

Supplementary Methods:

Gamma-band synchronization in visual cortex predicts speed of change detection

Thilo Womelsdorf*, Pascal Fries*, Partha P. Mitra, Robert Desimone

*These authors contributed equally to this work.

Neurophysiological recording techniques and LFP preprocessing:

Neuronal recordings were made from two hemispheres in two monkeys through chambers implanted over area V4 during surgeries conducted under aseptic conditions with isofluorane anesthesia. Before recording, four to eight tungsten microelectrodes (impedances of 1-2 M Ω) were advanced separately at a very slow rate (1.5 mm/s) to minimize deformation of the cortical surface by the electrode (“dimpling”). Electrodes tips were separated by 650 or 900 μ m. Data amplification, filtering and acquisition were done with a Multichannel Acquisition Processor from Plexon Incorporated. The signal from each electrode was passed through a headstage with unit gain and an output impedance of 240 Ω and then split to separately extract the spike and the LFP components. For spike recordings, the signals were filtered with a passband of 100-8000 Hz, further amplified and digitized with 40 kHz. A threshold was set interactively and spike waveforms were stored for a time window from 150 μ s before to 700 μ s after threshold crossing. The threshold clearly separated spikes from noise, but was chosen to include multi-unit activity. Offline, we performed a principal component analysis of the waveforms and plotted the first against the second principal component. Those waveforms that corresponded to artifacts were excluded. For multi-unit analyses, all other waveforms were accepted and the times of threshold crossing were kept and down-sampled to 1 kHz. For LFP recordings, the signals were filtered with a passband of 0.7-170 Hz, further amplified and digitized at 1 kHz. The powerline artifact was removed from the LFP using the following procedure: All signals had been recorded continuously for the entire duration of the recording session. For each time epoch of interest (and each recording channel), we first took a 10 second epoch out of the continuous signal with the epoch of interest in the middle. We then calculated the Discrete Fourier Transform (DFT) of the 10 s epoch at 50 Hz, 100 Hz and 150 Hz without any tapering. Since the powerline artifact is of a perfectly constant frequency, the 10 s epoch contains integer cycles of the artifact frequencies

and all the artifact energy is contained in those DFTs. We then constructed 50 Hz-, 100 Hz- and 150 Hz-sine waves with the amplitudes and phases as estimated by the respective DFTs and subtracted those sine waves from the 10 s epoch. The epoch of interest was then cut out of the cleaned 10 s epoch. Power spectra of the cleaned 10 s epochs demonstrated that all artifact energy was eliminated, leaving a notch of a bin width of 0.1 Hz (=1/10 s). The actual spectral analysis used the multi-taper method, with a spectral smoothing of ± 4 Hz (for frequencies between 8 and 20 Hz) or ± 16 Hz (for frequencies between 20 and 100 Hz). Thus, the notch typically became invisible.

Time-frequency spectral analysis based on multitapering:

For the assessment of power and coherence spectra, we used windows of ± 125 ms length that were moved over the data in steps of 10 ms from 500 ms before to 150 ms after the stimulus change. For each window, we applied multitaper methods to achieve optimal spectral concentration¹⁻³. Multitapering involves the multiplication of data segments with multiple tapers before Fourier transformation. Tapering effectively concentrates spectral estimates across a specified frequency band. We used two different sets of tapers: For the low frequency range (8 to 20 Hz), we chose a spectral concentration over ± 4 Hz, while for the high frequency range (20 to 100 Hz), we chose a spectral concentration over ± 16 Hz.

For each taper, the data epoch was multiplied with that taper and then Fourier transformed, giving the windowed Fourier transform, $\tilde{x}_k(f)$:

$$\tilde{x}_k(f) = \sum_t^N w_k(t) x_t e^{-2\pi jft}$$

where x_t , ($t=1,2,\dots,N$) is the time series of the signal under consideration and $w_k(t)$, ($k=1,2,\dots,K$) are K orthogonal taper functions.

The multitaper estimates for the spectrum $S_x(f)$ and the cross-spectrum $S_{yx}(f)$ are given by

$$S_x(f) = \frac{1}{K} \sum_l^K |\tilde{x}_k(f)|^2$$

$$S_{yx}(f) = \frac{1}{K} \sum_l^K \tilde{y}_k(f) \tilde{x}_k^*(f)$$

Spectra and cross-spectra are averaged over trials before calculating the coherency

$$C_{yx}(f) = \frac{S_{yx}(f)}{\sqrt{S_x(f)S_y(f)}}$$

Coherency is a complex quantity. Its absolute value is termed coherence and ranges from 0 to 1. A coherence value of 1 indicates that the two signals have a constant phase relationship (and amplitude covariation), a value of 0 indicates the absence of any phase relationship.

Coherence estimates have a positive bias that decreases with an increase in the amount of data. To correct for this, a non-linear transformation was applied to the coherence spectra¹, which will be referred to as a z-transformation. If C is the untransformed coherence estimate and ν is the degrees of freedom (two times the total number of tapers applied), then the variable

$$q = \sqrt{-(\nu - 2)\log(1 - |C|^2)}$$

has a Raleigh distribution with density

$$p(q) = qe^{-q^2/2}.$$

This density function does not depend on ν and furthermore has a tail that closely resembles a Gaussian. For certain values of a fitting parameter β , a further linear transformation

$$r = \beta(q - \beta)$$

leads to a distribution that closely resembles a standard normal Gaussian for $r > 2$. We therefore refer to r as the z-transformed coherence. A reasonable choice for β is 23/20.

Spike-field coherence was calculated with spikes and LFP from the same and from different electrodes. Restricting the SFC analysis to SFCs with spikes and LFP from different electrodes left the results unchanged.

Neural activity sorted according to reaction time:

We assessed power, coherence and spike rate for trials with the 25 % fastest and 25 % slowest reaction times of individual recording sessions. The mean RT for the subset of fast (slow) trials was 346 ms (490 ms) (median: 359 ms / 485 ms). This separation

ensured the least overlap of fast and slow reaction time bins across recording sites (such that slowest reaction times in the fast bin did not overlap with the fastest reaction times in the slow bin across sessions), while providing as much neural data as possible to calculate the spectral estimates. Moreover, percentile separation ensures an equal number of trials in each subset and thereby eliminates sampling biases¹.

We wanted to test whether the variable *gamma-band power* (or *gamma-band spike field coherence*, or *firing rate*) differed between the conditions *fast response* and *slow response*. Therefore, we calculated t-values for the difference between the two conditions and across recording sites (or pairs of sites), separately for all time windows in the sliding window analysis. We performed a correction for the multiple comparisons done, using a non-parametric permutation approach for significance testing^{4,5}. To this end, the following procedure was performed 10000 times: For each recordings site (or pair of sites), a random decision was made to either exchange the two conditions (50% probability) or leave them unchanged (50% probability). Subsequently, paired t-tests were determined across the sites (or pairs of sites) between the two conditions and for all time windows. The maximal and the minimal t-value across all time windows was kept, resulting in 10000 maximal and 10000 minimal t-values. From this empirical distribution of global maxima and minima, we determined the 2.5 % and the 97.5 % points, $t(\text{global}, 2.5 \%)$ and $t(\text{global}, 97.5 \%)$. For each time window, we then determined the t-value between the non-randomized conditions. The non-randomized t-value for a given time window was considered significant if it was larger than $t(\text{global}, 97.5 \%)$ or less than $t(\text{global}, 2.5 \%)$. This procedure corresponds to a two-sided test with a global false positive rate of 5 % and corrects for the multiple comparisons across the time interval.

Correlation analysis:

To determine the predictive capability of neuronal activity parameters for reaction time, we calculated the Pearson correlation coefficient between the trial-by-trial variations in reaction times and the trial-by-trial variations of power, coherence and spike rate. This was done separately for each analyzed time window and, in the case of power and coherence, separately for each analyzed frequency. Correlation coefficients were Fisher Z-transformed and those Fisher Z-values were transformed to z-scores. Z-scores were pooled across recording sites or pairs of sites according to:

$$z = \frac{1}{N} \sum_{i=1}^N z_i$$

with z_i being the z-score of the i-th recording site or pair of sites. Note that spike field coherence cannot be estimated directly for short time epochs of single trials. We therefore derived a new analysis method to obtain single-trial coherence estimates based on z-transformed coherence values as outlined in the following paragraph.

Estimation of spike-field coherence of single-trials:

Coherence cannot be estimated directly for single short data epochs (of two signals). The reason is that coherence is in itself a statistic about the distribution of phase differences between the two signals, that requires multiple estimates of this phase difference. Here, we aimed at estimating the correlation between variations of reaction times across trials and variations of spike-field coherence across trials. We furthermore aimed at doing this for short (250 ms) time windows, that cannot be reasonably cut into a sufficient number of pieces, nor allow the application of a sufficient number of multi-tapers to estimate coherence directly from the single trial data. In order to nevertheless obtain an estimate of single-trial coherence, we computed single-trial coherence pseudovalues (STCP). The rationale of this is that while we cannot determine the single-trial coherence directly, we can determine the coherence for all trials and we can also determine the coherence for all-but-one trials. For a linear function F of a sample S , the value of $F(S_i)$, i.e. the value of the function for the i-th observation, is identical to the pseudovalue

$$P = N \times F(S) - (N - 1) \times F(S_{(i)}),$$

With $S_{(i)}$ being the entire sample with the i-th observation left out. We accordingly determined the *STCP* for trial i as

$$STCP = N \times C(S) - (N - 1) \times C(S_{(i)}),$$

with C being the z-transformed coherence.

Reference List

1. Jarvis, M.R. & Mitra, P.P. *Neural Comput.* **13**, 717-749 (2001).
2. Mitra, P.P. & Pesaran, B. *Biophys. J.* **76**, 691-708 (1999).
3. Pesaran, B., Pezaris, J.S., Sahani, M., Mitra, P.P. & Andersen, R.A. *Nat. Neurosci.* **5**, 805-811 (2002).
4. Maris, E. *Psychophysiology* **41**, 142-151 (2004).
5. Nichols, T.E. & Holmes, A.P. *Hum. Brain Mapp.* **15**, 1-25 (2001).