



## Optimal band separation of extracellular field potentials

Cesare Magri<sup>a,b</sup>, Alberto Mazzoni<sup>c</sup>, Nikos K. Logothetis<sup>a,d</sup>, Stefano Panzeri<sup>c,\*</sup>

<sup>a</sup> Max Planck Institute for Biological Cybernetics, 38 Spemannstrasse, 72076 Tübingen, Germany

<sup>b</sup> Bernstein Center for Computational Neuroscience Tübingen, Germany

<sup>c</sup> Robotics, Brain and Cognitive Sciences Department, Italian Institute of Technology, Via Morego 30, 16163 Genova, Italy

<sup>d</sup> Division of Imaging Science and Biomedical Engineering, University of Manchester, Manchester M13 9PT, United Kingdom

### ARTICLE INFO

#### Article history:

Received 19 July 2011

Received in revised form 30 October 2011

Accepted 2 November 2011

#### Keywords:

Neural code  
Local Field Potential  
Visual cortex  
Mutual information  
Frequency bands

### ABSTRACT

Local Field Potentials (LFPs) exhibit a broadband spectral structure that is traditionally partitioned into distinct frequency bands which are thought to originate from different types of neural events triggered by different processing pathways. However, the exact frequency boundaries of these processes are not known and, as a result, the frequency bands are often selected based on intuition, previous literature or visual inspection of the data. Here, we address these problems by developing a rigorous method for defining LFP frequency bands and their boundaries. The criterion introduced for determining the boundaries delimiting the bands is to maximize the information about an external correlate carried jointly by all bands in the partition. The method first partitions the LFP frequency range into two bands and then successively increases the number of bands in the partition. We applied the partitioning method to LFPs recorded from primary visual cortex of anaesthetized macaques, and we determined the optimal band partitioning that describes the encoding of naturalistic visual stimuli. The first optimal boundary partitioned the LFP response at 60 Hz into low and high frequencies, which had been previously found to convey independent information about the natural movie correlate. The second optimal boundary divided the high-frequency range at approximately 100 Hz into gamma and high-gamma frequencies, consistent with recent reports that these two bands reflect partly distinct neural processes. A third important boundary was at 25 Hz and it split the LFP range below 50 Hz into a stimulus-informative and a stimulus-independent band.

© 2011 Elsevier B.V. All rights reserved.

### 1. Introduction

Local Field Potentials (LFPs) are a massed signal which captures multiple neural contributions, from dendrosomatic dipoles generated by synaptic activity to non-synaptic slow activity such as voltage-dependent membrane oscillations and spike afterpotentials (Logothetis, 2003). Since LFPs are sensitive to supra and subthreshold processes, investigating stimulus encoding by LFPs can offer additional insights into sensory representations beyond those offered by only measuring neuronal spike trains (Belitski et al., 2010; Logothetis, 2003; Montemurro et al., 2008; Nicoletis and Lebedev, 2009; Ray et al., 2008). Their stimulus tuning has been intensively investigated in recent years (Belitski et al., 2008; Berens et al., 2008; Frien et al., 2000; Kayser and König, 2004; Liu and Newsome, 2006; Siegel and König, 2003). LFPs exhibit a broadband spectral structure which is traditionally partitioned into distinct frequency bands, initially introduced in the human EEG literature, which are thought to originate from different types of neural events

triggered by different neural processing pathways (Buzsáki, 2006). However, the exact frequency boundaries of these processes are not known. Frequency bands are thus often selected based on intuition, previous results in the literature or visual inspection of the data. As a result, the band boundaries vary greatly across studies. For example, the gamma band is sometimes referred to as a narrow (a few Hz wide) frequency range (Llinas and Ribary, 1993; Miltner et al., 1999) and other times as a wide interval spanning tens of Hz (Chrobak and Buzsáki, 1998; Kruse and Eckhorn, 1996).

A rigorous approach to defining the boundaries of a stimulus-tuned LFP band has been proposed by Siegel and König (2003). Their method, based on maximizing an index of selectivity to a stimulus parameter, allows the identification of both the parts of the frequency spectrum that are best tuned to a specific stimulus, and of the boundaries that optimize the tuning of a single band to the stimulus. This approach, which was successfully applied to identify the optimal boundaries for orientation tuning of the gamma band in visual cortex (Siegel and König, 2003), requires a model of a tuning curve describing the relationship between an external correlate of interest and the induced changes in the LFP band considered. A complementary approach was developed by our group (Belitski et al., 2008) for investigating the relationship between the power

\* Corresponding author. Tel.: +39 10 70781437.

E-mail address: [stefano.panzeri@iit.it](mailto:stefano.panzeri@iit.it) (S. Panzeri).

at each frequency in the Fourier decomposition – the smallest frequency element in which we can break the LFP range – and complex visual stimuli. This analysis revealed which regions in the LFP spectrum are informative about naturalistic movie stimuli and which regions are, instead, stimulus-unrelated. However, these previous methods cannot determine how to partition the entire LFP range in such a way that the bands in the partition collectively carry the largest possible amount of information about a set of stimuli. In fact, by fine-tuning one band at a time one might fail to take into account the relationships between different frequency regions. For instance, tuning one band at a time may individuate a set of bands which are individually highly informative and that all carry very similar information, but may fail to individuate bands that express complementary aspects of stimulus tuning and thus carry more information collectively.

Here we introduce a new procedure for partitioning the LFP response into bands that collectively carry maximal mutual information about a set of external correlates. This procedure does not make any assumption on the type of relationship between the power in an LFP band and the external correlate of interest. Rather, because of the properties of mutual information, it naturally takes into account linear and non-linear correlations at any order as well as any relationship between the powers in different bands. The article is structured as follows. After defining the procedure, we investigate and address the problems related to its computational implementation. We then apply the technique to the electrophysiological signal recorded from the primary visual cortex of anesthetized macaques, and we determine a stable partition of LFPs into bands that carries an optimal amount of information about all visual features in the movie.

## 2. Materials and methods

### 2.1. Recording procedures, sensory stimuli and data extraction

We analyzed sensory evoked LFPs and MUA recorded from primary visual cortex (V1) of two anaesthetized monkeys. The recording procedures and stimulation paradigms for the different datasets are described shortly in the following and have been detailed previously (Belitski et al., 2008; 2010; Montemurro et al., 2008). All experiments were approved by the local authorities (Regierungspräsidium) and were in full compliance with the guidelines of the European Community (EUVD 86/609/EEC) for the care and use of laboratory animals. Prior to the experiments, form-fitting head posts and recording chambers were implanted during an aseptic and sterile surgical procedure.

Recordings from V1 of two adult rhesus monkeys (Macaca Mulatta) were obtained while the animals were anaesthetized (remifentanyl, 1 µg/kg/min), muscle-relaxed (mivacurium, 5 mg/kg/h) and ventilated (end-tidal CO<sub>2</sub> 33 mmHg, oxygen saturation >95%). Body temperature was kept constant and lactated Ringer's solution supplied (10 ml/kg/h). Vital signs (SpO<sub>2</sub>, ECG, blood pressure, endtidal CO<sub>2</sub>) were continuously monitored. Neuronal activity was recorded from opercular V1 (foveal and para-foveal representations) using microelectrodes (FHC Inc., Bowdoinham, Maine, 300–800 k Ohms), signals were high-pass filtered (1 Hz; digital two pole Butterworth filter), amplified using an Alpha Omega amplifier system (Alpha Omega Engineering) and digitized at 20.83 kHz. Binocular visual stimuli were presented at a resolution of 640 × 480 pixels (field of view: 30° × 23°, 24 bit true color, 60 Hz refresh) using a fiberoptic system (Avotec, Silent Vision, Florida). Stimuli consisted of 'naturalistic' complex and commercially available movies (30 Hz frame rate), from which 3.5–4.3 min long sequences were presented and repeated 30–40 times. We confirmed that the receptive fields of all multi-unit

activity from each recording site were well within area of visual stimulation (Rasch et al., 2008).

## 3. Information theoretic analysis

### 3.1. Definition of the main information theoretic quantities

To determine how well the power in a given LFP band encodes the sensory stimulus we computed Shannon's mutual information between the stimulus and the power of the LFP band (Shannon, 1948), which is defined as

$$I(S; B) = \sum_{b,s} P(s)P(b|s) \log_2 \frac{P(b|s)}{P(b)} \quad (1)$$

where  $P(s)$  is the probability of presenting stimulus  $s$ ,  $P(b|s)$  is the probability of observing a power value  $b$  given presentation of stimulus  $s$ , and  $P(b)$  is the probability of observing power value  $b$  across all trials to any stimulus.  $I(S; B)$  quantifies, in units of bits, the reduction of uncertainty about the stimulus that can be gained from the observation of a single-trial power value. One bit corresponds to a reduction of uncertainty of a factor of two.  $I(S; B)$  is zero only when the stimulus and the power are statistically independent, indicating that no knowledge about the stimulus can be gained by observing the response.

Eq. (1) can be extended to quantify the information about a set of stimuli gained by the simultaneous observation of the power in  $L$  different LFP bands, as follows:

$$I(S; B_1 \dots B_L) = \sum_{b,s} P(s)P(b_1 \dots b_L|s) \log_2 \frac{P(b_1 \dots b_L|s)}{P(b_1 \dots b_L)} \quad (2)$$

where in the above  $P(b_1 \dots b_L|s)$  is the probability of the joint observation of observing jointly the power values  $b_1 \dots b_L$  in the considered bands during a single-trial presentation of stimulus  $s$ .

### 3.2. Definition of stimulus set

Computation of information requires the definition of a set of sensory stimuli  $S$  about which the neural response carries information. Our dataset was made of recordings in response to repeated presentation of the same movie. To quantify the amount of visual information carried by the LFP power in this experiment, we used a feature-independent definition of stimulus, described next (de Ruyter van Steveninck et al., 1997; Strong et al., 1998). The movie presentation time was divided into non-overlapping time windows of length  $T=960$  ms and each window was labeled with a stimulus identification number ( $s=1, 2, \dots$ ). In this paper we present results obtained with a window of length  $T=960$  ms, but results were qualitatively similar when varying the length of the running window  $T$  in the range 480–1920 ms (not shown). The LFP power was quantified separately in each window (see below) and we computed the information between the LFP power and the section of the movie ("stimulus window") at which the response was collected. It is important to note that the response collected in the  $s$ -th window is potentially elicited by the whole temporal pattern of visual stimulation up to the last time point of stimulus window  $s$ . Therefore this definition of information carried by the response about which stimulus window was being presented takes into account the potential contributions of all stimulus features at all previous movie frames as well as the contribution of the visual features in the current movie frame. As such, the information calculation does not assume an instantaneous relationship between LFPs and stimuli, and takes into account the effects of all time lags between the neural response and any visual feature that provoked it. This definition of stimulus set has the advantage that it does not make any assumptions about which stimulus features or combinations of

them, either presented at the current time or in the period prior to the considered neural response, provokes the neuronal response. This also makes the information value formally invariant to any rigid time shift between stimulus windows and neural response corresponding to any assumption of a given latency value between stimuli and responses.

### 3.3. Computation of spectral power in each band for information calculation

Computation of information requires the evaluation of the single-trial spectral power of each band in each of the “stimulus windows” of length  $T$  defined above. Here we computed the single-trial power in each window using the multitaper method (Percival and Walden, 2006) with time-bandwidth parameter  $NW=2$ . This method of spectral estimation was chosen because it gives an excellent trade-off between bias and variance and because it was shown to work well on LFP data (Belitski et al., 2008; Pesaran et al., 2002). To compute the power in a given frequency band we integrated the spectral power between the band boundaries. However, we verified that the use of alternative methods for band power estimation within a window, such as averaging over the time points within the window the instantaneous Hilbert power of a signal band-passed within the defined bound boundaries with a Kaiser filter (as done e.g. in (Belitski et al., 2010; Kayser et al., 2009)), gave results which were essentially identical to that obtained with the summation of Multitaper power used here (see (Magri, 2009) for more details).

### 3.4. Information redundancy

The information redundancy between the power in different LFP bands was defined as the difference between the information provided by the joint signals  $b_1, \dots, b_L$  and the sum of the independent information values (Belitski et al., 2008; Hatsopoulos et al., 1998; Pola et al., 2003; Schneidman et al., 2003):

$$Red(B_1 \dots B_L) = I(S; B_1 \dots B_L) - \sum_{i=1}^L I(S; B_i) \quad (3)$$

Positive values of redundancy indicate that the joint knowledge of the power of all bands carries less information than the sum of the information provided by each individual band. Negative values of redundancy denote the presence of synergistic interaction between different bands, which make the total information greater than the sum of the information provided individually by each band. We normalized redundancy expressing it as percentage of the total information carried by the joint observation of all bands in the partition. This *percentage redundancy* was thus defined as

$$Red(B_1 \dots B_L) = 100 \times \frac{Red(B_1 \dots B_L)}{I(S; B_1 \dots B_L)} \quad (4)$$

We also computed the information carried by the power at a single Fourier coefficient rather within an LFP band. The calculation of the single frequency bin information was done exactly as described above, the only difference being that  $b$  was now the power of a single Fourier coefficient.

## 4. Computation of information

The computation of the information with Eqs. (1) and (2) requires the estimation of stimulus-conditional response probabilities  $P(b|s)$  and  $P(b_1, \dots, b_L|s)$ . These probabilities are not known a priori and must be measured experimentally from a finite number of trials. In this work we compared two complementary procedures for measuring the probability distributions and the corresponding information values, namely, the Direct Method and the Gaussian

Method described below. All information analyses were performed using the Information Breakdown Toolbox ([www.ibtb.org](http://www.ibtb.org); (Magri et al., 2009)).

### 4.1. The Direct Method

Computation of information according to the Direct Method (Strong et al., 1998) was as follows. We first discretized the response space by binning the response power into  $M$  equi-populated bins. The parameter  $M$  was varied between 3 and 32 in order to evaluate the effect of the discretization on the Direct information estimates. We then estimated the probabilities of the discrete neural responses by simply computing the fraction of trials in which each response value is observed, and then by inserting these response-probability estimates into the information Eqs. (1) and (2). The Direct Method, being based on empirically computing the probability histograms of discrete or discretized neural responses, does not make any assumption on the shape of the probability distributions. This makes the Direct Method widely applicable to many different types of data.

Even when each stimulus is repeated many times, as was the case here, the estimated probabilities suffer from finite sampling errors, which induce a systematic error (bias) in information estimates (Panzeri et al., 2007). To correct for the bias, we used a quadratic extrapolation procedure (Strong et al., 1998) to estimate and subtract out the bias of each information quantity. When computing the information conveyed by the joint observation of the LFP power in more than one band we additionally applied the “shuffling procedure” described in (Montemurro et al., 2007; Panzeri et al., 2007), which greatly reduces the bias of multidimensional information estimates and typically makes the residual bias negative. This implies that our estimates of the joint information carried by two signals (and thus of redundancy) tend to be slightly biased downward.

### 4.2. The Gaussian Method

An alternative approach to the Direct estimation of information is to use analytical models of the probability distributions; fit these distributions to the data; and then compute the information from these probability models. The Gaussian Method for computing information and entropies is the one based on fitting response probabilities to Gaussian functions.

Under the Gaussian hypothesis, mutual information is given by a simple function of the trial-to-trial variance of LFP responses (Cover and Thomas, 2006)

$$I(S; \mathbf{B}) = \frac{1}{2} (\log_2 [2\pi e |\sigma^2(\mathbf{B})|] - \log_2 [2\pi e |\sigma_s^2(\mathbf{B})|]) \quad (5)$$

where  $|\sigma^2(\mathbf{B})|$  and  $|\sigma_s^2(\mathbf{B})|$  are the determinants of the matrices of covariance computed across trials and stimuli, and across trials to stimulus  $s$ , respectively.

The Gaussian Method has two main advantages: it does not require data discretization prior to the information calculation, and it depends only on a few parameters that characterize the neural response (*i.e.*, the variance and covariance of the responses), which makes it more data-robust and less prone to sampling bias than the Direct calculation (Magri et al., 2009; Yu et al., 2010). The potential danger with this approach is that the estimates provided by Eq. (2) may be inaccurate if the underlying distributions are not close enough to Gaussians.

Although less severe than in the Direct case, the bias of the information calculation due to limited sampling is still present when using the Gaussian Method. When the underlying distributions are

Gaussian an exact expression for the bias of  $I(S;B)$  can be computed as follows (Goodman, 1963; Misra et al., 2005; Oyman et al., 2003):

$$\text{bias}[I(S; \mathbf{B})] = g_{\text{bias}}(N_{tr} \times N_{win}) - \sum_s p(s) g_{\text{bias}}(N_{tr}) \quad (6)$$

where  $N_{tr}$  is the number of available trials per stimulus (in our experiment the number of movie repetitions) and  $N_{win}$  is the number of stimuli (in our case the number of windows of length  $T$  into which the movie presentation time was subdivided). The function  $g_{\text{bias}}(\cdot)$  is defined as

$$g_{\text{bias}}(N) = \frac{1}{2 \ln 2} \left[ L \ln \left( \frac{2}{N-1} \right) + \sum_{j=1}^L \gamma \left( \frac{N-j}{2} \right) \right] \quad (7)$$

and  $\gamma$  is the polygamma function and  $L$  is the number of bands considered.

When using the Gaussian Method we took the cubic root of power; we fitted the distribution of this response to each stimulus to a Gaussian; and we finally computed the information through Eq. (5) subtracting the analytic Gaussian bias correction, Eq. (6). The reason for applying the cubic root transformation is that multitaper power estimates are asymptotically chi-square distributed (Percival and Walden, 2006) thus their cubic root is approximately Gaussian (Wilson and Hilferty, 1931). The cubic root operation, being monotonic, does not affect the underlying information values of the power, but it makes response probabilities much more Gaussian and thus facilitates information estimation with the Gaussian Method.

#### 4.3. Description of the optimal procedure for partitioning the LFP response

The procedure that we propose for an optimal partition of the LFP range into a given number of bands works as follows. First, we define the high and the low cutoffs frequencies that delimit the part of the spectrum of the extracellular signal which is taken to be the LFP. The high frequency cutoff, usually taken in the range between one and a few hundred Hz, is not consistently defined in the literature. Our approach here is to consider the effect of different values of this parameters upon the band partition. On the other hand, the low frequency cutoff of the LFP is usually constrained only by the high pass filter of the preamplifier in the recording setup, which in our case was 1 Hz (see Section 2). Because the power of the extracellular signal is negligible below this high-pass value, in the following for simplicity we will report our results using a low frequency cutoff of 0 Hz. However, the results would have been identical had we used a low frequency cutoff of 1 Hz.

Once the LFP cutoff frequencies have been specified, our method seeks to partition this range into a number of bands chosen to maximize the information carried by the LFP partition about an external set of correlates of interest, such as a set of sensory stimuli. This means that, among all possible partitions of the LFP range into the considered number of bands, we select the one which extracts the largest stimulus-information as defined in Eq. (2). Additionally, we propose to implement this procedure in an incremental way, *i.e.* we start by partitioning the LFP response into two bands and we then increase the number of bands considered. This is useful for two reasons. First, it provides a way to rank the boundaries by importance, under the assumption that the boundaries defining the optimal partition with small number of bands are more fundamental than boundaries appearing only in optimal partitions with a larger number of bands. Second, the incremental procedure permits to identify the boundaries that are most stable to the number of bands considered, *i.e.* if a boundary identifies an optimal partition into a given number of bands, the same boundary should

also belong to the optimal partition with one more band. Indications that the bands are stable to recursive refinement of the partition are important to support the notion that the identified band boundaries are true “separators” of different neural processes, or at least are stable boundaries of spectral regions which express complementary information about the stimulus.

## 5. Results

We recorded LFPs from 23 sites in the primary visual cortex (V1) from two anesthetized monkeys, using an array of electrodes, during the presentation of a 3.5- to 4.3-min-long sequence from commercially available movies. Each movie extract was repeated 30–40 times. The responses of the same site to different movies were analyzed separately. In some cases, we were able to measure the responses of a recording site to repeated presentations of a second or third different movie. In such cases, the responses of the same site to different movies were analyzed separately. This gave us a total of 57 cases, which we collected all together in our population analysis and we considered each of them, for simplicity, as a different recording site.

The spectrogram of the LFP (Fig. 1A) was typical of primary cortical recordings, in that it showed a broad band spectrum with significant power at frequencies in the whole range 0–100 Hz and with the power decreasing at higher frequencies after reaching an initial peak at low frequencies (in this case the peak was at 2 Hz). Presentation of the movie stimulus induced a marked increase in spectral power between 40 and 150 Hz over the power observed during spontaneous activity.

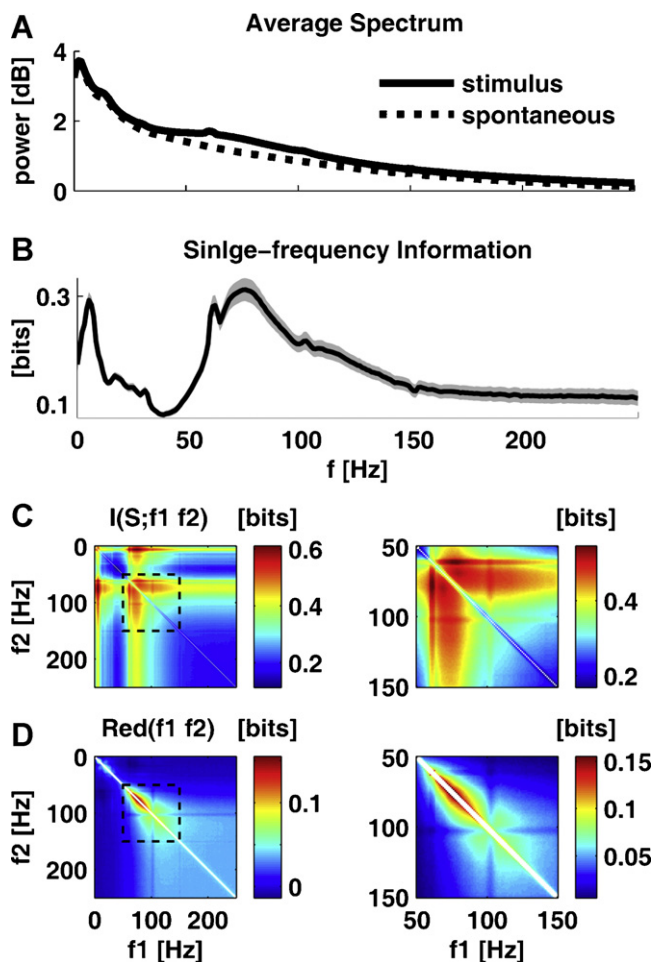
### 5.1. The information conveyed by single LFP frequencies

The dataset used in this study has been investigated in several previous works from our group (Belitski et al., 2008, 2010; Magri et al., 2009). In these studies we investigated the information about a naturalistic movie stimulus carried jointly by the LFP power at each frequency of the Fourier spectrum (Belitski et al., 2008; Magri et al., 2009), and the correlations between LFP power at different frequencies (Magri et al., 2009). Here we briefly summarize these results from these previous works, with the aim of pointing to the attention of the reader the facts and findings that are most relevant to the analysis presented in this study. We also facilitate the comparison between the present study and our previous work by re-computing our previously published information quantities with exactly the same information estimation algorithm (the Gaussian Method) used for the new optimal band partition results.

We begin this summary by reporting the information that LFP power carries about which scene was being presented. The average over all recordings is reported in Fig. 1B. We found two informative bands in the LFP spectrum: a low frequency range below 12 Hz (corresponding to the delta and theta bands) and a high frequency range (above 50 Hz) in the gamma and high-gamma bands. While the information at low frequencies was high only over a narrow band centered around 5 Hz, the information peak at high frequencies was broad and long-tailed toward frequencies higher than 100 Hz. Intermediate frequencies [12–50 Hz] did not carry much information.

Having established that both gamma, high-gamma and low LFP frequencies convey information, the next important question is to determine whether the different informative frequencies ranges are redundant or not, *i.e.* whether or not they carry the same or different stimulus information. We computed both the joint information carried by the power of pairs of LFP frequencies (Eq. (2)) and their redundancy (Eq. (3)). The information obtained by the combined knowledge of the power at low frequencies and the power at





**Fig. 1.** The information carried by LFP power at single frequencies. Results are averaged over a set of 57 V1 recording sites from anaesthetized macaques. Information estimates were computed using the Gaussian Method corrected with the Gaussian bias correction procedure and using the cube root of the LFP power intensities. (A) Average power spectrum of the LFPs recorded during the presentation of a naturalistic movie (solid line) and during spontaneous activity (dashed line). (B) Information about a naturalistic movie stimulus carried by LFP power at single frequencies. The area indicates the SEM. (C) Left: information about a naturalistic movie carried jointly by the LFP power at two frequencies. Right: inset of the information values in the region 50–150 Hz indicated by the rectangle in the left panel. (D) Left: the redundancy between the information carried by the LFP power at two frequencies. Right: inset of the redundancy values in the region 50–150 Hz indicated by the rectangle in the left panel.

gamma frequencies was nearly the sum of the information carried by the two frequencies separately (Fig. 1C). This means that the redundancy between the information carried by the power of high and low frequencies is nearly zero (Fig. 1D). This zero redundancy reflected the fact that, as shown in (Belitski et al., 2008; Magri et al., 2009), LFP frequencies below 40 Hz were totally decoupled from gamma and high-gamma oscillations (both in terms of similarity of stimulus selectivity (“signal correlations”), and of trial-to-trial covariations (“noise correlations”).

In contrast, frequencies in the gamma band and high-gamma band were highly redundant between each other (Fig. 1D). This was because (Belitski et al., 2008; Magri et al., 2009) gamma frequencies had a similar response profile across different movie scenes (in other words, they had high “signal correlation”). The redundancy was particularly pronounced between frequencies in the range 50–100 Hz, suggesting that all frequencies in this range reflect to a large extent the same network phenomenon. We found considerably smaller redundancy between frequencies in the gamma range

[50–100 Hz] and those in the high-gamma range, suggesting that gamma and high-gamma may reflect partly different neural phenomena.

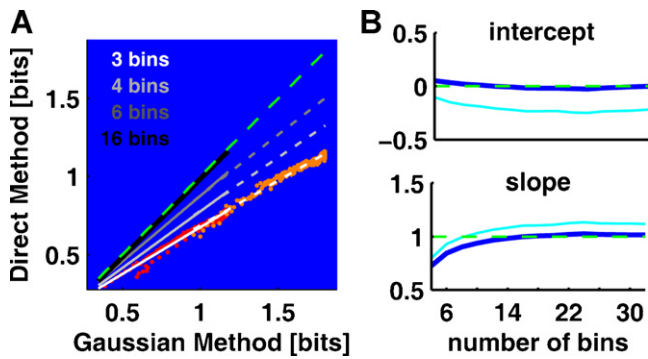
## 5.2. Validation of the methods used to estimate information carried by the power in the LFP bands

A problem limiting the use of the Direct Method in our study is that it can only be applied to compute the information carried jointly by one or maximum two LFP bands. The small number (30–40) of trials available does not allow an accurate sampling of the probability distributions required for calculating information with the Direct Method when several bands are considered. This is because, as reviewed in (Panzeri et al., 2007), the bias of the direct method grows exponentially with the number of dimensions of the response space (which in this case corresponds to the number  $L$  of bands considered in the information calculation). This bias problem is greatly alleviated by the Gaussian Method, because it requires extracting only a small number of parameters from the data. These parameters can be robustly computed using few trials even when several bands are considered at the same time. Before applying the Gaussian Method, however, we need to ensure that method is accurate on the dataset under analysis. To this aim, we compared information values obtained with the Gaussian and with the Direct Method when only one or two LFP bands were considered. If the two methods return comparable information results we can conclude that the Gaussian Method can be used safely for computing the information carried by the power of several LFP bands. The Gaussian Method was previously validated on single frequency bins (Magri et al., 2009), but not yet on power of extended bands.

We began this check by computing the stimulus-information carried by the LFP power within a frequency band  $B = [f-250 \text{ Hz}]$  (for 61 values of  $f$  uniformly distributed between 0 and 250 Hz) using both the Direct and the Gaussian Method. For the computation with the Direct Method we varied  $M$ , the number of bins used for discretizing the neural response, between 3 and 32. The latter is an upper limit to the number of bins for which we could still compute largely unbiased information estimates with the Direct Method given the available number of trials, see (Panzeri et al., 2007). We found (Fig. 2A) that the relation between the information values obtained with the two methods was approximately linear. When few ( $M=3, 4, 6$ ) bins were used with the Direct computation the linear regressor lay below the quadrant diagonal, meaning that the Direct Method used with small number of bins provides consistently lower information values than the ones estimated by the Gaussian Method. This likely happens because the use of coarse response discretization leads to an information loss. However, the slope,  $b$ , of the linear regressor increased with increasing number of bins used for the Direct computation, suggesting that the use of finer bins reduced the information loss of the Direct Method. For  $M > 13$ , the intercept  $a$  and the slope  $b$  (the parameters of the linear regressor) reached a plateau at  $a=0$  and  $b=1$  (Fig. 2B), demonstrating that for large number of bins  $M$  the Gaussian and the Direct Methods provide essentially identical results.

Fig. 2B shows what happens when the Gaussian analysis is performed without applying the cube root transformation prior to the computation of information: independently of the number of bins used for the Direct Method, the Gaussian and the Direct Method failed to converge to a common estimation value. This proves that the cube root transformation is a necessary step for reliably computing the information conveyed by LFP power using the Gaussian Method.

Finally, we computed the values of the stimulus information  $I(S; B_1 B_2)$  carried jointly by a partition of the LFP response into two-bands  $B_1 = [0 \text{ Hz}-f]$  and  $B_2 = [f-250 \text{ Hz}]$  for all possible values of the boundary parameter  $f$  between 0 and 250 Hz. Again, we



**Fig. 2.** Validation of the Gaussian Method to evaluate information. Data were taken from a representative case (channel 7 of session D04nm1). Unless otherwise stated, Gaussian information estimates were computed using the Gaussian Method corrected with the Gaussian bias correction procedure and using the cube root of the power values (see text); the Direct information values were computed after discretizing the power response into  $M$  equi-populated levels. (A) Comparison of the information values obtained using the Gaussian Method and the Direct Method for the information about a naturalistic movie stimulus carried by LFP power. The red scatter plot shows Gaussian and Direct estimates of the stimulus-information carried by the LFP power in a frequency band of width  $[f-250\text{ Hz}]$  for different values of  $f$  between 0 and 250 Hz. The orange scatter plot illustrates the stimulus-information carried jointly by the LFP power in two frequency bands of width  $[0\text{ Hz}-f]$  and  $[f-250\text{ Hz}]$  for the same values of the parameter  $f$  considered for the single-band case. For both scatter plots,  $M = 3$  bins were used for discretizing the power response for the Direct computation. The white solid line and the white dashed line illustrate the linear fit to the points in the single-band information and in the two-bands information scatter plots, respectively. The light grey, the dark grey and the black lines indicate the linear fit when  $M = 4, 6$  and 16 levels are used for the discretization, respectively (the corresponding scatter plots are not shown for these cases). The same convention is used as for the white line, with the solid and the dashed lines indicating the fit for the single-band and for the two-bands information values, respectively. Note that, due to the limited number of trials available, for  $M = 16$  bins it was not possible to compute the Direct information for the two-bands case, therefore only the solid line is shown. The green dashed line indicates the quadrant diagonal. (B) Asymptotic behavior of the slope parameter,  $b$ , and of the intercept parameter,  $a$ , of the linear fit between information values obtained with the Gaussian and the Direct Methods as a function of the number of bins,  $M$ , used for discretizing the power response for the Direct computation (blue line).  $M$  was varied between 3 and 32. The cyan lines indicate the values of  $a$  and  $b$  when no cube root transformation is applied to the LFP power for the computation of information with the Gaussian Method.

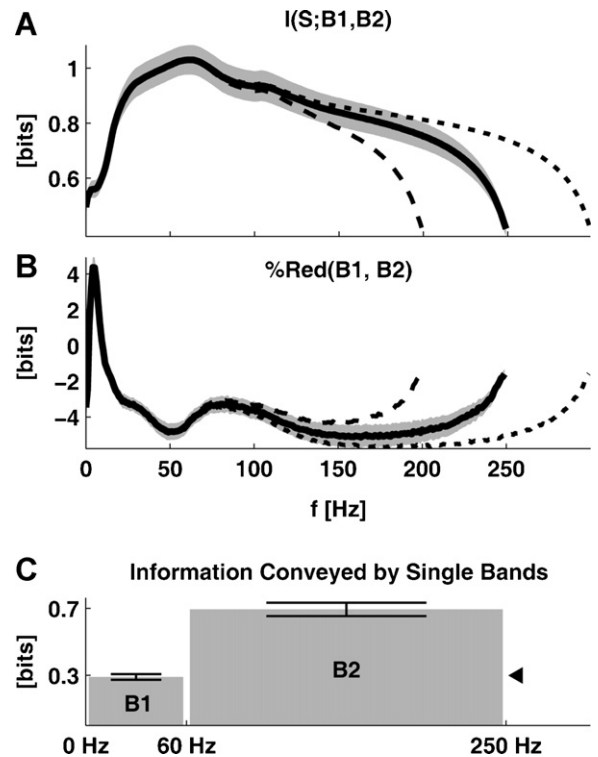
found (Fig. 2A) a linear relationship between the information values obtained using the Direct and the Gaussian Method, the slopes of the linear fitting curves being nearly identical to those obtained when only one band was considered. Note, however, that in this case we could not compute information for  $M > 6$  using the Direct Method because the number of trials available does not allow to properly sample the response space for large values of  $M$ .

Taken together these results demonstrate that the accuracy of the Gaussian approximation to compute information from LFP powers. When computing the information about a movie stimulus carried by the power in one or more LFP bands, differences between Gaussian and Direct information values arise mainly from the discretization used for the Direct procedure and that the two methods provide comparable results when enough bins are used for the discretization.

In what follows, unless otherwise stated, we will always use the Gaussian Method with the Gaussian bias correction (see Section 2) for computing information values.

### 5.3. Partitioning of LFP responses into two bands

We began by investigating what is the best way to partition the LFP spectrum into two frequency bands. Although the LFP is consistently defined in the literature as the low frequency range of the spectrum of the extracellularly recorded neural signal, there



**Fig. 3.** Partitioning the LFP into two bands. Results are averaged over a set of 57 V1 recording sites from anaesthetized macaques. Information estimates were computed using the Gaussian Method corrected with the Gaussian bias correction procedure and using the cube root of the power values. (A) The information about a naturalistic movie stimulus carried jointly by the power in the two LFP bands  $B1 = [0\text{ Hz}-f]$  and  $B2 = [f-250\text{ Hz}]$  as a function of the boundary parameter  $f$  varied between 0 and 250 Hz. The grey area indicates the SEM. The dashed and dotted lines illustrate the information obtained when varying the upper edge of  $B2$  from 250 to 200 and to 300 Hz, respectively. (B) The percentage redundancy between the information carried by the power in the two LFP bands  $B1$  and  $B2$  defined as in panel A. The grey area indicates the SEM. The dashed and dotted lines illustrate the redundancy obtained when varying the upper edge of  $B2$  from 250 to 200 and to 300 Hz, respectively. (C) The information about a naturalistic movie stimulus carried individually by LFP power in the band  $B1 = [0-60\text{ Hz}]$  and by the LFP power in the band  $B2 = [60-250\text{ Hz}]$ . The error bars indicate the SEM. The triangle marker indicates the information conveyed by the unpartitioned  $[0-250\text{ Hz}]$  LFP range.

is no agreement about the precise frequency range that should be considered as LFP activity. Some authors include frequencies only up to 100 Hz, while others include up to several hundred Hz. Here we started by assuming the LFP spectrum to range from 0 to 250 Hz. We later checked for the effect of considering other definitions of the LFP frequency range.

We computed the stimulus-information,  $I(S;B1\ B2)$ , carried by the LFP power when partitioned into the two bands  $B1 = [0\text{ Hz}-f]$  and  $B2 = [f-250\text{ Hz}]$ , as a function of the boundary parameter  $f$  (with  $0 < f < 250\text{ Hz}$ ). We found (Fig. 3A) that the amount of information carried jointly by the two bands was, on average across the population, maximal when splitting the bands at  $f = 60\text{ Hz}$ . The information, was however very high for a broad frequency region, extending approximately between 50 and 70 Hz. The distribution across recording sites of the value of the frequency boundary  $f$  at which information was maximal in each recording site had a median of 62 Hz and an interquartile range of 53–79 Hz. The information about the movie stimulus obtained after splitting the  $[0-250\text{ Hz}]$  LFP range into two bands at the  $f = 60\text{ Hz}$  optimal boundaries was  $1.03 \pm 0.05$  bits (here and hereafter, and unless otherwise stated, information values are reported as mean  $\pm$  SEM over population). In comparison, the information in the un-partitioned  $[0-250\text{ Hz}]$  LFP power was  $0.30 \pm 0.02$  bits. Thus, splitting the LFP response at the optimal boundary  $f = 60\text{ Hz}$  allowed the

extraction of more than three times the information carried by the un-partitioned LFP power. Taken together, these results clearly indicate that the most informative way to partition the LFP response into two bands is to separate low and high LFP frequencies with a boundary placed at approximately 55–65 Hz.

We noted that the information (Fig. 3A) did not fall monotonically after the maximum at 60 Hz, but rather a second local peak was present at approximately 105 Hz. Similarly, the redundancy (Fig. 3B) reached a second local minimum in the same frequency region.

To understand better the reasons why it was optimal to partition the LFP range by splitting it at  $f=60$  Hz, we investigated how the redundancy between the information carried by the bands depended upon the boundary value  $f$ . We found (Fig. 3B) that the redundancy between the two bands was in general small and was minimal in the range 50–60 Hz. The minimum was reached at 53 Hz and was  $-0.05 \pm 0.004$  bits. This suggests that splitting the LFP at the optimal boundary range (50–70 Hz) partitions the LFP into two regions with largely independent (even weakly synergistic) information content and stimulus tuning properties. These observations about LFP bands are consistent with the results obtained regarding the information content of the power of single LFP frequencies, see Fig. 1C and (Belitski et al., 2008).

Fig. 3C reports that the information carried by the high-frequency [60–250 Hz] band was  $0.69 \pm 0.04$  bits. In contrast, the low frequency [0–60 Hz] band carried much less information ( $0.29 \pm 0.02$  bits). This result, together with the above information redundancy findings, indicate that the large amount of information gained by partitioning the LFP range into the high and low frequency components arises mostly because this partitioning allows accessing the information in the low-power but highly informative high-frequency range. In the unpartitioned signal, instead, the relatively low power of the high-frequency range is masked by the high power of the less informative but very low-frequency range.

It is interesting to compare the amount of information carried by the power of an optimally defined LFP band with the amount of information that can be extracted by the power at a single frequency. We first compared the information in the entire high frequency band with the information in the individual frequencies in the high-frequency range. Consistent with previous results (Siegel and König, 2003), we observed that integrating LFP power over the range [60–250 Hz] greatly increased the total amount of information ( $0.69 \pm 0.04$  bits) compared to the highest information carried by a single LFP frequency in the gamma range ( $0.31 \pm 0.02$  bits). This can be understood by the fact that frequencies above 60 Hz carry largely redundant information with largely similar stimulus selectivity, high “signal correlation”, and almost independent trial-to-trial variability, i.e. relatively low “noise correlation” (Belitski et al., 2008; Magri et al., 2009). Integrating neural responses with similar stimulus selectivity and weakly correlated noise has been demonstrated to greatly increase information (Abbott and Dayan, 1999; Oram et al., 1998; Zohary et al., 1994). We then compared the information in the entire low frequency band [0–60 Hz] with the information in the individual frequencies in the low-frequency range. We found (Fig. 3C) that integrating the LFP power over the low [0–60 Hz] frequency range provided information ( $0.292 \pm 0.015$  bits) comparable to the maximum information obtained for single LFP frequencies in this range ( $0.289 \pm 0.017$  bits; Fig. 1B). The fact that essentially no information is gained by integrating the power in the low frequencies is likely due to the fact that low frequencies have a small redundancy (Fig. 1C) due to a very small similarity in their tuning to the movie (Belitski et al., 2008; Magri et al., 2009) and summing differently tuned responses does not lead to information gain.

As we mentioned earlier, the upper frequency cutoff defining the LFP range is an arbitrary parameter not consistently used in

the literature. We investigated whether the optimal band splitting depended on the upper frequency cutoff defining the LFP by varying this parameter in the range 200–300 Hz. We found (Fig. 3) that both the value of the optimal frequency cutoff and the amount of information obtained from the optimally split LFP were highly stable when varying this parameter.

#### 5.4. Partitioning of LFP responses into three bands

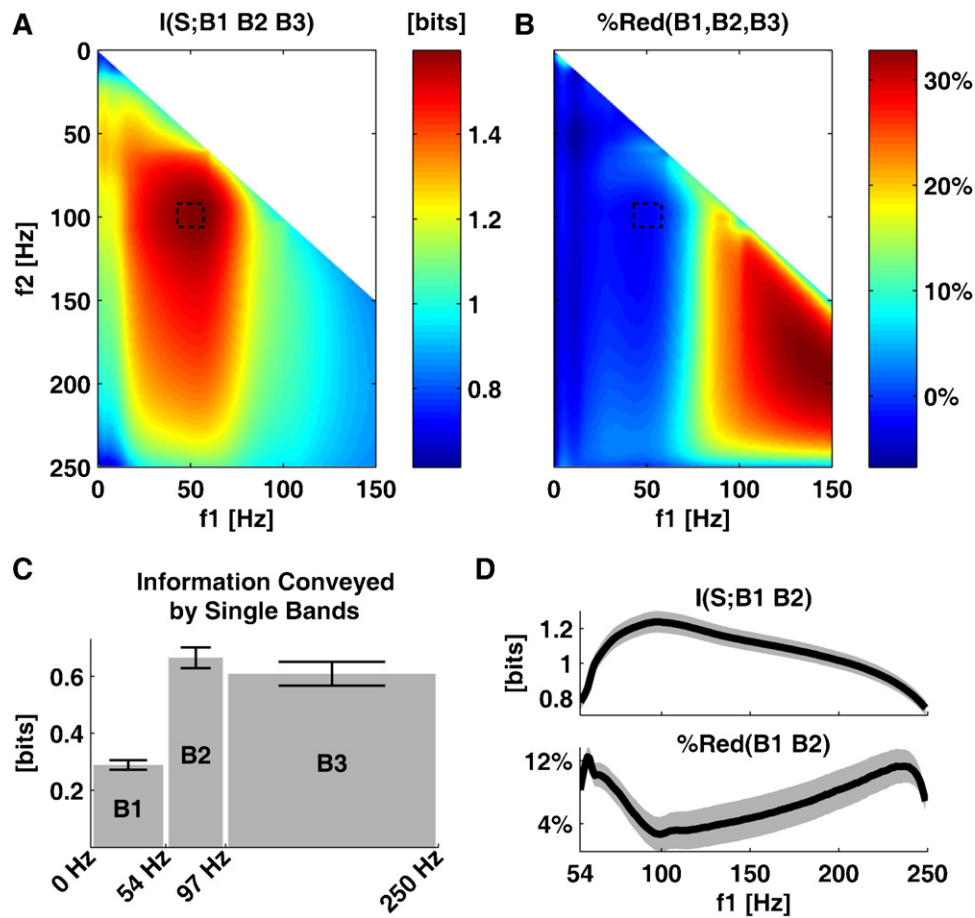
Next, we considered the problem of finding boundaries  $f_1$ ,  $f_2$  that optimally partition the LFPs into three bands,  $B_1=[0-f_1]$ ,  $B_2=[f_1-f_2]$  and  $B_3=[f_2-250\text{ Hz}]$ . Fig. 4A reports the information carried jointly by the three frequency ranges, computed as a function of the two boundaries parameters  $f_1$  and  $f_2$  (with  $0\text{ Hz} < f_1 < f_2 < 250\text{ Hz}$ ) and averaged over the population. The optimal partition was found at  $f_1 = 54$  Hz and  $f_2 = 97$  Hz. As for the case of two bands, the position of the information peak was robust to the choice of the upper boundary used for defining the LFP range (not shown). Again, we found that the redundancy (Fig. 4B) was negligible for  $f_1$  and  $f_2$  in the optimally informative region. Across the population the distribution of the position of the information peaks was centered around 52 Hz (median value; interquartile range: 43–57 Hz) for  $f_2$  and around 98 Hz (median value; interquartile range: 92–106 Hz).

The first optimal boundary (at approximately 55 Hz) to split the LFP range into three bands is essentially coincident with the optimal boundary for splitting the LFP into two bands. The second optimal boundary (at 97 Hz) for the three band partition corresponds to the local peak (Fig. 3B) observed when partitioning the LFP response into two bands. This second optimal boundary separates the high-frequency range [54–250 Hz] into two ranges [54–97 Hz] and [97–250 Hz]. The first range corresponds to the frequency range laying inside what is traditionally called the gamma range, and is consistent with the optimal gamma band definition [64–103 Hz] obtained for the cat visual cortex using an alternative procedure for functionally defining this band (Siegel and König, 2003). This also suggests, consistent with (Siegel and König, 2003), that widely used definition of the gamma band as the region around 40 Hz (the frequency at which LFP power peaks in this range, and neural synchronization is usually maximal) may not be the optimal choice when considering stimulus coding. The second range [97–250 Hz] corresponds to the range which has been termed the “high-gamma” band, and which was found (Ray and Maunsell, 2011) to correlate with spiking activity.

The fact that the position of the first optimal boundary is essentially unchanged when partitioning the LFP response into three bands indicates that this boundary is stable and that it defines bands which are likely to reflect true distinct neural phenomena. It also indicates that the most important step, when partitioning the LFP response according to our procedure, is to separate low and high LFP frequencies. The significance of this optimal boundary is also indicated by the extension of the peak of the information curve averaged over the population (Fig. 4A), which is narrower along the  $f_1$  direction compared to  $f_2$ , indicating a sharper tuning of the information curve to the first boundary. Additionally, the broadness of the  $f_2$  peak is consistent with the observation that the boundary found for delimiting the high gamma range should not be thought of as a strict boundary and that the lower and upper frequency limits of the high-gamma band appear to be somewhat variable depending on the task, cortical site, and subject (Crone et al., 2006; Ray and Maunsell, 2010).

We observed that the first optimal boundary was stable in that it appeared both when partitioning the LFP range into two or into three bands. To test for the stability of the second optimal boundary we partitioned the range [54–250 Hz] into two bands. The rationale for considering this range separately from the low frequencies





**Fig. 4.** Partitioning of LFP response into three bands. Results are averaged over a set of 57 V1 recording sites from anaesthetized macaques. Information estimates were computed using the Gaussian Method corrected with the Gaussian bias correction procedure and using the cube root of the power values. (A) The information about a naturalistic movie stimulus carried jointly by the power in the three LFP bands  $B1 = [0\text{ Hz}-f_1]$ ,  $B2 = [f_1-f_2]$  and  $B3 = [f_2-250\text{ Hz}]$  as a function of the boundary parameters  $f_1$  and  $f_2$  varied between 0 and 250 Hz with  $f_1 < f_2$ . The dashed rectangle indicates the interquartile range for the position of the information peaks across the population. (B) The percentage redundancy between the LFP power in the three LFP ranges  $B1$ ,  $B2$  and  $B3$ . The dashed rectangle indicate the same region boxed in panel A. (C) The information about a naturalistic movie stimulus carried individually by the LFP power in the band  $B1 = [0-54\text{ Hz}]$ , in the band  $B2 = [54-97\text{ Hz}]$  and in the band  $B3 = [97-250\text{ Hz}]$ . The error bars indicate the SEM. (D) Top: the information about a naturalistic movie stimulus carried jointly by the power in the two LFP bands  $B1 = [54\text{ Hz}-f]$  and  $B2 = [f-250\text{ Hz}]$  as a function of the parameter  $f$  varied between 54 and 250 Hz. Bottom: the redundancy between the information carried by the LFP power in  $B1$  and  $B2$ . The grey area indicates the SEM.

is that we have observed (Figs. 1D and 3B) that LFP power below 54 Hz conveys information which is independent from that of the higher frequencies. We found (Fig. 4D) that the information was maximal for  $f$  again in the range [90–110 Hz]. We also found that the redundancy between the information conveyed by the gamma and the high-gamma bands was positive but very small (reaching a minimum of  $3\% \pm 2\%$  at 99 Hz) consistent with the low values of the redundancy (Fig. 1C) observed between single gamma and high-gamma frequencies.

Among the bands in the optimal three-band partition, the optimal gamma band [54–100 Hz] was the one carrying the largest single-band stimulus information ( $0.66 \pm 0.03$  bits; Fig. 4C). The second most informative range was [97–250 Hz] ( $0.61 \pm 0.04$  bits) followed by the low (<54 Hz) LFP frequencies ( $0.29 \pm 0.02$  bits).

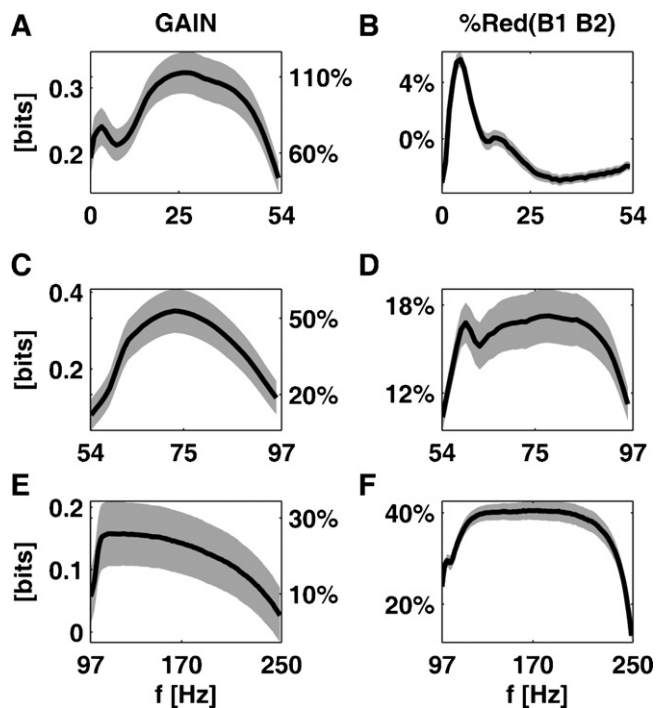
### 5.5. Determining additional optimal boundaries

An issue with the method proposed for determining the optimal boundaries is that the parameter space in which to search for the optimal boundaries increases exponentially with the number of boundaries considered. For example, a spectral window of length  $T=960$  ms has 241 Fourier coefficients between 0 and 250 Hz with our sampling rate. Identifying the optimal partition of the LFP into

three bands thus requires computing information for several hundred thousand possible combinations of the boundary parameters for each recorded channel. This is difficult to perform systematically with current computational means. Downsampling the frequency space is also not a solution to this problem, as a high frequency resolution becomes more and more crucial when considering many bands, since the optimal boundaries become closer and thus harder to discriminate.

In order to assess the utility of a further partition of the LFP spectrum into four bands, we followed a slightly different approach: we fixed the two previously obtained principal band boundaries and we computed the optimal way to further split each of the three principal bands, rather than performing a brute-force search across the entire four-dimensional frequency space. This choice was justified by our findings that the optimal boundaries at 54 and 97 Hz (which we call the principal boundaries) are stable with respect both to changes in the parameters of the analysis and to a refinement of the partition, and that the three principal bands ([0–54 Hz], [54–97 Hz] and [97–250 Hz]) convey largely independent information and, consequently, their further refining can be studied independently of the others. In what follows, the information gained by further splitting a band was quantified as the difference between the information carried by the partition of the





**Fig. 5.** Partitioning the LFP into four bands. Results are averaged over a set of 57 V1 recording sites from anaesthetized macaques. Information estimates were computed using the Gaussian Method corrected with the Gaussian bias correction procedure and using the cube root of the power values. The solid lines indicate the average over all available recordings and the grey areas indicates the SEM. (A) The information gain about a naturalistic movie stimulus obtained when partitioning the LFP range [0–54 Hz] into two bands  $B1 = [0 \text{ Hz} - f]$  and  $B2 = [f - 54 \text{ Hz}]$  as a function of the boundary parameter  $f$ . The right y-axis expresses the gain as the percentage of the information conveyed by the un-partitioned band. (B) The percentage redundancy between the information carried by the power in the LFP bands  $B1 = [0 \text{ Hz} - f]$  and  $B2 = [f - 54 \text{ Hz}]$  as a function of the boundary parameter  $f$ . (C) Same plot as in panel A but for a partition of the band [54–97 Hz]. (D) Same plot as in panel B but for a partition of the band [54–97 Hz]. (E) Same plot as in panel A but for a partition of the band [97–250 Hz]. (F) Same plot as in panel B but for a partition of the band [97–250 Hz].

selected band and the information carried by the un-partitioned band. This information gain was expressed either as an absolute difference (in bits) between these two terms, or was normalized as percentage gain with respect to the information carried by the un-partitioned band.

We first considered how to split the [0–54 Hz] low-frequency range. We found (Fig. 5A) that the information gain was small when the boundary between the two sub-bands was set below 15 Hz but that it increased rapidly reaching a global maximum of  $0.32 \pm 0.03$  bits at 27 Hz, corresponding to an information gain of  $111\% \pm 11\%$  with respect to the information carried by the un-partitioned band. This optimal split at 27 Hz corresponds roughly to partitioning the low frequency range into a stimulus-informative [0–27 Hz] and to the stimulus-unrelated frequencies [27–50 Hz] band (see the single frequency information analysis reported in Fig. 1B and in (Belitski et al., 2008; Magri et al., 2009)). The overall information gain of the further partition of the [0–54 Hz] band can be conceptualized as helping isolating the “noisy” frequencies from the stimulus-evoked ones.

We then considered how best to sub-partition the gamma [54–97 Hz] band into two sub-bands. Fig. 5C shows that the optimal split was obtained using a boundary at 73 Hz. The information gain of this optimal split over the information in the whole band was  $0.35 \pm 0.06$  bits (corresponding to  $53\% \pm 8\%$  more information). Thus, a further split of the gamma range was worthwhile because it increased the information, although with a less

important proportional gain with respect to the one obtained by further partitioning the low frequency range. The fact that the partition of the gamma range gave a lower proportional increase can be understood by observing that the optimally split sub-ranges within the gamma frequency domain were still almost 18% redundant (Fig. 5D), suggesting that the division at 73 Hz separates similar frequency regions with a slightly different tuning rather than identifying regions which qualitatively different stimulus coding properties.

The smallest information gain (both in absolute and in proportional terms) was obtained when partitioning the high-gamma [97–250 Hz] range, Fig. 5E. The optimal split of the high-gamma range was achieved with a boundary at 111 Hz, but it led to only a small ( $0.16 \pm 0.05$  bits, corresponding to  $26\% \pm 8\%$ ) information gain, indicating that little additional information can be gained by additionally partitioning the high frequencies. This finding is further supported by the very high values of redundancy (Fig. 5F) obtained for any choice of the boundary.

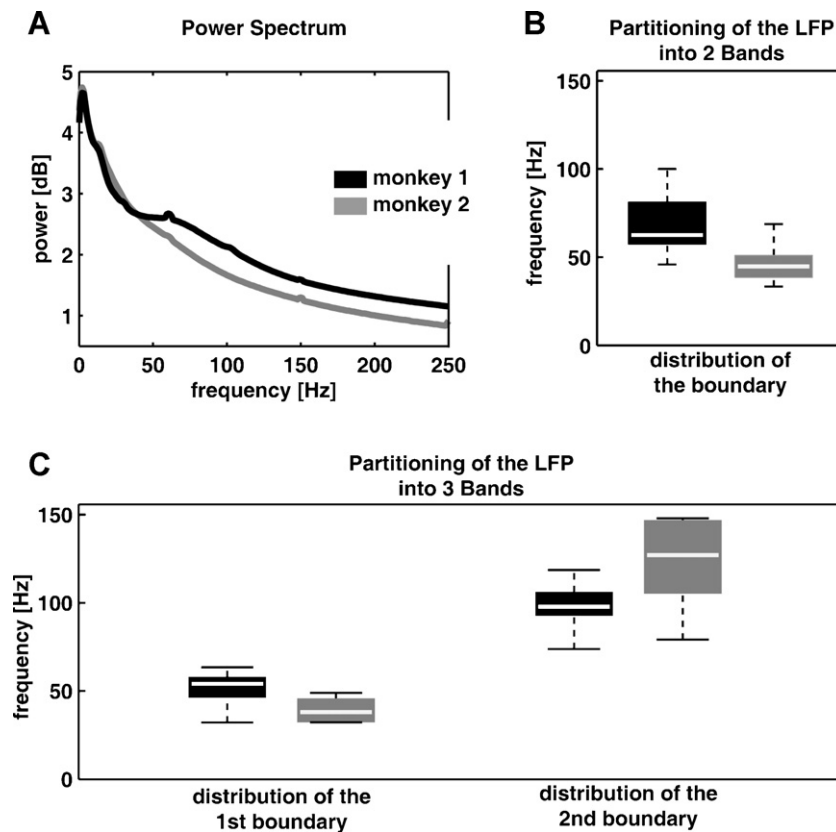
These results suggest that the most convenient way of adding a third optimal boundary is to partition the low frequency range into a [0–27 Hz] band and a [27–54 Hz] range since considering these two bands separately provides a proportionally large gain in information and it separates regions with a qualitatively different relation to the visual stimulus.

#### 5.6. Distribution of optimal band boundaries across sites and animals

In the above sections we partitioned the LFPs into optimal bands using information values averaged over all available sites and monkeys. Previous work has shown that the LFP spectrum can change depending on factors such as the details of the stimulus used (Ray and Maunsell, 2010), the task (Fries et al., 2001) or recording depth (Maier et al., 2011). In particular, it was shown that gamma activity can vary significantly in power and peak frequency across subjects and recording site (Lima et al., 2010; Muthukumaraswamy et al., 2010). Here we investigate the results obtained by applying our partitioning technique independently for each recording channel and for each monkey.

Results of the single-site two-bands optimal partition are reported in Fig. 6B. We found that the method was able to find optimal boundaries also at the single channel level. The position of the optimal boundary for the two-bands partition had a median of 62 Hz (range 45–100 Hz) for the first monkey and of 45 Hz (range 33–69 Hz) for the second monkey. The distribution of the boundaries in the first monkey was significantly higher than that of the second monkey ( $p < 0.05$ ; Wilcoxon rank sum test). Although the limited number of subjects used in the study does not allow to perform an exhaustive analysis of the dependency of the band boundaries on the spectral properties of each animal, it is tempting to suggest that the difference in the distribution of the first optimal boundary between the two monkeys may be attributed to the much larger gamma power observed in spectrum of the first monkey (Fig. 6A).

Furthermore, when partitioning the LFP range into three bands independently for each site of the first monkey, we found that the first optimal boundary was distributed between 32 and 63 Hz with a median of 54 Hz and that the second optimal boundary was between 74 and 118 Hz with a median of 98 Hz. For the second monkey the first boundary was between 32 and 49 Hz (median of 38 Hz), comparable to the range obtained when partitioning the LFPs recorded from this monkey into two bands (Fig. 6B), and the second boundary was between 79 and 147 Hz (median of 127 Hz). It is important to note that the partitioning algorithm determined that the first boundary was preserved when refining the partition from two to three bands even at the single recording site level, and



**Fig. 6.** Distribution of the optimal boundaries across sites and animals. Results are shown for each of the two monkeys used in this study. Quantities plotted in grey refer to monkey 1, quantities plotted in black to monkey 2. (A) Average power spectrum of the LFPs recorded during the presentation of a naturalistic movie. (B) Distribution of the optimal boundaries obtained for each recording site when partitioning the LFPs into two bands. (C) Distribution of the first and second optimal boundaries obtained for each recording site when partitioning the LFPs into three bands.

not only at the population level. This suggests that our algorithm is robust enough to analyze and partition single cases, and not only grand population averages.

## 6. Discussion

We developed a new procedure for partitioning the LFP response into a specified number of bands, which opens the way for a more objective, semi-automatic analysis of the role of cortical oscillatory activity in sensory function. The method selects the band boundaries that maximize the information about a set of external correlates carried jointly by all bands in the partition. We used this procedure to determine a maximally informative partitioning of primary visual cortical LFPs into two to four bands, and identified band boundaries which are compatible with our knowledge and intuition about the functional LFP spectrum in visual cortex. It is important to note that our method, though exemplified for LFPs, can be readily applied to a variety of broad-band neural signals, from intracellular recordings of membrane potential fluctuations to population spiking activity or to non-invasive neuroimaging signals such as EEG, magneto encephalography (MEG) and functional magnetic resonance imaging (fMRI). Our method can therefore become a useful tool for the endeavor of extracting information from these signals (Alenda et al., 2010; Fuhrmann Alpert et al., 2007; Ostwald et al., 2010; Panzeri et al., 2008; Quiroga and Panzeri, 2009; Schyns et al., 2011).

The implications of both the methodological advances and of the neurophysiological results are discussed in the following.

### 6.1. Potential applications of methods for partitioning neural responses into maximally informative bands

As illustrated by our application to visual cortical LFPs, our method is potentially relevant for sensory physiology. The spectrum of LFP oscillations and fluctuations in response to a stimulus is often broad and lacks a clear separation between spectral regions tuned to different stimulus features. The relationship between LFPs and stimuli is particularly difficult to untangle when considering complex stimuli made of many features varying at different time scales, such as for example naturalistic movies or natural sounds (Kayser and Konig, 2004; Kayser et al., 2009). Understanding how to partition the LFP into a specific number of bands that are collectively maximally informative about the stimulus sheds light on differences in stimulus tuning of distinct LFP frequency regions, and helps individuating frequency regions that must be considered separately because they carry complementary information about the external correlates. Being able to individuate objectively and almost automatically the optimal definitions of such bands will also facilitate homogeneity of analysis across future studies and will make it easier and more meaningful to compare different studies or preparations.

Although here we focused on maximizing the information about an external correlate, our method could be readily used to determine the optimal partition of broadband neural signals into bands which are maximally informative about internal (rather than external) correlates. For example, our procedure could be used to determine a set of LFP regressors providing an optimal prediction of the BOLD fMRI.

Finally, determining the partitions of the LFP spectra which are maximally informative about sensory stimuli is crucial for the development of brain machine interfaces (Andersen et al., 2004; Donoghue et al., 1998). The LFP is a particularly suited neural signal for BMIs because they have the potential to prolong the lifetime of electrode implants and to improve decoding performance when combined with spikes (Andersen et al., 2004; Belitski et al., 2008). It is widely recognized that the main bottleneck to the performance of BMIs is the relatively small amount of information that we are currently able to extract from neural activity (Donoghue et al., 1998). Our method provides a tool for maximizing the information about external correlates (such as, for example, movement direction) thus potentially enlarging bottlenecks in performance.

## 6.2. Significance of the methodological advances

One main advantage of using information as a measure of the relationship between the power of the band partition and the external correlate is that information captures all types of relationships (linear and non-linear and of any order). Additionally, by considering the information conveyed simultaneously by the power in all bands, our method does not only take into account the relationship between each individual band and the external correlate, but also the relationship between the different bands in the partition. These factors cannot be taken into account by single-frequency or single-band analyses.

We also considered the practical problems involved in the application of our method to empirical neurophysiological datasets. To alleviate the bias problem arising from the limited size of the datasets that it is normally possible to collect, we proposed a Gaussian Method and Gaussian bias correction for computing information. We validated this method on visual cortical LFPs and demonstrated that, when one or two bands are considered, differences between the information estimates obtained using the Gaussian Method and those obtained with the Direct Method stem from the discretization used for the Direct Method. This result indicates that the Gaussian Method-based procedure that we propose can be successfully applied to compute the information conveyed by the power in the partition of the LFP response. A second practical problem is that of implementing the bias corrections and estimation procedure required for the proper estimation of information from the data. We developed a fast toolbox (Magri et al., 2009), which included both Direct Method and Gaussian computation and can thus be used to implement and to validate our procedure, and made it publicly available ([www.ibtb.org](http://www.ibtb.org)). A third practical problem is that of computational time. Besides implementing a fast toolbox, we addressed this problem by proposing a procedure for incrementally refining the partition based on further splitting optimal bands. This can save a great deal of computing time when needing to partition the LFP into a large number of bands. This approach is justified by two key findings: the principal band boundaries are stable to refinement of the partition, and optimal partitions in most cases correspond to splits into largely independent bands.

The partitioning method proposed in this work also provides a straightforward way to quantitatively evaluate the relative gain of adding extra bands to the partition. This can be done by comparing the information obtained for a given number of bands with that gained by introducing one extra band. By evaluating the statistical significance of the information increase when adding extra bands, this method could be extended to determine the smallest number of bands that need to be used to describe all stimulus information available in the LFP. However, given the richness and complexity of LFPs recorded in response to complex naturalistic stimuli, we feel that it may not be possible to describe all the information

content of LFPs using only a very small number of bands such as that considered in the above analysis.

The above considerations, as well as the large number of different bands and partitions proposed in the neurophysiological and in the neuroimaging literature, raise the question of how to evaluate the information content of a number of bands much larger than the handful considered here. Both the bias of the information calculation and the difficulty of the evaluation of its significance grow rapidly with the cardinality of the space in which the stimulus-response probabilities are defined, in this study with the number of bands included in the analysis (Ince et al., *in press*). Evaluation of the exact information content of a large number of bands (of the order of ten or more) requires either collecting larger datasets or developing accurate stimulus response models described by even less parameters than the ones considered here. We note, however, that the information bias of the method considered here (Eqs. (6) and (7)) depends upon the number of bands but is largely independent of the exact band boundaries. This means that the partitioning method considered here may be still useful to determine the boundaries for optimal partition into larger number of bands, although in such case it may not allow to evaluate the precise information content of the optimal partition.

## 6.3. The optimal band partitions describing visual primary cortical encoding of naturalistic information

We used our method to compute the optimally informative band partitions of LFPs recorded in V1 of anesthetized during presentation of naturalistic movies.

The first, and most important, optimal boundary that appeared in our partition was at approximately 60 Hz, and separated low [0–60 Hz] and a high [60–250 Hz] LFP frequencies. The finding that this is the chief boundary is consistent with previous analyses (Belitski et al., 2008; Magri et al., 2009) reporting that the frequencies above 50–60 Hz are almost completely uncorrelated in both signal and noise to frequencies below this cutoff. The large amount of information gained over the unpartitioned LFP when splitting the LFPs at this frequency lends support to the hypothesis that the low- and the high-frequency ranges convey independent information about the stimulus because they originate from different neural pathways. In particular, theoretical and experimental studies suggested that low LFP frequencies reflect entrainment of local neural activity to the low frequency components in the input (Chandrasekaran et al., 2010; Lakatos et al., 2008; Mazzone et al., 2008) whereas high LFP frequencies in the gamma range and above convey information about the strength of the input because the latter modulates the locally generated rhythms (Brunel and Wang, 2003; Henrie and Shapley, 2005; Mazzone et al., 2008; Ray and Maunsell, 2010).

We found a second optimal boundary at approximately 100 Hz separating the high-frequency LFP band into two ranges corresponding to the gamma [60–100 Hz] and to the high-gamma [100–250 Hz] band, leading to a large increase (approx. 30%) in the amount of information that we can extract about the movie stimulus compared to the case in which only two bands are considered. Gamma oscillations are thought to originate to a large extent from the recurrent interactions among inhibitory neurons and among excitatory and inhibitory neurons (Bartos et al., 2007; Brunel and Wang, 2003). High-gamma power has been proposed to originate partly from broadband LFP activity and partly from brief bursts of power associated with spikes generated near the electrode (the so called “spike bleed-through”). In some conditions, for example when varying contrast, gamma, high-gamma and spiking activity are largely correlated and carry similar information about the stimulus parameters. However, in other conditions, for example when varying the stimulus size, gamma activity is



partially decoupled and has a different stimulus selectivity than that of high gamma and spiking activity, which remain highly correlated in most conditions (Gieselmann and Thiele, 2008; Ray and Maunsell, 2011). Our finding that the split between gamma and high gamma frequencies at approximately 100 Hz is the second most important split to maximize information extraction from the data, supports the hypothesis that gamma rhythm and high-gamma activity reflect partly different neural phenomena, and highlights the importance of distinguishing and studying separately these two frequency ranges when investigating stimulus selectivity of cortical activity.

We also found that adding a third boundary at approximately 25 Hz led to a considerable gain in the stimulus-information that could be extracted from the low LFP frequencies. This split divides the low frequency LFP range into two qualitatively different parts: a stimulus-modulated part and a stimulus independent part. A high information gain was obtained also when further splitting the gamma range [50–100 Hz] at approximately 70 Hz. In this case, however, the redundancy between the two sub-bands was still considerable. This suggests that the information gained by finer partition of the gamma range results from isolating part of the spectrum originating from the same neural phenomenon (most likely recurrent inhibitory–inhibitory and excitatory–inhibitory interactions) with slightly different tuning to the stimulus. This is compatible with theoretical models that predict that the peak frequency, the width and the stimulus modulation of the oscillation spectrum all change with the strength of the input to the network (Brunel and Wang, 2003; Mazzoni et al., 2008). According to this model, further partitioning the gamma range would allow discriminating these finer changes in LFPs in response to the stimuli, hence the information increase despite the presence of redundancy.

## Acknowledgements

This work was supported by the BMI project of IIT, by the Max Planck Society, by the Compagnia di San Paolo, and was part of the research program of the Bernstein Center for Computational Neuroscience, Tübingen, funded by the German Federal Ministry of Education and Research (BMBF; FKZ: 01GQ1002).

## References

- Abbott L, Dayan P. The effect of correlated variability on the accuracy of a population code. *Neural Comput* 1999;11:91–101.
- Alenda A, Molano-Mazón M, Panzeri S, Maravall M. Sensory input drives multiple intracellular information streams in somatosensory cortex. *J Neurosci* 2010;30:10872–84.
- Andersen RA, Musallam S, Pesaran B. Selecting the signals for a brain-machine interface. *Curr Opin Neurobiol* 2004;14:720–6.
- Bartos M, Vida I, Jonas P. Synaptic mechanisms of synchronized gamma oscillations in inhibitory interneuron networks. *Nat Rev Neurosci* 2007;8:45–56.
- Belitski A, Gretton A, Magri C, Murayama Y, Montemurro MA, Logothetis NK, et al. Low-frequency local field potentials and spikes in primary visual cortex convey independent visual information. *J Neurosci* 2008;28:5696–709.
- Belitski A, Panzeri S, Magri C, Logothetis NK, Kayser C. Sensory information in local field potentials and spikes from visual and auditory cortices: time scales and frequency bands. *J Comput Neurosci* 2010;29:533–45.
- Berens P, Keliris GA, Ecker AS, Logothetis NK, Tolias AS. Feature selectivity of the gamma-band of the local field potential in primate primary visual cortex. *Front Neurosci* 2008;2:199–207.
- Brunel N, Wang XJ. What determines the frequency of fast network oscillations with irregular neural discharges? I. Synaptic dynamics and excitation-inhibition balance. *J Neurophysiol* 2003;90:415–30.
- Buzsáki G. Rhythms of the brain. USA: Oxford University Press; 2006.
- Chandrasekaran C, Tureson HK, Brown CH, Ghazanfar AA. The influence of natural scene dynamics on auditory cortical activity. *J Neurosci* 2010;30:13919–31.
- Chrobak J, Buzsáki G. Gamma oscillations in the entorhinal cortex of the freely behaving rat. *J Neurosci* 1998;18:388–98.
- Cover T, Thomas J. Elements of information theory. New York: Wiley; 2006.
- Crone NE, Sinai A, Korzeniewska A. High-frequency gamma oscillations and human brain mapping with electrocorticography. *Prog Brain Res* 2006;159:275–95.
- de Ruyter van Steveninck RR, Lewen GD, Strong SP, Koberle R, Bialek W. Reproducibility and variability in neural spike trains. *Science* 1997;275:1805–8.
- Donoghue JP, Sanes JN, Hatsopoulos NG, Gaál G. Neural discharge and local field potential oscillations in primate motor cortex during voluntary movements. *J Neurophysiol* 1998;79:159–73.
- Frien A, Eckhorn R, Bauer R, Woelbern T, Gabriel A. Fast oscillations display sharper orientation tuning than slower components of the same recordings in striate cortex of the awake monkey. *Eur J Neurosci* 2000;12:1453–65.
- Fries P, Reynolds JH, Rorie AE, Desimone R. Modulation of oscillatory neuronal synchronization by selective visual attention. *Science* 2001;291:1560–3.
- Fuhrmann Alpert G, Sun FT, Handwerker D, D'Esposito M, Knight RT. Spatio-temporal information analysis of event-related BOLD responses. *Neuroimage* 2007;34:1545–61.
- Gieselmann MA, Thiele A. Comparison of spatial integration and surround suppression characteristics in spiking activity and the local field potential in macaque V1. *Eur J Neurosci* 2008;28:447–59.
- Goodman N. The distribution of the determinant of a complex Wishart distributed matrix. *Ann Math Stat* 1963;34:178–80.
- Hatsopoulos NG, Ojakangas CL, Paniński L, Donoghue JP. Information about movement direction obtained from synchronous activity of motor cortical neurons. *P Natl Acad Sci U S A* 1998;95:15706–11.
- Henrie JA, Shapley R. LFP power spectra in V1 cortex: the graded effect of stimulus contrast. *J Neurophysiol* 2005;94:479–90.
- Ince A, Mazzoni A, Bartels A, Logothetis N, Panzeri S. A novel test to determine the significance of neural selectivity to single and multiple potentially correlated stimulus features. *J Neurosci Meth*. doi:10.1016/j.jneumeth.2011.11.013, in press.
- Kayser C, König P. Stimulus locking and feature selectivity prevail in complementary frequency ranges of V1 local field potentials. *Eur J Neurosci* 2004;19:485–9.
- Kayser C, Montemurro MA, Logothetis NK, Panzeri S. Spike-phase coding boosts and stabilizes information carried by spatial and temporal spike patterns. *Neuron* 2009;61:597–608.
- Kruse W, Eckhorn R. Inhibition of sustained gamma oscillations (35–80 Hz) by fast transient responses in cat visual cortex. *P Natl Acad Sci U S A* 1996;93:6112–7.
- Lakatos P, Karmos G, Mehta AD, Ulbert I, Schroeder CE. Entrainment of neuronal oscillations as a mechanism of attentional selection. *Science* 2008;320:110–3.
- Lima B, Singer W, Chen NH, Neuenschwander S. Synchronization dynamics in response to plaid stimuli in monkey V1. *Cereb Cortex* 2010;20:1556–73.
- Liu J, Newsome WT. Local field potential in cortical area MT: stimulus tuning and behavioral correlations. *J Neurosci* 2006;26:7779–90.
- Linias R, Ribary U. Coherent 40-Hz oscillation characterizes dream state in humans. *P Natl Acad Sci U S A* 1993;90:2078–81.
- Logothetis NK. The underpinnings of the BOLD functional magnetic resonance imaging signal. *J Neurosci* 2003;23:3963–71.
- Magri C. An information theoretic analysis of brain oscillations. Milano: Department of Computer Sciences, University of Milano; 2009.
- Magri C, Whittingstall K, Singh V, Logothetis NK, Panzeri S. A toolbox for the fast information analysis of multiple-site LFP, EEG and spike train recordings. *BMC Neurosci* 2009;10:81.
- Maier A, Aura CJ, Leopold DA. Infragranular sources of sustained local field potential responses in macaque primary visual cortex. *J Neurosci* 2011;31:1971–80.
- Mazzoni A, Panzeri S, Logothetis NK, Brunel N. Encoding of naturalistic stimuli by local field potential spectra in networks of excitatory and inhibitory neurons. *PLoS Comput Biol* 2008;4:e1000239.
- Miltner WHR, Braun C, Arnold M, Witte H, Taub E. Coherence of gamma-band EEG activity as a basis for associative learning. *Nature* 1999;397:434–5.
- Misra N, Singh H, Demchuk E. Estimation of the entropy of a multivariate normal distribution. *J Multivariate Anal* 2005;92:324–42.
- Montemurro M, Senatore R, Panzeri S. Tight data-robust bounds to mutual information combining shuffling and model selection techniques. *Neural Comput* 2007;19:2913–57.
- Montemurro MA, Rasch MJ, Murayama Y, Logothetis NK, Panzeri S. Phase-of-firing coding of natural visual stimuli in primary visual cortex. *Current Biol* 2008;18:375–80.
- Muthukumaraswamy SD, Singh KD, Swettenham JB, Jones DK. Visual gamma oscillations and evoked responses: variability, repeatability and structural MRI correlates. *Neuroimage* 2010;49:3349–57.
- Nicolelis MA, Lebedev MA. Principles of neural ensemble physiology underlying the operation of brain-machine interfaces. *Nat Rev Neurosci* 2009;10:530–40.
- Oram MW, Foldiak P, Perrett DI, Sengpiel F. The 'Ideal Homunculus': decoding neural population signals. *Trends Neurosci* 1998;21:259–65.
- Ostwald D, Porcaro C, Bagshaw AP. An information theoretic approach to EEG-fMRI integration of visually evoked responses. *Neuroimage* 2010;49:498–516.
- Oyman O, Nibar RU, Bolcskei H, Paulraj AJ. Characterizing the statistical properties of mutual information in MIMO channels. *IEEE Trans Signal Process* 2003;51:2784–95.
- Panzeri S, Magri C, Logothetis NK. On the use of information theory for the analysis of the relationship between neural and imaging signals. *Magn Reson Imaging* 2008;26:1015–25.
- Panzeri S, Senatore R, Montemurro MA, Petersen RS. Correcting for the sampling bias problem in spike train information measures. *J Neurophysiol* 2007;98:1064–72.
- Percival DB, Walden AT. Wavelet methods for time series analysis. Cambridge Univ Pr; 2006.
- Pesaran B, Pezaris JS, Sahani M, Mitra PP, Andersen RA. Temporal structure in neuronal activity during working memory in macaque parietal cortex. *Nat Neurosci* 2002;5:805–11.

- Pola G, Thiele A, Hoffmann K, Panzeri S. An exact method to quantify the information transmitted by different mechanisms of correlational coding. *Network-Comp Neural* 2003;14:35–60.
- Quiroga RQ, Panzeri S. Extracting information from neuronal populations: information theory and decoding approaches. *Nat Rev Neurosci* 2009;10:173–85.
- Rasch MJ, Gretton A, Murayama Y, Maass W, Logothetis NK. Inferring spike trains from local field potentials. *J Neurophysiol* 2008;99:1461–1476.
- Ray S, Crone NE, Niebur E, Franaszczuk PJ, Hsiao SS. Neural correlates of high-gamma oscillations (60–200 Hz) in macaque local field potentials and their potential implications in electrocorticography. *J Neurosci* 2008;28:11526–36.
- Ray S, Maunsell JHR. Differences in gamma frequencies across visual cortex restrict their possible use in computation. *Neuron* 2010;67:885–96.
- Ray S, Maunsell JHR. Different origins of gamma rhythm and high-gamma activity in macaque visual cortex. *PLoS Biol* 2011;9:e1000610.
- Schneidman E, Bialek W, Berry MJ. Synergy, redundancy, and independence in population codes. *J Neurosci* 2003;23:11539–53.
- Schyns PG, Thut G, Gross J. Cracking the code of oscillatory activity. *PLoS Biol* 2011;9:e1001064.
- Shannon CE. A mathematical theory of communication. *AT&T Tech J* 1948;27:379–423.
- Siegel M, König P. A functional gamma-band defined by stimulus-dependent synchronization in area 18 of awake behaving cats. *J Neurosci* 2003;23:4251–60.
- Strong SP, Koberle R, de Ruyter van Steveninck RR, Bialek W. Entropy and information in neural spike trains. *Phys Rev Lett* 1998;80:197–200.
- Wilson EB, Hilferty MM. The distribution of chi-square. *P Natl Acad Sci U S A* 1931;17:684–8.
- Yu Y, Crumiller M, Knight B, Kaplan E. Estimating the amount of information carried by a neuronal population. *Front Comput Neurosci* 2010;4:10.
- Zohary E, Shadlen MN, Newsome WT. Correlated neuronal discharge rate and its implications for psychophysical performance. *Nature* 1994;370:140–3.