nature neuroscience

SUPPLEMENTARY INFORMATION

https://doi.org/10.1038/s41593-018-0171-8

In the format provided by the authors and unedited.

Investigating large-scale brain dynamics using field potential recordings: analysis and interpretation

Bijan Pesaran[1,2]*, Martin Vinck[3], Gaute T. Einevoll[4,5], Anton Sirota[6], Pascal Fries[3,7], Markus Siegel[8], Wilson Truccolo[9,10], Charles E. Schroeder[11,12] and Ramesh Srinivasan[13]

[1]Center for Neural Science, New York University, New York, NY, USA. [2]NYU Neuroscience Institute, New York University Langone Health, New York, NY, USA. [3]Ernst Strüngmann Institute (ESI) for Neuroscience in Cooperation with Max Planck Society, Frankfurt, Germany. [4]Faculty of Science and Technology, Norwegian University of Life Sciences, Ås, Norway. [5]Department of Physics, University of Oslo, Oslo, Norway. [6]Bernstein Center for Computational Neuroscience Munich, Munich Cluster of Systems Neurology (SyNergy), Faculty of Medicine, Ludwig-Maximilians-Universität München, Planegg-Martinsried, Germany. [7]Donders Institute for Brain, Cognition, and Behaviour, Radboud University Nijmegen, Nijmegen, Netherlands. [8]Centre for Integrative Neuroscience & MEG Center, University of Tübingen, Tübingen, Germany. [9]Department of Neuroscience and Institute for Brain Science, Brown University, Providence, RI, USA. [10]Center for Neurorestoration and Neurotechnology, U.S. Department of Veterans Affairs, Providence, RI, USA. [11]Translational Neuroscience Division, Center for Biomedical Imaging and Neuromodulation, Nathan Kline Institute, Orangeburg, NY, USA. [12]Department of Neurosurgery, Columbia College of Physicians and Surgeons, New York, NY, USA. [13]Department of Cognitive Sciences, Department of Biomedical Engineering, University of California, Irvine, CA, USA. *e-mail: bijan@nyu.edu

Investigating large-scale brain dynamics

[12]Department of Neurosurgery
Columbia College of Physicians and Surgeons
New York, NY, USA

[13]Department of Cognitive Sciences
Department of Biomedical Engineering
University of California
Irvine, CA USA

S1 – Activation

S1A - Spectral estimation	1691 words
S1B - Higher-order structure	446
S1C - Inverse modeling	816

S2 – Correlation

S2A – Mathematical model of coherence	402
S2B – Spike-field coherence	289
S2C - Volume conduction	432
S2D - Experimental considerations	1747

S3 - Communication

S3A - Time-delays	683

Supplementary References	110 references

Investigating large-scale brain dynamics

S1 - Activation

S1A - Spectral estimation

Issues related to statistical estimation are fundamental to the analysis of field potentials. Spectrum and autocorrelation function measures have different statistical properties. The autocorrelation function at each time lag reflects signal contributions from all frequencies. The statistical distribution of the autocorrelation function is, in general, unknown for stationary stochastic processes with temporal correlations. Temporal correlations mean neighboring time bins in the autocorrelation function have different statistical distributions and so cannot be averaged together. As a result, the statistical significance of temporal correlations is difficult to assess. In contrast, spectral estimates have a known distribution under often-reasonable stationarity and mixing assumptions that include temporal correlations.

Spectral estimation is a mature field and many spectral estimators have been developed whose mathematical properties are well-understood. Due to the Central Limit Theorem, estimates of spectral power at each frequency are distributed as a χ^2-distribution with a certain number of degrees of freedom (DOF) scaled by the true value of the spectrum[56]. The first spectral estimator to be developed was the periodogram[171]. The periodogram spectral estimate has two DOFs and is highly variable since the χ^2-distribution has a long tail. Estimating the periodogram using longer recordings by increasing T improves the spectral resolution and reduces the Rayleigh frequency. However, increasing T does not change the number of statistical DOFs for the periodogram estimator. Since the number of DOFs is unchanged, the periodogram estimator remains variable and does not converge with increasing data (it is not asymptotically convergent) and the resulting estimate of power remains poor. Stated more formally, the periodogram is an asymptotically unbiased estimator, i.e. as T becomes infinitely long, the expected value converges to the true power spectral density. However, the periodogram is not an asymptotically consistent estimator, i.e. as T becomes infinitely long, the periodogram estimate does not converge to the true power spectral density. This implies that as T becomes infinitely long, the variance does not go to zero. Trial-averaging increases the number of DOFs by two for each trial in the average and so decreases variability. However, trial-averaging will not improve estimation bias.

More sensitive spectral estimators use data tapers. For example, the Welch-averaging or multitaper estimates[172,173], both use data tapers to generate spectral estimates with better bias and variance properties than the periodogram[75]. Data tapers are functions that multiply the data prior to Fourier transformation. Data tapers work by defining orthogonal subsets of data recorded on the same trial, and the shape of each taper is chosen to reduce the influence of activity at one frequency on another, called spectral leakage. Since spectral leakage is reduced this leads to good bias properties. Multitaper estimates, in particular, employ data tapers that offer maximal spectral concentration within the frequency band of interest, Slepian functions, and formally minimize spectral leakage beyond a certain frequency resolution, W[173,174]. Each data taper increases the number of DOFs in the estimate by two and so reduces variability in the estimate. In addition, the longer the time interval, T, the more data tapers can be used for a given spectral concentration. Therefore, increasing T reduces variability. Welch-averaging and multitaper estimates converge as the number of degrees of freedom increase, a property known as asymptotic convergence. Stated more formally, multitapering (and Bartlett or Welch) only provides a consistent spectral estimator if we let K grow with recording length T, such that the effective smoothing bandwidth actually becomes 0.

Efficient estimators are important for assessing neural coding as decoding field potentials depends on estimating spectral quantities from single trials. Trial averaging can be used to further increase the total number of DOFs. Estimates based on 100 trials have 10 times more DOFs than estimates based on 10 trials. Multitaper estimates offer an additional advantage over Welch-averaging estimates because they can generate twice as many DOFs on each trial. Formally, multitaper estimates saturate the Cramer-Rao bound and provide an estimate with the least bias and variability for a given amount of data[175]. We can estimate time-varying spectral characteristics using different tapers, such as e.g. sliding Hanning windows, Gaussian windows (i.e. wavelets[176]) or multitapers.

Since the distribution of all spectral estimates is given by a χ^2-distribution that is scaled by the true value of the spectrum, large values of the spectrum have more variable estimates. Moreover, since field potential spectra generally have greater power at lower frequencies, plotting the spectrum on a linear amplitude scale

can be misleading. After taking the logarithm of the spectrum, the value of the spectrum enters as an additive term. This ensures that fluctuations are comparable in amplitude across frequencies. The logarithm also has the benefit of being a variance-stabilizing transformation for the χ^2-distribution, i.e. rendering the distribution more Gaussian. Therefore, best practice for plotting the raw, non-normalized power spectrum is to plot the log-spectrum and not the spectrum. For the coherence, the variance-stabilizing transformation for the coherence is the arc-tanh. This transformation or equivalent should be applied to the estimates before performing operations, such as estimating confidence intervals, that assume a Gaussian or symmetric distribution.

The above spectral estimators are non-parametric estimates. The spectrum can also be estimated by fitting a parametric model, such as an autoregressive model, to yield a parametric spectral estimate[177]. In general, parametric spectral estimators suffer problems of bias unless the underlying process model is valid. An important advantage of using a parametric model is that if the model is valid, time and frequency resolution using parametric spectral estimators can be very high and does not need to satisfy the time-frequency uncertainty principle[173].

Stationarity: Spectrum and correlation function estimates assume that neuronal activity is stationary. Stationarity means the signal statistics remain constant over all time and is never strictly satisfied. For data generated by multiple trial experiments, it is important to consider stationarity across trials. To limit the influence of event-related responses, it is advisable to subtract the mean response of the raw recordings aligned to the external event (Event Related Potential – ERP) before subsequent analysis. Under the assumption of linear additive models combining the ERP and the ongoing response, one can subtract a scaled and/or time-shifted version of the ERP from each trial that minimizes the time series residual[101,102,106] to account for variability in the ERP across trials. A general solution for the case of nonlinear interactions between ERP and ongoing activity remains a challenge.

Computational considerations: Due to the use of the Fast Fourier Transform, computations for spectral estimates scale with K*N*log(N) with K the number of tapers and N the number of time points. As a result, the spectral analysis of long recordings with large N can be impractical. One solution is to divide the recording into shorter segments according to the Welch method, analyzing each segment with a small number of tapers according to the multitaper method, and averaging the resulting spectral analyses. In this case, the benefits of multitaper methods, which scale linearly with the number of tapers, can be accrued without incurring an excessive computational burden. Dividing the recording into shorter sequences has the added benefit that stationarity assumptions can be better satisfied. To summarize, best practice is to avoid the periodogram, use multitaper methods with parameters T and W chosen to fit to the temporal and spectral extension of the investigated activity pattern, and to limit the number of tapers to 10 when computational considerations are a factor.

Hypothesis testing: Hypothesis testing is often performed following additional processing steps to suppress unwanted spectral features. **1 -** The $1/f^a$ trend can be removed by performing a linear regression of log power onto log frequency (i.e. fitting the $1/f^a$ curve) and then analyzing the residual spectral power, or by explicitly fitting a linear mixture of a $1/f^a$ curve and a Gaussian fitting the spectral peak[178]. However, care is warranted to avoid misestimation and distortion of the signal at smaller amplitudes. **2 -** The spectrum can be normalized to the total power across frequencies. However, after such normalization, changes in spectral power at a given frequency may reflect changes at other frequencies, or a change in the steepness of the $1/f^a$ curve. **3 -** The relative spectrum between conditions can reveal band-limited changes in spectral power between conditions. **4 -** To remove background noise artifacts, it is advisable to examine decomposition techniques such as ICA (see **S2D**).

Testing hypotheses by comparing the spectrum and coherency between experimental conditions and across different frequencies involves the multiple-comparison problem. Properly controlling for multiple comparisons is complicated by the presence of correlations in the spatial, temporal and frequency dimensions of spectral quantities. Classic ways of dealing with the multiple comparison problem like Bonferonni or false-discovery-rate (FDR)[179] correction can be overly conservative by assuming too many independent statistical tests. A better solution is to perform permutation statistics to correct for the effective number of independent statistical tests. Many permutation statistics are based on cluster mass[72]. Cluster mass procedures favor the detection of significant effects for frequencies with a large bandwidth (e.g. gamma-band) compared with lower

frequencies with a small bandwidth (e.g. delta or theta). Cluster mass procedures also favor detection of effects for larger brain areas. An alternative is to perform permutation statistics on the maximum difference between conditions (e.g. across frequencies) rather than cluster mass[180–182]. Maximum difference procedures can also control for the FDR instead of the false alarm rate and can be performed sequentially[180], allowing a large number of significant frequency bands to be detected.

In some cases, it is useful to test hypotheses using surrogate data. Surrogate data processes the experimental data to create new versions of the experimental data that satisfy certain constraints. For example, it is often useful to test the observed signals against the null hypothesis that a signal is constrained to be a stationary, stochastic linear Gaussian process[183,184] (e.g. an autoregressive process with the same spectrum), that has been observed after transformation by a monotonic function. To perform this test, the surrogate data is given by the observed field potentials for which phases at each frequency have been randomized. To generate the surrogate data, we take the Fourier transform of the field potential observations, shuffle the phase of the complex coefficients and then perform the inverse Fourier transform to obtain the phase-randomized surrogate data.

S1B - Higher-order structure:

Complex, non-linear structure involving interactions between different frequency ranges, cross-frequency coupling, is often present in brain signals. For example, coupling of high-frequency amplitudes to the phase of slow frequencies, called phase-amplitude coupling, exists in various brain systems and species[112,115,185–190]. Cross-frequency interactions reflect non-linearities because in linear systems, activity at one frequency is independent of other frequencies. Higher-order correlations are needed to characterize non-linear processes[56], so measures other than the power spectrum are needed.

Many measures of cross-frequency phase-amplitude coupling have been proposed[112,188] including a statistical estimator using Generalized Linear Models (GLMs)[191,192]. The GLM framework provides straightforward model-fitting algorithms and principled procedures for statistical inference of cross-frequency coupling and has revealed coupling between (~ 10 Hz) theta and (~ 50 Hz) gamma LFP activity and between (2 – 3 Hz) spike-wave discharges and (80 – 150 Hz) high-frequency oscillations in neocortex during human focal seizures[193]. Cross-frequency phase-amplitude coupling measures merit some methodological and conceptual caveats[59]. Proper bandwidth selection for both phase and power extraction is critical[59,194]. Neuronal non-stationarities, non-linearities in signal preprocessing and the presence of signal harmonics can all lead to phase-amplitude coupling, without an underlying interaction between distinct neuronal processes in different frequency ranges. It is also important to assess how two processes (phase of the slow and power of the fast rhythms) reflect independent or interrelated neuronal processes. Intrinsic cellular mechanisms may play a role. For most pyramidal cells, low-pass filtering properties of dendrites leads to low-frequency signals as a simple result of integration of fast time scale input giving rise to cross-frequency coupling[195]. Similarly, since high-frequency field activity (>50 Hz) can reflect spiking, phase-amplitude coupling between high-gamma band activity (>50 Hz) and lower frequencies may reflect spike-field coupling rather than the interaction between two distinct rhythms[196].

Beyond phase and amplitude, power correlations also measure functional connectivity[9,197,198]. The alignment of oscillatory phases between neuronal populations may modulate their communication by aligning rhythmic excitability fluctuations[8]. Similarly, the temporal alignment of amplitude bursts across neuronal populations may modulate their interactions by aligning temporal windows for efficient input and output. However, common fluctuations of signal power may reflect direct and indirect interactions as well as shared input to neuronal populations. One way to suppress spurious power correlations due to field spread or the undesired consequences of referencing is to subtract signals shared across multiple field potential recordings, a procedure called orthogonalization, and then estimate correlations[198,199]. Brain-wide power correlations of orthogonalized MEG and EEG signals reveal rich spatial structure that is frequency-specific, reflects known functional and anatomical connectivity, and is correlated with BOLD fMRI correlation structure[197–200]. Power correlations between orthogonalized LFPs can also yield frequency-specific functional connectivity between brain regions while mitigating potential volume conduction or referencing confounds[190].

S1C - Inverse modeling

Due to the ill-posed nature of the inverse model, the interpretation of time and nature of current generators that result from inverse modeling should be made with great caution and, when possible, should include complementary information such as anatomy and spiking activity. In current-source density (CSD) analysis, an observed CSD sink in, say, cortical layer-3 can both be due to synaptic excitation onto the apical dendrites of layer-5 and layer-2/3 pyramidal cells and even synaptic inhibition onto the soma regions of these two pyramidal-cell populations[201]; both of these scenarios have been observed empirically[202,203].

In general, CSD estimates suffer problems of estimation bias and variance that could, if uncorrected, result in flawed interpretations. To reduce estimation bias and variance, recordings need to be scaled to account for variations in electrode impedance and amplifier gain. To further reduce estimation variance, one should select appropriate spatial smoothing parameters. Minimizing both bias and variance requires care because increasing spatial smoothing to reduce variance can distort the resulting CSD estimate and increase bias. Best practice is to increase the amount of smoothing and reduce variance to a degree that does not excessively distort the resulting estimate.

LFP recordings of oscillatory dynamics that are made using amplifiers that filter low-frequency signals, termed alternating-current (AC)-coupled recordings, leads to situations in which actual dipoles alternate with artificial ones[143,146]. Since a simple solution to dissociate one from the other does not exist, recordings for CSD analysis should be made with as little filtering as possible.

The spatial differentiation when performing CSD analysis increases spatial resolution of the field-generating currents. Differentiation also eliminates volume conduction of remote sources positioned orthogonally to the array axis/axes, and strongly attenuates remote sources positioned in the axis. However, because the signal of interest is typically partially shared between the differentiated electrodes, CSD calculation often reduces the signal-to-noise ratio.

CSD estimation methods such as the inverse CSD (iCSD) method[145] and the kernel CSD (kCSD) method[204] avoid the assumption of horizontal homogeneity and permit irregular arrangements of electrode contacts. Finally, all these CSD analysis methods assume that the recorded extracellular potentials stem from transmembrane currents only. However, diffusion of ions may also give rise to extracellular potentials[3,205], and, if present, this should be corrected for in the CSD analysis[6].

For EEG analysis, the current source density at the scalp level is estimated by taking the surface Laplacian, or second spatial derivative, along the two-dimensional scalp surface[6]. This is called source reconstruction because the surface Laplacian provides an estimate of radial current density in the skull, which peaks over the sources of the scalp EEG improving the spatial resolution of EEG. Another approach is to directly estimate the intracranial current source distribution from M/EEG recordings and a volume conduction model of the head.

The inverse problem of source reconstruction is often particularly ill-posed as typically available measurements do not provide enough constraints, and additional assumptions are needed[206]. In some cases, the assumption of a single isolated source may be reasonable e.g., an epilepsy focus, or early sensory evoked potentials. In other cases, a popular approach is to regularize the contribution of many sources by constraining, for example, the total energy of the dipole density[206]. One issue is that regularization procedures generally lack a biophysical basis. Newer methods based on Bayesian inference permit model comparisons based on a variety of explicit assumptions[207]. Alternatively, adaptive spatial filtering ('beamforming') techniques can be used to reconstruct source-level activity from sensor-level M/EEG[32]. Advantages of beamforming techniques are that they are adaptive to the recorded data and that activity can be reconstructed for specific source-locations of interest without the need of a complete source-model to explain the sensor level data. Finally, the recent development of the magnetrode offers ground-truth *in-vivo* magnetic field measurements that might help validating MEG source modeling and reconstruction[208].

In addition to source-reconstruction techniques, a variety of decomposition schemes have been proposed to dissociate neuronal populations or brain regions contributing to field potentials. Decomposition schemes explicitly define the mathematical projection from high-dimensional space of activity to the lower-dimensional measurement space. Principal component analysis (PCA)[209] and independent component analysis

(ICA)[159,210,211] are blind-source separation methods which assume zero correlation or independence, respectively, between different current sources. These assumptions are not true in general, especially if neuronal signals come from densely interconnected cortical columns or regions. In addition, while the independence assumption may hold, non-linear interactions of currents according to the cable equation will result in strictly non-linear summation in the resultant extracellular current[212]. Laminar population analysis (LPA)[213] and dynamic causal modeling (DCM)[214–217] decompose field potentials into multiple sources by fitting generative models to the field potential responses. These model-based decompositions do not necessarily assume independence or lack of correlation between current sources but depend on the validity of the underlying model. In addition, blind-source and model-based decompositions assume the measurement projection is stationary. Each of these assumptions should be carefully examined before interpreting any resulting decomposition.

S2 – Correlation

S2A - Mathematical model of coherence

Consider a field potential, $x(t)$, composed of a signal, $r(t)$, and additive, uncorrelated noise, $\eta_x(t)$. Consider another field potential, $y(t)$, composed of a signal, $u(t)$, that is uncorrelated with $x(t)$, another signal from $x(t)$ but delayed in time by τ, $r(t-\tau)$ and weighted by a coupling coefficient a and additive uncorrelated noise, $\eta_y(t)$. Assume that $\eta_x(t)$ and $\eta_y(t)$ are also uncorrelated. We can write the following equation for each signal:

$$x(t) = r(t) + \eta_x(t)$$

$$y(t) = u(t) + ar(t-\tau) + \eta_y(t)$$

The coherency between $x(t)$ and $y(t)$, $C_{xy}(f)$ is a complex-valued quantity defined as the ratio of the cross-spectrum between x and y, $S_{xy}(f)$, and the square root of the product of the spectrum of each signal, $S_x(f)$, and $S_y(f)$:

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

The coherence is the real-valued magnitude of the coherency.

The spectrum is defined as:

$$S_x(f) = E[\tilde{X}^*(f)\tilde{X}(f)]$$

$$S_y(f) = E[\tilde{Y}^*(f)\tilde{Y}(f)]$$

The cross-spectrum is defined as:

$$S_{xy}(f) = E[\tilde{X}^*(f)\tilde{Y}(f)]$$

$E[]$ denotes the expected value operation. $\tilde{X}(f)$ denotes the Fourier transform of $x(t)$. $\tilde{Y}(f)$ denotes the Fourier transform of $y(t)$. We denote the Fourier transform of the signal $u(t)$ as $\tilde{U}(f)$ and the Fourier transform of the signal $r(t)$ as $\tilde{R}(f)$. Substituting $t+\tau$ for t we can write the Fourier transform of $r(t-\tau)$ as $exp(-2\pi if\tau)\tilde{R}(f)$. Finally, we denote the Fourier transform of the noise $\eta_x(t)$ and $\eta_y(t)$ as $\widetilde{\eta_x}(f)$ and $\widetilde{\eta_y}(f)$, respectively and their spectra as $N_x(f)$ and $N_y(f)$.

Taking the Fourier transform of $x(t)$ and $y(t)$ yields:

$$\tilde{X}(f) = \tilde{R}(f) + \widetilde{\eta_x}(f)$$

$$\tilde{Y}(f) = \tilde{U}(f) + a * exp(-2\pi i f \tau)\tilde{R}(f) + \widetilde{\eta_y}(f)$$

Using the definition of the spectrum and cross-spectrum, that the expected value operation is linear, and how each signal and each noise process are uncorrelated and the noise processes are also uncorrelated, we get:

$$S_x(f) = S_R(f) + N_x(f)$$

$$S_y(f) = S_U(f) + a^2 S_R(f) + N_y(f)$$

$$S_{xy}(f) = a * exp(-2\pi i f \tau) S_R(f)$$

The coherency is given by:

$$C_{xy}(f) = \frac{a * exp(-2\pi i f \tau)}{\sqrt{\left(a^2 + \frac{S_U(f) + N_y(f)}{S_R(f)}\right)\left(1 + \frac{N_x(f)}{S_R(f)}\right)}} \qquad (1)$$

Simplifying $N_x(f)$=$N_y(f)$=0 yields:

$$C_{xy}(f) = \frac{a * exp(-2\pi i f \tau)}{\sqrt{a^2 + \frac{S_U(f)}{S_R(f)}}} \qquad (2)$$

This simplified equation (2) reveals three important features of the coherency. **1 –** The coherence is proportional to the strength of linear association reflected in the coupling coefficient, a. **2 –** Increasing uncorrelated signal power, $S_U(f)$, reduces the coherence and increasing correlated signal power, $S_R(f)$, increases the coherence. Note, however, that this is primarily true when the ratio, $\frac{S_R(f)}{S_U(f)}$, is comparable to 1. **3 –** The impact of a time-delay manifests as a linear change in phase across frequency. To estimate the time-delay, regress the phase of the coherency around the frequency of interest, f, and divide the slope of the regression by 2π to obtain the time-delay.

The more detailed equation (1) reveals how increased uncorrelated noise in either process also reduces the coherence according to a signal to noise ratio.

S2B - Spike-Field Coherence

SFC is usually preferred over the spike-triggered LFP because SFC decomposes correlations by time-scale unlike the spike-triggered LFP. which is sensitive to temporal autocorrelations present in spiking and LFP activity. The decomposition permits efficient estimation, less averaging is needed to reduce variance, broadly-applicable measures of confidence are available for common estimators, and normalized comparisons can be made. Compared with field potential spectral estimates, SFC spectral estimators typically require more smoothing in frequency and more trial-averaging. This is because the number of DOFs in the SFC estimate is equal to the smaller of the number of DOFs in the spiking activity and the field activity. Spike trains are sparse and so always have fewer DOFs than field potentials[75]. Since the asymptotic properties of SFC estimators are not often satisfied, statistical tests based on permutation tests are particularly important for assessing significance. Moreover, unlike for field potentials, the number of DOFs in the spectral analysis of spike trains does not increase linearly with the number of trials. The correction to the number of DOFs depends on the total number of spikes in the analysis – ie number of trials times duration per trial times mean firing rate[75]. Therefore, if the total number of spikes is less than twice the number of DOFs then additional increases in the number of DOFs will not reduce the variance of the spectral estimate. A rule of thumb is that detecting the presence of coherence on the order of 0.1 requires hundreds of spikes and detecting changes in coherence on

the order of 0.01 requires a thousand spikes or more. In some cases, these issues can be avoided by using procedures that do not perform a spectral analysis of the spike train (see **C2 - SNR confound).**

S2C - Volume conduction

Volume conduction presents a number of challenges. In rodents, the situation is particularly complicated because the hippocampus contains signal generators with out-of-phase coherence. An example occurs when volume conduction of multiple phase-shifted theta dipoles generated across para-hippocampal regions results in interference patterns where theta frequency power and phase depend on the relative position of the recording site and current generating dipoles[86]. This introduces neocortical theta coherence at non-zero phase delays[86,218].

One strategy is uses the fact that, if neural tissue is not capacitive, volume conduction occurs instantaneously. Therefore, we can construct a measure of functional connectivity based on the imaginary part of the cross-spectral density, and ignore the real component that captures the instantaneous common mode. Doing so yields several measures such as the imaginary part of coherency[88], phase lag index (PLI)[89] and weighted phase lag index (WPLI)[90].

These techniques can help but they are subject to important limitations. Imaginary coherence is insensitive to zero-phase lag coherence, and volume conduction introduces only zero phase-lag coherence. However, imaginary coherence is a biased estimate of coherence and is phase sensitive. It favors phase relations close to +90° or – 90°, whereas it ignores true coherence at 0° or 180°, i.e. phase relations of particular theoretical and physiological importance. PLI and WPLI are also sensitive to the phase difference, albeit less so. Therefore, changes across experimental conditions, or across space, time or frequency, should be interpreted as a change in correlation only if the relative phase does not change. Unfortunately, the relative phase in turn can be affected by changes in the common-mode contribution. These concerns are particularly significant in the case of M/EEG. Best practice is to first establish that the common mode contribution does not change and that the relative phase is constant across conditions, space, time and frequency. Statistical testing between conditions or changes over time during performance of a task within a single electrode pairing also avoids these concerns.

Correlation measures using the imaginary cross-spectral density effectively filter the signals in time. We can also perform spatial filtering to suppress volume conduction by making use of a bipolar reference, computing the CSD or Laplacian and applying source reconstruction / beam-forming techniques described above. The CSD or Laplacian method removes very large-scale, regional or global coherence as well as volume conduction. Correlations due to volume conduction can survive source reconstruction[91,92] and re-referencing procedures. Great care should be taken that the re-referencing actually removes the common reference. The widely used common-average reference maintains the average of all recorded signals as a common reference and thereby introduces artifacts to which correlations are particularly sensitive.

S2D - Experimental considerations

Signal conditioning: Historically, signal acquisition in differential mode with respect to a ground and reference was necessary to suppress artifacts and avoid saturating the recording amplifier. Modern amplifiers have wide dynamic range making it possible to acquire the signal in single-ended (non-differential) mode with respect to a signal ground. Nevertheless, high-quality signal acquisition still depends on effective signal conditioning. Amplifier gain and analog-to-digital conversion settings must acquire the signal without introducing quantization noise during digitization. Low bias current, high input impedance, low noise, and low drift are essential performance characteristics of the front-end (head-stage) and back-end (before analog-to-digital conversion) amplifiers. Amplifiers with low input impedance distort signals recorded by high-impedance electrodes. Without high-pass front-end filtering, ~ 0.1 Hz, electrode electrochemistry can give rise to direct current components and amplifier saturation depending on electrode surface materials. A low-pass, anti-aliasing filter below the Nyquist frequency, half the sampling rate, is needed to reject high-frequency noise sources. Since noise sources and their suppression can alter the signal, time spent systematically isolating and removing potential noise sources and signal conditioning typically pays returns.

Referencing: The choice and placement of the recording electrode and reference electrode plays a fundamental role[219,220]. The ideal reference does not pick up any non-common brain-related sources. In practice, the choice of reference involves a trade-off between suppressing noise and suppressing the signal of interest. In rodents, the surface of the cerebellum is a common choice due to low spatial coherence of afferents to Purkinje cells. In larger mammals, including intracranial recordings in humans, the cerebrospinal fluid is also a suitable reference. This said, the best choice of reference can depend on the frequency of the signal of interest. A reference that is suitable for low frequency signals may not be suitable for high frequency signals. Defining different references for different phenomena may be warranted. For example, when measuring synchronized up and down states, a choice of reference in the corpus callosum or near the dura may strongly attenuate the signal, and a cerebellar reference might make more sense. However, a cerebellar reference may be a poor choice when measuring signals with poor SNR, such as gamma activity, due to pick up of EMG activity or when local cerebellar high frequency activity contaminates the reference signal[221]. More generally, any local reference in the brain of small animals can contaminate the measured signal due to pickup of local or volume conducted signals by the reference electrode. Therefore, when cerebellar electrode cannot be used and common-mode artifacts are not saturating the amplifiers, single-ended, non-differential recordings which do not employ a local reference electrode should be considered.

EEG recordings on the scalp present a challenge as current sources are measured at both recording and reference electrode. The labels of reference and recording electrode are arbitrary. The potential over the entire head should average to zero, as the head is essentially a closed object with very little current leaving the head via the neck. Hence, best practice is to use many electrodes and reference each electrode to the average. Note that this average reference is still a common reference, which inflates correlation between any two signals referenced to it. Compared to electric field recordings, magnetic field recordings have the advantage that they do not require referencing. This benefit has long been available at a coarse scale with MEG in human subjects, and has recently become available at the micrometer-scale with magnetrodes[208] for invasive in-vivo recordings inside the neuropil.

The impact of local referencing on the measured signal can be tested and better understood via analysis of bipolar electrode measurements. The forward model teaches us that geometric and anatomical factors significantly contribute to field potentials and, consequently, need to be considered when designing an experiment. Consider the case of two electrodes with similar impedances in a bipolar electrode arrangement, an arrangement that is often effective. The distance and orientation of the bipolar electrodes with respect to any distant, confounding sources will affect how much the noise is suppressed. The distance and orientation with respect to the signal source dipole will affect how well the signal is measured. How the bipolar electrodes span the current-generating somato-dendritic locus is also important, and can alter recording magnitude and phase. In cortex, bipolar electrodes spanning layers will often best capture local sources that are likely due to pyramidal neurons. Consequently, if the bipolar reference is in different depths, it will preserve the signals better. Similarly, if bipolar electrodes can be located perpendicular to the orientation of the remote source that will give the best suppression. When the bipolar reference is in the horizontal plane, instead of spanning layers, the resultant field potentials may attenuate different types of network phenomena differently. For example, highly spatially coherent alpha waves might be attenuated very strongly, whereas gamma waves that are less spatially coherent might be hardly suppressed. Closely-placed electrodes in poorly-chosen configurations can lead to signal rejection, attenuation and distortion. This can be observed when recording with an electrode in white matter and another in the hippocampus. Hippocampal theta activity can be significantly attenuated and/or phase distorted depending on electrode placement and separation[86]. Therefore, experimental designs that place electrodes in different anatomical configurations are often needed to constrain inferences.

Environmental noise: Noise problems arise when recording from patients in clinical settings and other natural settings. Even in a research lab, environmental noise can be challenging. Ideally, the impedance of the reference and recording electrodes and associated signal conditioning should be similar to ensure similar sensitivity to noise sources and high suppression of common modes. When the reference is too distant, noise sources may not appear as common modes and blind-source separation techniques may be appropriate. Electrodes are sensitive to environmental noise due to capacitive and inductive coupling[222]. Capacitively-coupled artifacts are worst when the electrode impedance is high and can be minimized by using a Faraday cage, shielding and grounding the cables, routing signal and reference wires away from noise sources, and by

using a headstage to reduce the signal impedance. Inductively-coupled artifacts can be minimized by twisting wires to reduce the cross-sectional area of any loops. Ground loops, mixed analog and digital grounds, badly designed power supplies, loose connections and relative movement of wires are additional sources of poor quality signals. Optical-isolation, direct-current power supplies and batteries can help debug environmental noise sources.

Physiological artifacts: Muscle activity or electromyography (EMG) in field potential recordings typically originates from the eyes, face, neck, and body. EMG artifacts due to heart beat are present in EEG/MEG and can also be present in ECoG and, depending on the location of the reference, LFP. Cardiac pulsation near major vasculature elements can also impact ECoG and LFP. For EEG and MEG, EMG artifacts can be substantial above 20 Hz[223]. Near the ears, eyes, and neck, cortical signals are as much as 200 times smaller[224]. Transient artifacts due to head movements, eye blinks, lateral eye-movement, or jaw clenching, as well as ocular muscle activity due to micro- and macro-saccades cause broad-band artifacts with different spatio-temporal patterns[35,225]. Cranial muscle activity is modulated by cognitive and affective factors[226]. The micro-saccade rate is also modulated by sensory stimulation and cognitive factors[35,227]. Thus, EMG artifacts substantially complicate analyses. EEG gamma-band responses to visual or auditory stimuli caused by modulated micro-saccade rates can masquerade as neuronal responses[227,228]. Co-fluctuations of gamma-band power between scalp or brain regions that are caused by co-fluctuation of muscle activity may be mistaken as correlated neuronal activity [197]. Changes in muscle tone may also mask or un-mask neuronal activity and be misinterpreted.

Stimulation artifacts: Neuronal stimulation techniques such as electrical microstimulation, optogenetic stimulation, transcranial magnetic stimulation (TMS) and transcranial direct/alternating current stimulation (tDCS/tACS) also introduce artifacts. Intracortical microstimulation and optical stimulation generate artifacts but if the recording amplifier does not saturate, artifacts will be brief, and can often be detected, modeled and suppressed effectively[229]. Shielding the recording cables can reduce capacitive-coupling to the stimulation current. Light shielding and particular electrode materials can reduce photoelectric coupling. TMS causes longer-lasting artifacts in EEG due to muscle stimulation, sensory responses, electrode polarization and movement[230]. TMS artifacts can be partly avoided by experimental design and suppressed using interpolation and Independent Component Analysis (ICA)[230–232]. tDCS and tACS induce artifacts in concurrent EEG and MEG that are continuous and several orders of magnitude larger than neuronal signals. Physiological processes such as heartbeats and respiration modulate the strength of these artifacts, inducing complex broad 10 Hz artifacts around the stimulation frequency and its harmonics[233]. Linear filtering reduces these artifacts, but cannot completely remove them[233]. Thus, techniques to robustly remove tACS/tDCS artifacts from concurrent M/EEG are generally lacking.

Artifact suppression: Adaptive power line noise-removal techniques can suppress artifacts with less distortion than conventional notch filters. Simultaneous electrocardiograms (ECG) and electrooculograms (EOG) can help identify and suppress heartbeat and ocular artifacts. Regression methods can automatically suppress artifacts[234,235]. As stated above, ICA has become an important tool for suppressing artifacts. ICA refers to a broad class of blind source-separation algorithms used to decompose linear mixtures of data[236,237]. The two most widely used algorithms in the EEG literature are the FastICA[238] and InfoMax ICA[232,239]. ICA components can identify artifacts[225,240]. However, ICA often yields non-normal mixtures with outliers. Adaptive spatial filtering techniques such as beamforming[32] can also suppress EMG artifacts[35]. Increasing the number of channels being acquired permits other effective common mode rejection procedures. Some caveats are needed. First, no known modern artifact correction technique is perfect for removing EMG artifacts, and no non-invasive recording technique is completely immune to muscle artifact[241].This is especially the case for the neck and face muscle when making EEG and MEG recordings. Muscle artifacts exhibit broadband frequency spectra with substantial relative power above 15 Hz; therefore analyses of the delta (1-4 Hz), theta (4-8 Hz), alpha (8-13 Hz) and mu (11-14 Hz) bands are typically more robust to muscle artifact contamination than the beta (14-20 Hz) or gamma band (20-50 Hz)[241]. Second, identification of an artifact component is inherently subjective. Some EMG artifacts are easy to identify like eye blinks and eye movements, but many are subtle. Third, subtraction of ICA artifact components reduces the rank of the data, potentially influencing further analysis, and it can introduce artifactual correlations. Complete removal of distant sources is unlikely without detailed volumetric imaging. Good recording and analysis practices remain critical. Subjects should remain still

and minimize jaw clenching, and the electrode array should be positioned tightly to avoid movement. EEG and MEG results above 20 Hz should always be critically assessed with EMG artifacts in mind.

S3 - Communication

S3A - Time-delays

The presence of non-zero time lags in cross-correlation functions and non-zero phase in the coherency may provide initial indicators for the directed interaction between two areas, a time delay[242] (but see **Box 2)**. However, several caveats are warranted. First, the phase of the coherence at a particular frequency itself cannot measure a time delay. A simple model in which one signal contains a time delayed version of another shows that the correct way to infer time-delays involves estimating gradients in relative phase as function of the frequency of spectral coherence (see **S2A**). Second, time delays derived from correlations may not match the actual interaction. This is because in a network, the intuition that causes precede their effects often fails. For example, in the sensorimotor cortex of non-human primates[94], estimated Granger causality (see **C3-Communication**) can be associated with either a positive or negative time delay as assessed via cross-correlation or spectral coherence functions. In linear vector-autoregressive models, time-delays and phase gradients can fail to detect directionality in as many as 20-40%[243] of the cases. In non-linear models, time-delay analyses can also suggest the wrong directionality if the driven system generates a prediction of the driver's input[244]. To mitigate these concerns, multiple convergent sources of evidence should be marshaled in support of time delays and communication, see below.

The GC influence measure assumes that we can measure the systems without measurement noise. In practice, most sources of noise are additive. The extent to which the noise sources are correlated across channels depends on the forward model. Correlated noise can arise from distant sources that have electric field spread to both sensors or a common reference. These common-mode influences can be largely eliminated by using locally referenced data, i.e. bipolar derivations or CSD. Locally referenced data may still be affected by strong physiological or artificial noise sources. Uncorrelated noise can arise for example from mechanical or electric artifacts, or from nearby physiological or non-physiological noise sources (see **Box 2**). Uncorrelated and correlated noise can severely distort the estimation of GC-influences, and create non-zero GC-influences that were not present in the noise-free case[88,93,243]. In case of additive noise, it is extremely difficult if not impossible to precisely reconstruct the underlying GC-influence values, although improved estimation can be achieved using a Kalman filtering approach[245,246].

A more tractable approach is to identify the direction of greatest GC-influence, i.e. whether GC(X → Y) is stronger than GC(Y → X), which is biologically informative[163,247]. Using GC-influence to identify the dominant G-causal direction is robust to the addition of uncorrelated noise. Assuming biologically realistic coherence values, uncorrelated noise tends to decrease G-causality similarly in both directions[243], and simulations show that the dominant GC direction remains unaltered in a very high fraction of bivariate vector-autoregressive models[243]. For very high coherence values and/or the presence of correlated noise, adding a moderate amount of uncorrelated noise to X can sometimes increase GC(Y → X), although it follows from its definition that GC(Y → X) eventually converges to zero as we add more noise.

Dependent, correlated noise, like that caused by volume conduction, poses major problems to standard GC measures even when identifying the dominant G-causal direction[88,243,248]. A promising approach is to analyze time-reversed signals[218,248,249]. Haufe et al. used time-reversed series as a surrogate distribution, testing whether the original difference in GC values (GC(X → Y)- GC(Y → X)) exceeded the same GC difference after time-reversal[248]. Vinck et al. used a more conservative approach by testing whether time-reversal of signals reverses the dominant G-causal direction[243.] This is based on the logic that the influence of noise on the power spectra and the real-valued component of the cross-spectral density is invariant to time-reversal, whereas the delay-sensitive part that changes with time-reversal, the imaginary component of the cross-spectral density, is not affected by correlated noise. Recent work shows that this time-reversal test strongly reduces the amount of false alarms both in the presence of uncorrelated and correlated noise[243] and has been applied to LFP recordings[243,250]. The main disadvantage is that, as for imaginary coherence, time-reversal misses weak asymmetries in interaction with a small imaginary component of the cross-spectrum[243].

Supplementary References

171. Schuster, A. On the investigation of hidden periodicities with application to a supposed 26 day period of meteorological phenomena. *J Geophys Res* **3,** 13–41 (1898).
172. Welch, P. The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms. *IEEE Trans Audio Electroacoust.* **15,** 70–73 (1967).
173. Thomson, D. J. Spectrum estimation and harmonic analysis. *Proc IEEE* **70,** 1055–1096 (1982).
174. Percival, D. B. & Walden, A. T. *Spectral Analysis for Physical Applications*. (Cambridge University Press, 1993).
175. Mitra, P. P. & Pesaran, B. Analysis of Dynamic Brain Imaging Data. *Biophys J* **76,** 691–708 (1999).
176. Kaiser, G. *A Friendly Guide to Wavelets*. **3,** (Springer Science & Business Media, 2010).
177. Mitra, P. P. & Pesaran, B. Analysis of dynamic brain imaging data. *Biophys. J.* **76,** (1999).
178. Hermes, D., Miller, K. J., Wandell, B. A. & Winawer, J. Stimulus Dependence of Gamma Oscillations in Human Visual Cortex. *Cereb. Cortex* **25,** 2951–2959 (2015).
179. Benjamini, Y. & Hochberg, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J Roy Stat Soc Ser B* **57,** 289–300 (1995).
180. Korn, E. L., Troendle, J. F., McShane, L. M. & Simon, R. Controlling the number of false discoveries: application to high-dimensional genomic data. *J Stat. Plan. Inference* **124,** 379–398 (2004).
181. Westfall, P. H. & Young, S. S. *Resampling-Based Multiple Testing: Examples and Methods for P-Value Adjustment*. (John Wiley & Sons, 1993).
182. Phillips, J. M., Vinck, M., Everling, S. & Womelsdorf, T. A long-range fronto-parietal 5- to 10-Hz network predicts 'top-down' controlled guidance in a task-switch paradigm. *Cereb Cortex* **24,** 1996–2008 (2014).
183. Schreiber, T. & Schmitz, A. Surrogate time series. *Phys. D Nonlinear Phenom.* **142,** 346–382 (2000).
184. Theiler, J., Eubank, S., Longtin, A., Galdrikian, B. & Doyne Farmer, J. Testing for nonlinearity in time series: the method of surrogate data. *Phys. D Nonlinear Phenom.* **58,** 77–94 (1992).
185. Bragin, A. *et al.* Gamma (40-100 Hz) oscillation in the hippocampus of the behaving rat. *J Neurosci* **15,** 47–60 (1995).
186. Lakatos, P. *et al.* An oscillatory hierarchy controlling neuronal excitability and stimulus processing in the auditory cortex. *J Neurophysiol* **94,** 1904–11 (2005).
187. Osipova, D., Hermes, D. & Jensen, O. Gamma power is phase-locked to posterior alpha activity. *PLoS One* **3,** e3990 (2008).
188. Tort, A. B. L. *et al.* Dynamic cross-frequency couplings of local field potential oscillations in rat striatum and hippocampus during performance of a T-maze task. *Proc Natl Acad Sci USA* **105,** 20517–22 (2008).
189. Spaak, E., Bonnefond, M., Maier, A., Leopold, D. A. & Jensen, O. Layer-specific entrainment of γ-band neural activity by the α rhythm in monkey visual cortex. *Curr Biol* **22,** 2313–8 (2012).
190. von Nicolai, C. *et al.* Corticostriatal coordination through coherent phase-amplitude coupling. *J Neurosci* **34,** 5938–48 (2014).
191. McCullagh, P. & Nelder, J. A. *Generalized Linear Models*. (Chapman and Hall, 1989).
192. Kramer, M. A. & Eden, U. T. Assessment of cross-frequency coupling with confidence using generalized linear models. *J Neurosci Meth* **220,** 64–74 (2013).
193. Truccolo, W. *et al.* Neuronal ensemble synchrony during human focal seizures. *J Neurosci* **34,** 9927–44 (2014).
194. Dvorak, D. & Fenton, A. A. Toward a proper estimation of phase-amplitude coupling in neural oscillations. *J Neurosci Meth* **225,** 42–56 (2014).

195. Vaidya, S. P. & Johnston, D. Temporal synchrony and gamma-to-theta power conversion in the dendrites of CA1 pyramidal neurons. *Nat Neurosci* **16,** 1812–20 (2013).

196. Ray, S. & Maunsell, J. H. R. Different origins of gamma rhythm and high-gamma activity in macaque visual cortex. *PLoS Biol* **9,** e1000610 (2011).

197. Siems, M., Pape, A.-A., Hipp, J. F. & Siegel, M. Measuring the cortical correlation structure of spontaneous oscillatory activity with EEG and MEG. *Neuroimage* **129,** 345–355 (2016).

198. Hipp, J. F., Hawellek, D. J., Corbetta, M., Siegel, M. & Engel, A. K. Large-scale cortical correlation structure of spontaneous oscillatory activity. *Nat Neurosci* **15,** 884–90 (2012).

199. Brookes, M. J., Woolrich, M. W. & Barnes, G. R. Measuring functional connectivity in MEG: a multivariate approach insensitive to linear source leakage. *Neuroimage* **63,** 910–20 (2012).

200. Hipp, J. F. & Siegel, M. BOLD fMRI Correlation Reflects Frequency-Specific Neuronal Correlation. *Curr Biol* **25,** 1368–74 (2015).

201. Schroeder, C. E. *et al.* Localization of ERP generators and identification of underlying neural processes. *Electroencephalogr. Clin. Neurophysiol. Suppl.* **44,** 55–75 (1995).

202. Brankack, J., Stewart, M. & Fox, S. E. Current source density analysis of the hippocampal theta rhythm: associated sustained potentials and candidate synaptic generators. *Brain Res.* **615,** 310–27 (1993).

203. Martín-Vázquez, G., Makarova, J., Makarov, V. A. & Herreras, O. Determining the true polarity and amplitude of synaptic currents underlying gamma oscillations of local field potentials. *PLoS One* **8,** e75499 (2013).

204. Halnes, G. *et al.* Effect of Ionic Diffusion on Extracellular Potentials in Neural Tissue. *PLoS Comp Biol* **12,** e1005193 (2016).

205. Halnes, G., Mäki-Marttunen, T., Pettersen, K. H., Andreassen, O. A. & Einevoll, G. T. Ion diffusion may introduce spurious current sources in current-source density (CSD) analysis. *J Neurophysiol* **118,** 114–120 (2017).

206. Hauk, O. Keep it simple: a case for using classical minimum norm estimation in the analysis of EEG and MEG data. *Neuroimage* **21,** 1612–21 (2004).

207. Baillet, S. & Garnero, L. A Bayesian approach to introducing anatomo-functional priors in the EEG/MEG inverse problem. *IEEE Trans Biomed Eng* **44,** 374–85 (1997).

208. Caruso, L. *et al.* In Vivo Magnetic Recording of Neuronal Activity. *Neuron* (2017). doi:10.1016/j.neuron.2017.08.012

209. Di, S., Baumgartner, C. & Barth, D. S. Laminar analysis of extracellular field potentials in rat vibrissa/barrel cortex. *J Neurophysiol* **63,** 832–40 (1990).

210. Makarov, V. A., Makarova, J. & Herreras, O. Disentanglement of local field potential sources by independent component analysis. *J Comput Neurosci* **29,** 445–57 (2010).

211. Lęski, S., Kublik, E., Swiejkowski, D. A., Wróbel, A. & Wójcik, D. K. Extracting functional components of neural dynamics with Independent Component Analysis and inverse Current Source Density. *J Comput Neurosci* **29,** 459–73 (2010).

212. Chizhov, A. V, Sanchez-Aguilera, A., Rodrigues, S. & de la Prida, L. M. Simplest relationship between local field potential and intracellular signals in layered neural tissue. *Phys Rev E, Stat. nonlinear, soft matter Phys.* **92,** 062704 (2015).

213. Einevoll, G. T. *et al.* Laminar population analysis: estimating firing rates and evoked synaptic activity from multielectrode recordings in rat barrel cortex. *J Neurophysiol* **97,** 2174–90 (2007).

214. Moran, R. J. *et al.* Dynamic causal models and physiological inference: a validation study using isoflurane anaesthesia in rodents. *PLoS One* **6,** e22790 (2011).

215. David, O. & Friston, K. J. A neural mass model for MEG/EEG: coupling and neuronal dynamics. *Neuroimage* **20,** 1743–55 (2003).

216. Bastos, A. M. *et al.* A DCM study of spectral asymmetries in feedforward and feedback connections between visual areas V1 and V4 in the monkey. *Neuroimage* **108,** 460–475 (2015).

217. Bastos, A. M. *et al.* Canonical Microcircuits for Predictive Coding. *Neuron* **76,** 695–711 (2012).

218. Vinck, M. *et al.* Cell-Type and State-Dependent Synchronization among Rodent Somatosensory, Visual, Perirhinal Cortex, and Hippocampus CA1. *Front Sys Neurosci* **9,** 187 (2015).

219. Cogan, S. F. Neural Stimulation and Recording Electrodes. *Annu Rev Biomed Eng* **10,** 275–309 (2008).

220. Shirhatti, V., Borthakur, A. & Ray, S. Effect of Reference Scheme on Power and Phase of the Local Field Potential. *Neural Comput* **28,** 882–913 (2016).

221. de Solages, C. *et al.* High-frequency organization and synchrony of activity in the purkinje cell layer of the cerebellum. *Neuron* **58,** 775–88 (2008).

222. Morrison, R. *Grounding and Shielding: Circuits and Interference*. (John Wiley and Sons, Inc, 2016).

223. Whitham, E. M. *et al.* Scalp electrical recording during paralysis: quantitative evidence that EEG frequencies above 20 Hz are contaminated by EMG. *Clin Neurophysiol* **118,** 1877–1888 (2007).

224. Fitzgibbon, S. P. *et al.* Surface Laplacian of scalp electrical signals and independent component analysis resolve EMG contamination of electroencephalogram. *Int. J. Psychophysiol.* **97,** 277–84 (2015).

225. Muthukumaraswamy, S. D. High-frequency brain activity and muscle artifacts in MEG/EEG: a review and recommendations. *Front Hum Neurosci* **7,** 138 (2013).

226. Whitham, E. M. *et al.* Thinking activates EMG in scalp electrical recordings. *Clin Neurophysiol* **119,** 1166–75 (2008).

227. Yuval-Greenberg, S., Tomer, O., Keren, A. S., Nelken, I. & Deouell, L. Y. Transient induced gamma-band response in EEG as a manifestation of miniature saccades. *Neuron* **58,** 429–41 (2008).

228. Yuval-Greenberg, S. & Deouell, L. Y. Scalp-recorded induced γ-band responses to auditory stimulation and its correlations with saccadic muscle-activity. *Brain Topogr* **24,** 30–9 (2011).

229. Wagenaar, D. A. & Potter, S. M. Real-time multi-channel stimulus artifact suppression by local curve fitting. *J Neurosci Meth* **120,** 113–120 (2002).

230. Ilmoniemi, R. J. & Kicić, D. Methodology for combined TMS and EEG. *Brain Topogr* **22,** 233–48 (2010).

231. Rogasch, N. C. *et al.* Removing artefacts from TMS-EEG recordings using independent component analysis: importance for assessing prefrontal and motor cortex network properties. *Neuroimage* **101,** 425–39 (2014).

232. Bell, A. J. & Sejnowski, T. J. An information-maximization approach to blind separation and blind deconvolution. *Neural Comput* **7,** 1129–59 (1995).

233. Noury, N., Hipp, J. F. & Siegel, M. Physiological processes non-linearly affect electrophysiological recordings during transcranial electric stimulation. *Neuroimage* (2016). doi:10.1016/j.neuroimage.2016.03.065

234. Gratton, G., Coles, M. G. & Donchin, E. A new method for off-line removal of ocular artifact. *Electroencephalogr Clin Neurophysiol* **55,** 468–84 (1983).

235. Schlögl, A. *et al.* A fully automated correction method of EOG artifacts in EEG recordings. *Clin Neurophysiol* **118,** 98–104 (2007).

236. Makeig, S., Bell, A. J., Jung, T.-P. & Sejnowski, T. J. Independent Component Analysis of Electroencephalographic Data. in *Advances in Neural Information Processing Systems* 145–151 (1996).

237. Jung, T.-P. *et al.* Extended ICA Removes Artifacts from Electroencephalographic Recordings. in *Advances in Neural Information Processing Systems* 894–900 (1998).

238. Hyvärinen, A. & Oja, E. A Fast Fixed-Point Algorithm for Independent Component Analysis. *Neural Comput* **9,** 1483–1492 (1997).

239. Delorme, A. & Makeig, S. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics

including independent component analysis. *J Neurosci Meth* **134,** 9–21 (2004).

240. Delorme, A., Sejnowski, T. & Makeig, S. Enhanced detection of artifacts in EEG data using higher-order statistics and independent component analysis. *Neuroimage* **34,** 1443–9 (2007).
241. Whitham, E. M. *et al.* Scalp electrical recording during paralysis: quantitative evidence that EEG frequencies above 20 Hz are contaminated by EMG. *Clin Neurophysiol* **118,** 1877–88 (2007).
242. Rosenberg, J. R., Halliday, D. M., Breeze, P. & Conway, B. A. Identification of patterns of neuronal connectivity--partial spectra, partial coherence, and neuronal interactions. *J. Neurosci. Methods* **83,** 57–72 (1998).
243. Vinck, M. *et al.* How to detect the Granger-causal flow direction in the presence of additive noise? *Neuroimage* **108,** 301–18 (2015).
244. Matias, F. S. *et al.* Modeling positive Granger causality and negative phase lag between cortical areas. *Neuroimage* **99,** 411–8 (2014).
245. Nalatore, H., Ding, M. & Rangarajan, G. Denoising neural data with state-space smoothing: method and application. *J Neurosci Meth* **179,** 131–41 (2009).
246. Nalatore, H., Ding, M. & Rangarajan, G. Mitigating the effects of measurement noise on Granger causality. *Phys Rev E, Stat. nonlinear, soft matter Phys.* **75,** 031123 (2007).
247. Michalareas, G. *et al.* Alpha-Beta and Gamma Rhythms Subserve Feedback and Feedforward Influences among Human Visual Cortical Areas. *Neuron* **89,** 384–97 (2016).
248. Haufe, S., Nikulin, V. V, Müller, K.-R. & Nolte, G. A critical assessment of connectivity measures for EEG data: a simulation study. *Neuroimage* **64,** 120–33 (2013).
249. Winkler, I., Panknin, D., Bartz, D., Muller, K.-R. & Haufe, S. Validity of Time Reversal for Testing Granger Causality. *IEEE Trans Signal Proc* **64,** 2746–2760 (2016).
250. Jávor-Duray, B. N. *et al.* Early-onset cortico-cortical synchronization in the hemiparkinsonian rat model. *J Neurophysiol* **113,** 925–36 (2015).