



Comparison of undirected frequency-domain connectivity measures for cerebro-peripheral analysis

Joachim Gross^{a,b,c,1}, Daniel S. Kluger^{a,b,1,*}, Omid Abbasi^a, Nikolas Chalas^{a,b},
Nadine Steingraber^a, Christoph Daube^c, Jan-Mathijs Schoffelen^d

^a Institute for Biomagnetism and Biosignal Analysis, University of Münster, Münster, Germany

^b Otto-Creutzfeldt-Center for Cognitive and Behavioral Neuroscience, University of Münster, Münster, Germany

^c Centre for Cognitive Neuroimaging, University of Glasgow, Glasgow, UK

^d Radboud University, Donders Institute for Brain, Cognition and Behaviour, Nijmegen, NL, the Netherlands

ARTICLE INFO

Keywords:

Cerebro-peripheral connectivity
Spectral analysis
Functional connectivity
Phase coupling

ABSTRACT

Analyses of cerebro-peripheral connectivity aim to quantify ongoing coupling between brain activity (measured by MEG/EEG) and peripheral signals such as muscle activity, continuous speech, or physiological rhythms (such as pupil dilation or respiration). Due to the distinct rhythmicity of these signals, undirected connectivity is typically assessed in the frequency domain. This leaves the investigator with two critical choices, namely a) the appropriate measure for spectral estimation (i.e., the transformation into the frequency domain) and b) the actual connectivity measure. As there is no consensus regarding best practice, a wide variety of methods has been applied. Here we systematically compare combinations of six standard spectral estimation methods (comprising fast Fourier and continuous wavelet transformation, bandpass filtering, and short-time Fourier transformation) and six connectivity measures (phase-locking value, Gaussian-Copula mutual information, Rayleigh test, weighted pairwise phase consistency, magnitude squared coherence, and entropy). We provide performance measures of each combination for simulated data (with precise control over true connectivity), a single-subject set of real MEG data, and a full group analysis of real MEG data. Our results show that, overall, WPPC and GCMI tend to outperform other connectivity measures, while entropy was the only measure sensitive to bimodal deviations from a uniform phase distribution. For group analysis, choosing the appropriate spectral estimation method appears to be more critical than the connectivity measure. We discuss practical implications (sampling rate, SNR, computation time, and data length) and aim to provide recommendations tailored to particular research questions.

1. Introduction

The analysis of cerebro-peripheral connectivity (CPC) has recently gained significant interest. We define CPC as significant connectivity between electrophysiological recordings (such as MEG, EEG, or LFP) and simultaneously recorded peripheral signals (Baillet, 2017; Hari et al. 2018, Gross, 2019). The application of CPC is highly versatile and offers a direct way to study body-brain interactions. The underlying physiological mechanisms as well as resulting signal processing requirements vary and depend on the peripheral signal under investigation. For example, a prominent early application of cerebro-peripheral connectivity was the investigation of connectivity between brain and muscle activity (Conway et al. 1995, Salenius et al., 1997), which has led to important insights into the role of neural rhythms in physiolog-

ical and pathological motor control (Bourguignon et al., 2017, 2019; A. Schnitzler & Gross, 2005; Schoffelen et al., 2005). The underlying assumed physiological mechanism of interaction for this cortico-muscular coherence is that the rhythmic activity in primary motor regions provides a rhythmic drive to the spinal motor neuron pool, which can be detected as rhythmicities in the electromyographic (EMG) signal. More recently, using the envelope of speech as a peripheral signal, this type of analysis has also proven useful for studying continuous speech processing due to the fact that brain signals are temporally synchronised to the speech envelope (Gross et al., 2013 b; Lakatos et al., 2019; Meyer et al., 2019; Obleser & Kayser, 2019; Zoefel, 2018). More generally, cerebro-peripheral connectivity can be studied to elucidate the ongoing coupling between any peripherally recorded signal and brain activity (Gross, 2019; Park et al., 2014; Rebollo et al., 2018) and even

* Corresponding author at: University of Münster, Institute for Biomagnetism and Biosignal Analysis, Malmedyweg 15, 48149 Münster, Germany.
E-mail address: daniel.kluger@wwu.de (D.S. Kluger).

¹ Both authors contributed equally

modulations of such connectivity measures as a function of a secondary peripheral signal. Examples for relevant peripheral signals are eye movements, pupil size (Pfeffer et al. 2021), heart beat (Park et al. 2014), respiration (Kluger & Gross, 2020), speech (Gross et al., 2013b), movement or muscle activity (Bourguignon et al., 2017, 2019; A. Schnitzler & Gross, 2005; A. Schoffelen et al., 2005), skin conductance or temperature, and blood pressure. Some of these signals (such as respiration, heartbeat, speech, tremor) are distinctively rhythmic, thus favouring analysis in the spectral domain. However, there is no consensus in the literature regarding a standard method to quantify cerebro-peripheral connectivity in the spectral domain. Instead, a large variety of methods has been used. In practice, spectral cerebro-peripheral connectivity analysis consists of two steps that can each be conducted in several ways: First, spectral estimation is performed where time series are transformed into the frequency domain (as complex-valued numbers). Spectral estimation is most often performed by using Fourier transformation, wavelet transformation, or bandpass-filtering (Bruns, 2004; Gross, 2014; Le Van Quyen & Bragin, 2007).

In a second step, connectivity measures can be estimated. Again, a large number of methods have been suggested (Bastos and Schoffelen, 2015; Marzetti et al., 2019) and some of them have been compared in previous studies (David et al., 2004; Kreuz et al., 2007; Quiñero et al., 2002). It is noteworthy that MEG/EEG connectivity is often discussed in the context of cerebro-cerebral connectivity, i.e. connectivity between different brain areas. This brings about complications that are absent in the case of cerebro-peripheral connectivity. Most importantly, estimation of non-invasive MEG/EEG time series from two regions of interest in the brain is never perfect and leads to leakage effects that contaminate the connectivity estimate (Schoffelen & Gross, 2009). This is typically circumvented using connectivity measures that exclude common zero-lag components in both time series (such as imaginary coherence). In the case of cerebro-peripheral connectivity, the estimation of time series in the brain is still not optimal but the second signal is a peripheral recording that does not share any spurious signal components with the brain signal that result from imperfect source reconstruction. Therefore, analyses of cerebro-peripheral connectivity do not require connectivity measures to exclude shared zero-lag signals.

Depending on the differences of multiple methods for spectral decomposition and estimation of effect size, the investigator's choice could affect the results of the analysis. Here, we aim to investigate the sensitivity of cerebro-peripheral connectivity analysis to the choice of spectral estimation and connectivity measures. We realise that such an investigation depends on the signals that are used and on the implementation of the spectral estimation and connectivity methods. Therefore, we cannot expect to provide authoritative guidance on the 'optimal' analysis approach that generalises to all possible applications. Still, we can expect to learn lessons that could be valuable to the community in the planning of similar studies and the analysis of cerebro-peripheral data.

A second contribution is to make our analysis scripts publicly available on GitHub (<https://github.com/IBiomag/>) so that a similar comparison can be performed for different simulated or real data and different methods can be added and evaluated.

Since we anticipate non-trivial interactions between different spectral estimation methods and different connectivity measures, we analyse all combinations of a set of six standard spectral estimation methods (comprising fast Fourier and continuous wavelet transformation, bandpass filtering, and short-time Fourier transform using Matlab's spectrogram function) and six connectivity measures (phase-locking value, Gaussian-Copula mutual information, Rayleigh test, weighted pairwise phase consistency, magnitude squared coherence, and entropy). We start our investigation by using simulated data where the connectivity between signals is precisely controlled. We then proceed to a single-subject real data set and finally to a full group analysis of an exemplary data set.

2. Material and methods

2.1. Data simulation

The simulated data is constructed by applying a fourth-order Butterworth bandpass filter (3–6 Hz) to a 1-minute simulated white noise signal (sampling rate: 100 Hz) with a mean of 0 and a standard deviation of 1. Two time series are then constructed by adding white noise (independently for each time series and with a mean of 0 and a standard deviation of 1) to the filtered noise. Therefore, the resulting time series show linear dependencies in the frequency range between 3–6 Hz that are evident as phase synchronisation and amplitude correlation. The degree of coupling can be adjusted through the amplitude of the added noise (see dedicated analyses below).

In what follows, the dependency between the time series will be quantified by applying all combinations of the six spectral estimation methods and the six undirected connectivity measures, which will be described in detail next.

2.2. Real data

We used MEG data recorded with a 275 whole-head sensor system (OMEGA 275, VSM Medtech Ltd., Vancouver, Canada) at a sampling frequency of 1200 Hz. The study was approved by the ethics committee of the University of Münster and conducted in accordance with the Declaration of Helsinki. Written informed consent was obtained before the measurement and participants received monetary compensation after the experiment.

Twenty native German-speaking participants (11 males, mean age 24.9 ± 2.6 years, range 20–32 years) listened to nine 1-min long audio recordings of their own voice in which they answered general questions such as 'What does a typical weekend look like for you?'. Speech data was captured at a sampling rate of 44.1 kHz using a microphone placed at a distance of 155 cm from the participant's mouth.

Prior to data analysis, MEG data were visually inspected. No jump artifacts or bad channels were detected. A discrete Fourier transform (DFT) filter was applied to eliminate 50 Hz line noise from the continuous MEG data.

The wideband amplitude envelope of the speech signal is computed using the method presented in (Chandrasekaran et al., 2009). Nine logarithmically spaced frequency bands between 100–10000 Hz were constructed by bandpass filtering (third-order Butterworth filters). Then we compute the amplitude envelope for each frequency band as the absolute value of the Hilbert transform and downsample them to 1200 Hz. This is standard preprocessing for this type of analysis (see e.g. Gross et al. 2013). We average them across bands and use the computed wideband amplitude envelope for all further analysis. Finally, MEG and speech envelope are downsampled to 256 Hz. In the preprocessing and data analysis steps, custom-made scripts in Matlab R2020 (The Mathworks, Natick, MA, USA) in combination with the Matlab-based Field-Trip toolbox (Oostenveld et al., 2011) are used following current MEG guidelines (Gross et al., 2013a).

For source localisation we align individual T1-weighted anatomical MRI scans with the digitized head shapes using the iterative closest point algorithm. Then, we segment the MRI scans and generate single-shell volume conductor models (Nolte, 2003) using the implementation in FieldTrip, and use this to create forward models. Next, the linearly constrained minimum variance (LCMV) algorithm is used to compute time series of voxels taken from a parcel showing medium connectivity (L_PFop located within the left inferior parietal lobule) of the volumetric HCP brain atlas (Glasser et al., 2016). The parcel selection is not relevant for the purpose of this study (which is focused on methods differences given two time series), but we ensure that the parcel shows significant connectivity to the speech envelope. The final time series representing activity from L_PFop is the first component of a singular value decomposition (SVD) of time series from all dipoles in this parcel.

2.3. Spectral estimation

Six different methods are used to perform a complex-valued spectral transformation of the time series in the frequency band. All methods except the wavelet transform use a frequency resolution of 0.5 Hz. For the subsequent connectivity estimation and evaluation, we focus on the frequency band between 1 and 10 Hz.

- 1-3) The first three methods use the Fast Fourier transform (FFT) based implementation in FieldTrip (Oostenveld et al., 2011). The first method uses Hanning tapers while the second and third methods use discrete prolate spheroidal sequences (DPSS) in a multi-taper approach with ± 1 Hz and ± 2 Hz smoothing, respectively. In all three cases a 2s window with 50% overlap is used.
- 4) This uses the continuous wavelet transform implemented in Matlab with Morlet wavelets (cwtfilterbank.m with wavelet parameters 3 and 20). It uses L1-normalization so that equal amplitude oscillatory components at different scales have equal magnitude in the spectral estimate. The Matlab function cwt.m performs the actual transformation into the frequency domain.
- 5) A series of bandpass filters (windowed sinc FIR filter) is applied with edge frequencies that are 1 Hz below and above the center frequency. The center frequency changes from 1-10 Hz in steps of 0.5 Hz. The Hilbert transform is then applied for each filtered signal to obtain the complex-valued spectral estimate.
- 6) This spectral estimate is computed from Matlab's spectrogram function in analogy to method 1. It also uses a 2s window with 50% overlap.

It should be noted that the number of complex valued data points returned from these methods is very different. Methods 1-3 and 6 are based on the FFT and return about one spectrum per second. Methods 4 and 5 instead return one spectrum per data sample and therefore provide many more, albeit largely redundant, data points. This has implications for computation time (see Table S1) but also for the degrees of freedom (df) in the spectral estimates. Since df might affect the performance of the connectivity measure, we quantify it numerically (see Supplementary Fig. S5). Specifically, we compute 100 instantiations of simulated data (see section 2.1 and Fig. 1). For each instantiation and each spectral estimate, we compute df separately for each frequency. Df is computed as the first e-folding of the autocorrelation of the time series (i.e. the first lag where autocorrelation has decreased by e). Another established method is the computation of df as the lag with the first zero crossing of the autocorrelation function. However, results are qualitatively the same for both methods.

We compute df for each frequency separately since for non-windowed methods (continuous wavelet transform and bandpass filtering) autocorrelation is expected to decrease (and df to increase) with higher frequencies. The results are presented as a mean across 100 instantiations in Supplementary Fig. S5.

For the windowed spectral estimates (FT0, FT1, FT2, SG), df corresponds to the number of windows times the number of tapers leading - as expected - to highest df for FT2 (multitaper estimate with +-2Hz smoothing). For continuous wavelet transform (CWT) and bandpass filtering (BF) df increases with frequency (because the autocorrelation decreases). For comparison, the df in the original data before spectral estimation is 6000 and corresponds to the number of samples.

2.4. Connectivity measures

We use six undirected spectral connectivity measures (see also corresponding Matlab code):

- 1) Phase-locking value (PLV; Lachaux et al., 1999): This is defined as the length of the vector average of the normalized (unit length) phase differences between time series x and y. It is suitable to identify unimodal deviations from a uniform distribution of phase differences.

- 2) Gaussian-Copula mutual information (GCMI; Ince et al., 2017): We compute mutual information between two bivariate time-series (real and imaginary part of x and y) using the original implementation (<https://github.com/robince/gcml>). GCMI uses the Copula transform to allow for efficient and robust analytical solutions to the computation of entropy of Gaussian variables. GCMI does not make specific assumptions on the distribution of the data.
- 3) Rayleigh test (R-test; Berens, 2009): The Rayleigh test is defined for circular (phase) data and tests for significant deviation from a uniform phase distribution. Here, it is applied to the phase difference. It is computed as the squared length of the vector average of the normalized (unit length) phase differences between time series x and y. Similar to PLV it is suitable to identify unimodal deviations from a uniform distribution of phase differences.
- 4) Weighted pairwise phase consistency (WPPC; Vinck et al., 2010): This measure does not directly test for a deviation of a phase distribution from a uniform distribution. Instead, it computes the pairwise difference of phases from this distribution. The rationale for this approach is that a preferred phase in the phase distribution would also lead to a cluster in the pairwise difference. However, in contrast to PLV, WPPC is not biased by the sample size. Sample size bias means that the expected value of a connectivity measure such as PLV or coherence depends on the number of samples used for the computation. In the absence of phase synchronization (i.e. when a distribution of phase differences is uniform) the PLV is larger than 0 for small sample sizes and converges to 0 as sample size increases. This is not the case for WPPC. We compute WPPC with code based on the FieldTrip implementation.
- 5) Magnitude squared coherence (COH): Coherence is a standard measure of association corresponding to a frequency domain correlation coefficient. It is computed by dividing the magnitude squared cross-spectral density between x and y by the product of the individual power spectra. COH is sensitive to bivariate dependencies of phase and amplitude in contrast to e.g. PLV that uses only phase information.
- 6) Entropy (ENT; Shannon, 1948): We used entropy to quantify the deviation of the distribution of phase differences from a uniform distribution. In contrast to the other measures, this is sensitive to more than just unimodal phase difference distributions. Here, the computation uses a binning of phase differences into 20 bins following the standard equation

$$H = -\sum(P_i * \log(P_i)) \quad (1)$$

where H is entropy and P_i the probability of a given phase i.

2.5. Surrogate data and normalisation

For each connectivity measure, surrogate data are computed by randomly shifting the spectral estimates of one of the time series with respect to the other with a circular wrapping around the edges of the time series (using circshift.m in Matlab). This temporal shifting of data is an established technique for creating surrogate data because it destroys any true synchronisation in the data (Andrzejak et al., 2003) while preserving the signals' autocorrelation structure. Temporal shifting does not change connectivity between two sine functions (it just changes the phase difference). However, in our simulated data connectivity arises from bandpass filtered noise (3-6 Hz) that is added to two independent noise time series (see 2.1), leading to more complex phase dynamics (compared to a sine function). Temporal shifting destroys simulated synchronization in this case. We perform this shifting procedure 200 times (unless otherwise stated) to create a null distribution of 200 surrogate data points for each connectivity measure. Next, we normalise each raw connectivity measure by subtracting the mean and dividing by the standard deviation of the surrogate distribution for each frequency (Lancaster et al., 2018; Schreiber and Schmitz, 2000). This normalises

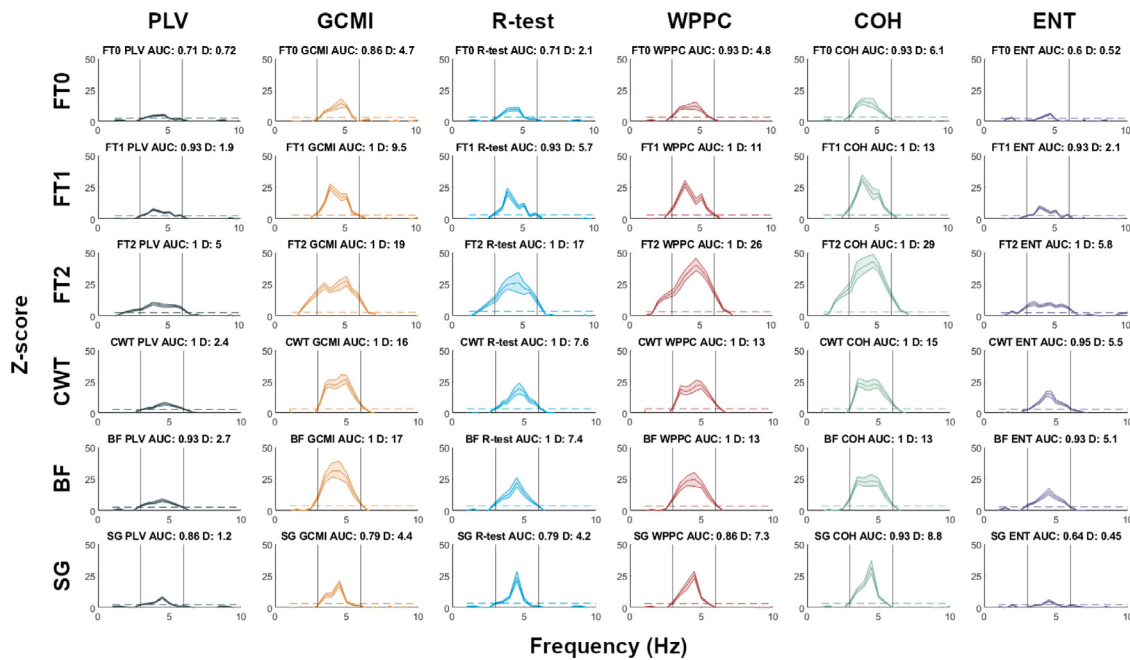


Fig. 1. Connectivity spectra for all combinations of spectral estimates and connectivity measures. Connectivity is estimated for simulated data with a ground truth effect between 3-6 Hz (indicated by vertical lines) with an SNR of 1/20. The solid line shows the connectivity spectrum of a single trial z-scored with the mean and standard deviation of 200 time-shifted versions. The shaded area quantifies the uncertainty of the normalization and is based on the 95 percent bootstrap confidence interval of mean and standard deviation of the surrogate distribution. The dashed line represents the 99th percentile of the surrogate distribution. Each row is based on the same spectral estimate corresponding to the six methods in the same order as described in the methods section. Each column shows results from the same connectivity measure in the same order as described in the methods section. The title of each panel shows the spectral estimation method, the connectivity measure, the area under curve value (AUC), and the D-value defined in the methods section. FT0: FFT with Hanning taper; FT1: multitaper with ± 1 Hz smoothing; FT2: multitaper with ± 2 Hz smoothing; CWT: continuous wavelet transform; BF: bandpass filter; SG: spectrogram; PLV: phase locking value; GCMi: gaussian copula mutual information; R-test: Rayleigh test; WPPC: weighted pairwise phase consistency; COH: coherence; ENT: entropy. The color code for connectivity measures is used throughout the manuscript.

the connectivity measure and transforms it into units of standard deviations of the surrogate distribution. It therefore directly reflects how different the connectivity is from the surrogate distribution. This useful normalisation makes measures more comparable to each other and is used throughout the manuscript.

To allow comparison of connectivity measures we define a performance measure D that quantifies the ‘average distance’ of the normalized connectivity estimate from the 99th percentile of the surrogate distribution. This is computed as the mean of all connectivity values exceeding the 99th percentile of the surrogate distribution in the frequency band of simulated connectivity (3-6 Hz). This measure quantifies how much the estimated connectivity deviates from the surrogate distribution. In addition, we use the area under curve (AUC) which is an established performance measure for classifiers ranging from 0 to 1. It corresponds to the area under the receiver operating characteristic (ROC) which shows true positive rate as a function of false positive rate for various thresholds.

2.6. Data and code availability

We will make the Matlab code and underlying data publicly accessible in full through GitHub (<https://github.com/IBiomag/>).

3. Results

3.1. Comparison of combinations of spectral and connectivity estimates

First, we provide in Fig. 1 an illustration of all combinations of spectral and connectivity measures for the simulated data described above (here with added noise with standard deviation of 1). For all of these combinations we plot the normalized connectivity spectrum (with the

95 percent bootstrap confidence interval) in the frequency range 0-10 Hz and the 99th percentile of the surrogate distribution (dashed line).

All combinations of methods show a clear peak within the frequency band where connectivity was simulated (3-6 Hz). At the same time, it is clearly evident that results differ substantially in the shape of the spectrum and how far peaks are separated from the 99th percentile of the surrogate distribution (i.e., sensitivity for the true effect). First, for the same spectral estimate, different connectivity measures show markedly different sensitivity in detecting synchronisation in the data (compare panels within a row). That is, given the same information, the use of this information is significantly different between connectivity measures. Second, for the same connectivity measure, different spectral estimates lead to very different results (compare panels for a given column). Recall that synchronisation between time series x and y was simulated in the frequency band 3-6 Hz. Ideally, the spectrum in this band should exceed the 99th percentile line leading to a high D-value.

From this simulation (based on 500 separate repetitions) we can already make several interesting observations. By comparing the different rows (spectral estimation methods), we note that the single taper FFT-based spectral estimates (FT0, SG) perform worse than the other methods (see Fig. 1, top and bottom row and note the individual scaling of each graph). An increased spectral smoothing with multitapers leads to an improved performance of all connectivity measures (higher D-values indicating larger separation from the surrogate distribution). However, this comes at the cost of a reduced spectral resolution which we will see in the analysis of real data (Fig. 5, third row from the top). Therefore, multitapers offer advantages for the detection of synchronisation (when the amount of spectral smoothing is similar to the spectral extent of underlying connectivity) while they might be disadvantageous when trying to resolve different spectral peaks. Besides the FT2 method, the

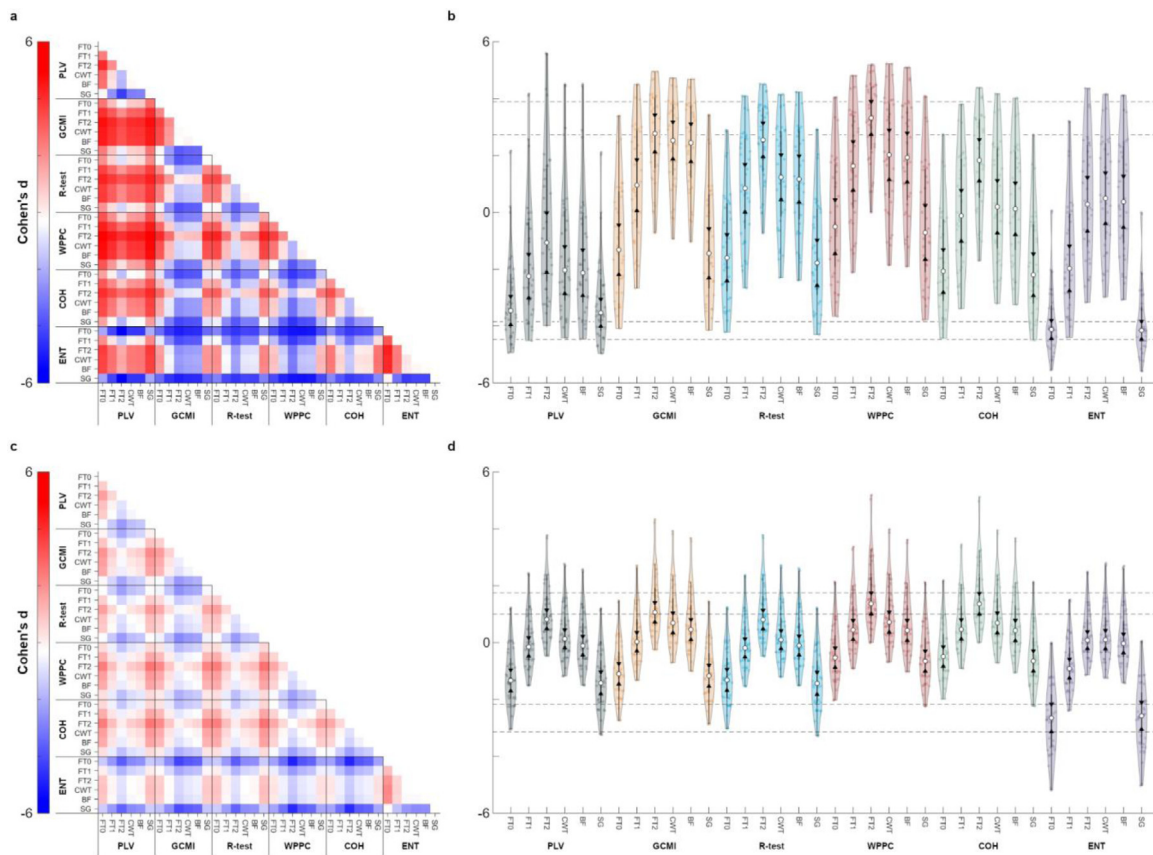


Fig. 2. Pairwise comparisons of spectral estimates and connectivity measures. a, To assess performance differences within the simulated data, we compare each combination of spectral estimate and connectivity measure with any other combination, resulting in a 36×36 symmetrical matrix. We compute Cohen's d as a measure of effect size separating the D-values from the respective 500 simulation iterations of any two combinations. Positive values indicate higher D values for the row (vs the column) combination. b, Violin plot shows the distribution of effect sizes for each of the 36 combinations (grouped according to connectivity measures). White dots mark the respective median of each combination, black triangles indicate box plot notches for comparison across combinations. As a reference, top dashed lines indicate box plot notches for WPPC with FT2 estimation, which show the best median performance overall. Similarly, bottom dashed lines indicate box plot notches for entropy with SG estimation whose performance was lowest overall. c, Same as panel a, but comparing the AUC instead of D values. d, Same as panel b, but for AUC instead of D values.

continuous wavelet transform, and bandpass filtering perform very well (Fig. 1, second and third row from the bottom).

A comparison of connectivity measures (different columns) reveals best performance for WPPC (shown in red) followed by GCMI (yellow). In contrast, ENT (purple) and PLV (grey) show relatively poor performance. Overall, simulation-based connectivity spectra suggest that the combination of FT2 and WPPC shows the best performance.

In order to make performance differences between methods more accessible, we provide pairwise comparisons of all 36 possible combinations (6 spectral estimates x 6 connectivity measures). Specifically, we compute Cohen's d as a measure of effect size separating the D values from the 500 simulations of each combination (see Fig. 2a). Not counting the main diagonal of the symmetrical 36×36 matrix, we gain 35 effect sizes for each combination of spectral estimate and connectivity measure. The respective distributions are shown in Fig. 2b and afford easy comparison between methods. Overall, pairwise comparisons corroborate the previous impression that WPPC with FT2 outperformed most of the other combinations: Judging by the box plot notches in Fig. 2b and d, only GCMI (with FT2, CWT, or BF) and the R-test (with FT2) reach a similar performance. Moreover, the performance for entropy combined with FT0 or SG is particularly subpar, paralleled only by PLV combined with the same estimates. Finally, pairwise comparisons support the initial impression of lowered performance of FT0 and SG in all combinations, irrespective of the connectivity measure (see Fig. 2b).

Fig. 2b also helps to address the question to what extent the performance of connectivity measures is driven by the degrees of freedom afforded by a given spectral estimate (section 2.3). Performance of all connectivity measures increase together with df from SG, FT0 via FT1 to FT2. However, BF has higher df compared to FT2 but leads to generally lower performance across connectivity measures. In summary, it seems that the increased df from multitapered spectral estimates (compared to single tapers) increases performance whereas a further df increase for non-windowed, continuous spectral methods (CWT and BF) does not lead to a further increase in connectivity performance.

3.2. Effect of SNR

Next, we aim to quantify the effect of different levels of signal-to-noise ratio (SNR) on performance. This is motivated by the hypothesis that different connectivity measures are differentially sensitive to varying SNR levels. Indeed, this can be seen in Fig. 3 which follows the arrangements of rows and columns from Fig. 1. Towards the right of the figure, SNR is increasing. A differential SNR-effect on performance is quite prominent in the comparison of the third and fifth column. While WPPC (shown in red) is the most sensitive measure in the middle column (SNR parameter = 1.5) it is outperformed by GCMI (yellow) for the highest SNR ($N = 0.5$, rightmost column). This indicates that performance of GCMI increases more strongly with SNR than for other measures. This

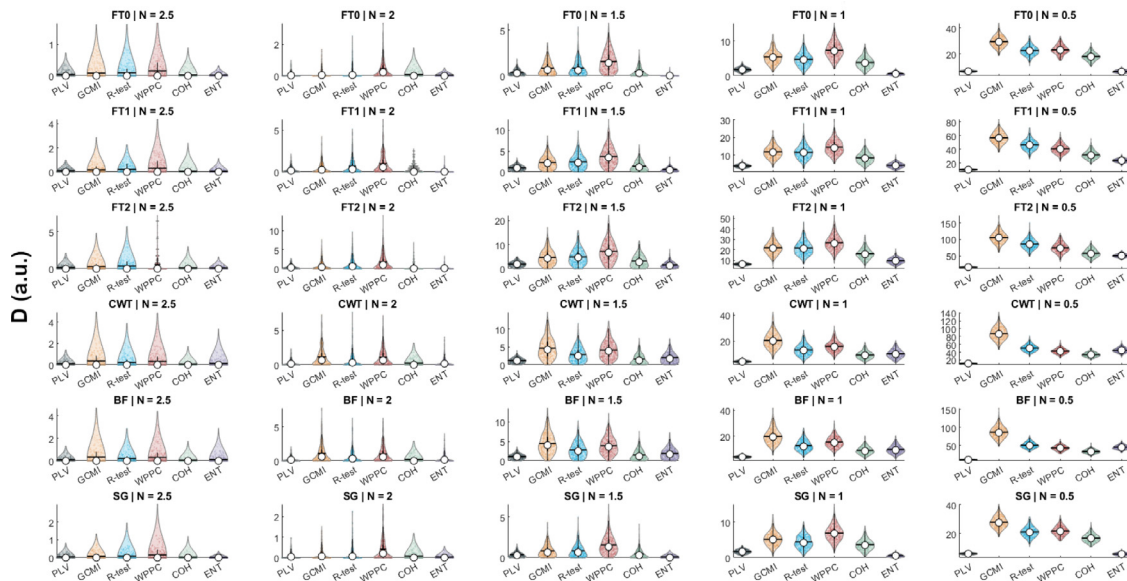


Fig. 3. Effect of SNR. Each subplot shows a violin plot for each of the six connectivity measures (same order as in Fig. 1) of the D-value across 500 repetitions of the simulation. The noise factor (N) specifies the standard deviation of the noise added to the signal. Rows correspond to spectral estimates in the same order as described in the methods section. Columns correspond to different SNRs which increases from left (N = 2.5) to right (N = 0.5). Note that GCMI outperforms the other measures as SNR increases (from middle to right), indicating stronger SNR benefits for GCMI than for other measures.

high performance for high-SNR data was also described in the original GCMI publication (Ince et al., 2017). While all measures benefit to some extent from SNR-increases (albeit none as much as GCMI), this benefit is considerably lower for PLV (grey) and entropy (purple) compared to the other measures. Interestingly, the SNR dependence of performance increase is rather similar across spectral estimation methods (e.g. the order of connectivity measures according to performance in the rightmost column is almost identical across spectral estimation methods (rows)). Still, the absolute D-values are very different across rows and show best performance for FT2 and BF and, as before, worst performance for FT0 and SG.

3.3. Deviation from unimodal phase distribution

Ideally, connectivity measures should be sensitive to any deviation of the phase distribution from a uniform distribution. Here, we test the specific case of a bimodal phase distribution. For the first half of the time series we simulate a zero-degree phase synchronization while the second half uses a simulation of a 180-degree phase difference between both signals. This results in a bimodal phase distribution with deviation from a uniform distribution at opposite sides of the circular phase space. Clearly, all connectivity measures except entropy (shown in purple) fail to capture this more complex phase dependency (see Fig. 4). Given the definition of these measures, this result is not surprising: In all measures (except entropy) the opposite phase differences across the unit circle lead to cancellation and result in a non-detectable phase synchronization. Entropy instead quantifies any deviation from a uniform distribution in phase bins across the unit circle and therefore captures this bimodal phase distribution. However, as we can see from the previous section, this sensitivity to more complex deviations from a uniform distribution leads to a reduced sensitivity for unimodal phase distributions (see Figs. 1 and 2).

3.4. Real data

Next, we compare the same combinations of spectral estimation and connectivity methods in real data. Before proceeding to group analysis, we study speech envelope to MEG connectivity spectra in a single 9-min long data set. Fig. 5 shows the results following the same computations

and plotting format as in our simulated data. Results are generally consistent with our findings from simulated data (see Fig. 1). Overall, best performance can be seen for FT2 and WPPC (third row from the top, red) followed by GCMI (yellow) and the Rayleigh test (blue). Interestingly, this computation on real data shows that the spectral structure is mostly determined by the spectral estimate and not so much by the connectivity method (i.e., spectra in a row are more similar than spectra in a column). Obviously, there is more spectral structure in real data than in the simulated data where only a single spectral peak was evident. Not surprisingly, this spectral structure is mostly lost in FT2 due to the spectral smoothing of +/- 2Hz. Instead, the highest complexity of spectral structure can be seen using the continuous wavelet transform (CWT, third row from the bottom) and still leads to high sensitivity (large D-values) compared to FT2. CWT is therefore probably most appropriate when preservation of the spectral structure is important for the research question at hand. However, the ‘true’ spectral structure of the data is unknown so we cannot evaluate and compare the performance of spectral estimation measures in this regard.

3.5. Group statistics

In the previous sections we have exclusively used single simulated or real data sets to compare performance of different spectral estimation and connectivity techniques. In our final analysis we will now extend this approach to group analysis. For data from 20 participants, we repeat the computations shown in Fig. 5. Specifically, the different connectivity spectra were computed and normalised with their individual surrogate distribution (as for the simulated data). To identify significant effects at the group level we statistically compare connectivity spectra against the 95th percentile of the surrogate distribution. The computation used non-parametric cluster-based permutation tests as implemented in FieldTrip with 2000 randomizations.

Fig. 6 shows spectra of t-values for the different combinations of spectral estimates and connectivity measures. First, comparing spectral estimates, we find that the multi-taper spectral estimate with smoothing of +/-2Hz (third row from the top) performs best, followed by the band-pass filter (second row from the bottom). The comparison of connectivity measures (different columns) shows markedly smaller differences in group results than in the single data sets. Surprisingly, PLV (grey) per-

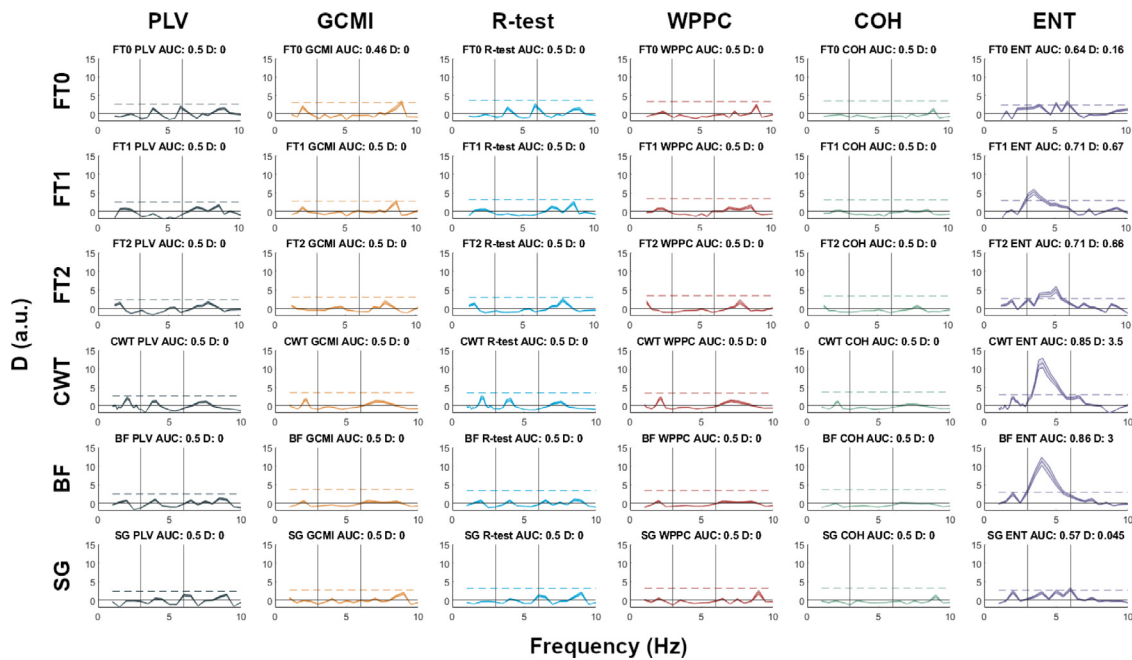


Fig. 4. Deviation from unimodal phase difference distribution. The layout is the same as in Fig. 1 with rows showing spectral estimates and columns showing connectivity measures. The underlying data lead to a bimodal phase distribution that is only detected by the entropy measure (purple, rightmost column).

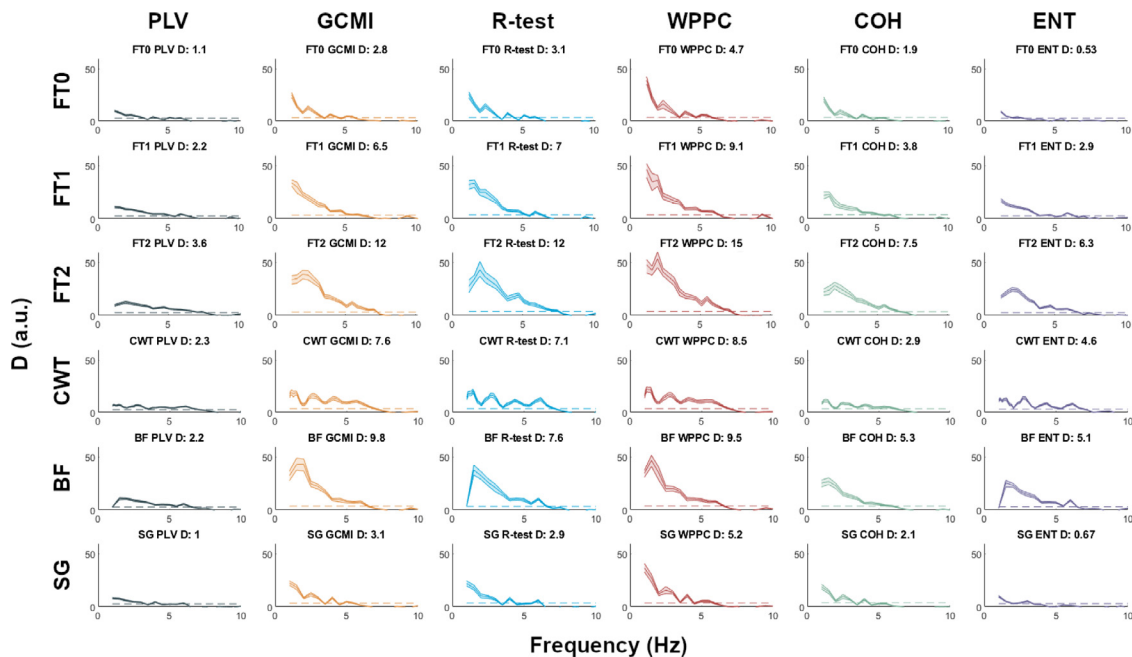


Fig. 5. Results for 9-min long MEG recording. The layout is the same as in Fig. 1 with rows showing spectral estimates and columns showing connectivity measures. Results are generally consistent with those from simulated data (see above). Note that the spectral structure is mostly determined by the spectral estimate and not so much by the connectivity method (i.e., spectra in a row are more similar than spectra in a column).

forms much better in group statistics compared to the single simulated and real data sets. Overall, in our group analysis, the choice of spectral estimation method appears to be more important than the connectivity measure.

4. Discussion

In this study we aim to demonstrate how the sensitivity to detect cerebro-peripheral connectivity is affected by different combinations of spectral estimates and connectivity measures. Results from simulated

and real data reveal conclusively that the selection of methods can facilitate or preclude the detection of significant connectivity, both at the individual and the group level.

Spectral estimates and connectivity measures interact with each other in non-trivial ways. For a given spectral estimate the available information about the underlying synchrony is utilized by different connectivity measures in markedly different ways. More precisely, if phase synchronization exists in the data (as in our simulated data) the distance of estimated connectivity from the surrogate distribution varies considerably across connectivity measures.

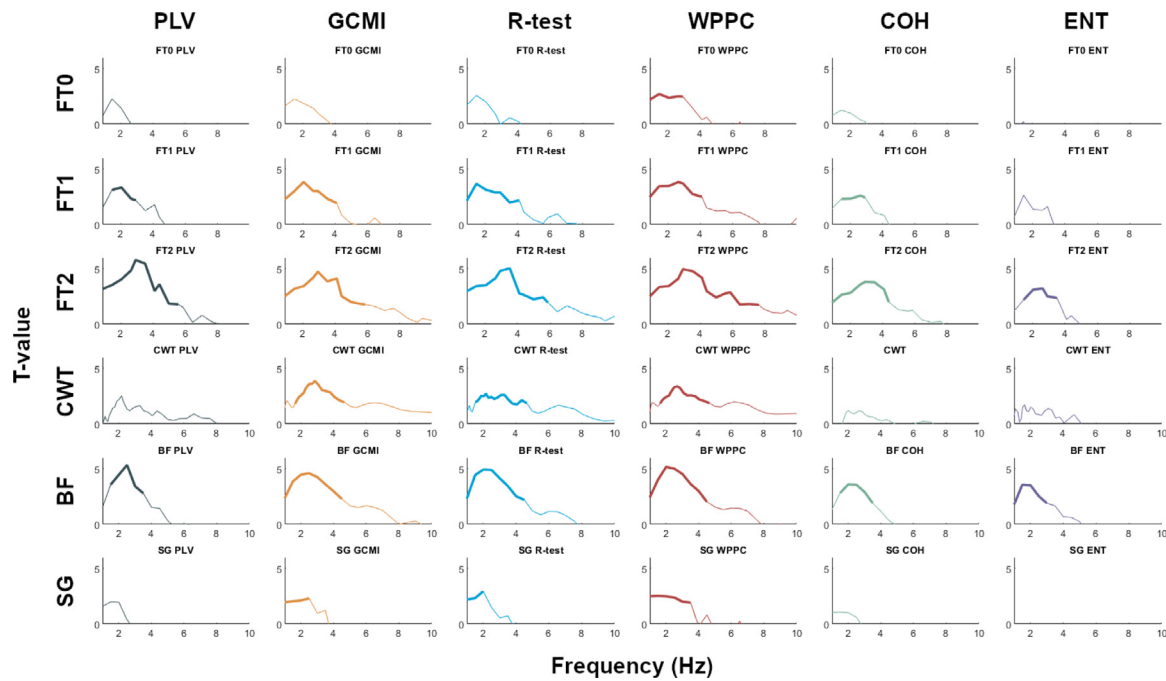


Fig. 6. Group statistics. The layout is the same as in Fig. 1 with rows showing spectral estimates and columns showing connectivity measures. T-values are plotted between 1-10 Hz. Cluster-corrected significant frequency bands are marked by increased line width. Multi-taper spectral estimate with smoothing of +/-2Hz (third row from the top) performs best, followed by the bandpass filter (second row from the bottom). Note that PLV (grey, leftmost column) performs much better in group statistics compared to the single simulated and real data sets.

Regarding spectral estimation methods, we compare different Fourier-based techniques that mostly differ in their spectral smoothing, with wavelet spectral estimates and those based on bandpass filtering followed by Hilbert transformation. Overall, highest performance was observed for FT2, the multi-taper approach with +/- 2 Hz spectral smoothing. CWT and BF performed also well and in general better than FT0 and SG. Conceptually, Fourier-based methods, Hilbert transformation, and wavelet transformation are very different, but it has been shown that - given well-chosen parameter settings - these three approaches can lead to converging results (Bruns, 2004). In our analysis, we use implementations with standard parameter settings. This might in part explain the difference in performance between FT0 and SG on the one hand, and between CWT and BF on the other hand. Both FT0 and SG reflect overlapping 2-second window FFT-based estimates, with a single Hanning taper applied to each data window. In the simulations, this resulted in 59 degrees of freedom for the spectral amplitude and phase estimates, one for each window. In comparison, both CWT and BF result in a single amplitude and phase estimate per original time point, which, even considering the large amount of redundancy for consecutive time points, likely leads to more stable estimates. Multi-taper based spectral estimation (Percival & Walden, 1993) trades spectral resolution for reduced variance in the spectral estimates, thus increasing sensitivity. This is also referred to as spectral smoothing, and is achieved by applying a set of tapers to the data, the number of which is determined by the time-bandwidth product NW , i.e. the length of the data segments (N) multiplied by the specified smoothing parameter (W). The number of tapers used is then typically $2NW-1$. In our case, as both FT1 and FT2 were implemented using 2-second long overlapping data windows, the smoothing increased the degrees-of-freedom for the spectral estimates by a factor of 3 and 7, respectively.

In general, we can expect that an analysis is optimal when the effective resolution of its spectral estimate is adjusted to the expected bandwidth of significant phase synchronization (which is unknown in real data). For example, if phase synchronization exists in a 4 Hz wide frequency band (e.g. 8-12 Hz) then multi-taper smoothing of +/- 2 Hz

should be optimal. This is largely what we observe here. However, additional factors apparently contribute to performance. For example, our simulation contains significant synchronization over a 3 Hz bandwidth. Therefore, if spectral smoothing were the only factor determining analysis performance then we would expect the +/- 1 Hz and +/- 2 Hz smoothing to perform equally well. The fact that +/- 2 Hz multitaper analysis performs better than other spectral estimates with less or no spectral smoothing indicates that the smoothing itself improves analysis sensitivity, albeit at the cost of reduced spectral resolution. Spectral resolution should be highest for CWT where different wavelets capture spectral structure even at low frequencies. Indeed, this point is nicely illustrated in Fig. 5. Whereas CWT-based connectivity spectra show separate peaks at low frequencies, these are largely merged into one for the +/- 2 Hz multitaper estimate. Since in real data the underlying spectral structure is unknown, it might be advisable to use two approaches, the FT2 computation for optimal sensitivity and CWT for optimal spectral resolution. Alternatively, longer data segments can be defined for the spectral transformation, which would then still allow for leveraging increased sensitivity of the multi-taper framework. For instance, increasing the window length from 2 seconds to 4 seconds would allow for a reduction of the smoothing parameter from 2 to 1 without compromising the number of tapers applied.

We non-exhaustively compare six different connectivity metrics aimed at capturing band-limited phase synchronization between signals. In most cases the weighted pairwise phase consistency (WPPC) outperform the other methods. The main exception is the improved performance of Gaussian-Copula based mutual information (GCMi) for data with high SNR. In general, GCMi and R-test also perform very well. Performance for coherence (COH) is overall quite good (particularly in the simulations), while performance for phase locking value (PLV) and entropy (ENT) is lowest overall. The entropy measure, however, is the only metric that proves sensitive to more complex distributions of phase differences. Here, we test the challenging case of a bimodal distribution of phase differences, with the modes of the distribution 180 ° apart, that

leads to cancellation in most methods and a failure to detect this more complex phase synchronization.

(Weighted) PPC (Vinck et al., 2010) has been proposed as a metric that provides a bias-free estimate of phase synchronization, as opposed to the more traditionally used phase locking value or coherence coefficient. Its improved performance could result from this reduced bias, possibly due to a reduction in variance of the surrogate distribution, as well as a shift towards zero. Our implementation of GCMI uses both amplitude and phase information for the estimation of the connectivity, just like WPPC and COH. R-test, PLV, and entropy only use the phase information. Obviously, the sensitivity of a particular metric is in part determined by the actual functional statistical relationship between the measured signals. If the relationship is mainly expressed in terms of the phase difference, then 'phase only' metrics will be sufficient. If the relationship is in part also expressed in terms of the amplitude correlations, then 'phase and amplitude' metrics will be more sensitive. Non-linear relationships might be more easily captured with GCMI or entropy.

Another point of practical importance for the design of cerebro-peripheral connectivity studies is the required data length. We compare performance of different combinations of spectral estimates and connectivity measures for data length between 1-9 min (see Supplementary Fig. S3). In almost all cases, the mean distance of estimated connectivity relative to the surrogate distribution increases continuously with increasing data length. However, this depends on the stationarity of the underlying connectivity. In general, statistical analysis will benefit from long recordings (see e.g. Daube et al., 2019), particularly if subtle experimental effects are to be detected and the nature of the connectivity is stationary over the observed period of time. It is worth noting that the validity and performance of cerebro-peripheral connectivity analysis does not only depend on the optimal measure but can also be affected by common artifacts in both signals. For example, if the peripheral signal is ECG then magnetocardiographic signals recorded with the MEG sensors will lead to significant connectivity that needs to be disambiguated from genuine cerebro-peripheral coupling. Note, when using EMG as peripheral signal, a suboptimally placed ground electrode in combination with nonlinearities in the recording hardware may lead to residual ECG in the EMG due to suboptimal common mode rejection. Similar issues arise with ocular MEG artifacts, when eye movement is used as peripheral signal.

Furthermore, we can speculate to what extent our findings generalise to other scenarios. For example, how are results affected by signal amplitude? Answering this question requires dedicated new simulations. It should be noted that we use measures that are based on signal phase and amplitude (GCMI, COH) or only signal phase (PLV, R-TEST, WPPC, ENT). Even those measures that use only the phase are sensitive to changes in signal-to-noise ratio. Our results do not generalise directly to the case of phase-amplitude coupling. While some measures (such as GCMI) can be applied across all possible types of connectivity (such as phase-phase, amplitude-amplitude, phase-amplitude) a comprehensive comparison of phase-amplitude measures needs to include measures that have not been used in the present study (such as the modulation index; Tort et al. 2010).

In summary, our analysis of cerebro-peripheral connectivity reveals that results depend significantly on the combination of spectral estimation and connectivity measures. Our analysis of simulated and real data provides some observations that might assist scientists in this field in making a more informed choice of analysis methods given their respective priorities. We hope that this leads to further advances in the exciting field of cerebro-peripheral connectivity analysis.

Data availability statement

All Matlab code as well as the data that support the findings of this study will be publicly accessible in full through GitHub (<https://github.com/IBiomag/>) upon acceptance. This data archive includes the code

necessary for data simulation, analysis of simulated as well as real data, and plotting of the results.

Declaration of Competing Interest

None.

Credit authorship contribution statement

Joachim Gross: Conceptualization, Methodology, Formal analysis, Resources, Writing – original draft, Writing – review & editing, Supervision, Project administration, Funding acquisition. **Daniel S. Kluger:** Conceptualization, Methodology, Investigation, Writing – original draft, Writing – review & editing, Visualization. **Omid Abbasi:** Methodology, Investigation, Data curation, Formal analysis, Writing – original draft, Writing – review & editing. **Nikolas Chalas:** Writing – original draft, Writing – review & editing. **Nadine Steingraber:** Writing – original draft, Writing – review & editing. **Christoph Daube:** Methodology, Formal analysis, Writing – original draft, Writing – review & editing. **Jan-Mathijs Schoffelen:** Conceptualization, Methodology, Formal analysis, Writing – original draft, Writing – review & editing, Project administration.

Acknowledgments

This work was supported by the Interdisciplinary Center for Clinical Research (IZKF) of the medical faculty of Münster (Gro3/001/19) and the DFG (GR 2024/5-1). This work was further supported by The Netherlands Organisation for Scientific Research (NWO Vidi: 864.14.011) to JMS. The authors would like to thank Karin Wilken, Ute Trompeter, and Hildegard Deitermann for their invaluable assistance during data collection.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at [doi:10.1016/j.neuroimage.2021.118660](https://doi.org/10.1016/j.neuroimage.2021.118660).

References

- Andrzejak, R.G., Kraskov, A., Stögbauer, H., Mormann, F., Kreuz, T., 2003. Bivariate surrogate techniques: necessity, strengths, and caveats. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* 68, 066202. doi:10.1103/PhysRevE.68.066202.
- Baillet, S., 2017. Magnetoencephalography for brain electrophysiology and imaging. *Nat. Neurosci.* 20, 327–339. doi:10.1038/nn.4504.
- Bastos, A.M., Schoffelen, J.-M., 2015. A Tutorial Review of Functional Connectivity Analysis Methods and Their Interpretational Pitfalls. *Front. Syst. Neurosci.* 9, 175. doi:10.3389/fnsys.2015.00175.
- Berens, P., 2009. circstat : a MATLAB toolbox for circular statistics. *J. Stat. Softw.* 31. doi:10.18637/jss.v031.i10.
- Bourguignon, M., Jousmäki, V., Dalal, S.S., Jerbi, K., De Tiège, X., 2019. Coupling between human brain activity and body movements: Insights from non-invasive electromagnetic recordings. *Neuroimage* 203, 116177. doi:10.1016/j.neuroimage.2019.116177.
- Bourguignon, M., Piitulainen, H., Smeds, E., Zhou, G., Jousmäki, V., Hari, R., 2017. MEG Insight into the Spectral Dynamics Underlying Steady Isometric Muscle Contraction. *J. Neurosci.* 37, 10421–10437. doi:10.1523/JNEUROSCI.0447-17.2017.
- Bruns, A., 2004. Fourier-, Hilbert- and wavelet-based signal analysis: are they really different approaches? *J. Neurosci. Methods* 137, 321–332. doi:10.1016/j.jneumeth.2004.03.002.
- Chandrasekaran, C., Trubanova, A., Stillitano, S., Caplier, A., Ghazanfar, A.A., 2009. The natural statistics of audiovisual speech. *PLoS Comput. Biol.* 5, e1000436. doi:10.1371/journal.pcbi.1000436.
- Conway, B.A., Halliday, D.M., Farmer, S.F., Shahani, U., Maas, P., Weir, A.I., Rosenberg, J.R., 1995. Synchronization between motor cortex and spinal motoneuron pool during the performance of a maintained motor task in man. *J Physiol (Lond)* 489 (3), 917–924. doi:10.1113/jphysiol.1995.sp021104, Pt.
- Daube, C., Ince, R.A.A., Gross, J., 2019. Simple Acoustic Features Can Explain Phoneme-Based Predictions of Cortical Responses to Speech. *Curr. Biol.* 29, 1924–1937. doi:10.1016/j.cub.2019.04.067, e9.
- David, O., Cosmelli, D., Friston, K.J., 2004. Evaluation of different measures of functional connectivity using a neural mass model. *Neuroimage* 21, 659–673. doi:10.1016/j.neuroimage.2003.10.006.

- Glasser, M.F., Coalson, T.S., Robinson, E.C., Hacker, C.D., Harwell, J., Yacoub, E., Ugurbil, K., Andersson, J., Beckmann, C.F., Jenkinson, M., Smith, S.M., Van Essen, D.C., 2016. A multi-modal parcellation of human cerebral cortex. *Nature* 536, 171–178. doi:10.1038/nature18933.
- Gross, J., Baillet, S., Barnes, G.R., Henson, R.N., Hillebrand, A., Jensen, O., Jerbi, K., Litvak, V., Maess, B., Oostenveld, R., Parkkonen, L., Taylor, J.R., van Wassenhove, V., Wibral, M., Schoffelen, J.-M., 2013. Good practice for conducting and reporting MEG research. *Neuroimage* 65, 349–363. doi:10.1016/j.neuroimage.2012.10.001.
- Gross, J., Hoogenboom, N., Thut, G., Schyns, P., Panzeri, S., Belin, P., Garrod, S., 2013b. Speech rhythms and multiplexed oscillatory sensory coding in the human brain. *PLoS Biol.* 11, e1001752. doi:10.1371/journal.pbio.1001752.
- Gross, J., 2014. Analytical methods and experimental approaches for electrophysiological studies of brain oscillations. *J. Neurosci. Methods* 228, 57–66. doi:10.1016/j.jneumeth.2014.03.007.
- Gross, J., 2019. Magnetoencephalography in cognitive neuroscience: a primer. *Neuron* 104, 189–204. doi:10.1016/j.neuron.2019.07.001.
- Hari, R., Baillet, S., Barnes, G., Burgess, R., Forss, N., Gross, J., Hämäläinen, M., Jensen, O., Kakigi, R., Mauguière, F., Nakasato, N., Puce, A., Romani, G.-L., Schnitzler, A., Taulu, S., 2018. IFCN-endorsed practical guidelines for clinical magnetoencephalography (MEG). *Clin. Neurophysiol.* 129, 1720–1747. doi:10.1016/j.clinph.2018.03.042.
- Ince, R.A.A., Giordano, B.L., Kayser, C., Rousselle, G.A., Gross, J., Schyns, P.G., 2017. A statistical framework for neuroimaging data analysis based on mutual information estimated via a gaussian copula. *Hum. Brain Mapp.* 38, 1541–1573. doi:10.1002/hbm.23471.
- Kluger, D.S., Gross, J., 2020. Depth and phase of respiration modulate cortico-muscular communication. *Neuroimage* 222, 117272.
- Kreuz, T., Mormann, F., Andrzejak, R.G., Kraskov, A., Lehnertz, K., Grassberger, P., 2007. Measuring synchronization in coupled model systems: a comparison of different approaches. *Physica D: Nonlinear Phenomena* 225, 29–42. doi:10.1016/j.physd.2006.09.039.
- Lachaux, J.P., Rodriguez, E., Martinerie, J., Varela, F.J., 1999. Measuring phase synchrony in brain signals. *Hum. Brain Mapp.* 8, 194–208. doi:10.1002/(sici)1097-0193(1999)8:4<194::aid-hbm4>3.0.co;2-c.
- Lakatos, P., Gross, J., Thut, G., 2019. A new unifying account of the roles of neuronal entrainment. *Curr. Biol.* 29, R890–R905. doi:10.1016/j.cub.2019.07.075.
- Lancaster, G., Iatsenko, D., Pidde, A., Ticcinelli, V., Stefanovska, A., 2018. Surrogate data for hypothesis testing of physical systems. *Phys. Rep.* 748, 1–60. doi:10.1016/j.physrep.2018.06.001.
- Le Van Quyen, M., Bragin, A., 2007. Analysis of dynamic brain oscillations: methodological advances. *Trends Neurosci.* 30, 365–373. doi:10.1016/j.tins.2007.05.006.
- Marzetti, L., Basti, A., Chella, F., D'Andrea, A., Syrjäälä, J., Pizzella, V., 2019. Brain functional connectivity through phase coupling of neuronal oscillations: a perspective from magnetoencephalography. *Front. Neurosci.* 13, 964. doi:10.3389/fnins.2019.00964.
- Meyer, L., Sun, Y., Martin, A.E., 2019. Synchronous, but not entrained: exogenous and endogenous cortical rhythms of speech and language processing. *Lang. Cogn. Neurosci.* 1–11. doi:10.1080/23273798.2019.1693050.
- Nolte, G., 2003. The magnetic lead field theorem in the quasi-static approximation and its use for magnetoencephalography forward calculation in realistic volume conductors. *Phys. Med. Biol.* 48, 3637–3652. doi:10.1088/0031-9155/48/22/002.
- Obleser, J., Kayser, C., 2019. Neural entrainment and attentional selection in the listening brain. *Trends Cogn. Sci.* 23, 913–926. doi:10.1016/j.tics.2019.08.004, Regul Ed.
- Oostenveld, R., Fries, P., Maris, E., Schoffelen, J.-M., 2011. FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput. Intell. Neurosci.*, 156869 doi:10.1155/2011/156869, 2011.
- Park, H.-D., Correia, S., Ducorps, A., Tallon-Baudry, C., 2014. Spontaneous fluctuations in neural responses to heartbeats predict visual detection. *Nat. Neurosci.* 17, 612–618. doi:10.1038/nn.3671.
- Percival, D.B., Walden, A.T., 1993. Spectral Analysis for Physical Applications. Cambridge University Press, Cambridge doi:10.1017/CBO9780511622762.
- Pfeffer, T., Keitel, C., Kluger, D.S., Keitel, A., Russmann, A., Thut, G., Donner, T.H., Gross, J., 2021. Coupling of pupil- and neuronal population dynamics reveals diverse influences of arousal on cortical processing. *BioRxiv* doi:10.1101/2021.06.25.449734.
- Quiñero, R., Kraskov, A., Kreuz, T., Grassberger, P., 2002. Performance of different synchronization measures in real data: a case study on electroencephalographic signals. *Phys. Rev. E* 65. doi:10.1103/PhysRevE.65.041903.
- Rebollo, I., Devauchelle, A.-D., Béranger, B., Tallon-Baudry, C., 2018. Stomach-brain synchrony reveals a novel, delayed-connectivity resting-state network in humans. *elife* 7 doi:10.7554/eLife.33321.
- Salenius, S., Portin, K., Kajola, M., Salmelin, R., Hari, R., 1997. Cortical control of human motoneuron firing during isometric contraction. *J. Neurophysiol.* 77, 3401–3405. doi:10.1152/jn.1997.77.6.3401.
- Schnitzler, A., Gross, J., 2005. Normal and pathological oscillatory communication in the brain. *Nat. Rev. Neurosci.* 6, 285–296. doi:10.1038/nrn1650.
- Schoffelen, J.-M., 2009. Source connectivity analysis with MEG and EEG. *Hum. Brain Mapp.* Gross, J., 30, 1857–1865. doi:10.1002/hbm.20745.
- Schoffelen, J.-M., Oostenveld, R., Fries, P., 2005. Neuronal coherence as a mechanism of effective corticospinal interaction. *Science* 308, 111–113. doi:10.1126/science.1107027.
- Schreiber, T., Schmitz, A., 2000. Surrogate time series. *Physica D: Nonlinear Phenomena* 142, 346–382. doi:10.1016/S0167-2789(00)00043-9.
- Shannon, C.E., 1948. A mathematical theory of communication. *Bell Syst. Tech. J.* 27, 379–423. doi:10.1002/j.1538-7305.1948.tb01338.x.
- Tort, A.B.L., Komorowski, R., Eichenbaum, H., Kopell, N., 2010. Measuring phase-amplitude coupling between neuronal oscillations of different frequencies. *J. Neurophysiol.* 104, 1195–1210. doi:10.1152/jn.00106.2010.
- Vinck, M., van Wingerden, M., Womelsdorf, T., Fries, P., Pennartz, C.M.A., 2010. The pairwise phase consistency: a bias-free measure of rhythmic neuronal synchronization. *Neuroimage* 51, 112–122. doi:10.1016/j.neuroimage.2010.01.073.
- Zoefel, B., 2018. Speech entrainment: rhythmic predictions carried by neural oscillations. *Curr. Biol.* 28, R1102–R1104. doi:10.1016/j.cub.2018.07.048.