NeuroImage 73 (2013) 95–112

Contents lists available at SciVerse ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

Traveling waves and trial averaging: The nature of single-trial and averaged brain responses in large-scale cortical signals

David M. Alexander [a,*], Peter Jurica [b], Chris Trengove [a,1], Andrey R. Nikolaev [a,1], Sergei Gepshtein [c], Mikhail Zvyagintsev [d], Klaus Mathiak [d], Andreas Schulze-Bonhage [e], Johanna Ruescher [e,f], Tonio Ball [e,g], Cees van Leeuwen [a,1]

[a] Laboratory for Perceptual Dynamics, KU Leuven, Belgium
[b] RIKEN Brain Science Institute, ABSP Lab, 2-1 Hirosawa, Wako-shi, Saitama, 351-0198, Japan
[c] Vision Center Laboratory & Systems Neurobiology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
[d] Department of Psychiatry, Psychotherapy and Psychosomatics, Medical School, RWTH Aachen University, Pauwelsstraße 30, 52074 Aachen, Germany
[e] Epilepsy Center, University Medical Center Freiburg, Hugstetter Str. 49, 79095, Freiburg, Germany
[f] Faculty of Biology, University of Freiburg, Hauptstr. 1, 79104, Freiburg, Germany
[g] Bernstein Center Freiburg, University of Freiburg, Hansastr. 9a, 79104, Freiburg, Germany

ARTICLE INFO

Article history:
Accepted 13 January 2013
Available online 24 January 2013

Keywords:
Cortical topography
Evoked response
Phase
Traveling waves

ABSTRACT

Analyzing single trial brain activity remains a challenging problem in the neurosciences. We gain purchase on this problem by focusing on globally synchronous fields in within-trial evoked brain activity, rather than on localized peaks in the trial-averaged evoked response (ER). We analyzed data from three measurement modalities, each with different spatial resolutions: magnetoencephalogram (MEG), electroencephalogram (EEG) and electrocorticogram (ECoG). We first characterized the ER in terms of summation of phase and amplitude components over trials. Both contributed to the ER, as expected, but the ER topography was dominated by the phase component. This means the observed topography of cross-trial phase will not necessarily reflect the phase topography within trials. To assess the organization of within-trial phase, traveling wave (TW) components were quantified by computing the phase gradient. TWs were intermittent but ubiquitous in the within-trial evoked brain activity. At most task-relevant times and frequencies, the within-trial phase topography was described better by a TW than by the trial-average of phase. The trial-average of the TW components also reproduced the topography of the ER; we suggest that the ER topography arises, in large part, as an average over TW behaviors. These findings were consistent across the three measurement modalities. We conclude that, while phase is critical to understanding the topography of event-related activity, the preliminary step of collating cortical signals across trials can obscure the TW components in brain activity and lead to an underestimation of the coherent motion of cortical fields.

Introduction

A wealth of evidence links cross-trial averaged, evoked response (ER) measures to various stages of perception, cognition and action. However, the mechanisms responsible for the ER are poorly understood. Here we relate ERs to cortical traveling waves (TWs). The functional significance of TWs, at the columnar scale (Eckhorn et al., 2004; Nauhaus et al., 2009), Brodmann area (Freeman and Barrie, 2000; Prechtl et al., 1997; Rubino et al., 2006; Takahashi et al., 2011) and in large scale cortex (Alexander et al., 2006b; Ito et al., 2007; Klimesch et al., 2007; Massimini et al., 2004) is the subject of a growing literature. The goal of this study is to analyze the composition of the ER, considered as a large-scale pattern of TW activity.

The ER is the trial-average of the measured time-series. ER measurements across the entire scalp are typically used in source localization techniques, for both event-related potentials and event-related fields (Liu et al., 1998; Pascual-Marqui et al., 1994). An underlying assumption is that ERs reflect the magnitude and location of brain activity and its time-course. Areas or intervals with low ER magnitude therefore tend to be ignored. The ER, along with other cross-trial measures such as coherence (Hipp et al., 2011; Lachaux et al., 1999), targets brain activity that consistently summates across trials. Brain activity that does not consistently summate

* Corresponding author at: Laboratory for Perceptual Dynamics, KU Leuven, Tiensestraat 102, Box 3711, B-3000 Leuven, Belgium.
E-mail addresses: david.murray.alexander@gmail.com (D.M. Alexander), juricap@gmail.com (P. Jurica), Chris.Trengove@ppw.kuleuven.be (C. Trengove), Andrey.Nikolaev@ppw.kuleuven.be (A.R. Nikolaev), gepshtein@salk.edu (S. Gepshtein), mzvyagintsev@ukaachen.de (M. Zvyagintsev), kmathiak@ukaachen.de (K. Mathiak), andreas.schulze-bonhage@uniklinik-freiburg.de (A. Schulze-Bonhage), johanna.ruescher@uniklinik-freiburg.de (J. Ruescher), tonio.ball@uniklinik-freiburg.de (T. Ball), cees.vanleeuwen@ppw.kuleuven.be (C. van Leeuwen).
[1] Postal address: Laboratory for Perceptual Dynamics, KU Leuven, Tiensestraat 102, Box 3711, B-3000 Leuven, Belgium.

1053-8119/$ – see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.neuroimage.2013.01.016

across trials is not considered functionally important; sites of low phase consistency are generally considered to reflect activation uncorrelated with the task. This activity is assumed to behave randomly across trials (Britten et al., 1992; Ray and Maunsell, 2011).

Though technically challenging, it is possible to fit the equivalent current dipole to individual trial data and then compute the average source location, or range of locations, indicated by these single trial fits (Liu and Ioannides, 1996). The resulting evidence suggests that the sequence of activations in the trial-averaged signal does not accurately reflect the sequence of activations within individual trials. In the present research we consider the sequence of activations in the cortex by analyzing TWs within trials.

The issue of trial-averaging also arises in discussions of the relative importance of amplitude and phase to the ER. Two alternative ideas of ER generation have emerged in the literature. In the evoked model, changes in amplitude contribute to the ER, both directly and by selectively enhancing the amplitude of some phase components (Mäkinen et al., 2005; Shah et al., 2004). In the phase resetting model, only the phase of the signal will affect the ER, via changes in cross-trial phase locking initiated by experimental events (Gruber et al., 2005; Makeig et al., 2002). In this discussion of ER generation, the relative effects of phase and amplitude on evoked responses are often considered only for a limited number of recording sites (Barry et al., 2004; Gruber et al., 2005; Fell, 2007 cf. Makeig et al., 2002). Consistent with this approach, oscillatory components that have been isolated using independent components analysis also summate in a manner that suggests a topography of activity arising from localized cortical sources (Makeig et al., 2002). However, if there are large-scale topographical relationships in the way phase and amplitude interact, they may be missed when only sites with maximum ER magnitude are considered. In particular, in the trial averaged signals, TWs are often not clearly in evidence. However, this does not preclude TWs from being prominent in the unaveraged signal. This we investigate here. The topography of phase might be explained better by TWs at the single trial level than by any component of the trial-averaged signal.

An alternative to cross-trial measurement is to characterize the amplitude and phase prior to collation of signals across trials. This may be achieved by considering the topography of these components at the single-trial level. Useful information can be extracted from single trials even without analyzing the Fourier components of the signal, for example, Arieli et al. (1996) considered the topography of local field potentials in cat visual cortex. They found that the spatial pattern of within trial activity dominated the signal; the ER was merely a relatively small component. The preceding within-trial activity was an excellent predictor for within-trial activity at the latency of the ER; a much better predictor than was the trial-averaged ER topography. Thus the within trials topography of activity is highly structured, not simply noise on top of an ER (Arieli et al., 1996).

Within trials topographical information has also been critical for detecting TWs, either by amplitude deflections or by phase estimates (Alexander et al., 2006b; Eckhorn et al., 2004; Massimini et al., 2004; Rubino et al., 2006). The present research tested the prevalence of TWs as an instance of highly structured activity patterns within trials, also in order to assess the strength of such patterns against trial-averaged components such as ERs and phase coherence. Whereas the former collates signals across space but within trials, the latter, more commonly used measures, collates signals for individual sites but across trials.

Theory and simulation of cortical mechanisms have predicted the existence of large-scale TWs (Nunez and Srinivasan, 2006; Wright et al., 2001). Essentially, the global resonances that account for the 1/f spectra of cortical activity are also associated with TW dynamics. Using the electroencephalogram (EEG) and magnetoencephalogram (MEG), global cortical waves have been shown to arise at a variety of frequencies, from the sub-delta through to gamma bands (Alexander et al., 2006b, 2009; Ito et al., 2007; Massimini et al., 2004; Ribary et al., 1991; Sauseng et al., 2002). These waves are typically of long wavelength, with a spatial period of the order of 10 to 20 cm. Over one temporal cycle of the wave, wave peaks typically traverse the entire EEG/MEG recording array.

The functional significance of TWs has been established by noting their close correspondence with the latency topography of known visual and auditory ERP components, such as the P1–N1 complex, P2–N2 complex, as well as the P3b (Alexander et al., 2006b, 2009; Anderer et al., 1996; Fellinger et al., 2012; Klimesch et al., 2007). For these event-related potential (ERP) components, the latency, temporal frequency and task-dependency of evoked TW components are consistent with latency, temporal frequency and task-dependency of the corresponding ERPs.

ERPs aside, more evidence needs to be brought to bear on the functional relevance of large-scale TWs at the single-trial level. Some progress has been made toward this goal in analyzing wave activity differences across defined brain states, such as rest (Ito et al., 2005, 2007), deep sleep (Massimini et al., 2004) and working memory (Fellinger et al., 2012; Sauseng et al., 2002). Single-trial TWs have also been used to uncover genetic differences in brain activity (Alexander et al., 2007), differences across age groups and clinical groups and in correlations with clinical symptoms (Alexander et al., 2006a,b, 2008, 2009).

Large-scale patterns of activity in the scalp EEG are partly a function of blurring due to volume conduction of the dura, skull and other tissues; to a lesser extent blurring effects also apply to MEG and electrocorticogram (ECoG) measurements. However, a number of analyses clearly indicate that TWs measured in the EEG are not an artifact of volume conduction. TWs can still be detected in the EEG using sparse electrode arrays with a minimum electrode separation of 10 cm (Alexander et al., 2009); the spatial resolution of the measurement is matched to the spatial resolution of the signal to discount blurring artifact. TWs in EEG can also be detected by measuring latency delays rather than spatial patterns, per se (Alexander et al., 2006b; Fellinger et al., 2012; Manjarrez et al., 2007; Massimini et al., 2004; Nauhaus et al., 2009; Patten et al., 2012). Since volume conduction effects are essentially instantaneous, to explain away the motion of the apparent waves requires a more complicated hypothesis than volume conduction alone provides (see Ray and Maunsell, 2011, for one such hypothesis; c.f. Nauhaus et al., 2012). To further address the issue of blurring by volume conduction, in the present research we analyzed data from a range of imaging modalities: MEG, EEG and ECoG. We chose these modalities because they have different effective spatial resolutions: approximately 4 cm, 10 cm and 1 cm for MEG, EEG and ECoG, respectively (Bullock et al., 1995; Srinivasan et al., 2007). Establishing the ubiquity of TWs using each of these measurement modalities would add further evidence against the argument that TWs arise as an artifact of volume conduction.

The results of this study show that phase plays the major role in the ER topography, more-so than amplitude, consistently across imaging modalities. We also observed that the topography of trial-averaged phase, while correlated with ER topography, co-occurs with ubiquitous episodes of TW activity at the single-trial level. In the three data sets analyzed here, the topography of within-trial phase was better described by TWs than by trial-averaged phase, suggesting a loss of information in the latter case. Linking TWs back to ER topography, we show that at some event-related times and frequencies, the ER topography can be approximated as a trial-average of TW components estimated at the single trial level. The existence of smooth spatial gradients of phase within trials, i.e. TWs, is entirely consistent with spatial and temporal variations in cross-trial phase locking. We propose that ER magnitudes are partly the product of TWs that summate and cancel differentially across measurement sites.

Materials and methods

We re-analyzed data from three previously published studies. These data were chosen because they involve a range of simple motor and/or sensory tasks recorded under controlled conditions.

The recordings involve three different imaging modalities: MEG, EEG and ECoG. These modalities have different degrees of spatial resolution (EEG<MEG<ECoG), allowing analysis of spatial patterns of activity under different conditions of blurring. The Apparent Motion task and the Dot Lattice task were carefully designed to avoid eye movement artifacts, with short inter-stimulus intervals and subjects being required to fixate their gaze.

MEG Apparent Motion task

Twenty human subjects (age range 22–36 years, mean age 27.1; 12 females) engaged in an audio–visual perceptual task, while their brain activity was recorded via MEG. All the subjects were right-handed, had no audiological abnormalities, and had normal or corrected-to-normal vision. Written informed consent was obtained from all subjects prior to participation in the study. The study was approved by the ethics committee of the University of Tübingen, Germany. The task required the subject to choose the direction of motion of an audio–visual apparent motion stimulus. The visual stimulus was located on the horizontal meridian with the distance of 15° of visual angle at either side of the screen center. The apparent motion illusion was elicited by presenting the stimuli for 67 ms at the one side, and then after 67 ms delay, for 67 ms at the opposite side. The auditory stimuli were the white noise bursts processed in such a way that the sound was spatially perceived at the position of the visual stimuli (+15° or −15°). Subjects were instructed to trigger the stimuli by pressing a button with either the left or right index finger, with random choice for each trial. The stimulus moved from the side indicated by the subject and then to the other side (i.e. either left-to-right or right-to-left; these were the 'predictable-left' and 'predictable-right' conditions). In some blocks of trials the direction of the stimulus motion was randomized and not due to the subject's choice (i.e. 'unpredictable-left' or 'unpredictable-right' conditions). We applied analyses only on the trials in which the side of first stimulus coincided with the button pressed: all trials in the predictable and approximately half of trials in the unpredictable condition. Further details of the experiment can be found elsewhere (Zvyagintsev et al., 2008).

Neuromagnetic responses were recorded in a magnetically shielded booth using a 151-sensor whole-head gradiometer (CTF Systems Inc., Vancouver, Canada). Measurements were performed with subjects in a sitting position. The MEG signals were sampled at 312.5 Hz. Time zero in each trial was designated as the time of the button press. Data were analyzed in each trial from −500 ms to +500 ms.

EEG Dot Lattice task

Thirteen healthy subjects (ages 19–36, mean age 23.6, 9 females) took part in the experiment. The stimuli were multistable dot lattices, each of which appeared to be grouped into strips of dots (Kubovy, 1994). The perceived organization of a dot lattice depends on its aspect ratio (AR), which is the ratio of the two shortest inter-dot distances. Four different values of AR: 1.0, 1.1, 1.2, and 1.3 were used. The lattices were presented at four different orientations (22.5, 67.5, 112.5, or 157.5°). The four aspect ratios and four orientations yielded 16 different stimuli. Within each experimental block each of the 16 conditions was presented 10 times in a random order. Four such blocks were presented to every observer.

The diameter of the dots was 0.2° of visual angle. The shortest distance between dot centers at AR = 1.0 was 0.6° of visual angle. Subjects sat 1.15 m from the screen in a dimly lit room. Each trial consisted of four intervals: fixation, stimulus, blank screen, and response screen. The duration of the stimulus interval and the blank-screen interval were both fixed at 300 ms. A response screen (four icons depicting the orientations of most likely dot groupings) was presented until a response was received. The subjects' task was to report the orientation of the perceived grouping by clicking on the corresponding icon. Further details of the experiment can be found elsewhere (Nikolaev et al., 2008).

EEG was recorded using a 256-channel Geodesic Sensor Net (Electrical Geodesics Inc., Eugene, OR). Data were digitized at 250 Hz. All channels were referenced to the vertex electrode (Cz). Impedance was kept below 50 kΩ. All channels were preprocessed on-line using 0.1 Hz high-pass filtering and 100 Hz low-pass filtering. The outer ring of electrodes (i.e. those over the forehead, cheek and neck) was removed from the analysis, leaving only the 184 electrodes over symmetrically positioned cranial sites. The channels were then re-referenced to average reference. Average reference was preferred over the original vertex reference since the typical range of measured phase over the whole array partly cancels out the effects of reference on the measured TWs. Time zero in each trial was designated as the time of the stimulus presentation. Data were analyzed in each trial from −100 ms to +400 ms. Epochs with eye-movement artifact were remove from analysis (Nikolaev et al., 2008).

ECoG Finger Movement task

A female patient (aged 55 years) suffering from intractable pharmacoresistant epilepsy with a right fronto-polar focal cortical dysphasia took part in this study after having given her informed consent. The study was approved by the university clinic's ethics committee. The patient was strongly right-handed according to a modified Oldfield questionnaire and showed no clinical signs of pareses or other movement disorders. A platinum electrode array (4 mm electrode diameter, 112 contacts, 7.1 mm inter-electrode distances) was subdurally implanted above the left fronto-parieto-temporal region for pre-neurosurgical diagnostics. Electrical stimulation was performed with the stimulator INOMED NS 60 (INOMED, Germany) to demarcate the eloquent brain areas. The intensity of stimulation was gradually increased up to 15 mA or to the induction of sensory and motor phenomena. All sites with arm or hand motor responses were located outside the ictal onset zone. A structural MRI data set with full head coverage was acquired on the day after electrode implantation using a T1 MPRAGE sequence.

The subject was instructed to perform self-paced index finger flexions of either the left or the right hand with inter-movement intervals of at least 4 s. Here the behavioral task analyzed differed from the task reported in a previous study using the same subject (Ball et al., 2009). The ECoG was recorded using a clinical AC EEG-System (IT-Med, Germany) digitized at 256 Hz and band-pass filtered (0.032 Hz to 97 Hz). Further details of the subject recording can be found elsewhere (Ball et al., 2009). ECoG data were re-referenced to average reference prior to data analysis. Onsets of index finger movement were determined in the electromyogram (EMG) of the M. flexor digitorum superficialis, pars indicis. Time zero in each trial was designated as the time of movement onset. Data were analyzed in each trial from −3000 ms to +3000 ms, to incorporate most of the inter-movement interval. Only the condition of movement contra-lateral to implanted electrodes was analyzed.

Numerical methods

We let T, S, L, and F denote—respectively—the sample times, measurement sites (MEG sensors or EEG or ECoG channels), trials and frequencies, as sequences of lengths N_T, N_S, N_L, and N_F. Let f_S denote the sample frequency. The raw time-series signal is a 3-dimensional data set, $x_{T\times S\times L}$. The Fourier components of the signal were estimated for a sequence of logarithmically-spaced center frequencies ranging from 0.5 to 28.0 Hz using 2 cycle Morlet wavelets. The use of two cycles in the Morlet wavelets enables the phase and amplitude to be estimated from very short time windows, at the expense of fine frequency resolution (Alexander et al., 2006b; Herrmann et al., 2005). This range of

frequencies was chosen as the maximum range that had good signal to noise ratio in all three data sets.

The Fourier components are denoted $X_{T\times S\times L\times F}$. From the Fourier components we define the following quantities having the same indices T, S, L, and F:

$$\Phi = \arg(X/||X||) \tag{1}$$

which is the phase angle i.e. $e^{i\Phi} = X/||X||$,

$$b = \mathrm{Re}(X) \tag{2}$$

which is the band-pass time-series and

$$p = \mathrm{Re}(X/||X||) \tag{3}$$

which is the amplitude-normalized band-pass time-series.

Comparing the spatial organization of different quantities

The present analyses gauge the spatial pattern of various quantifications of the activity, e.g. the inter-trial phase coherence (ITPC). We compare the topographical similarity of the various quantities. Since all the statistics used herein quantify the pattern over the topography, for brevity we leave 'the topography of' implicit in the naming scheme. So ITPC should be read as *the topography of the ITPC*. A glossary of the statistics used is provided in Table 1.

The statistics used in this paper follow a general strategy: we compare a vector of observed signals with a vector of model signals. The components of each vector may correspond to the set of all measurement sites at time sample, or at a number of points in time in a window centered around a time sample. In the case where the vector components belong to one time sample, they are indexed by elements of T, F and S, and the comparison of two vectors over sites S is made for each $(t,f)\in T\times F$. This is written as

$$r_{t,f}(H,K) = \rho(H_S, K_S) \tag{4}$$

where ρ denotes the correlation function for two vectors H and K. The S index is implicit on the left-hand side of the expression. $r_{t,f}$ is the correlation value. If the signals at the individual sites are real-valued, H and K are vectors of length N_S. In case of complex-valued phases at each site, H and K are vectors of length $2N_S$, created from each complex-valued vector of length N_S by appending the vector of cosines to the vector of sines of the phases. In the case where the vector components are drawn from a range of time samples in a window, the components are indexed by elements of T, F, S and C and the correlation over sites S and times C (relative to time t; as defined below) is calculated for each $(t,f)\in T\times F$. This is written as

$$r_{t,f}(H,K) = \rho(H_{S\times C}, K_{S\times C}) \tag{5}$$

where the indices S and C are implicit on the left-hand side of the expression. We compare spatial patterns over a range of times when the model signal is in the time domain and we compare spatial patterns at a single sample when the model signal is in the frequency domain.

Evoked response measures

The average of the time-series over trials is $\overline{x}_{T\times S} \equiv \langle x_{T\times S\times L}\rangle_L$, also known as the evoked response (ER). We compare the evoked response to two model signals. These are constructed from trial-averages of band-pass and amplitude-normalized band-pass time-series, at each frequency f. We compare the evoked response with the model signals over all sites and over one temporal cycle at the frequency of interest (e.g. 100 ms for 10 Hz). This time window of width $1/f$, centered at sample t, is defined as follows. Let C_f denote a sequence of time offsets at frequency f: $C_f = (c_1, c_2, \ldots, c_{N_c})$ where $\{c_1, c_2, \ldots, c_{N_c}\} = (-{}^1\!/_2 f\,,\, {}^1\!/_2 f)\cap(\mathbb{Z}/f_s)$, and $\mathbb{Z}$ denotes the integers. Thus $N_{C_f}\approx f_S/f$. At each $(t,f)\in T\times F$ we define windows of the trial averaged signals around t: $\hat{x}_{S\times C_f} = (\overline{x}_{t+c,s})_{s\in S, c\in C_f}$. The evoked response is then normalized to have a mean of zero for each site by subtracting the mean over the window C_f:

$$\tilde{x}_{S\times C_f} = \hat{x}_{S\times C_f} - \left\langle \hat{x}_{S\times C_f}\right\rangle_{C_f} \times 1_{C_f} \tag{6}$$

$\tilde{x}_{S\times C_f}$ is the evoked response, base-lined over one cycle. Hereafter we refer to this response as the *normalized ER*. Similarly, for the band-pass signal, we average b across trials, and define windows of $\overline{b}$ around each t as $\hat{b}_{S\times C_f}$ and subtract the mean for each site to get $\tilde{b}_{S\times C_f}$. This is the trial-average of the band-pass signal, hereafter referred to as the *band-pass ER*. Similarly from the phase-only signal p we define $\tilde{p}_{S\times C_f}$, which is the trial-average of the amplitude-normalized band-pass signal, hereafter referred to as the *phase-only ER*.

The correlation of the normalized ER to the band-pass ER is denoted $r_{t,f}\left(\tilde{x},\tilde{b}\right)$, hereafter referred to as the *band-pass fit*. The correlation of the normalized ER to the phase-only ER is denoted $r_{t,f}(\tilde{x},\tilde{p})$, hereafter referred to as the *phase-only fit*. The measure band-pass fit answers the following question: how much of the variance in the ER is explained as the trial average of the band-pass signal when the two signals are compared as a pattern over all sites and over one temporal cycle? (to be exact, the variance explained is the correlation squared). The measure phase-only fit answers an analogous question: how much of the variance in the ER is explained as the trial average of the amplitude-normalized band-pass signal when the two signals are compared as a pattern over all sites and over one temporal cycle?

Spatial organization of inter-trial coherence

The phase-only ER and the band-pass ER are in the time domain and therefore have the advantage that they can be compared directly to the evoked potential (here, normalized ER). This last quantity has the advantage that it is more widely researched than measures based on Fourier components such as ITPC, and is therefore more readily interpretable

Table 1
Glossary of the central measures used in this study.

Glossary		
Name	Notation	Description
Average-phase fit	$r_{t,f}(\Phi_q,\Phi_{!q})$	The measured phase correlated with the average-phase model.
Average-phase model	$\overline{\Phi}_{S,!q}$	The topography of trial-averaged phase.
Band-pass ER	$\tilde{b}_{S\times C_f}$	The topography of the trial-averaged, band-pass filtered signal.
Band-pass fit	$r_{t,f}\left(\tilde{x},\tilde{b}\right)$	The normalized ER correlated with the band-pass ER.
ITBC	$\overline{B}_S$	The topography of the inter-trial band coherence.
ITPC	$\overline{P}_S$	The topography of the inter-trial phase coherence.
ITPC fit	$r_{t,f}\left(\overline{B},\overline{P}\right)$	The inter-trial band coherence correlated with the inter-trial phase coherence.
Measured phase	$\Phi_{S,q}$	The topography of the complex-valued phase.
Normalized ER	$\tilde{x}_{S\times C_f}$	The topography of the evoked response (ER), i.e. trial-averaged time-series.
Phase-only ER	$\tilde{p}_{S\times C_f}$	The topography of the trial-averaged, amplitude-normalized, band-pass filtered signal.
Phase-only fit	$r_{t,f}(\tilde{x},\tilde{p})$	The normalized ER correlated with the phase-only ER.
TAA	$\overline{A}_S$	The topography of the trial-averaged amplitude.
TAA fit	$r_{t,f}\left(\overline{B},\overline{A}\right)$	The inter-trial band coherence correlated with the trial-averaged amplitude.
Wave activity	$r_{t,f}(\Phi_q,\Theta_q)$	The estimated phase correlated with the wave map.
Wave ER	$\overline{w}_{S\times C_f}$	The topography of the trial-averaged wave maps.
Wave fit	$r_{t,f}(\tilde{x},\overline{w})$	The normalized ER correlated with the wave ER.
Wave map	$\Theta_{S,q}$	The traveling wave, represented as a topography of complex-valued phase.

as the time-course of the magnetic field generated by an equivalent current dipole or (as the case may be) a spatio-temporal pattern of ERPs. Here we can work in the time-domain, and nevertheless take the phase of the signal into account by comparing to the ER over one cycle at the frequency of interest. The disadvantage of the band-pass fit and the phase-only fit, however, is that the effect of amplitude at different frequencies on the evoked potential is not directly assessed; it is inferred by comparing the relative effects on trial-averaged signals that include an amplitude component (band-pass ER) or do not include an amplitude component (phase-only ER).

For this reason we performed a supplementary analysis in which we directly compared the spatial pattern of the Fourier components averaged across trials. Here we can work directly with the Fourier components so we consider only the values at single samples. Even so, the phase and amplitude values are not perfectly localized in time, since they can be considered to be a weighted average over the time-window used in the estimate of the Fourier components (Herrmann et al., 2005). In the present study we used two-cycle Morlet wavelets to estimate the Fourier components. Since the Gaussian component of the Morlet wavelet assigns a heavier weighting to the middle cycle of the two cycles of the time-series window, the window in the phase estimates approximates that of the one cycle window used for the band-pass ER and the phase-only ER in the first analyses.

In order to directly compare the effects of amplitude and phase on the trial-averaged band-pass signal, we use two measures of inter-trial coherence. At each $(t,f)\in T\times F$ we compute the *inter-trial phase coherence*, ITPC (Delorme and Makeig, 2004; Makeig et al., 2004; or phase-locking factor; Tallon-Baudry et al., 1996). We first normalize the lengths of each Fourier component to unity, then compute their complex average, and take the magnitude of the resultant:

$$\overline{P}_S = \left| \left\langle \frac{X_{S\times L}}{|X_{S\times L}|} \right\rangle_L \right|. \tag{7}$$

Analogously, the *inter-trial band coherence* (ITBC; or linear inter-trial coherence, Delorme and Makeig, 2004) equals the magnitude of the trial-averaged Fourier component at each $(t,f)\in T\times F$.

$$\overline{B}_S = \left| \left\langle \frac{X_{S\times L}}{A_{\mathrm{rms}}} \right\rangle_L \right| \tag{8}$$

where A_{rms} is the RMS of the amplitude at each measurement site over all samples for that subject/condition, i.e.

$$A_{\mathrm{rms}} = \left\langle |X_{T\times S\times L}|^2 \right\rangle_{T\times L}^{1/2}. \tag{9}$$

The ITBC is therefore a trial-average that includes (RMS normalized) amplitude information and ITPC is a trial-average that excludes amplitude information. We also calculate the mean amplitudes over trials (RMS normalized):

$$\overline{A}_S = \left\langle \frac{|X_{S\times L}|}{A_{\mathrm{rms}}} \right\rangle_L. \tag{10}$$

This we simply call the *trial-averaged amplitude*, TAA. Throughout this work we reserve the term amplitude for signal deviations observable prior to trial-averaging and reserve the term magnitude for deviations in trial-averaged signals, such as ERs.

For each $(t,f)\in T\times F$ we compare TAA directly to ITBC over sites as $r_{t,f}\left(\overline{B},\overline{A}\right)$, hereafter referred to as the *TAA fit*. We also compare ITPC to the ITBC over sites: $r_{t,f}\left(\overline{B},\overline{P}\right)$, hereafter referred to as the *ITPC fit*. Using these two statistics we can directly compare the contributions of either phase or amplitude to the trial-averaged band-pass signal as it varies over all measurement sites.

Models of phase organization

In this paper we show that, at task relevant times and frequencies, the phase component of the signal can explain a larger proportion of the variance in the spatial pattern of the ERs than the amplitude component. The question naturally arises, given that the trial-averaged phase signal can constitute a reasonable model of the spatial pattern of ERs, what does this tell us about phase at the single trial level? Is phase a variable that is ordered at sites with large magnitude ER but random at sites with low magnitude ER? This would rule out large-scale TWs i.e. globally ordered phase gradients, at the single trial level. On the other hand, if we could demonstrate the ubiquity of TWs at the single trial level, this would militate against the strong version of the statement that phase is random at sites of low magnitude ER. A corollary of such a demonstration is that the phases in cross-trial measures as the ITPC can appear to be random because they have been collated across trials rather than spatial locations.

To pit these hypotheses against each other, we model the spatial pattern of single samples of phase, comparing the accuracy of two alternative models. The first model is derived from trial-averaged phase (i.e. at time-locked samples) and the second is derived from phase gradient information available within single samples. Here again we use single samples of phase, since these samples of phase are calculated across two cycles of the time signal using Morlet wavelets.

At each $(t,f)\in T\times F$, the model spatial patterns of phase are compared directly to the spatial vector of single-sample phase at trial q, denoted $\Phi_{S,q}$ where $\Phi_{S,q}=(\Phi_{s,q})_{s\in S}$, hereafter referred to as the *measured phase*. To compare the measured phase to the trial average, we calculate the trial average excluding trial q:

$$\overline{\Phi}_{S,!q} = \left\langle \left(\Phi_{S\times L\backslash q}\right) \right\rangle_{L\backslash q} \tag{11}$$

where $L\backslash q$ is the sequence of trials with trial q excluded. The spatial vector of the trial averaged phase is referred to as the *average-phase model*. We then compute the correlation of the average-phase model to the measured phase, $r_{t,f}\left(\Phi_q,\overline{\Phi}_{!q}\right)$, hereafter referred to as the *average-phase fit*.

We also fit the measured phase to a TW model. The model TW is denoted $\Theta_{S,q}$, hereafter referred to as the *wave map*. The fit between the wave map and the measured phase is denoted $r_{t,f}(\Phi_q,\Theta_q)$, and hereafter referred to as the *wave activity*. The wave activity was calculated as follows. First, the measured phase was spatially unwrapped. Next, we obtain the wave map as the single gradient vector best characterizing the phase gradient across the measurement array. This wave map is then correlated with the measured phase to give the wave activity.

Traveling wave model

We convert the phase into unwrapped phase by taking the discrete spatial derivative and then reintegrating spatially. The chief reason for this is that the spatially unwrapped phases allow the representation of a wave as a gradient over a field of scalars.

Nearest neighbor edges on the measurement array e_{ij}; $i,j\in S$, are determined by Delaunay triangulation. For the phase at some t, f, and q, the phase difference on each edge is defined as $\delta\Phi_{ij}=\arg(\Phi_i)-\arg(\Phi_j)+2k\pi$ with integer k such that $\delta\Phi_{ij}\in[-\pi,\pi)$. We assign unwrapped phase values Ψ_i, $i\in S$, using nearest neighbor phase relationships $\Psi_i=\Psi_j+\delta\Phi_{ij}$. Edges are added sequentially, in ascending order of $|\delta\Phi_{ij}|$, to an initially edgeless graph with sites as vertices. If $|\delta\Phi_{ij}|<<\pi$ is not everywhere true, unwrapping errors can arise, where $\Psi_i-\Psi_j=\delta\Phi_{ij}+2k\pi$ for some integer $k\neq 0$. However, the number of unwrapping errors is small (compared to the number of e_{ij}) in data where phase varies smoothly. Unwrapping errors may either reflect legitimate discontinuities in the phase or measurement noise (Spagnolini,

1995). Adding edges in ascending order of $|\delta\Phi_{ij}|$ places unwrapping errors in regions of highest $|\delta\Phi_{ij}|$, i.e. at legitimate discontinuities and high measurement noise.

Exploration of phase dynamics using *k*-means clustering (Ito et al., 2007) and principal component analysis (Alexander et al., 2006b) indicated that much of the variance in the unwrapped phases could be explained by linear gradients of phase with the form

$$\Psi_S = \cos\left(\varphi\left(P_S - \left(o_\theta, o_\varphi\right)\right)\right) \tag{12}$$

where (o_θ, o_φ) is the location of the origin (bullseye) of the phase gradient on the surface of a sphere, P_S gives the coordinates of each measurement site projected onto the surface of a sphere and $\varphi(\ )$ gives the polar angle in the spherical coordinate system. This wave model is linear when expressed in Cartesian coordinates. The principal component analysis of the present data sets confirmed this linear wave model as explaining a large amount of the variance in unwrapped phases (see Results section).

We therefore estimated the *wave map* (Alexander et al., 2006b, 2009) as a linear gradient in the volume defined by the measurement sites, by solving a regression equation of the form

$$\Psi_S = \beta_0 + \beta_{AP}C_{AP} + \beta_{IS}C_{IS} + \beta_{LR}C_{LR} \tag{13}$$

where β represents the regression coefficients and (C_{AP}, C_{IS}, C_{LR}) are the Cartesian coordinates of the measurement sites. *AP* refers to the anterior–posterior axis of the measurement array, *IS* to the inferior–superior axis, and *LR* to left–right axis. The relative phases are thus fit to a gradient described by a single three dimensional vector, namely (β_{AP}, β_{IS}, β_{LR}), plus a constant phase term.

The wave map, Θ_q, computed from the measured phase at trial q is then:

$$\Theta_{S,q} = \exp\left(i\left(\beta_\varphi + \beta_{AP}C_{AP} + \beta_{IS}C_{IS} + \beta_{LR}C_{LR}\right)\right) \tag{14}$$

where β_ϕ re-introduces the correct phase offset that was removed in the spatial unwrapping of phase. Note that this formulation differs slightly from Alexander et al. (2006b, 2008, 2009), where the wave map (or phase gradient) was expressed in the metric of a linear gradient of relative phases i.e. without converting to a complex number.

We were interested in whether the ER can be understood as an average of TW components. To test this idea, we computed the average of the traveling wave components over trials, and compared this average to the normalized ER. Since the wave map is defined for a single sample, we first extended the wave map over a window of one cycle to allow comparison with the normalized ER:

$$\theta_{S\times C_f,q} = \left(\Theta_{S,q} + 2\pi fc\right)_{c\in C_f}. \tag{15}$$

Some examples of the wave map extended over one cycle are given in Fig. 3, third column. The *wave ER* is defined as the trial-average of the real part of $\theta_{S\times C_f}$.

$$\overline{w}_{S\times C_f} = \left\langle \mathrm{Re}\left(\theta_{S\times C_f\times L}\right)\right\rangle_L. \tag{16}$$

The *wave fit* is analogous to the phase-only fit, and is denoted $r_{t,f}(\tilde{x},\overline{w})$. The measure wave fit answers the following question: how much of the variance in the ER is explained as the trial average of the model traveling waves when the two signals are compared as a pattern over all sites and over one temporal cycle?

Permutation methods to test over-fitting and family-wise error

The measures average-phase fit, $r_{t,f}(\Phi_q,\overline{\Phi}_{!q})$, and wave activity, $r_{t,f}(\Phi_q,\Theta_q)$, both compare the phase signal from a sample within a single trial to a model signal. However, the information used to construct the two model signals is quite different. The average-phase model pools information from N_L-1 trials, and the mean phase is calculated independently at each measurement site, prior to expression as a vector over sites. The wave activity pools information across the measurement array, from within a single sample of measured phase, to produce a vector that describes the direction and magnitude of the wave flow across the measurement array. This wave map is then correlated with the empirical phase. The average-phase model thus pools information across trials, whereas the wave map describes the 'average' behavior across the measurement array, within a single sample.

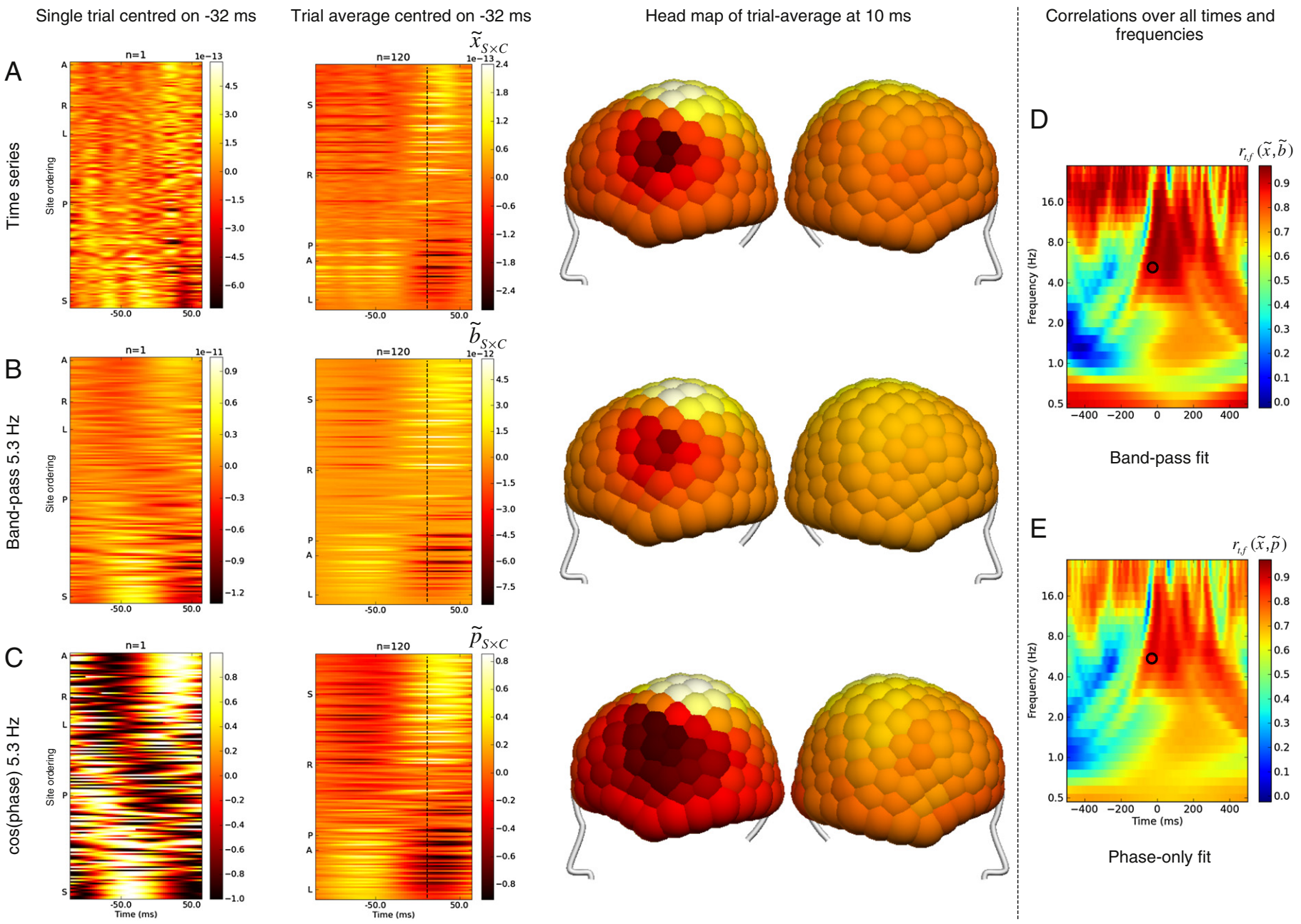
The calculation of the wave activity, however, involves a step of estimating the model from the empirical phase, prior to calculating the fit of the model back to the empirical phase. As a simple check that the wave activity was not over-fitting the data, we re-analyzed one subject from each measurement modality. In this extra analysis we randomized the phase at each site and sample, in order to assess the correlations expected from the fitting procedure alone.

Each correlation of the form $r_{t,f}(H,K_a)$ was paired with another measure $r_{t,f}(H,K_b)$ against which it was compared, where H is the empirical signal and K is the model signal. To gauge whether these pairs of correlations differed from each other, for each $(t,f)\in T\times F$ we performed a t-test or Mann–Whitney U-test. We used the former for subject-wise comparisons (e.g. $r_{t,f}(\tilde{x},\tilde{p})$ vs. $r_{t,f}(\tilde{x},\tilde{b})$) and the latter for trial-wise comparisons (e.g. $r_{t,f}(\Phi_q,\Theta_q)$ vs. $r_{t,f}(\Phi_q,\overline{\Phi}_{!q})$ within subjects), since the wave activity values were not normally distributed. Due to the large number of statistical tests that this entails ($N_T\times N_F$ for each pair of measures, for each subject/condition, for each measurement modality), the overall statistical significance for each time by frequency matrix was assessed by analysis of the spatial clustering of contiguous significant regions in the time/frequency matrix (Alexander et al., 2008, 2009; Friston et al., 1991, 1996). For measures in the present study all significant clusters were extremely large or encompassed all (t,f), and passed the permutation tests for family-wise error. The statistical differences between pairs of measures were not central to the arguments presented in this paper.

Results

We present the results by numerical method rather than by imaging modality; so MEG, EEG and ECoG data are considered together, with MEG used to illustrate the details of each analysis. We first quantify the relationship between the ER and two surrogates for the ER: the band-pass ER and the phase-only ER. The band-pass fit provides only slightly better explanatory power than the phase-only fit at

Fig. 1. Illustration of the method for calculating the contribution of band-pass ER and phase-only ER to the normalized ER. This data is from MEG subject A6, condition predictable-left. A. Time-series data, *x*, in units of Tesla. B. Band-pass filtered data, *b*, at 5.3 Hz, also in Tesla. C. Cosine of the phase, *p* (amplitude-normalized band-pass). The first column shows single trial examples of these data. Each quantity is shown over one cycle at the frequency of interest i.e. from −126 to +62 ms (x-axis) and over all sites (y-axis). Here the center of the time-window is at −32 ms relative to the subject's button press. Some sites are labeled to indicate the approximate ordering: 'A' is the most-anterior site, 'P' posterior, 'I' inferior, 'S' superior, 'L' left, and 'R' right. The sites are ordered by values of the wave map calculated from the phases of the center sample. This is done solely for representational convenience, to highlight any smooth spatio-temporal patterns in phase; it is not relevant to the calculation of the final statistic. The second column shows normalized ER, band-pass ER and phase-only ER i.e. $\tilde{x}_{S\times C}$, $\tilde{b}_{S\times C}$ and $\tilde{p}_{S\times C}$. Here the sites are ordered by calculating the wave map from the trial-averaged phase values at the center sample. The black dashed line indicates the sample at +10 ms. The third column shows a projection of the trial-averaged data at the +10 ms sample onto the measurement array coordinates. The units of the color scale are the same as in the second column. D and E. The fourth column shows the values of the band-pass fit and phase-only fit for all *t*, *f* and the values at −32 ms, 5.2 Hz are indicated with a black circle.


Single trial centred on -32 ms
Trial average centred on -32 ms
Head map of trial-average at 10 ms
Correlations over all times and frequencies
A
Time series
B
Band-pass 5.3 Hz
C
cos(phase) 5.3 Hz
D
E
Band-pass fit
Phase-only fit
Site ordering
Frequency (Hz)
Time (ms)
n=1
n=120

event-related peaks in these measures. We conclude that amplitude plays a lesser role in ER magnitudes than phase when the ER magnitudes are considered as a spatio-temporal pattern. Since the ER can be well modeled by phase-only signals, we then compared the single-trial phase to two model signals: the average-phase model and the wave map estimated from single samples of phase. The latter explains a larger amount of variance in the measured phase than the former. We also show that the ER can be well modeled as an average of traveling wave components. We conclude that regional and temporal variations in phase coherence in trial-averaged measures are consistent with intermittent but ubiquitous epochs of TWs within trials.

Evoked response as a spatial pattern

Here we assess the contribution of phase and amplitude to the pattern of ER magnitudes over the whole measurement array and over one temporal cycle (see Materials and methods section). Fig. 1 (head maps) shows the spatial pattern of the phase-only ER for a single MEG subject (A6) at 5.3 Hz, 10 ms. The sample at +10 ms is shown rather than the center of the window at −32 ms, because at the center of the window, the apparent standing wave is reversing in sign, for this subject's ER. The spatial distribution of ER values shows the pattern of magnetic field strengths due to motor dipole at the time of the button press (compare with Zvyagintsev et al., 2008, Fig. 2). Fig. 1 (head maps) also shows that the spatial pattern of the phase-only ER and the band-pass ER at +10 ms closely matches that of normalized ER. Indeed the entire time-course of the normalized ER is well approximated by the phase-only ER and the band-pass ER, as shown in Fig. 1 (second column). In the phase-only ER, no traveling wave is apparent, but instead the dipole-related activity has the appearance of a standing wave that reverses in sign over the 188 ms window. Supplementary Video 1 illustrates the time-course of the phase-only ER at 5.3 Hz, for another subject.

In general, the time-course of the normalized ER, over the entire scalp, is reasonably approximated by the band-pass ER and the phase-only ER, provided that an appropriate time window and frequency is chosen for the latter measures. This approximation is best near at peaks in the band-pass fit and in the phase-only fit. The band-pass and phase-only fits for subject A6 are given in Fig. 1, last column. In this plot it can be seen that the value of the band-pass fit is large at −32 ms, 5.3 Hz, reflecting our choice of center sample and frequency for Fig. 1, columns 1 to 3.

Fig. 1 (first column) shows data from a single trial, for subject A6's MEG data (from the predictable-left condition), in the form of a time by site matrix. The plots are of the raw time-series, the band-pass signal filtered with a center frequency at 5.3 Hz, and the cosine of phase at 5.3 Hz. A wave travels from anterior-right to posterior-superior sites, seen most clearly in the cosine of the phase. This trial was chosen because its behavior is quite different to that revealed in the trial-average for the same subject. We explore this aspect of the data further, in the second half of the results.

For all subjects, we compared the pattern of ER magnitudes with model signals composed of either the band-pass ER or the phase-only ER i.e. the band-pass fit and the phase-only fit. Plots of these measures for another MEG subject are shown in Fig. 2. Typically for this task, the trial-average signals reveal ongoing patterns of activity in the alpha/beta bands, prior to the button press at time zero, that well approximate the normalized ER. This match is more prominent in the band-pass fit plot, resulting in a negative value (blue region) in the difference plot for the alpha/beta bands prior to time zero. The event-related peaks in the band-pass fit around time zero (button press) are also typical for this experiment. These theta-band peaks are related to the motor dipole (Zvyagintsev et al., 2008). The alpha band peaks from +50 ms to +180 ms are related to visual and auditory dipoles elicited by the stimulus (Zvyagintsev et al., 2008). At all these event-related peaks, the phase-only ER explains only slightly less of the variance in normalized ER than does the band-pass ER.

Subject-average plots of the phase-only fit and the band-pass fit are shown in Supplementary Fig. 1. As in the single subject examples, a pattern of time-locked responses on these measures begins with a theta band component around the time of the button press, related to the motor dipole. The audio–visual stimulus then evokes a series of events in the alpha band from 80 ms onwards, also consistent with previous dipole modeling of visual (80 ms) and auditory evoked components (150 ms) (Zvyagintsev et al., 2008). Ongoing alpha and beta band activity was also apparent, prior to the button press, which is substantially larger for the band-pass fit.

Spatial patterns of evoked response in three measurement modalities

At these task relevant times (+10, +80 and +150 ms) for the MEG task, we calculated the peak frequency of the band-pass fit, given in Table 2. The values of the band-pass fit and the phase-only fit at these times and frequencies are also given. At these event-related peaks in the band-pass fit, the band-pass ER correlated better with the normalized ER than did the phase-only ER. However, the peaks in the band-pass fit and the phase-only fit were of similar magnitude, with an additional correlation of +0.5 to +0.9 when the amplitude portion of the signal was included in the calculation. This observation was true for all times and frequencies relevant to the motor- and stimulus-related MEG signal i.e. in the interval −150 ms to +200 ms (Supplementary Fig. 1A, third panel).

A similar observation applies to all three data sets. Fig. 2 also shows a typical subject for the EEG data (subject KH, AR = 1.0). Typically for this task, the trial-average signals reveal ongoing patterns of activity in the beta band, prior to the stimulus presentation at time zero, that well approximate the normalized ER. This is more prominent in the band-pass fit plot, resulting in a negative value (blue region) in the difference plot in the beta band prior to time zero. The event-related peaks at +90 ms and +170 ms are typical for this experiment and are related to the P1 and N1 ERPs (Nikolaev et al., 2008). The phase-only ER explains only slight less of the variance in normalized ER than does the band-pass ER, for these evoked components.

The grand-average EEG plots for the phase-only fit and the band-pass fit are shown in Supplementary Fig. 1. The two main peaks in the band-pass fit occur at about +100 ms and +152 ms. The latencies are those of the P1 and N1 ERPs (Nikolaev et al., 2008). The values (and peak frequencies) of the band-pass fit and the phase-only fit at these times and frequencies are also given in Table 2. The peaks in the band-pass fit were greater than the phase-only fit, but of similar magnitude, with only +0.4 additional correlation when the amplitude portion of the signal was included in the calculation. This general observation held for all times and frequencies relevant to the stimulus-related EEG signal i.e. in the interval +50 ms to +250 ms (Supplementary Fig. 1B, third panel).

For the ECoG data (Fig. 2C), there is ongoing activity in the alpha band, briefly disrupted by the motor event at time zero, that well approximates the normalized ER. This ongoing alpha activity is more prominent in the plot of the band-pass fit, resulting in a negative value (blue region) in the difference plot for the alpha band. The event-related peaks in the band-pass fit were at −50 ms and +850 ms, in the low frequency range. These latencies are similar to the low frequency peaks in movement-related power modulation seen in the primary motor cortex for center-out arm movements (Ball et al., 2009), and here we refer these events as motor related potentials one and two. The values (and peak frequencies) of the band-pass fit and the phase-only fit at these times and frequencies are given in Table 2. The peaks in the band-pass fit were greater than the phase-only fit, but of similar magnitude, with only +0.6 to +0.7 additional correlation when the amplitude portion of the signal was included in the

calculation. This observation held for all task-related frequencies relevant to the event-related ECoG signal i.e. 0.5 Hz to 4.0 Hz.

Previous research has shown that the pattern of magnitudes in the trial-locked time-series is a function of both single-trial phase and amplitude (Barry, 2009; Fell, 2007; Min et al., 2007; Sauseng et al., 2007). Here we have examined this question as a function of spatial pattern of ERs over the whole measurement array, and over one temporal cycle. At all event-related peaks examined, the amplitude component added surprisingly little to the overall pattern of ER magnitudes. We conclude that considering ER magnitudes as a global topographic pattern emphasizes the role of phase, at the expense of amplitude. It is worth noting that the relative contribution of the amplitude was

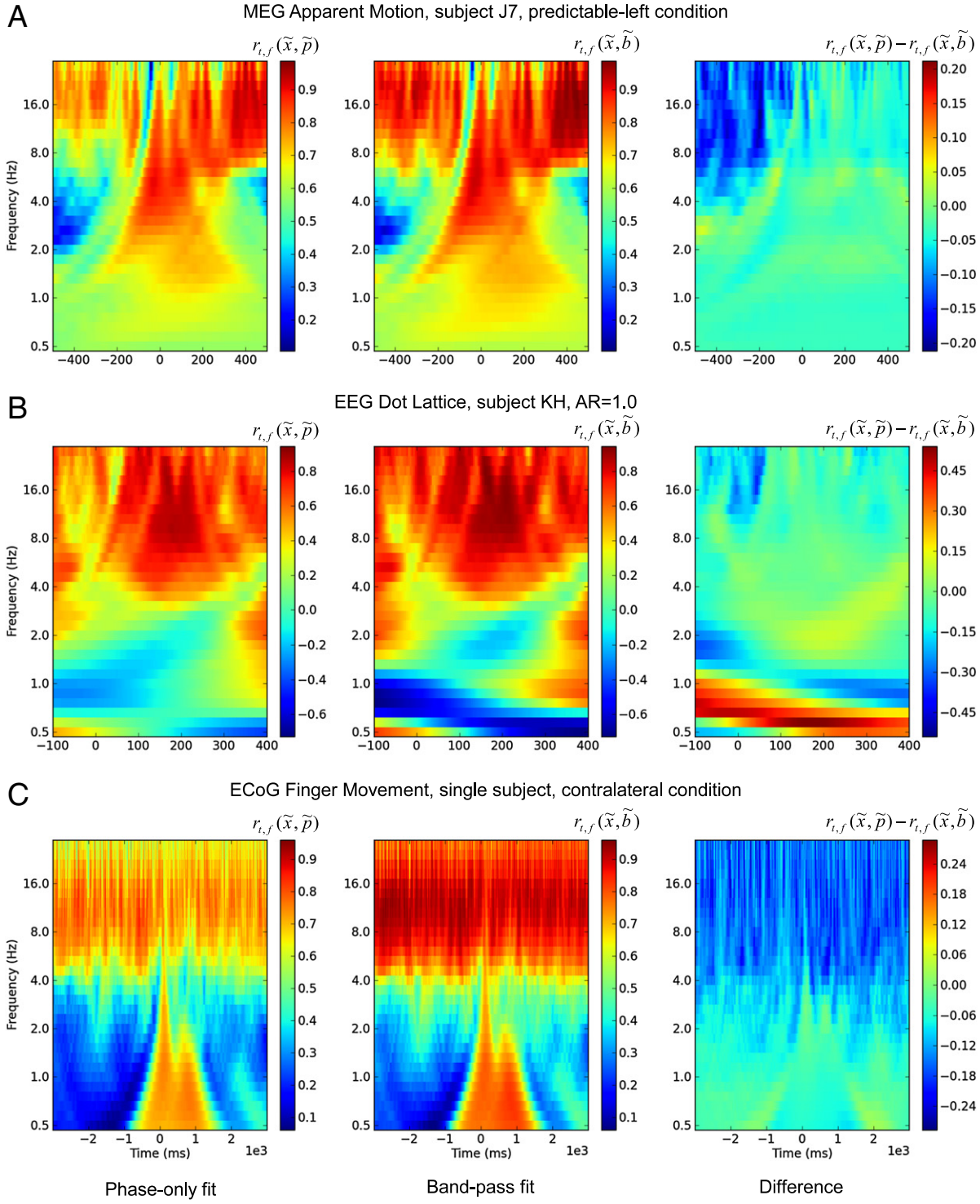


Fig. 2. Single subject correlations of normalized ER with either the phase-only ER or band-pass ER, over all times and frequencies. The first column shows the values of the phase-only fit, $r_{t,f}(\tilde{x}, \tilde{p})$. The second column shows the values of the band-pass fit, $r_{t,f}(\tilde{x}, \tilde{b})$. The third column shows the pixel-wise difference between the two correlations i.e. $r_{t,f}(\tilde{x}, \tilde{p}) - r_{t,f}(\tilde{x}, \tilde{b})$. Statistical analysis of this difference is not included, since only one difference value per subject, per condition is calculated. A. Results for one MEG subject, J7, predictable-left condition. B. Results for one EEG subject, HK, AR = 1.0. C. Results for the ECoG subject, for the contralateral movement condition.

greater for ongoing activity in the alpha and/or beta bands than it was for the event-related latencies, for all tasks and imaging modalities.

To confirm these findings, we also compared the topography of the inter-trial phase coherence, ITPC, with the topography inter-trial band coherence, ITBC. Likewise, we compared the spatial pattern of the trial-averaged amplitude, TAA, with the ITBC. The findings are shown in Supplementary Fig. 2, and support the previous analysis. The values of the ITPC fit and the TAA fit at event-related peaks in the band-pass fit are given in Table 2 and show ratios in the range 1.1 to 2.4. At almost all times and frequencies, the ITPC fit was better than the TAA fit. The exception occurred in the EEG data, where in the theta and delta bands the ITPC fit and the TAA fit did not significantly differ across subjects. These findings support our conclusion that both the single-trial phase and the single trial amplitudes contribute to the topography of the (trial-averaged) band-pass signal, but overall, the phase contributes more.

One point of note: there is less event-related structure (fewer peaks and troughs) in ITPC than in the phase-only ER because the former is based on the complex-valued phase, rather than the real part of the signal. This means that the spatial pattern of ITPC is approximated by the magnitude of the phase-only ER; so the former quantity looks similar to the head map shown in Fig. 1C (second column), but taking only the magnitudes and thus ignoring the sign.

Spatial pattern of evoked response is mostly due to phase

In summary, the overall pattern of ER magnitudes had a larger contribution from single-trial phase than from single-trial amplitude. This is true whether the single-trial amplitudes are considered as an additional component within the band-pass signal (i.e. the amplitude component of $\tilde{b}_{S\times C}$ in $r_{t,f}(\tilde{x},\tilde{b})$) or as a separate component (i.e. trial-averaged amplitude $\overline{A}$ in $r_{t,f}\left(\overline{B},\overline{A}\right)$). It is true across the three imaging modalities assessed here. The result is of immediate importance for our understanding of source localization techniques, since these techniques likewise use trial-averaged signals. The amplitude, by definition, is that part of the signal which contributes to any expected decrease in the signal strength as the inverse-square of distance from a localized source. Phase, by definition, has spatially uniform amplitude. The phase component can only contribute to ER magnitudes via selective summation and cancelation over trials. Since we conclude that the phase contributes more than the amplitude, it follows that the pattern of ER magnitudes is primarily a function of interference effects, in phase, across trials (c.f. Acacio de Barros and Suppes, 2009).

Measures of TW propagation often estimate the phase of the signal as a preliminary analytical step; whether via Fourier methods (Alexander et al., 2006b; Ito et al., 2005, 2007; Patten et al., 2012) or as peak and trough measurements in the time-domain (Klimesch et al., 2007; Manjarrez et al., 2007; Massimini et al., 2004). Phase estimates are therefore a sufficient basis to characterize TW behavior in large-scale measurements of cortical activity. Based on the results of this section, we may also propose that ER topography can be modeled as a trial-average of phase components of the signal, at appropriately chosen event-related times and frequencies. This provides us with two model signals to describe large-scale cortical activity, the trial-averaged phase and the wave map, which can both be expressed in the same units as the signal to be explained: the measured phase. Our interest was in how well these model signals describe within-trial activity. We directly compared the correlations of the trial-averaged phase and wave map to the measured phase, which we describe in the next sections.

The spatial pattern of trial-averaged phase

We next took the trial-average of the phase, considered as a topography over measurement sites, and asked: how well does it explain the spatial pattern of phase at single trials? Fig. 3A (right) shows a head-map of the measured phases in the MEG at 198 ms, 2.0 Hz for one trial. Fig. 3B (right) shows the average-phase model at this time and frequency. The patterns of measured phase and average-phase model are very similar; this was the reason for choosing this example, which is typical for high values of the average-phase fit. Fig. 3D (right) shows the correlation of measured phases, during this trial, to the average-phase model, over all times and frequencies. The fit rises and falls in an intermittent fashion at different times and different frequencies, indicating that sometimes the single trial phase matches the average-phase model, but often not.

We applied the measure average-phase fit to all subject's data. Fig. 4A (second column) shows the trial-average of the average-phase fit for another MEG subject. The grand-average plots for the average-phase fit in the MEG data are shown in Supplementary Fig. 3A (second column). The average-phase fit increases across a broad range of frequencies during the motor and stimulus period of the task. Fig. 4B (second column) shows the trial-average of the average-phase fit for one subject's EEG data. Grand-average plots for the EEG on this measure are shown in Supplementary Fig. 3A (second column). There is a maximum at about 150 ms, 6 Hz corresponding to latency of the P1 -N1 complex of the dot-lattice task. Fig. 4C (second column) shows the trial average of the average-phase fit for the ECoG subject's

Table 2
Values of the various measures at times and frequencies of peaks in the band-pass fit. The latency, and peak frequency at that latency, of the most prominent event-related peaks in band-pass fit are shown for the three measurement modalities/experimental paradigms. The names provided for each component correspond to components at the same latency in the originally published studies (Ball et al., 2009; Nikolaev et al., 2008; Zvyagintsev et al., 2008). The values for phase-only fit, band-pass fit, ITPC fit and TAA fit are subject-averages, except for ECoG Finger Movement where there was only one subject. The values for wave activity and average-phase fit are grand-averages, or simply a trial average for the ECoG Finger Movement. p-Values for these measures are from t-tests on the subject-wise averages (MEG n = 20, EEG n = 13), except in the case of ECoG Finger Movement, where p-values are from Mann–Whitney U-tests on individual trial values (n = 137).

	Time (ms)	Frequency (Hz)	Phase-only fit	Band-pass fit	Sig.	ITPC fit	TAA fit	Sig.	Wave activity	Average-phase fit	Sig.
MEG Apparent Motion											
Motor dipole	10	10.6	0.75	0.84	$p<0.001$	0.88	0.36	$p<0.001$	0.51	0.09	$p<0.001$
Auditory dipole	80	7.0	0.83	0.88	$p<0.001$	0.94	0.69	$p<0.001$	0.50	0.26	$p<0.001$
Visual dipole	150	12.1	0.85	0.90	$p<0.001$	0.93	0.45	$p<0.001$	0.49	0.09	$p<0.001$
EEG Dot Lattice											
P1 ERP	100	16.0	0.82	0.86	$p<0.001$	0.92	0.45	$p<0.001$	0.61	0.27	$p<0.001$
N1 ERP	152	8.0	0.89	0.91	$p<0.001$	0.87	0.60	$p<0.001$	0.66	0.40	$p<0.001$
ECoG Finger Movement											
Motor related potential 1	−50	0.66	0.73	0.79	n/a	0.97	0.92	n/a	0.34	0.26	$p<0.001$
Motor related potential 2	850	0.66	0.76	0.83	n/a	0.95	0.91	n/a	0.34	0.27	$p<0.001$

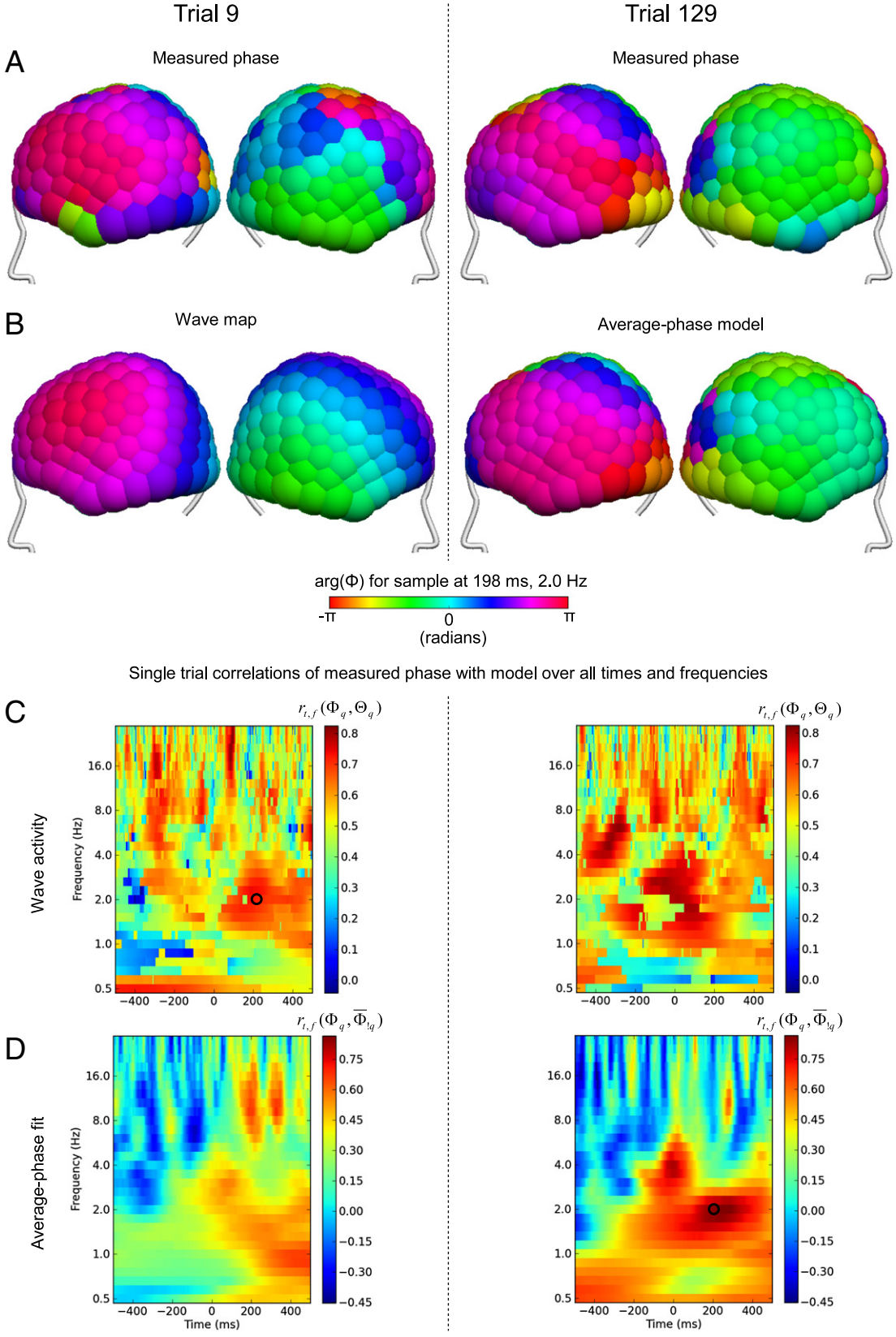


Fig. 3. Illustration of the methods for calculating wave activity and average-phase fit. Left side of the figure illustrates the method of fitting estimated phase from a single trial (#9), right side of figure is for another trial (#129). All data are from MEG subject O1, predictable-left condition. A. Head maps of measured phase at two different trials, at $t = 198$ ms, $f =$ 2.0 Hz. Color scale shows the phase angle. B. Model phase angle, for either the wave map for trial 9 (left) or the average-phase model for trial 129 (right). C. Correlations of the measured phase to the wave map, i.e. the wave activity, $r_{t,f}(\Phi_q, \Theta_q)$, for all times and frequencies in the two trials. In the left figure, the black circle indicates the time and frequency corresponding to the values shown in the head maps (A and B) on the left side of the figure. D. Correlations of the measured phase to the average-phase model, i.e. the average-phase fit, $r_{t,f}(\Phi_q, \overline{\Phi}_{!q})$, for all times and frequencies in the two trials. The black circle in the right figure indicates the time and frequency corresponding to the values shown in the head maps (A and B) on the right side of the figure.

data. There is a peak in the average-phase fit in the lower frequency range (0.5 to 1.0 Hz) at about 300 ms, related to the motor task. In all three measurement modalities then, it is during the event-related period that the spatial pattern of the trial-averaged phase best matches the spatial pattern of the single-trial phase, as might be expected. The values of the average-phase fit, at the time and frequency of the peaks in the band-pass fit, are given in Table 2.

Traveling wave model

Spatial variation of phase coherence has been attributed to spatial variation in the degree of order vs. randomness of phase at the different measurements sites (Ray and Maunsell, 2011). This view is derived from cross-trial measures of phase, and does not contradict the analyses presented in the previous sections. However, measurement of TWs at

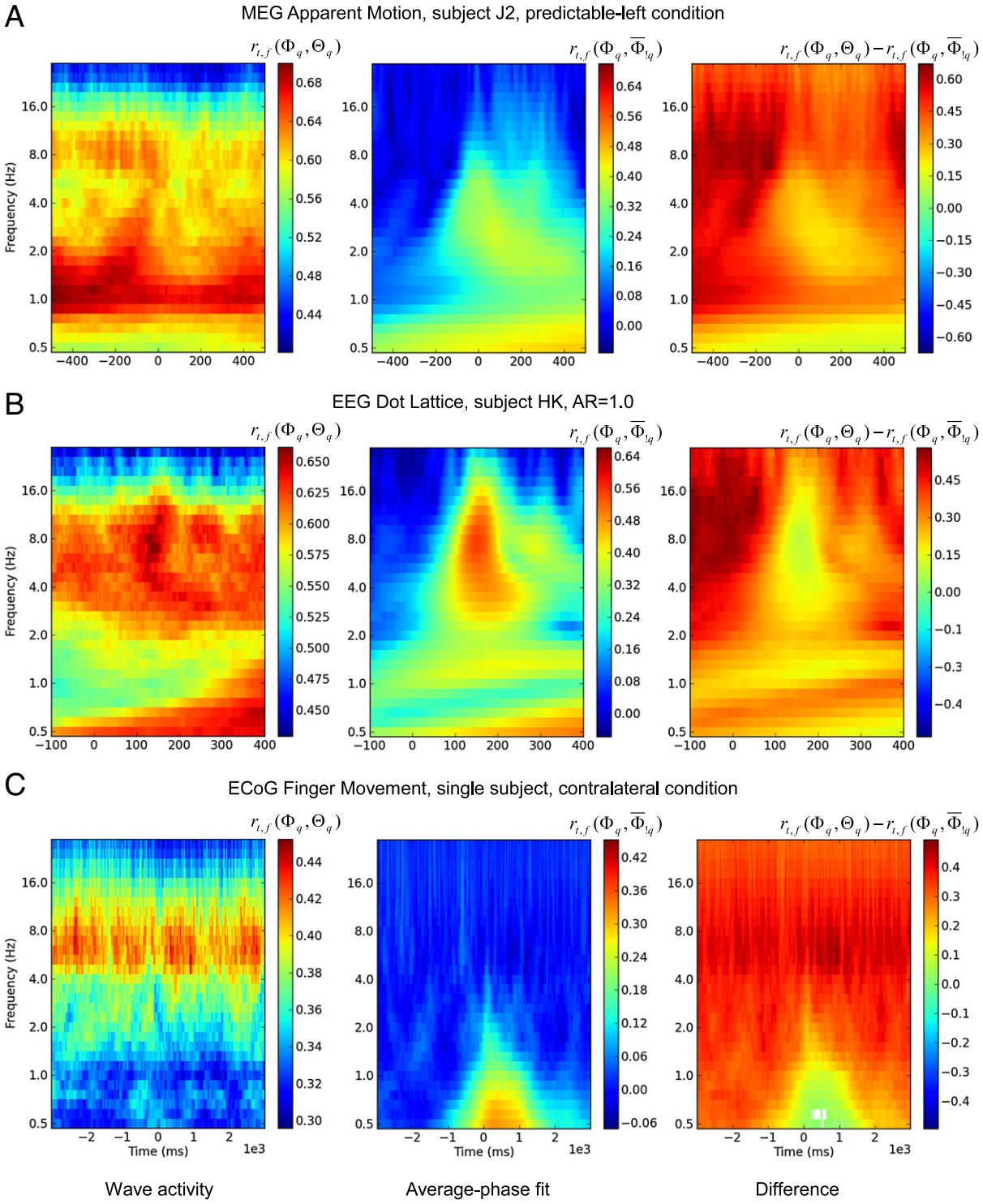


Fig. 4. Trial-averages of wave activity and average-phase fit. The first column shows the values of correlations of the measured phase with the wave map, i.e. wave activity, $r_{t,f}(\Phi_q,\Theta_q)$, averaged over trials. The second column shows the values of correlations of the measured phase with the average-phase model, i.e. the average-phase fit, $r_{t,f}(\Phi_q,\overline{\Phi}_{!q})$, averaged over trials. The first and second columns are not on the same color scale. The fourth column shows the difference between the mean wave activity and the mean average-phase fit, i.e. units of $r_{t,f}(\Phi_q,\Theta_q)-r_{t,f}(\Phi_q,\overline{\Phi}_{!q})$. Pixel-wise significance tests were carried out (Mann–Whitney U-test, $p<0.05$ shown in non-white). A. Results for MEG subject J2, predictable-left condition ($n=134$). B. Results for EEG subject HK, condition AR = 1.0 ($n=84$). C. Results for ECoG subject, contralateral movement condition ($n=137$).

the single trial level suggest that epochs of global phase gradients are common events in the cortex (Alexander et al., 2006b). With TWs, we see a globally coherent event across the measurement array, not a mixture of ordered and random phases depending on the site. These two views can be made consistent by noting that TW models collate signal across space, whereas phase coherence collates signals across trials or times. This means that phase coherence does not necessarily indicate the degree of order vs. randomness of phase, when phase is considered in terms of spatial organization within a single time-sample. In this section we show that episodes of large-scale phase gradients occur with ubiquity, including at event-related times and frequencies when there is spatial variation of (cross-trial) phase coherence.

In the present data, TWs were generally not apparent in the ERs i.e. in the trial-average. The normalized ER shown in Fig. 1 is typical in this respect. We observed that only for a few MEG subjects were TWs apparent in the normalized ER and in the phase-only ER. The situation was different at the single-trial level, however. We calculated the correlation of the TW model, the wave map, to the measured phase, giving the wave activity. By showing that the wave activity is generally greater than the average-phase fit, we establish that TWs exist at the single-trial level, and that the wave activity captures components of the signal that are not present in the trial-average phase.

Previous work on large-scale patterns in the EEG has shown that the phase vectors form spatial patterns that are often of long wavelength (Alexander et al., 2006b; Ito et al., 2007); here long-wavelength means with a spatial period approximately equal to the scale of the measurement array. These long wavelength patterns explained more than 50% of the variance in the spatially unwrapped phases, according to principal component analysis (Alexander et al., 2006b). The first three eigenvectors were well approximated by the coordinate system of the measurement array. That is, the long-wavelength eigenvectors consisted of an anterior–posterior component, an inferior–posterior component and a left–right component, or an arbitrary rotation of these three components. The principal component analysis was applied to spatially unwrapped phase in the present data sets, reproducing these basic conclusions. The spatial organization of the principal components is shown in Supplementary Fig. 4. The percentage of variance explained by the long wavelength components was 62% for MEG (first four eigenvectors), 69% for EEG (first four eigenvectors) and 56% for the ECoG data (first three eigenvectors). Each of these sets of eigenvectors was approximated by the coordinate axes of the measurement array, or by an arbitrary rotation of the array coordinates. Subsequent eigenvectors beyond these long wavelength ones were of shorter wavelength (period doubled or greater).

The weighted sum of these long-wavelength eigenvectors describes a general class of spatially smooth gradients in phase (see Materials and methods section). Smooth gradients of phase across the measurement array are equivalent to globally coherent TW behavior (Alexander et al., 2006b). A large proportion of the variance in unwrapped phases in the MEG, EEG and ECoG is therefore captured by TW components of the signal. These results from the principal component analysis mean that, at the phenomenological level, the TWs often have the appearance of a phase trajectory passing through the 3D coordinate system of the measurement array. Specifically, regions of equal phase, in this description of TWs, lie at the intersection of (a) the surface defined by the measurement array and (b) parallel planes through the volume defined by the measurement array (see Eq. (12), Materials and methods section). These bands of equal phase can be seen in the wave map illustrated in Fig. 3B (left). The wave activity measure in this research is a sensor-level measure of single-trial phase organization; here we are not directly concerned with the underlying sources that generate this spatio-temporal pattern of phases.

Examples of TW events detected by the present method are shown in Fig. 5. It can be seen from these examples that a high value of the wave activity—calculated from the spatial vector of phase—is indicative of a traveling wave that has a concomitant *latency* gradient across the head (Alexander et al., 2006b). This is indicated by the diagonal lines of equal phase in the time by site plots, third column. In this representation of the wave map, lines of equal phase in the traveling wave form a slightly curved diagonal because the 'site ordering' on the y-axis does not show the spatial coordinates in exactly equal units of linear distance along the wave trajectory. For the MEG traveling wave event highlighted in Fig. 5A, the wave trajectory begins slightly to the anterior-right of the superior site, and ends in the posterior-left-inferior region. This trajectory is indicated by the ordering of 'S', 'R', 'A', 'P' and 'L' along the y-axis. In the EEG traveling wave event highlighted in Fig. 5B, the wave trajectory is primarily posterior to anterior, but also with a slight right to left component. In the ECoG traveling wave highlighted in Fig. 5C, the wave trajectory is anterior-superior to inferior. For reference, the electrode positions A1, A8, H1 and H8 are shown in Supplementary Fig. 5, along with the results of the electrical stimulation mapping that was performed in the patient. An example wave in the 938MEG is also shown in Supplementary Video 2. Supplementary Videos 3 and 4 provide example TWs in the EEG and ECoG.

A general observation was made in the present data sets about the relationship between the wave trajectories and sites of the maximum ER magnitude. In general, the bulls-eye of the wave does not arise at the site of the maximum ER magnitude. In other words, the site of maximum activity does not appear as a 'source' for the waves, with phase latency increasing with distance from the source. Instead, the waves begin their motion at a more distal site, and progress onto the site of the maximum ER magnitude. For example, the waves in ECoG generally traversed the array from one edge to the other at the time of the motor response, but the motor related potential itself was maximal near the center of the array (see Supplementary Fig. 5). This observation from the present data sets is consistent with previous results in the EEG, for example: TWs associated with the P3b ERP are generally aligned along an anterior–posterior axis, while the P3b is maximal at centro-parietal sites (Alexander et al., 2009).

We also note in passing that common sources of noise in EEG recordings, such as muscle artifact and eye-movement artifact, do not conform to the pattern of activity that we detect with the TW measure. These TWs reveal a band of activation that traverses the array over one temporal cycle. Eye-movement artifacts, for example, have more the appearance of a standing wave with a localized peak around the orbits.

Fig. 3C (left) shows the typical time-course of the wave activity at different frequencies during a single trial in the MEG data. The wave activity has a peak at around 200 ms, 2 Hz, among other peaks. An important point to note is that these events are intermittent, as has been described for other types of single-trial phase coherency events (Eckhorn et al., 2004; Ito et al., 2007; Nikolaev et al., 2010). When viewed at the single trial level, there is no obvious relationship to event-related paradigms, much like single trial 'ERs'. Only when the wave activity values are collated together over trials do the time- and frequency-locked components become apparent. An example of trial-averaged wave activity for a single MEG single subject is shown in Fig. 4A (first column). The grand-average wave activity in the MEG is shown in Supplementary Fig. 3A (first column). The figure shows ongoing delta TWs, as well as ongoing alpha TWs that are partly disrupted by the subject's button press and subsequent stimulus events. Ongoing delta-band waves are less disrupted, and there is a peak in the delta band at about 500 ms, 1.7 Hz once the stimulus is complete.

Fig. 4B (first column) shows a typical EEG subject's trial-averaged wave activity. The grand-average for the wave activity in the EEG is shown in Supplementary Fig. 3B (first column). The wave activity is strongest in the alpha and theta bands, and there is an event-related peak at about 150 ms, 6 Hz corresponding to the latency of the N1 ERP of the dot-lattice task. This peak is very similar in timing and frequency to the peak in the average-phase fit for the EEG data. The ECoG subject's trial-average wave activity is shown in Fig. 4C (first column). The dominant feature is TWs in the alpha band. There is a lack of an

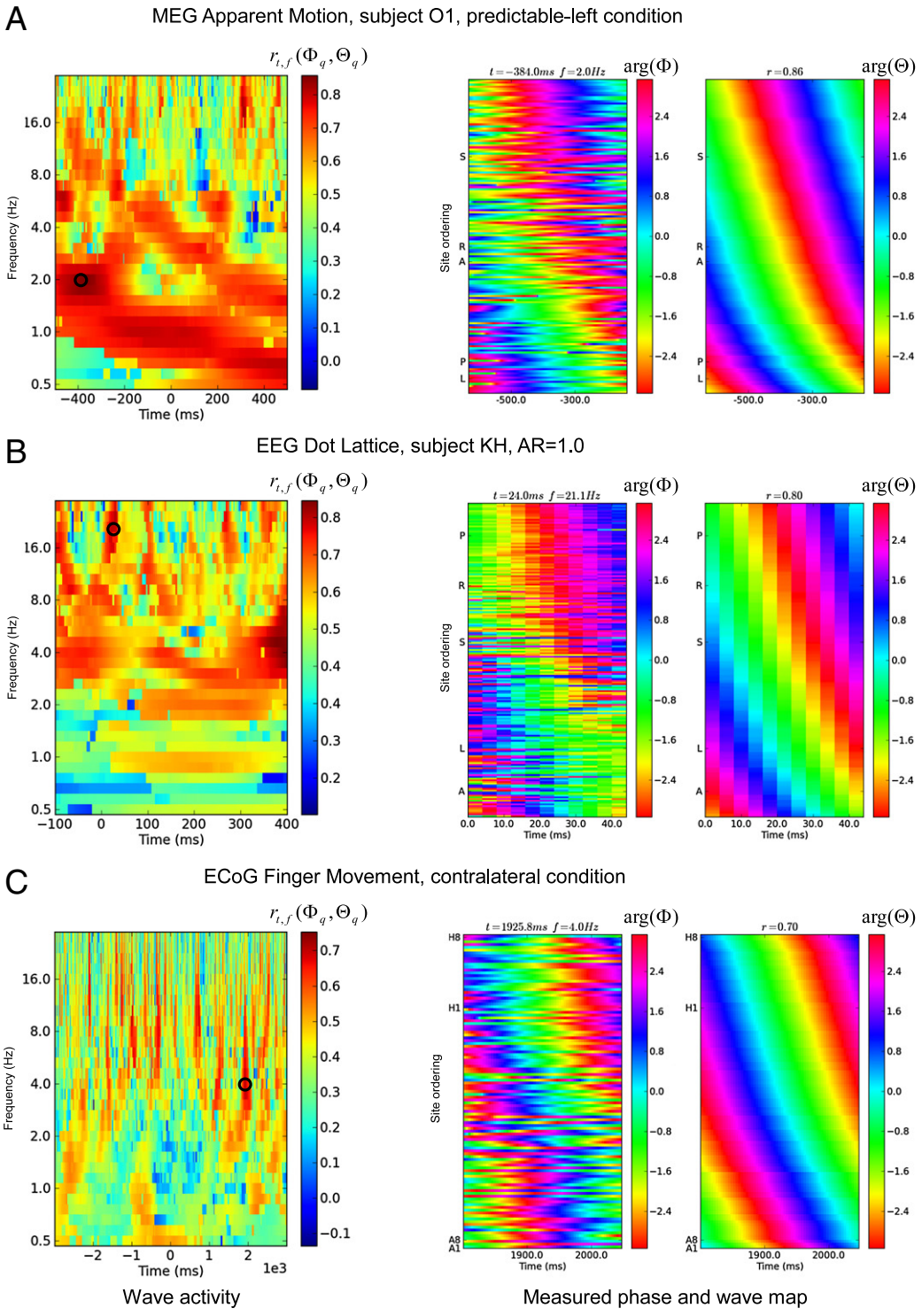


Fig. 5. Single trial examples of correlations of empirical phase to the wave map, the wave activity, $r_{t,f}(\Phi_q,\Theta_q)$. The left half of the figure shows one example trial from each of the data sets. The right half of the figure shows an example traveling wave from a peak in wave activity in the trial. Both the empirical phase angle, arg(Φ) (middle), and the wave map phase angle, arg(Θ) (right), are shown. Note that the wave map is calculated for the empirical phase from the central sample in the time-window, not over the entire window. The full cycle of the empirical phase (and extrapolation of the wave map, as $\arg(\theta_{S\times C_f})$) are included for interpretive convenience. A. MEG: A peak in wave activity occurs at 384 ms, 2.0 Hz, indicated by the black circle in the first plot. When represented in time by site plots (second and third plots), it can be seen that the wave trajectory is from the superior site, to the right and anterior sites, then the posterior and left sites. The sensor/electrode labels are as given in Fig. 1. B. EEG: A peak in wave activity occurs at 24 ms, 21.1 Hz, indicated by the black circle in the first plot. In the time by site plots (second and third plots), it can be seen that the wave intersects with the measurement array in the following order: posterior, right, superior, left, anterior. The sensor/electrode labels are as given in Fig. 1. C. ECoG: A peak in wave activity occurs at 1926 ms, 4.0 Hz, indicated by the black circle in the first plot. In the time by site plots (second and third plots), it can be seen that the wave travels from posterior-ventral, before reaching posterior-dorsal and then ending anteriorly. The labeled electrodes are on the corners of the array (see Ball et al., 2009, Fig. 3 and Supplementary Fig. 5), and A1 is anterior-dorsal, A8 is anterior-ventral, H1 is posterior-dorsal and H8 is posterior-ventral.

event-related peak in the wave activity associated with the motor activity. The values of the wave activity for all measurement modalities, at the time and frequency of peaks in the band-pass fit, are given in Table 2.

Trial-averages and traveling waves

For the EEG signal in the dot-lattice task, trial-average and TW measures yielded similar timing of event-related peaks. This result supports previous findings of relationships between the characteristics of event-related wave activity and ERP components — in latency, temporal frequency and task-specificity (Alexander et al., 2006b; Anderer et al., 1996; Klimesch et al., 2007). For the MEG signal in the Apparent Motion task, event-related peaks in the average-phase fit were coincident with event-related drops in the wave activity, such as during the auditory and visual evoked activity at +80 and +150 ms, respectively. In the ECoG, the wave activity was sensitive to ongoing alpha activity, but the average-phase fit was not. The obverse relationship held for the activity time-locked to the finger movement. Even so, a lack of an event-related peak in the wave activity does not imply that this TW component was small. The wave activity for ECoG subject was greater than the average-phase fit at times and frequencies of both peaks in the band-pass fit (Table 2). Indeed, the mean values for the wave activity in the ECoG subject were larger than the values for the average-phase fit at almost all times and frequencies (Fig. 4C, third plot). A few pixel-wise comparisons near the event-related peak in the average-phase fit did not reach significance. This was also true for the grand-average values in the MEG and EEG subjects (Supplementary Fig. 3, third column). These observations mean that, in general, the TW model explained more of the single-trial phase than did the trial-averaged model.

A small caution has to be used in interpreting this result, since the wave activity and the average-phase fit are calculated from measured phase using different methods with differing degrees of freedom. The simplicity of the two models and lack of free parameters renders overfitting of the data unlikely. As an additional check on this conclusion, we recalculated the results shown in the first column of Fig. 4 using randomized phases at each sample i.e. spatial randomization of the measured phase. This check was required because the calculation of the wave activity is a two-step process, in which the data is first fit to a model and then the correlation between the model and the data is calculated. For the ECoG subject, for example, this surrogate mean (of 137 trials) wave activity was calculated 46,050 times (at 1535 samples and 30 frequencies). The range of the mean wave activity values produced from randomized phases in the surrogate data, for the three subjects given in Fig. 4, was 0.07 to 0.13. This compares with the mean wave activity values in Fig. 4, first column, at event-related peaks in the mean wave activity, of 0.45 to 0.70. The fitting procedure does not account for the observed high values of the wave activity.

We conclude that for all our data sets, the wave activity explains more variance in the measured phase than does the average-phase fit. We suggest that the trial-averaging of phase removes much of the spatially organized component that is present in the single trial phase. Trial-averaging results in summation over cyclic components of the signal, i.e. the phase, and can account for much of the observed spatial variation in ER magnitude, as we concluded at the end of the previous section. To this interim conclusion we now add the following: the observation that cross-trial phase locking is high at some measurements sites and low at others is entirely consistent with TWs occurring intermittently but with ubiquity at the single trial level. In other words, cross-trial phase locking can vary from site to site, but this pattern co-exists with spatially smooth—ordered—phase gradients at the single-trial level.

Since TWs are ubiquitous at the single trial level, we were interested in exploring the relationship between the TWs and cross-trial measures such as the ER. Can the topography of the ER be explained as the average of TW components? The fit of the normalized ER to the wave ER is shown in Fig. 6, for one subject each in the MEG, EEG and ECoG. Comparison of these wave fits to the band-pass fits shown in Fig. 2 (first column), shows the qualitative similarity of the two measures. The wave fit is highest at times and frequency of peaks in the band-pass fit. This is confirmed in the grand-average plots for the MEG and EEG subjects, shown in Supplementary Fig. 6. In the MEG data, the initial theta-band peak in the wave fit is related to the motor dipole and the alpha band peak at ~+250 ms is consistent with the timing and frequency of late visual and auditory evoked components in response to the completion of the stimulus (Zvyagintsev et al., 2008). In the EEG data, there is a peak in the wave fit at the time and frequency associated with the P1 and N1 ERPs. Here the P1 component of the ER is relatively enhanced (c.f. Fig. 2B, first column). The main features in the wave fit in the ECoG data are the low-frequency peaks related to the event-related motor potentials (Ball et al., 2009) and peak in the alpha band at ~−600 ms whose significance is unknown. These peaks in the wave fit explained a maximum of 47, 59 and 37% of the variance in the normalized ER, for MEG, EEG and ECoG, respectively, for the typical subjects shown in Fig. 6. The wave fit measure was not sensitive to the ongoing alpha and beta band activity present in the band-pass fit measure. So while the wave ER explained less variance in the evoked signal than the band-pass fit or the phase-only fit, it revealed a relative enhancement of time-locked components of the signal.

Using the measures band-pass fit, phase-only fit and wave fit, we have shown that successively more of the original signal present in the ER can be thrown away, while still preserving the topographic pattern of activity seen for some ERs. The band-pass fit only makes use of signal from a specific center-frequency, ignoring other frequency components. The phase-only fit does not make use of amplitude information, instead, reconstructs the ER magnitude topography purely as an interference pattern in phase. The wave fit represents one temporal cycle of signal from each trial with four parameters, giving the trajectory of the wave in the three dimensions of the measurement array plus the phase offset of the wave. When averaged across trials, in appropriate form for comparison to the normalized ER, the TWs pick out a qualitatively similar set of event-related features in the normalized ER to those seen when using the band-pass ER as a comparison signal.

Discussion

Summary and conclusions

Two related issues were addressed in this paper. The first concerned the relative contributions of phase and amplitude to ERs, where the latter is considered as a pattern of magnitudes over all sites and over one temporal cycle at the frequency of interest. Quantitative methods in brain science often make use of large-scale patterns of activity for source localization, network analysis of phase coherence, or detection of TWs; this motivated our present approach. We found the magnitude pattern of ERs to be a function of phase, more than amplitude. Task-related changes to ER topography were not primarily the result of location-specific amplification of brain activity, although this did play a role as well. The primary process was a shift in the way phases at different locations constructively and destructively interfere when summed over trials.

From the conclusion that phase plays a primary role in the topography of ER magnitudes, the second issue arose: what is lost from the phase signal when we average over trials? Previous research has demonstrated the existence of globally organized TWs within single trials (Alexander et al., 2006b, 2009). This phenomenon seemed, however, to be inconsistent with attributing low ER magnitudes to low amplitude signal or to random phase. We showed that at time/frequency peaks where the spatial pattern of phase coherence was closely related to the ER, TW episodes were ubiquitous at the single trial level, though

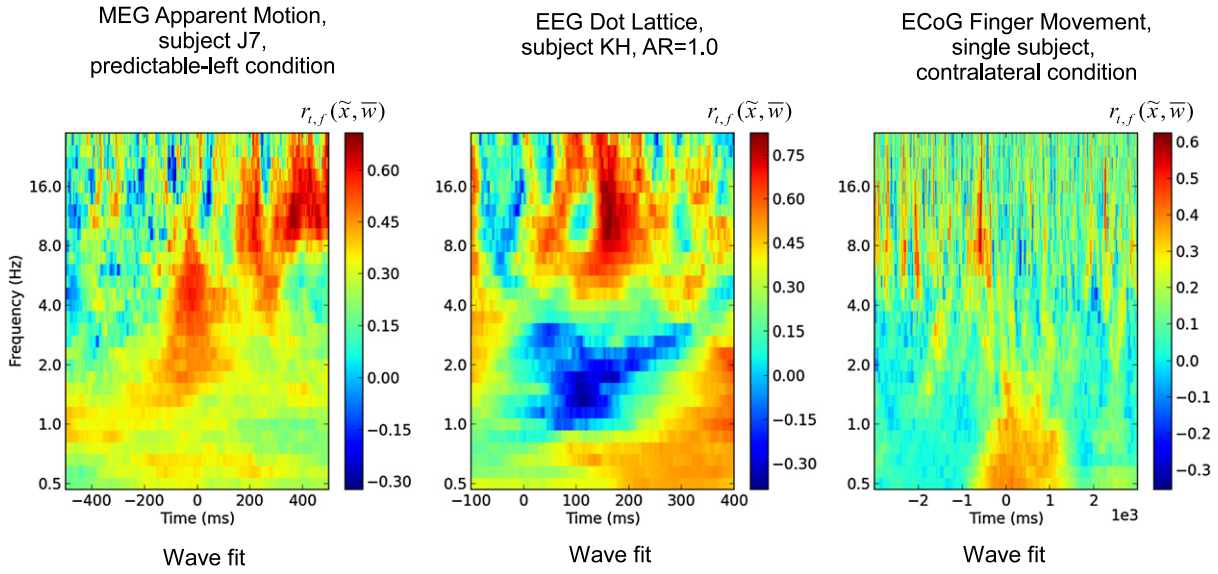


Fig. 6. Single subject correlations of normalized ER with the wave ER i.e. the wave fit $r_{tf}(\tilde{x},\overline{w})$, over all times and frequencies. The left panel shows the results for one MEG subject, J7, predictable-left condition. The middle panel shows the results for one EEG subject, HK, AR = 1.0. The right panel shows the results for the ECoG subject, for the contralateral movement condition.

intermittent. Further, the TWs accounted for more of the variance in the within-trial phase of the signal better than did trial-averaged phase at task-relevant time and frequencies, and indeed at most times and frequencies. This suggests that information about phase organization between spatially distant sites is destroyed when cross-trial averages are calculated.

We also showed that the topography of the ER was well modeled as the trial-average of traveling wave components. This, even though TWs were not generally apparent in the trial-averaged topographies for the present data sets.

The results reported here were consistent across imaging modalities with differing spatial resolutions. This is important because large-scale patterns of activity in the EEG are distorted by blurring of the signal via volume conduction effects. For this reason we confirmed importance of phase to the ER topography and confirmed the ubiquity of TWs using two additional measurement modalities: MEG and ECoG. Similarly, the MEG results allow us to be confident that the EEG and ECoG results do not depend on our choice of average reference. The inclusion of three measurement modalities also allows us to rule out eye-movement and muscle artifacts as the cause of the ubiquity of the TWs, since each modality is differentially prone to these issues. Likewise, the broad range of frequencies at which we found TWs means that any particular non-cortical explanation (i.e. beta-band muscle artifact, delta-band eye-movement artifact) cannot account for the ubiquity of the TWs.

It is commonly assumed that signal not captured by the topography of the event-locked response can be treated as random background noise (Arieli et al., 1996; Gruber et al., 2005). Ray and Maunsell (2011, p. 12676) state this assumption with clarity: "If the oscillations are perfectly uniform and synchronized across cortex, the amplitude should not decrease with distance. However, in our data the phase consistency … [was] … more random at larger interelectrode distances, leading to more reduction in the overall magnitude." This account of phase consistency takes "more random" to be equivalent to less "perfectly uniform and synchronized across cortex". However, there are several ways for phase consistency to diverge from perfectly uniform across the cortex.

One possibility is that there are a variety of TW components in the signal and the observed low phase consistency is due to phase cancelation in the trial-average. Previous studies using single-trial measures of TWs have shown that the trajectories of the waves have a distribution with several modes (Alexander et al., 2009; Massimini et al., 2004; Rubino et al., 2006; Takahashi et al., 2011). That is, the traveling waves can move in a variety of directions. When several modes are present, trajectories with opposing directions will summate in the manner of an interference pattern. The result is that the traveling wave information is lost, although the residue—the ER, ITPC or phase consistency—has a maximum that is localized in space and time. The magnitude of this event-locked response decreases with increasing distance from the site of maximum, just as in the case of phase consistency described by Ray and Maunsell (2011). This can be easily demonstrated by making an analogy with Moiré patterns—the component images (analogous to individual trials) of these interference patterns do not have the same amplitude patterns as the Moiré pattern itself (analogous to the inter-trial measure). Supplementary Video 5 shows how four TWs, each with different trajectories, can summate to give the spatio-temporal pattern typical of ERs.

For some cognitive tasks/states, one TW mode can be dominant, for example anterior to posterior waves in deep sleep (Massimini et al., 2004). TWs are also observed in cross-trial measures during rest (Nolte et al., 2008), in local field potentials during low contrast visual stimulation (Nauhaus et al., 2009) and perceptual interference tasks such as Stroop (Klimesch et al., 2007). These observations are consistent across a range of methods i.e. for mean delay map, phase slope, spike-triggered average or filtered ERP topography. A case in point is the latency topography of the P3b ERP, which varies across age groups (Anderer et al., 1996). There is a posterior–anterior latency gradient in early adolescence but an anterior–posterior latency gradient in older adults (Alexander et al., 2006b, 2008). However, neither mode is dominant during late adolescence. This effect has been proposed to explain the absence of a clear P3b latency gradient for this age group (Alexander et al., 2009). In the present research, we generalize this observation by showing the ubiquity of TWs under and variety of measurement and task conditions. TWs are common-place within trials whether or not they are apparent in the trial-average. We further suggest that averaging over trials destroys functional information about the sequence of activations in moment-to-moment cognition.

Our conclusions about trial-averaging extend, in principle, to measures such as pair-wise phase coherence. Pair-wise coherence measures assume that the mean phase-offset—over trials or over several seconds of temporally contiguous data—reflects functional connectivity. However, TW events are usually short-lasting (Alexander et al., 2006b; Gabriel and Eckhorn, 2003; Massimini et al., 2004), intermittent (Eckhorn et al.,

2004; Gong et al., 2007; Ito et al., 2007; Nikolaev et al., 2010) and can have a variety of trajectories (Alexander et al., 2009; Manjarrez et al., 2007; Massimini et al., 2004; Rubino et al., 2006; Takahashi et al., 2011). Under these conditions, mean phase-offsets will underestimate the scope of coherent TW events due to destructive interference across trials.

Outlook

Recent findings have shown network coupling effects via endogenous fields in cortical slices (Anastassiou et al., 2011; Fröhlich and McCormick, 2010). The fields play a causal role in network activity, but any causal effects from these fields arise in real-time, without access to cross-trial information. It is therefore timely to uncover differences and correspondences between field properties captured by single-trial and inter-trial quantities. In vivo, the measured fields reflect mass activity in the cortex; the rise and fall of activity across millions of neurons in an organized pattern.

What role might TWs play in the moment-by-moment organization of cognition? In short, they 'carry' the cortical activation to task-relevant sites in time for the ER. If we accept that ERs arise, in the main, as a cross-trial interference pattern in phase, this implies that ERs reflect the consistency in timing of brain activity, at particular sites, across trials. Specifically, the results from the present data sets show that the topography of ER magnitudes largely reflects the topography of phase consistency across trials. This observation is not surprising and is consistent with the phase-resetting model of ERP genesis. However, the topography of phase consistency across trials does not convey the same information as the topography of phase organization within trials; this was the second observation of the present research.

A corollary of these observations is that within-trial TWs carry a peak in activation (or inhibition) toward the sites of high cross-trial phase consistency. The wave arrives just in time to produce the pattern embodied in the cross-trial signal as the ER. This implies that the observed phase-resetting occurs in the form of a coherent TW event passing over the site of the maximum ER magnitude. A promising line of future research would be to show that the TW activations directly preceding the latency of the localized phase-reset event are also functionally relevant to the ER or predict task performance. Given the intermittent nature of TW activity, future research also needs to demarcate the experimental conditions under which TWs arise on a moment-by-moment basis, and when they are absent.

In the context of visual and motor cortex, TWs have been proposed as a mechanism for integrating spatially disparate signals (Rubino et al., 2006; Sato et al., 2012). We have also hypothesized that TWs interact with the functional mapping within cortical areas to enable rapid generalization of cortical computations (Alexander et al., 2011). The proposed generalization mechanism works by broadcasting invariant components of the cortical signal, using TWs, across the functionally varying cortical topography. At the at 10 cm scale and larger, TWs may also act as a mechanism to integrate cognitive events across the entire cerebral cortex. Alterations in TW patterns could provide a marker for cognitive dysfunction, and such has been observed for ADHD and schizophrenia (Alexander et al., 2008, 2009). Some recent and intriguing work on the role of cortical phase in perception has shown that visual thresholds can be altered by the ongoing phase dynamics (Busch et al., 2009; Mathewson et al., 2009; VanRullen et al., 2011). This suggests a general mechanism by which TWs function, allowing distal cortical sites to alter the thresholds within task relevant areas. Since the waves are spatio-temporal, the distal sites are able to alter thresholds in the near-future, providing a natural account of sequencing of activations across the cortical medium.

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.neuroimage.2013.01.016.

Acknowledgments

DMA, CT, ARN and CvL were supported by an Odysseus grant from the Flemish Organization for Science, FWO.

References

Acacio de Barros, J., Suppes, P., 2009. Quantum mechanics, interference, and the brain. J. Math. Psychol. 53, 306–313.
Alexander, D.M., Arns, M.W., Paul, R.H., Rowe, D.L., Cooper, N., Esser, A.H., Fallahpour, K., Stephan, B.C.M., Heesen, E., Breteler, R., Williams, L.M., Gordon, E., 2006a. EEG markers for cognitive decline in elderly subjects with subjective memory complaints. J. Integr. Neurosci. 5, 49–74.
Alexander, D.M., Trengove, C., Wright, J.J., Boord, P.R., Gordon, E., 2006b. Measurement of phase gradients in the EEG. J. Neurosci. Methods 156, 111–128.
Alexander, D.M., Williams, L.M., Gatt, J.M., Dobson-Stone, C., Kuan, S.A., Todd, E.G., Schofield, P.R., Cooper, N.J., Gordon, E., 2007. The contribution of apolipoprotein E alleles on cognitive performance and dynamic neural activity over six decades. Biol. Psychol. 75, 229–238.
Alexander, D.M., Hermens, D.F., Keage, H.A.D., Clark, C.R., Williams, L.M., Kohn, M.R., Clarke, S.D., Lamb, C., Gordon, E., 2008. Event-related wave activity in the EEG provides new marker of ADHD. Clin. Neurophysiol. 119, 163–179.
Alexander, D.M., Flynn, G.J., Wong, W., Whitford, T.J., Harris, A.W.F., Galletly, C.A., Silverstein, S.M., 2009. Spatio-temporal EEG waves in first episode schizophrenia. Clin. Neurophysiol. 120, 1667–1682.
Alexander, D.M., Trengove, C., Sheridan, P.E., van Leeuwen, C., 2011. Generalization of learning by synchronous waves: from perceptual organization to invariant organization. Cogn. Neurodyn. 5, 113–132.
Anastassiou, C.A., Perin, R., Markram, H., Koch, C., 2011. Ephaptic coupling of cortical neurons. Nat. Neurosci. 14, 217–223.
Anderer, P., Semlitsch, H.V., Saletu, B., 1996. Multichannel auditory event-related brain potentials: effects of normal aging on the scalp distribution of N1, P2, N2 and P300 latencies and amplitudes. Electroencephalogr. Clin. Neurophysiol. 99, 458–472.
Arieli, A., Sterkin, A., Grinvald, A., Aertsen, A., 1996. Dynamics of ongoing activity: explanation of the large variability in evoked cortical responses. Science 273, 1868–1871.
Ball, T., Schulze-Bonhage, A., Aertsen, A., Mehring, C., 2009. Differential representation of arm movement direction in relation to cortical anatomy and function. J. Neural Eng. 6, 016006.
Barry, R.J., 2009. Evoked activity and EEG phase resetting in the genesis of auditory Go/NoGo ERPs. Biol. Psychol. 80, 292–299.
Barry, R.J., Rushby, J.A., Johnstone, S.J., Clarke, A.R., Croft, R.J., Lawrence, C.A., 2004. Event-related potentials in the auditory oddball as a function of EEG alpha phase at stimulus onset. Clin. Neurophysiol. 115, 2593–2601.
Britten, K.H., Shadlen, M.N., Newsome, W.T., Movshon, J.A., 1992. The analysis of visual motion: a comparison of neuronal and psychophysical performance. J. Neurosci. 12, 4745–4765.
Bullock, T.H., McClune, M.C., Achimowicz, J.Z., Iragui-Madoz, V.J., Duckrow, R.B., Spencer, S.S., 1995. EEG coherence has structure in the millimeter domain: subdural and hippocampal recordings from epileptic patients. Electroencephalogr. Clin. Neurophysiol. 95, 161–177.
Busch, N.A., Dubois, J., VanRullen, R., 2009. The phase of ongoing EEG oscillations predicts visual perception. J. Neurosci. 29, 7869–7876.
Delorme, A., Makeig, S., 2004. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J. Neurosci. Methods 134, 9–21.
Eckhorn, R., Gail, A., Bruns, A., Gabriel, A., Al-Shaikhli, B., Saam, M., 2004. Neural mechanisms of visual associative processing. Acta Neurobiol. Exp. (Wars) 64, 239–252.
Fell, J., 2007. Cognitive neurophysiology: beyond averaging. NeuroImage 37, 1069–1072.
Fellinger, R., Gruber, W., Zauner, A., Freunberger, R., Klimesch, W., 2012. Evoked traveling alpha waves predict visual-semantic categorization-speed. NeuroImage 59, 3379–3388.
Freeman, W.J., Barrie, J.M., 2000. Analysis of spatial patterns of phase in neocortical gamma EEGs in rabbit. J. Neurophysiol. 84, 1266–1278.
Friston, K.J., Frith, C.D., Liddle, P.F., Frackowiak, R.S., 1991. Comparing functional (PET) images: the assessment of significant change. J. Cereb. Blood Flow Metab. 11, 690–699.
Friston, K.J., Holmes, A., Poline, J.B., Price, C.J., Frith, C.D., 1996. Detecting activations in PET and fMRI: levels of inference and power. NeuroImage 4, 223–235.
Fröhlich, F., McCormick, D.A., 2010. Endogenous electric fields may guide neocortical network activity. Neuron 67, 129–143.
Gabriel, A., Eckhorn, R., 2003. A multi-channel correlation method detects traveling γ-waves in monkey visual cortex. J. Neurosci. Methods 131, 171–184.
Gong, P., Nikolaev, A.R., van Leeuwen, C., 2007. Intermittent dynamics underlying the intrinsic fluctuations of the collective synchronization patterns in electrocortical activity. Phys. Rev. E 76, 011904.
Gruber, W.R., Klimesch, W., Sauseng, P., Doppelmayr, M., 2005. Alpha phase synchronization predicts P1 and N1 latency and amplitude size. Cereb. Cortex 15, 371–377.
Herrmann, C.S., Grigutsch, M., Busch, N.A., 2005. EEG Oscillations and Wavelet Analysis. MIT Press, Cambridge, MA.
Hipp, J.F., Engel, A.K., Siegel, M., 2011. Oscillatory synchronization in large-scale cortical networks predicts perception. Neuron 69, 387–396.
Ito, J., Nikolaev, A.R., van Leeuwen, C., 2005. Spatial and temporal structure of phase synchronization of spontaneous alpha EEG activity. Biol. Cybern. 92, 54–60.

Ito, J., Nikolaev, A.R., van Leeuwen, C., 2007. Dynamics of spontaneous transitions between global brain states. Hum. Brain Mapp. 28, 904–913.
Klimesch, W., Hanslmayr, S., Sauseng, P., Gruber, W.R., Doppelmayr, M., 2007. P1 and traveling alpha waves: evidence for evoked oscillations. J. Neurophysiol. 97, 1311–1318.
Kubovy, M., 1994. The perceptual organization of dot lattices. Psychon. Bull. Rev. 1, 182–190.
Lachaux, J.P., Rodriguez, E., Martinerie, J., Varela, F.J., 1999. Measuring phase synchrony in brain signals. Hum. Brain Mapp. 8, 194–208.
Liu, L., Ioannides, A.A., 1996. A correlation study of averaged and single trial MEG signals: the average describes multiple histories each in a different set of single trials. Brain Topogr. 8, 385–396.
Liu, A.K., Belliveau, J.W., Dale, A.M., 1998. Spatiotemporal imaging of human brain activity using functional MRI constrained magnetoencephalography data: Monte Carlo simulations. Proc. Natl. Acad. Sci. U. S. A. 95, 8945–8950.
Makeig, S., Westerfield, M., Jung, T.-P., Enghoff, S., Townsend, J., Courchesne, E., Sejnowski, T.J., 2002. Dynamic brain sources of visual evoked responses. Science 295, 690–694.
Makeig, S., Debener, S., Onton, J., Delorme, A., 2004. Mining event-related brain dynamics. Trends Cogn. Sci. (Regul. Ed.) 8, 204–210.
Mäkinen, V., Tiitinen, H., May, P., 2005. Auditory event-related responses are generated independently of ongoing brain activity. NeuroImage 24, 961–968.
Manjarrez, E., Vázquez, M., Flores, A., 2007. Computing the center of mass for traveling alpha waves in the human brain. Brain Res. 1145, 239–247.
Massimini, M., Huber, R., Ferrarelli, F., Hill, S., Tononi, G., 2004. The sleep slow oscillation as a traveling wave. J. Neurosci. 24, 6862–6870.
Mathewson, K.E., Gratton, G., Fabiani, M., Beck, D.M., Ro, T., 2009. To see or not to see: prestimulus α phase predicts visual awareness. J. Neurosci. 29, 2725–2732.
Min, B.-K., Busch, N.A., Debener, S., Kranczioch, C., Hanslmayr, S., Engel, A.K., Herrmann, C.S., 2007. The best of both worlds: phase-reset of human EEG alpha activity and additive power contribute to ERP generation. Int. J. Psychophysiol. 65, 58–68.
Nauhaus, I., Busse, L., Carandini, M., Ringach, D.L., 2009. Stimulus contrast modulates functional connectivity in visual cortex. Nat. Neurosci. 12, 70–76.
Nauhaus, I., Busse, L., Ringach, D.L., Carandini, M., 2012. Robustness of traveling waves in ongoing activity of visual cortex. J. Neurosci. 32, 3088–3094.
Nikolaev, A.R., Gepshtein, S., Kubovy, M., Leeuwen, C., 2008. Dissociation of early evoked cortical activity in perceptual grouping. Exp. Brain Res. 186, 107–122.
Nikolaev, A.R., Gepshtein, S., Gong, P., van Leeuwen, C., 2010. Duration of coherence intervals in electrical brain activity in perceptual organization. Cereb. Cortex 20, 365–382.
Nolte, G., Ziehe, A., Nikulin, V.V., Schlögl, A., Krämer, N., Brismar, T., Müller, K.-R., 2008. Robustly estimating the flow direction of information in complex physical systems. Phys. Rev. Lett. 100, 234101.
Nunez, P.L., Srinivasan, R., 2006. A theoretical basis for standing and traveling brain waves measured with human EEG with implications for an integrated consciousness. Clin. Neurophysiol. 117, 2424–2435.
Pascual-Marqui, R.D., Michel, C.M., Lehmann, D., 1994. Low resolution electromagnetic tomography: a new method for localizing electrical activity in the brain. Int. J. Psychophysiol. 18, 49–65.
Patten, T.M., Rennie, C.J., Robinson, P.A., Gong, P., 2012. Human cortical traveling waves: dynamical properties and correlations with responses. PLoS One 7, e38392.
Prechtl, J.C., Cohen, L.B., Pesaran, B., Mitra, P.P., Kleinfeld, D., 1997. Visual stimuli induce waves of electrical activity in turtle cortex. PNAS 94, 7621–7626.
Ray, S., Maunsell, J.H.R., 2011. Network rhythms influence the relationship between spike-triggered local field potential and functional connectivity. J. Neurosci. 31, 12674–12682.
Ribary, U., Ioannides, A.A., Singh, K.D., Hasson, R., Bolton, J.P., Lado, F., Mogilner, A., Llinás, R., 1991. Magnetic field tomography of coherent thalamocortical 40-Hz oscillations in humans. Proc. Natl. Acad. Sci. U. S. A. 88, 11037–11041.
Rubino, D., Robbins, K.A., Hatsopoulos, N.G., 2006. Propagating waves mediate information transfer in the motor cortex. Nat. Neurosci. 9, 1549–1557.
Sato, T.K., Nauhaus, I., Carandini, M., 2012. Traveling waves in visual cortex. Neuron 75, 218–229.
Sauseng, P., Klimesch, W., Gruber, W., Doppelmayr, M., Stadler, W., Schabus, M., 2002. The interplay between theta and alpha oscillations in the human electroencephalogram reflects the transfer of information between memory systems. Neurosci. Lett. 324, 121–124.
Sauseng, P., Klimesch, W., Gruber, W.R., Hanslmayr, S., Freunberger, R., Doppelmayr, M., 2007. Are event-related potential components generated by phase resetting of brain oscillations? A critical discussion. Neuroscience 146, 1435–1444.
Silberstein, R.B., Cadusch, P.J., 1992. Measurement processes and spatial principal components analysis. Brain Topogr. 4, 267–276.
Shah, A.S., Bressler, S.L., Knuth, K.H., Ding, M., Mehta, A.D., Ulbert, I., Schroeder, C.E., 2004. Neural dynamics and the fundamental mechanisms of event-related brain potentials. Cereb. Cortex 14, 476–483.
Spagnolini, U., 1995. 2-D phase unwrapping and instantaneous frequency estimation. IEEE Trans. Geosci. Remote Sens. 33, 579–589.
Srinivasan, R., Winter, W.R., Ding, J., Nunez, P.L., 2007. EEG and MEG coherence: measures of functional connectivity at distinct spatial scales of neocortical dynamics. J. Neurosci. Methods 166, 41–52.
Takahashi, K., Saleh, M., Penn, R.D., Hatsopoulos, N.G., 2011. Propagating waves in human motor cortex. Front. Hum. Neurosci. 5.
Tallon-Baudry, C., Bertrand, O., Delpuech, C., Pernier, J., 1996. Stimulus specificity of phase-locked and non-phase-locked 40 Hz visual responses in human. J. Neurosci. 16, 4240–4249.
VanRullen, R., Busch, N.A., Drewes, J., Dubois, J., 2011. Ongoing EEG phase as a trial-by-trial predictor of perceptual and attentional variability. Front. Psychol. 2, 60.
Wright, J.J., Robinson, P.A., Rennie, C.J., Gordon, E., Bourke, P.D., Chapman, C.L., Hawthorn, N., Lees, G.J., Alexander, D., 2001. Toward an integrated continuum model of cerebral dynamics: the cerebral rhythms, synchronous oscillation and cortical stability. Biosystems 63, 71–88.
Zvyagintsev, M., Nikolaev, A.R., Mathiak, K.A., Menning, H., Hertrich, I., Mathiak, K., 2008. Predictability modulates motor-auditory interactions in self-triggered audio–visual apparent motion. Exp. Brain Res. 189, 289–300.