikes to the distribution of the analogous values obtained by taking randomly an equal number of pseudospikes 1000 times. The peak-to-trough amplitude was calculated within the interval covering one cycle of the averaged oscillation. STA was considered significant if its peak-to-trough amplitude was above 95% of the randomly taken STAs triggered by pseudospikes. The spikes used to compute STAs occurred in the early delay period (from 300 to 700 ms after S1 onset) when the feature-related coherence maximum was observed and when the multiunit response to S1 had already faded.

Spike-triggered average of feature-related oscillations (STA) with individual AE + or AE- cells activity as the trigger.

The STAs were calculated in the same way as described in the previous section, but using as triggers the spiking activities of each of the AE + and AE- LIP cells recorded from the sites which exhibited significant feature-related LFP coherence difference (13 AE- and 17 AE + cells). The AE + /AE- classification of a particular LIP cell was based on the Attentional Enhancement Index,

$$AEI = (R2 - R1)/(R1 + R2),$$

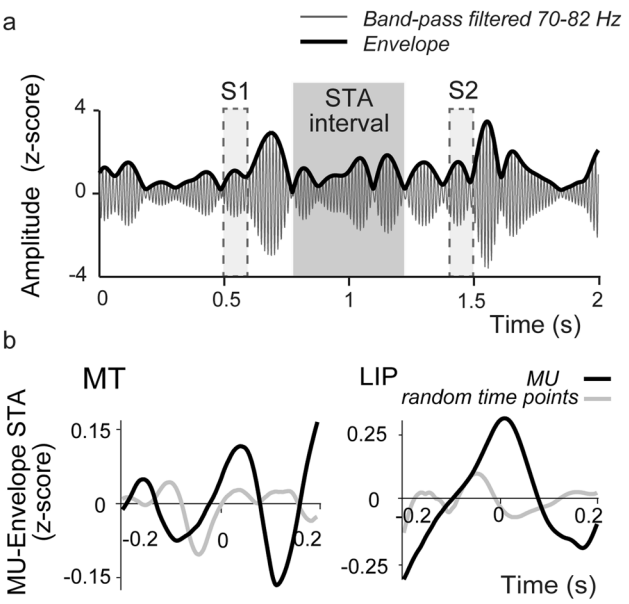

Fig. 4 Schematic depiction of LFP data preparation to obtain an envelope of a feature-related oscillatory LFP activity with examples of MU-triggered STAs. **a** Example of LFP during a single trial, transformed for STA analysis. Thin black trace represents LFP filtered in the frequency range where significant feature-related coherence difference was found and the thick trace shows the magnitude of the Hilbert transform of the filtered LFP (Envelope). Intervals corresponding to S1 and S2 presentations are shown as dashed rectangles and the grey area shows the interval used to calculate STA. **b** Examples of STAs calculated with MUs simultaneously recorded in MT and LIP as triggers for each area and the oscillation envelope shown as the averaged signal. STAs are shown for the recordings made with S1 being of the preferred orientation. Black lines represent STA triggered by MU, grey lines STA triggered by pseudospikes

where R1 and R2 are responses to stimuli S1 and S2 in those match trials, where both stimuli were presented within the receptive field of the cell (for details see Levichkina et al. 2017).

Results

We analysed local field potentials (LFPs) and spike trains which were simultaneously recorded from retinotopically matching sites of areas LIP and MT ($n = 36$) while the monkey was performing the attention demanding DMS task (Fig. 2b). As the monkey had to match stimuli by both location and pattern, its attention had to be directed to both the stimulus feature and where it was presented (Fig. 2b). We had reported earlier that in this task, feedback LIP signals led to increased responses in topographically corresponding, attended locations in MT and reduced responses in unattended locations (Saalmann et al. 2007). Such spatial attention feedback was facilitated by coherent oscillations between the two areas in the beta-to- low gamma frequencies

(25–45 Hz) during the late delay period and around the presentation of the second stimulus. The prominent characteristics of this LIP to MT feedback were its focal attention on the location without any feature selectivity (Saalmann et al. 2007) and its reliance on the AE + cells of LIP (Levichkina et al. 2017). This suggests that the transfer of relevant featural information from AE- to AE + cells occurs earlier than onset of S2, and thus requiring a separate process as argued elsewhere (Levichkina et al. 2017). To reveal any such process, we first determined the range of LFP frequencies which showed significant feature-related or attention-related coherence between the two areas by applying the Aversen statistical technique to multiple comparisons across frequencies (Bokil et al. 2007) (see Fig. 5a and Materials and Methods for details). Figure 5a shows frequency ranges sensitive to feature discrimination or to focal attention, namely the frequencies where significant coherence differences were found between the relevant conditions. Thus the range of frequencies that may be related to feature information was obtained from comparisons between trials with different S1 stimuli. The range of frequencies that may be related to spatial attention was obtained from comparisons between periods of low and high attentional load within a single match trial. As Fig. 5a shows, feature-related coherence differences were found in 17 pairs of recordings and attention-related coherence in 19 pairs. It can be seen that feature-related coherence tends to be in the higher gamma frequency bands (45–200 Hz, blue lines), whereas enhanced attention-related coherence is largely seen in the beta and low gamma frequencies (13–45 Hz, red lines), accompanied by suppression of coherence in the high frequency range (above 47 Hz, grey lines).

Our next goal was to select the frequency bands which can represent task-related changes of coherence between LIP and MT and investigate the time cource of task-related coherences. Thus we next analysed the temporal course of LFP coherence between MT and LIP sites just in the ranges of frequencies where either significant feature-related or significant attention-related coherence occurred. The blue trace in Fig. 5b shows the course of the feature-related LFP coherence between MT and LIP sites, which rises and reaches its maximum during the delay period, approximately 400 ms after the offset of the first stimulus, S1 ($n = 17$ sites, where significant feature-related coherence was found). Thereafter, feature-related coherence comes down to baseline approximately 200 ms before the onset of the second stimulus, S2. On the other hand, attention-related coherence, as shown by the red trace, begins to rise only about 300 ms prior to S2, just as the feature-related coherence declines ($n = 19$, sites where significant attention-related coherence was observed). For much of the course of the trial, there is also a general negative correlation between the two coherences (Pearson

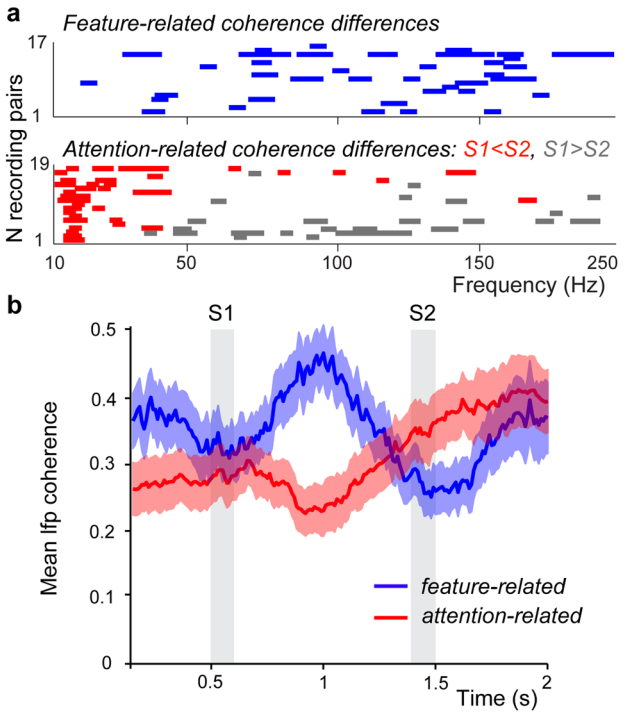

Fig. 5 MT-LIP coherence specificity. **a** Frequency bands where significant LFP coherence differences were observed between MT and LIP recording sites. Feature-related differences (calculated using responses to the optimum and orthogonal orientations of the first (S1) grating, when attention has not been captured by the location or grating orientation) are shown in blue. Attention-related coherence differences (calculated from responses to grating of optimum orientation on RF as S2) are shown in red when coherence was enhanced in the Attention 2 interval (Case A in Fig. 1b) and in grey when it was suppressed in the Attention 2 interval (Cases C and D in Fig. 1b). **b** Mean dynamics of significant LFP coherences averaged across all recording pairs with the same trial structure ($N = 14$ for feature-related and $N = 17$ for attention-related). Feature-related coherence dynamics is shown as dark blue line, attention-related (enhancement) as dark red line. Light-coloured areas represent ± SE. Grey rectangles (S1 and S2) designate periods of stimulus presentation

$r = -0.54$, $p < 0.001$), when calculated over the whole length of the trial).

We also checked the possibility whether the feature-related coherence seen in high gamma frequencies could be caused by 'spike leakage', an artifact due to the LFP reflecting lower frequency components of action potential waveforms (Ray and Maunsell 2011; Scheffer-Teixeira et al. 2013; Waldert et al. 2013). We did this by comparing the time course of the feature-related coherence in these significant gamma frequencies with the time course of multiunit responses (see Materials and Methods for details). Figure 6 (left panel) shows the results for three of the recorded pairs. As can be seen in these three example cases and also when analysed across all those recording sessions where there was both significant LFP coherence and a multiunit response

(MU) to the first grating, there was a significant latency difference between the spike response maxima and the feature-related coherence maxima in both areas (LIP: $N = 14$ recording sites, $p < 0.001$; MT: $N = 15$, $p < 0.001$, Wilcoxon Signed Rank test). In MT, the peak spike response preceded the coherence maximum by 279 ms and in LIP by 341 ms. Time-courses of attention-related coherence are also demonstrated for the same pairs as above for comparison (Fig. 4, right panel). In two of these examples, LIP recording sites contained both AE + and AE- cells (top and bottom panels). In one (middle panels) no AE + cells were present in LIP, which is reflected in the MU responses as absence of the ramping up prior to S2 and no appreciable enhancement of the response to S2. The signature dynamics of the attention-related coherence can be seen in in the late delay period, but it is weaker in the middle-panel example. The clear separation in time of the maximum spike responses and the coherence maxima supports the idea that the high gamma frequencies are not due to spike leakage. A recent study done in area MT also shows that synchrony seen at certain high gamma frequencies (180–220 Hz) represents sensory signals rather than a reflection of spike leakage (Khamechian et al. 2019), adding further credence to the importance of high gamma frequencies in information processing.

Our results suggest that the feature-related synchrony between MT and LIP in the early part of the delay and occurring after the vigorous response to the S1 stimulus is likely to be the first stage of a two-stage model of attention. Thus, the relevant feature information may be retained as a working memory buffer at a site where it can be readily used for directing top-down attention (Kusunoki et al. 2000), consistent with the suggestion that area LIP constructs such a saliency or priority map for this purpose (Bisley and Goldberg 2010). We therefore investigated whether neuronal spiking was related to the amplitude of the LFP oscillations involved in featural coherence in the period of its maxima. It has been noted that interareal communication may not directly evoke responses of cortical neurons, but may just change their excitability by inducing coherence between the two sites (Singer 1994). This has indeed been observed in the attentional modulation of MT by LIP in the delay period just prior to the attentional enhancement seen in the response to the visual stimulus (Saalmann et al. 2007). According to this view, a saliency map relying on initial featural input can be seen as a "patchwork" of synchronized and desynchronized cell assemblies. As synchronised input is more likely to elicit spikes, we analysed the amplitude modulation of the coherent oscillations and the phase-locking of spiking to that amplitude modulation of the LFP. This was done by calculating spike-triggered average (STA) of the amplitude envelopes of the oscillations occurring in the relevant frequency bands (Fig. 4). The envelope helps to reveal the amplitude modulation of the oscillation and to relate neuronal spiking to the changes of that amplitude (Fig. 4a). Using multiunit spikes as the trigger we found that STAs in the majority of the recording sites in both MT (13/17) and LIP (17/17) were significant compared to STAs triggered by pseudospikes around the period of maximum coherence (Details in Materials and Methods, illustrated in Fig. 4b). On 12 occasions we were able to reliably classify the LIP cells recorded from these 17 recording sites as either AE + or AE-. That required the sites to have detectable cell spikes and also LFP that demonstrated feature-selective coherence with the MT site with its receptive field at the topographically corresponding locus. In addition, since LIP cells could belong to different cell types, the probability of finding the cell of interest is relatively low. Interestingly, all AE + cells derived from the same LIP sites as the envelope ($n = 9/9$) triggered significant STAs, while only a small fraction of AE- cells had such a relationship ($n = 3/13$). In other words, spiking activity of the AE + cells is modulated by the amplitude of the coherent feature-related oscillations. This implies that AE + cells, largely responsible for generating spatial attention and attention-related feedback, are the likely recipients of the feature-related information retained during early delay period ($p = 0.0001$, Exact Fisher test). AE + cells synchronize their activities at relatively low frequencies while providing attentional feedback to early visual areas (Levichkina et al. 2017), which is probably why the amplitude modulation of the feature-related oscillation, obviously occurring at a lower frequency rate than the oscillation itself, may be effective in driving these AE + cells.

Discussion

Our results suggest that neuronal activity in LIP not only processes spatial information, as traditionally known, but also featural information. Though this has been known (Toth and Assad 2002, Swaminathan et al. 2013; Ibos and Freedman 2016; Ogawa and Komatsu 2009; Levichkina et al. 2017), our results show that the process continues even after the termination of the direct spike rate response to the visual stimulus, almost entirely as coherent activity in high gamma frequencies. We show that such coherence is not only within LIP but also with MT, which transmits the visual signals to LIP.

The presence of such feature-related synchronised oscillatory activity in the feedforward MT to LIP connection during the delay period may be the site of the earliest emergence of working memory in the brain, prior to its well-documented presence in the prefrontal cortex. It is known that spiking activity in MT does not reflect working memory

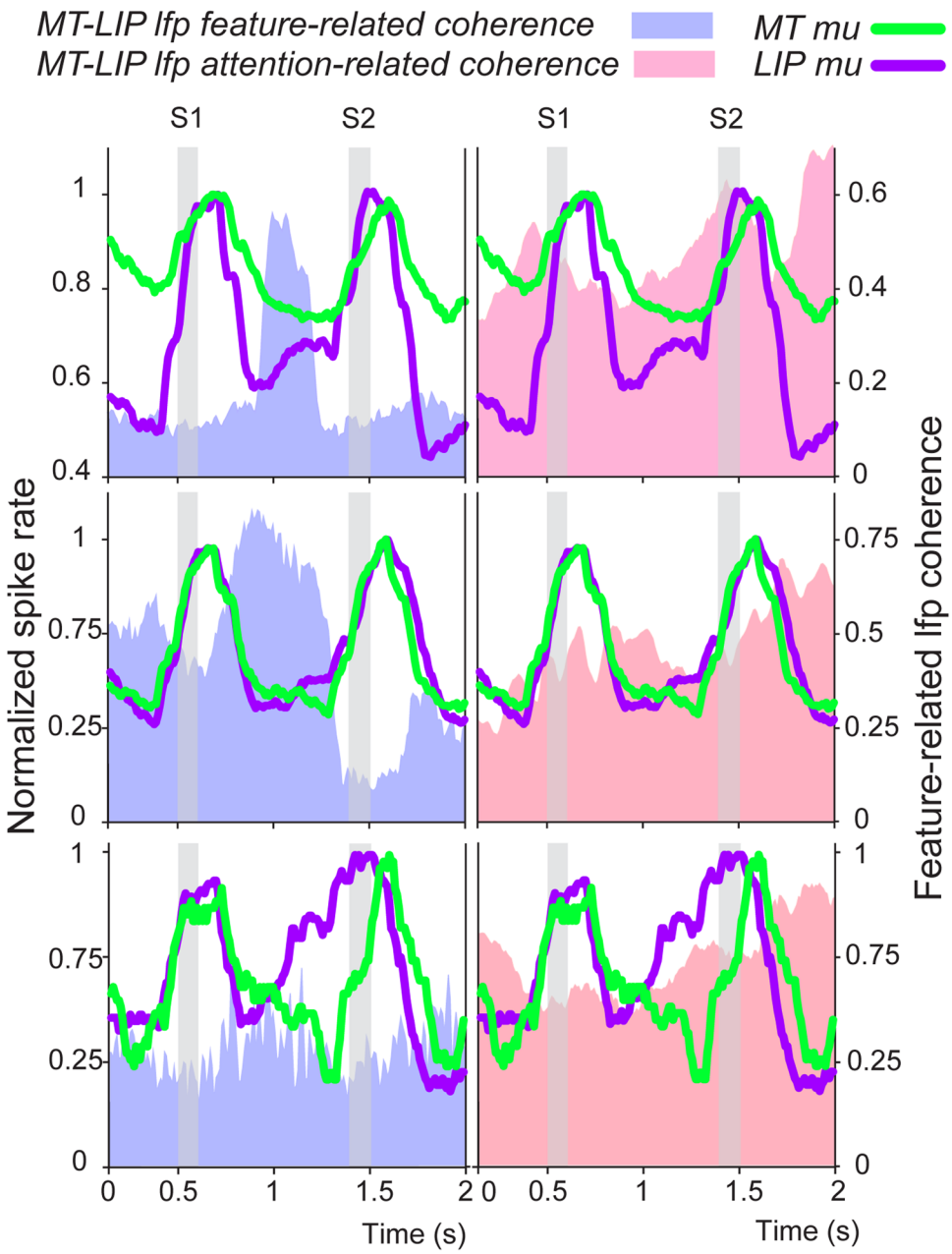

Fig. 6 Time courses of multiunit (mu) responses and two types of coherence between MT and LIP, feature-related coherence and attention-related coherence. Normalized multiunit responses are shown for simultaneously recorded LIP (purple trace) and MT (green trace) pairs, showing spike rates for the match trials when both S1 and S2 had the preferred orientation. Each row represents responses of one pair of retinotopically matched recording sites in MT and LIP. The spike rate is normalized to maximal response, with its scale on the left. Corresponding time courses of feature-related LFP coherence are shown in the left column (light blue filled area) and attention-related related LFP coherence in the right column (light red filled area), with the scale on the right. The same multiunit responses are shown in both columns for comparison with the time courses of each type of coherence

trace during the delay period (Saalmann et al. 2007; Mendoza-Halliday et al. 2014), while higher-order areas such as lateral prefrontal cortex (LPFC) possess neurons that are active during delay period (Mendoza-Halliday et al. 2014). However since feature sensitivity appears in MT prior to these higher-order areas, our results suggest that the short-term retention of the sensory input prior to its being used for attention-related activity in the frontoparietal network may be at least partly in the form of coherent LFP oscillations between MT and LIP. Such coherence-dependent representation without elevation of neuronal firing rates in MT is consistent with the elevation of the LFP power observed in MT during the delay period in the absence of increase in the firing rates of individual neurons (Mendoza-Halliday et al. 2014). Our finding is also consistent with the suggestion that cortical synchronization may play a role in working memory, especially in retention of sensory information for short periods (Lee et al. 2005; Payne and Kounios 2009; Liebe et al. 2012; Hawellek et al. 2017; Nikolaev and van Leeuwen 2019; Noguchi and Kakigi 2020). Given the temporal sequence of the coherencies in our results, we propose that the feature-related LFP coherence drives or at least precedes the oscillatory activity among LIP's attention-related sites, which in turn drives the oscillations in topographically

corresponding sites in MT, as seen here and in our earlier study (Saalmann et al. 2007). AE + cells were previously identified as the ones involved in providing spatial attention feedback to area MT (Saalmann et al. 2007; Levichkina et al. 2017). Modulation of the AE + cells by the amplitude of feature-related oscillation demonstrates how temporarily stored featural information can be used to create a spotlight of attention. This suggestion of ours is further supported by the dependency we see of spiking in both areas on the amplitude of the feature-related LFP oscillations during the early part of the delay when retention of featural information is needed to guide subsequent top-down attention. The ability of feature-related oscillation to modulate activity of attention-related LIP cells shows a saliency map can be used to guide attention.

We suggest that the transient storage of featural information can be LFP coherence-dependent, while more sustained representation of the item in memory such as by reverberating circuits, may be associated with spiking activity in higher-order visual areas such as dorsolateral prefrontal cortex (Funahashi et al. 1989) or by an intermediate layer of cells within LIP itself. It was beyond the scope of the present study to investigate also the involvement of the higher order prefrontal parts of the dorsal visual stream in the transient storage of featural information or the transfer of the information between LIPs's saliency map and the prefrontal executive areas of the frontoparietal attention network. It is therefore an open question whether in our task, the transition from feature selectivity in the early delay period to attention related activity in the late delay period involves the prefrontal spiking activity or takes a more direct route within LIP, but possibly with some degree of prefrontal modulation.

Another intriguing possibility is that the oscillatory activity shown by LIP AE + cells may be directly boosted by either gamma-frequency activity of feature selective AE- cells in MT or by the gamma oscillations of AE- cells of LIP itself. Our current sample size was not large enough to yield a clear outcome. However, given the limited temporal overlap between the two coherences, a direct MT source seems less likely than an intermediate set of cells as discussed above. Nevertheless, in future studies, a fruitful approach to resolve this will be to collect data with multielectrode arrays from different layers of both areas and thus maximizing the probability to observe cells with feedforward vs feedback connectivity and also within-area interactions. Analysis of spike to field coherence between the two cortical areas can reveal the directionality of frequency specific interactions (Pesaran 2010).

The two sets of coherent oscillations we found that operate at two different ranges of frequencies, one at beta/low gamma and the other at higher gamma, demands both a mechanistic explanation as well as a possible functional role. The biophysical characteristics of the cells and the circuitry they are embedded in leave them with a range of frequencies that display optimal resonance (Hutcheon and Yarom 2000). Thus it is plausible that the feature-selective oscillations happening in LIP's input layers arise from cells with a particular morphology, possibly smaller stellate cells, whereas cells in the deeper, output layers providing feedback to MT are larger pyramidal cells, which, due to their lower input resistance and longer time constants are likely to have a lower resonance frequency. Since the purpose of an attentional feedback to early visual areas is likely to be to select the object location for more detailed processing of all features associated with the object (Vidyasagar 1999; Bullier 2001), it is more meaningful for the feedback to be simply spatial and acting on all feature domains at the object locus. In fact, the LIP feedback to MT seems to be purely spatial and makes little distinction between the different feature-selective cells in the corresponding retinotopic location (Saalmann et al. 2007). Human psychophysical experiments also reinforce the point that gating functions of the dorsal stream prioritise spatial locations rather than specific features (Verghese et al. 2013). The different frequencies for the two functions may also be functionally fortuitous, consistent with recent evidence that the brain may use different synchronising frequencies for different functions (Khamechian et al. 2019).

In conclusion, our findings suggest that feature-selective coherence may represent object-related information stored in a saliency map which is used to drive the activity of the cells directing spatial attention. Such a system would enable serial allocation of attention as in classical visual search situations (Levichkina et al. 2017). It would also be useful in situations where such working memory needs to be retained at a site such as the LIP to readily focus spatial attention, e.g., when external featural information may not be continuously available due to occlusion of objects or due to saccades made to different locations in the visual scene. Our study also typically underscores the importance of considering simultaneously all three domains where information processing needs to be considered – temporal, spectral and spatial (Zich et al. 2020).

Acknowledgements We thank Dr. Ivan Pigarev for taking part in some of the early studies. We are grateful to Drs. Andrew Metha and Chris French for critical comments on the manuscript.

Author contributions YS and TRV conceptualised and performed the original experiments and collected the data. EL, TRV and MK conceptualised the present model. EL developed the analytical tools for studying the model. EL and MK did the bulk of the new analysis. EL and TRV wrote the original draft. All authors critically reviewed and edited the final manuscript. TRV acquired funding and administered the project.

Funding This work was supported by project grants (251600, 454676 and 628668) from the Australian National Health and Medical Research

Council to T.R.V. E.L. was partly supported by ARC Centre of Excellence in Integrative Brain Function.

Data availability All data that are not presented in manuscript are available from the corresponding author upon reasonable request.

Declarations

Conflict of interest The authors declare no conflict of interests or competing interests.

Ethics approval The study was conducted as per the guidelines of the National Health and Medical Research Council Australian Code of Practice for the Care and Use of Animals for Scientific Purposes and approved by the University of Melbourne Animal Experimentation Ethics Committee.

References

Baizer JS, Ungerleider LG, Desimone R (1991) Organization of visual inputs to the inferior temporal and posterior parietal cortex in macaques. J Neurosci 11(1):168–190

Bisley JW, Goldberg ME (2010) Attention, intention, and priority in the parietal lobe. Annu Rev Neurosci 33:1–21

Bokil H, Purpura K, Schoffelen J-M, Tomson D, Mitra P (2007) Comparing spectra and coherences for groups of unequal size. J Neurosci Methods 159:337–345

Bullier J (2001) Integrated model of visual processing. Brain Res Rev 36:96–107

Buschman TJ, Miller EK (2007) Top-down versus bottom-up control of attention in the prefrontal and posterior parietal cortices. Science 315:1860–1962

Buzsáki G (2006) Rhythms in the brain. Oxford University Press

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

Buzsáki G, Logothetis N, Singer W (2013) Scaling brain size, keeping timing: evolutionary preservation of brain rhythms. Neuron 80:751–764

Fries P (2005) A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. Trends Cogn Sci 9:474–480

Funahashi S, Bruce CJ, Goldman-Rakic PS (1989) Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex. J Neurophysiol 61:331–349

Gregoriou GG, Gotts SJ, Zhou H, Desimone R (2009) Long-range neural coupling through synchronization with attention. Science 176:35–45

Hawellek DJ, Wong YT, Pesaran B (2016) Temporal coding of reward-guided choice in the posterior parietal cortex. Proc Nat Acad Sci 113:13492–13497

Herrington TM, Assad JA (2010) Temporal sequence of attentional modulation in the lateral intraparietal area and middle temporal area during rapid covert shifts of ateetntion. J Neurosci 30(9):3287–3296

Hutcheon B, Yarom Y (2000) Resonance, oscillation and the intrinsic frequency preferences of neurons. Trends Neurosci 23:216–222

Ibos G, Freedman DJ (2016) Interaction between spatial and feature attention in posterior parietal cortex. Neuron 91:931–943

Jarvis MR, Mitra PP (2001) Sampling properties of the spectrum and coherency of sequences of action potentials. Neural Comput 13:717–749

Khamechian MB, Kozyrev V, Treue S, Esghaei M, Daliri MR (2019) Routing information flow by separate neural synchrony frequencies allows for "functionally labeled lines" in higher primate cortex. Proc Natl Acad Sci USA 116:12506–12515

Kusunoki M, Gottlieb J, Goldberg ME (2000) The lateral intraparietal area as a salience map: the representation of abrupt onset, stimulus motion, and task relevance. Vision Res 40:1459–1468

Lee H, Simpson GV, Logothetis NK, Rainer G (2005) Phase locking of single neuron activity to theta oscillations during working memory in monkey extrastriate visual cortex. Neuron 45:147–156

Levichkina E, Saalmann YB, Vidyasagar TR (2017) Coding of spatial attention priorities and object features in the macaque lateral intraparietal cortex. Physiol Rep 5:e13136

Liebe S, Hoerzer GM, Logothetis NK, Rainer G (2012) Theta coupling between V4 and prefrontal cortex predicts visual short-term memory performance. Nat Neurosci 15:456–462

Luczak A, McNaughton BL, Harris KD (2016) Packet-based communication in the cortex. Ann Rev Neurosci 16:745–755

Maloney RT, Jayakumar J, Levichkina E, Pigarev IN, Vidyasagar TR (2013) Information processing bottlenecks in macaque posterior parietal cortex: an attentional blink? Exp Brain Res 288:365–376

Mendoza-Halliday D, Torres S, Martinez-Trujillo JC (2014) Sharp emergence of feature-selective sustained activity along the dorsal visual pathway. Nat Neurosci 17:1255–1262

Mitra PP, Bokil H (2008) Observed brain dynamics. Oxford University Press

Mitra PP, Pesaran B (1999) Analysis of dynamic brain imaging data. Biophys J 76:691–708

Nikolaev AR, van Leeuwen C (2019) Scene buildup from latent memory representations across eye movements. Front Psychol 9:2701

Noguchi Y, Kakagi R (2020) Temporal codes of visual working memory in the human cerebral cortex brain rhythms assovciated with high memory capacity. Neuroimage 222:117294

Ogawa T, Komatsu H (2009) Condition-dependent and condition-independent target selection in the macaque posterior parietal cortex. J Neurophysiol 101:721–736

Payne L, Kounios J (2009) Coherent oscillatory networks sustain short-term memory retention. Brain Res 1247:126–132

Pesaran B (2010) Neural correlations, decisions, and actions. Curr Opin Neurobiol 20:166–171

Pigarev IN, Nothdurft HC, Kastner S (1997) A reversible system for chronic recordings in macaque monkeys. J Neurosci Methods 77:157–162

Pigarev IN, Saalmann YB, Vidyasagar TR (2009) A minimally invasive and reversible system for chronic recordings from multiple brain sites in macaque monkeys. J Neurosci Methods 181:151–158

Quiroga RQ, Nadasdy Z, Ben-Shaul Y (2004) Unsupervised spike detection and sorting with wavelets and superparamagnetic clustering. Neural Comput 16:1661–1687

Ray S, Maunsell JHR (2011) Network rhythms influence the relationship between spike-triggered local field potential and functional connectivity. J Neurosci 31:12674–121685

Saalmann YB, Pigarev IN, Vidyasagar TR (2007) Neural mechanisms of visual attention: how top-down feedback highlights relevant locations. Science 316:1612–1615

Scheffer-Teixeira R, Belchior H, Leao RN, Ribeiro S, Tort ABL (2013) On high-frequency field oscillations (>100Hz) and the spectral leakage of spiking activity. J Neurosci 33:1535–1539

Sereno AB, Maunsell JHR (1998) Shape selectivity in primate lateral intraparietal cortex. Nature 395:500–503

Singer W (1994) Putative functions of temporal correlations in neocortical processing. In: Davis J, Koch C (eds) Large-Scale Neuronal Theories of the Brain. MIT Press, pp 201–237

Swaminathan SK, Masse NY, Freedman DJ (2013) A comparison of lateral and medial intraparietal areas during a visual categorization task. J Neurosci 33:13157–13170

Toth LJ, Assad JA (2002) Dynamic coding of behaviourally relevant stimuli in parietal cortex. Nature 415:165–168
Verghese A, Anderson AJ, Vidyasagar TR (2013) Space, color, and direction of movement: how do they affect attention? J Vision 13(8):20
Vidyasagar TR (1999) A neuronal model of attentional spotlight: parietal guiding the temporal. Brain Res Rev 30:66–76
Waldert S, Lemon RN, Kraskov A (2013) Influence of spiking activity on cortical local field potentials. J Physiol 591:5291–5303
Wolfe JM (1994) Guided Search 2.0 A revised model of visual search. Psychon Bull Rev 1:202–238
Zich C, Quinn AJ, Mardell LC, Wrad NS, Bestmann S (2020) Dissecting transient burst events. Trends Cogn Sci 24:784–788