Localizing Bicoherence from EEG and MEG

Forooz Shahbazi Avarvand[1], Sarah Bartz[2], Christina Andreou[3,4], Wojciech Samek[1], Gregor Leicht[3], Christoph Mulert[3], Andreas K. Engel[2], Guido Nolte[2]

[1] *Department of Video Coding & Analytics, Fraunhofer Heinrich Hertz Institute, Berlin, Germany*
[2] *Department of Neurophysiology and Pathophysiology, University Medical Center Hamburg-Eppendorf, Hamburg, Germany*
[3] *Department of Psychiatry, University Medical Center Hamburg-Eppendorf, Hamburg, Germany*
[4] *Center for Psychotic Disorders, University of Basel Psychiatric Clinics, Basel, Switzerland*

Abstract

We propose a new method for the localization of nonlinear cross-frequency coupling in EEG and MEG data analysis, based on the estimation of bicoherences at the source level. While for the analysis of rhythmic brain activity, source directions are commonly chosen to maximize power, we suggest to maximize bicoherence instead. The resulting nonlinear cost function can be minimized effectively using a gradient approach. We argue, that bicoherence is also a generally useful tool to analyze phase-amplitude coupling (PAC), by deriving formal relations between PAC and bispectra. This is illustrated in simulated and empirical LFP data. The localization method is applied to EEG resting state data, where the most prominent bicoherence signatures originate from the occipital alpha rhythm and the mu rhythm. While the latter is hardly visible using power analysis, we observe clear bicoherence peaks in the high alpha range of sensorymotor areas. We additionally apply our method to resting-state data of subjects with schizophrenia and healthy controls and observe significant bicoherence differences in motor areas which could not be found from analyzing power differences.

1. Introduction

EEG and MEG are two non-invasive and widely used techniques, which allow for the study of brain activity at a high temporal resolution at the drawback of a lower spatial resolution. While many methodological approaches and experimental studies have focused on linear behavior of signals, i.e. the cross-correlation functions in the time domain or cross-spectra in the frequency domain [Nolte et al., 2004, Nunez et al., 1997, Ansari-Asl et al., 2006], nonlinear aspects of brain dynamics, especially cross-frequency coupling, have become a recent focus of interest. One of the measures to study cross-frequency coupling is bicoherence. Formally speaking, bicoherence is the extension of coherence to the next statistical order: while coherence reflects the general second order statistical properties of stationary multivariate data, bicoherence is constructed from third order statistical moments.

The rather formal explanation of what bicoherence is makes it difficult to understand it intuitively. Like coherence, it is a measure of phase-phase coupling, where the phases of each segment or time point are weighted with their respective amplitudes. In contrast to coherence, phases can be taken from signals at different frequencies, such that bicoherence reflects a form of cross-frequency coupling. For this reason, it also provides a measure of the deviation of signals from linear dynamics since linear systems do not contain cross-frequency coupling. In fact, bicoherence in general measures the coupling of three different signals at three different frequencies with the dependence of bicoherence on both phases and amplitudes of all these signals adding to the confusion.

Because of the rather complicated construction of bicoherence it is perhaps not surprising that it received little attention from the scientific community so far. Bicoherence has been used in several, but rather few, clinical- and research applications, including studies examining intracranial EEG in sleep, wakefulness and seizures [Bullock et al., 1997] and studies regarding the differentiation of anesthetic levels [Watt et al., 1996]. Further examples of bicoherence analyses can be found in research of neurogenic pain, epilepsy and movement disorders [Sarnthein et al., 2003].

A much more prominent measure of cross-frequency-coupling is Phase-Amplitude-Coupling (PAC) where the amplitude of a high frequency oscillation is related to the phase of a low frequency oscillation [Canolty et al., 2006, Osipova et al., 2008, Tort et al., 2010]. At first sight it seems that the concepts of bicoherence, being a weighted measure of phases, and phase-amplitude are very different. However, using instructive examples it was pointed out by Hyafil [2015] that PAC is related to bicoherence. In this paper, we will go beyond specific examples and show how PAC can be calculated from bicoherence in general terms.

Both PAC and bicoherence are sensitive to non-sinusoidal waveshapes. This phenomenon typically leads to two types of approaches: While it elicits warnings and suggestions for the usage of PAC on how to avoid 'spurious' phase-amplitude coupling [Jensen et al., 2016], an increasing interest emerges in these non-sinusoidal properties of waveshapes, also known as higher harmonics, as a potentially informative signature of brain dynamics [Cole and Voytek, 2017, Cole et al., 2017]. In human EEG resting-state recordings, higher harmonics of alpha oscillations are a strong neural source of bicoherence, and rather than dismissing this as uninteresting, we study this phenomenon. Like other measures of coupling between different neural sites or sensors, bicoherence is prone to artifacts of volume conduction. In such cases, neuronal coupling between different brain areas is confused with phenomena caused by an incomplete demixing of source signals. This problem was tackled by Chella et al. [2014], who introduced the antisymmetric part of bicoherence as a measure insensitive to these artifacts. A corresponding modification for PAC doesn't exist, which we consider as a conceptual advantage of bicoherence.

The main goal of this paper is to localize bicoherence in the brain. In particular, the proposed method estimates the dipole direction at each voxel, such as to maximize the resulting univariate bicoherence magnitude. In section 2, the mathematical details of this method are presented. After recalling the definition of bicoherence in section 2.1 we will derive general relations between PAC and bicoherence in section 2.2. In section 2.3 we will present the mathematical details of the localization. In section 3 we present results for various applications: we show typical results for univariate bicoherence in EEG resting state data, we illustrate bicoherence for simulated higher harmonics, for simulated phase amplitude coupling and for empirical LFP data containing phase amplitude coupling, we demonstrate the idea of the maximization procedure in simulations, we add empirical evidence to justify some technical details, we localize the first harmonic of alpha rhythms in EEG resting state data, and we show differences between patients with schizophrenia and healthy controls. Finally, section 4 covers the discussion and conclusions of the study.

2. Methods

2.1. Definition of bicoherence

Cross-bispectra are the general third order statistical moments of data in the frequency domain, defined as

$$D_{ijk}(f_1, f_2) = \langle z_i(f_1) z_j(f_2) z_k^*(f_1 + f_2) \rangle \,, \tag{1}$$

where $z_i(f)$ is the Fourier coefficient of the data in channel (or source) i at frequency f during some time segment, and $\langle \cdot \rangle$ denotes the expectation value, which is approximated by an average over segments. The frequency of the third signal, z_k in the above formula, is constrained to be the sum of the first two frequencies, because for stationary data, i.e. data like resting state data without intrinsic clock as given e.g. by a stimulus, all other choices for the third frequency result in vanishing cross-bispectra. Analogous to coherence, which is the normalized version of a cross-spectrum, bicoherence is the normalized version of a cross-bispectrum

$$B_{ijk}(f_1, f_2) = \frac{D_{ijk}(f_1, f_2)}{N_{ijk}(f_1, f_2)} \tag{2}$$

While the appropriate normalization for coherence, taken as the square root of the product of powers, is undebated, the situation is less clear for bicoherence. For this paper, two normalizations are relevant: The 'classical normalization', essentially constructed from 2-norms, reads

$$N_{ijk}^{two}(f_1, f_2) = \left\langle |z_i(f_1)|^2 \right\rangle^{1/2} \left\langle |z_j(f_2)|^2 \right\rangle^{1/2} \left\langle |z_k(f_1 + f_2)|^2 \right\rangle^{1/2}. \tag{3}$$

Including the power of the third signal in the normalization appears to be the natural generalization from second- to third order statistics. On the other hand it has the drawback that in this case the absolute value of bicoherence is not constrained by one, and the interpretation of bicoherence is rather difficult. A different normalization was suggested by Shahbazi et al. [2014] as

$$N_{ijk}^{three}(f_1, f_2) = \left\langle |z_i(f_1)|^3 \right\rangle^{1/3} \left\langle |z_j(f_2)|^3 \right\rangle^{1/3} \left\langle |z_k(f_1 + f_2)|^3 \right\rangle^{1/3}, \tag{4}$$

which bounds the absolute value of bicoherence by one. In the context of this paper, a drawback of the latter normalization in Eq. 4 is that it cannot be calculated in source space from corresponding low order statistical moments in sensor space. (In contrast, e.g. using a linear inverse method, a cross-spectrum in source space can be calculated easily by applying the spatial filter on the cross-spectrum in sensor space. This includes the normalization, constructed from the diagonal elements of the cross-spectrum, to arrive at coherence.) This normalization must therefore always be calculated using the entire raw data in source space which would be extremely slow. As will be explained in more detail below, we will use both normalization. For technical reasons we will use Eq. 3 to find a source direction and then we will recalculate bicoherence using Eq. 4. To avoind confusion, we recall that in this paper we only calculate univariate bicoherence.

2.2. Bicoherence versus Phase-Amplitude Coupling

Bicoherence is a measure of phase coupling between signals at three different frequencies, for which the phases are weighted with their respective amplitudes. Conceptually, this appears to be very different from the frequently used phase-amplitude coupling (PAC), which is constructed to measure the relationship between the amplitude of a high frequency signal and the phase of a low frequency signal.

As pointed out by Hyafil [2015], the strong relation between the two measures can be illustrated by the phenomenon of 'beating' in acoustics: the superposition of two high frequency sinusoids with a small frequency difference, leads to a slow variation of amplitudes depending on the phase difference at a specific time point. If this phase difference is coupled to the phase of a low frequency signal, we observe both, phase-amplitude coupling and coupling between all three phases as measured by bicoherence. Following this line of reasoning, we derive a formal relation between the two coupling measures, showing that the described relation between PAC and bicoherence is generally valid. Furthermore, we will give arguments for the superiority of bicoherence.

Formal definitions of PAC vary in the literature, but the principal goal is the same. In [Canolty et al., 2006] the essential quantity is

$$P = \left| \langle |x_H(t)| \, e^{i\Phi_L(t)} \rangle_t \right| = \left| \langle |x_H(t)| \, \frac{x_L(t)}{|x_L(t)|} \rangle_t \right|, \tag{5}$$

where $x_L(t)$ and $x_H(t)$ are the (complex[1]) Hilbert transforms of a signal filtered in a low and high frequency band. The corresponding phases and magnitudes at time t are denoted by $\Phi_{L/H}(t)$ and $\left|x_{L/H}(t)\right|$, and $\langle\cdot\rangle$ denotes average over time. From P and corresponding values of that quantity for surrogate data, P_s , where the low and a high frequency parts are shifted relative to each other by a random delay, PAC can be expressed as a z-score

$$PAC = \frac{P - mean(P_s)}{std(P_s)} \tag{6}$$

[1]In mathematics only the imaginary part is called Hilbert transform.

where $mean(P_s)$ and $std(P_s)$ are the mean and standard deviations of P_s calculated of N surrogate data sets.

One cannot relate P of Eq. 5 exactly to a bispectrum because P is not a third order statistical moment. To formulate an exact relation a slight modification is necessary:

$$\tilde{P} = \left| \langle x_L(t) |x_H(t)|^2 \rangle_t \right| , \tag{7}$$

where the phase of $x_L(t)$ is now weighted with the corresponding amplitude, and the amplitude of $x_H(t)$ is squared.

For ease of notation we derive the relation for the univariate case. The bivariate case is completely analogous. Let $x(t)$ be the original data for integer times $-T \le t \le T$ with $\hat{x}(f)$ being its Fourier transform, and let $F_L(f)$ and $F_H(f)$ be the low and high pass filters in the frequency domain. Then we can write the Hilbert transform of the filtered data

$$x_{L,H}(t) = 2 \sum_{f=1}^{f_N - 1} F_{L,H}(f)\hat{x}(f) \exp\left(\frac{i2\pi f t}{2T+1} \right) \tag{8}$$

with f_N being the Nyquist frequency. For simplicity of notation we assumed that the data is zero mean and that the amplitude vanishes at the Nyquist frequency. The Hilbert transform is then just the inverse Fourier transformation omitting half of the frequencies (negative or above Nyquist frequency depending on convention) and multiplying the signal of the other frequencies by a factor 2 such that the real part of the Hilbert transform is identical to the original filtered data. These factors of 2 cancel with any normalization and will be omitted in the following. In the derivation below, we will make use of the relation

$$\sum_{t=-T}^{T} \exp\left(\frac{i2\pi f t}{2T+1} \right) = (2T+1)\delta_{f,0}, \tag{9}$$

where $\delta_{i,j}$ is the Kronecker delta symbol.

In addition, we will assume an infinite number of time points, i.e. the limit $T \to \infty$, and ergodicity, i.e. that infinite time averages are identical to ensemble averages. In this case, any such time average can trivially also be written as an ensemble average of that time average, i.e. we not only have one time series of infinite length, but we have an infinite number of them taken from an ensemble. This can be formally written as $\langle \cdot \rangle_t = \langle \langle \cdot \rangle_t \rangle$. This simplifies the derivation because now we can calculate expectation values of the quantities in the Fourier domain. Recalling that $|x|^2 = xx^*$, we can put things together:

$$\begin{aligned}
\tilde{P} &= \left| \langle x_L(t) |x_H(t)|^2 \rangle_t \right| = \left| \langle \langle x_L(t) \left| x_H(t) \right|^2 \rangle_t \rangle \right| \\
&= \left| \sum_{f_1, f_2, f_3} F_L(f_1) F_H(f_2) F_H(f_3) \langle \hat{x}(f_1)\hat{x}(f_2)\hat{x}^*(f_3) \rangle \frac{1}{2T+1} \sum_t \exp\left\{ \frac{i2\pi t(f_1 + f_2 - f_3)}{2T+1} \right\} \right| \\
&= \left| \sum_{f_1, f_2, f_3} F_L(f_1) F_H(f_2) F_H(f_3) \langle \hat{x}(f_1)\hat{x}(f_2)\hat{x}^*(f_3) \rangle \delta_{f_1+f_2, f_3} \right| \\
&= \left| \sum_{f_1, f_2} F_L(f_1) F_H(f_2) F_H(f_1 + f_2) \langle \hat{x}(f_1)\hat{x}(f_2)\hat{x}^*(f_1 + f_2) \rangle \right| \\
&= \left| \sum_{f_1, f_2} F_L(f_1) F_H(f_2) F_H(f_1 + f_2) D(f_1, f_2) \right|
\end{aligned} \tag{10}$$

Note, that $\tilde{P}$ is calculated from the bispectrum $D(f_1, f_2) = \langle \hat{x}(f_1)\hat{x}(f_2)\hat{x}^*(f_1 + f_2) \rangle$ using the filters F_L and F_H. In general, such filters cause smearing across frequencies with details depending on the width of filters, such that narrow band features, like the alpha rhythm and its higher harmonics, might be difficult to be identified in PAC.

In [Hyafil, 2015] it is argued, that bicoherence should not be used because it lacks specificity as it depends not only on the phases of the low frequency signal but also on the amplitudes. However, also measures of PAC based on coherence between low frequency signal and high frequency amplitude [Osipova et al., 2008] contain a weighted phase. In fact, the difference of these coupling measures using weighted or unweighted phases is analogous to the difference between coherence and phase locking value (PLV). While there are theoretical differences between these measures, in practice these differences are very small [Nolte et al., 2004]. Even though phases and amplitudes are more intuitive than complex numbers that doesn't mean that the separation of phases and amplitudes accurately reflects brain dynamics. E.g., formulating brain dynamics as dynamics of phases is an effective simplification and approximation of weakly coupled nonlinear systems, but we do not see a reason why the brain should completely ignore an amplitude of a signal even if its phase is more relevant. We would hence question the need to be specific, and we see a couple of advantages of bicoherence over PAC:

1. Bicoherence can be calculated in principle with an arbitrary frequency resolution, whereas PAC requires wide band filters for the high frequency signal [Aru et al., 2015], in general leading to smearing across frequencies as was shown here formally. This makes it more difficult to detect higher harmonics of narrow band signals as such, whether or not those higher harmonics are the objective of the study (as is done here) or are considered as a confounder.
2. Even though bicoherence was studied as a univariate measure in this paper, the full multivariate formulation allows to use it as a coupling measure of brain interactions between different brain regions. For EEG and MEG, such measures are prone to 'artifacts of volume conduction', where incomplete demixing of independent source is falsely interpreted as brain interaction. This problem can be solved using antisymmetric combinations of cross-bispectra [Chella et al., 2014], but a corresponding solution for PAC does (so far) not exist.
3. Bicoherence is much faster to calculate because the PAC requires filtering of the raw data not only for each high frequency but in addition this filter should depend on the low frequency [Aru et al., 2015] and the data need to be refiltered many times. Also, our approach to calculate the cross-bispectrum only once in sensor space and then to map it into source space using a linear inverse method was only possible because the cross-bispectrum is of low statistical order. Already a generalization to a fourth order moment would not be practical and an analogous approach using PAC would not be possible.
4. As a multivariate measure, bicoherence is more general than PAC as it quantifies interactions in general at three different sites. PAC appears to be a special case by setting two of the three sites to be equal.

2.3. Localization of bicoherence in source space

The proposed method aims at the localization of cross-frequency interaction from EEG/MEG data, by determining dipole directions of single voxels as the directions which maximize the absolute value of bicoherence. For this case we omit the voxel index j and only consider one source direction denoted by α_k for $k = 1, 2, 3$. Source activity is then estimated from the activity at channel i i.e. x_i

$$z(f) = \sum_k \alpha_k y_k(f) = \sum_{k,i} \alpha_k A_{ik} x_i(f), \tag{11}$$

where A_{ik} is the weight matrix, constructed from an inverse method such as eLORETA, to project the channel data to a voxel at a specific direction k.

For the univariate bispectrum

$$D(f_1, f_2) =< z(f_1)z(f_2)z(f_1 + f_2)^* > \tag{12}$$

we then get

$$D(f_1, f_2) = \sum_{pqr} \alpha_p \alpha_q \alpha_r E_{pqr}(f_1, f_2) \tag{13}$$

where

$$E_{pqr}(f_1, f_2) = \sum_{lmn} A_{lp} A_{mq} A_{nr} G_{lmn}(f_1, f_2) \tag{14}$$

and

$$G_{lmn}(f_1, f_2) = E[x_l(f_1) x_m(f_2) x_n^*(f_1 + f_2)]. \tag{15}$$

Here, G is the cross-bispectrum in sensor space and E is the cross-bispectrum in source space for a given voxel for all three orthogonal dipole directions. The strategy here is to calculate G from the raw data once in sensor space for a given pair of frequencies, and then calculate E for a given voxel. Such an approach is not possible for the normalization using the 3-norms, and therefore we use the 2-norms arriving at

$$B(f_1, f_2) = \frac{D(f_1, f_2)}{(P(f_1)P(f_2)P(f_1 + f_2))^{1/2}} \tag{16}$$

where

$$P(f_1) = \alpha^T \hat{C}(f_1) \alpha \tag{17}$$

with $\hat{C}(f_1)$ being the 3×3 cross-spectrum at the specific voxel. Let C(f) be the cross-spectral matrix in sensor space, i.e., $C(f) = E[x(f)x^\dagger(f)]$ in $\mathbb{C}^{N \times N}$ with N being the number of channels and $\dagger$ denoting transposition and complex conjugation. Then

$$\hat{C}(f_1) = A^T C(f_1) A \tag{18}$$

and we maximize the square of the absolute value of $B(f_1, f_2)$, i.e., $L(f_1, f_2, \alpha)$

$$L(f_1, f_2, \alpha) = \frac{|D(f_1, f_2)|^2}{P(f_1)P(f_2)P(f_1 + f_2)} = \frac{\left|\sum_{pqr} \alpha_p \alpha_q \alpha_r E_{pqr}(f_1, f_2)\right|^2}{\alpha^T \hat{C}(f_1)\alpha \alpha^T \hat{C}(f_2)\alpha \alpha^T \hat{C}(f_1 + f_2)\alpha} \tag{19}$$

with respect to the orientation α. This optimization cannot be solved analytically. We employ a gradient approach for which the gradient can be calculated analytically with details presented in the appendix.

Once a source direction was found using the 2-norms for normalization, bicoherence can be recalculated using the 3-norms. In principle it is conceivable, that optimization with respect to the 3-norm leads to a different direction. However, such optimization is computationally too costly and we found in many examples, which will partly be shown below, that the choice of norm has almost no effect on the orientation.

3. Applications and Results

3.1. Data

We apply our method to EEG recordings provided by the Department of Psychiatry of the University Medical Center Hamburg-Eppendorf, and we recall the description of the data from Andreou et al. [2015a,b]. The data comprises continuous resting-state recordings (5-10 minutes, sampled at 1kHz, eyes closed) of 22 patients with first-episode schizophrenia and 24 healthy controls. Patients were recruited through the Psychosis Center of the Department of Psychiatry, while controls were taken from the general public according to predefined inclusion/exclusion criteria regarding their medical history[Leicht et al., 2015]. The data was recorded using 64 Ag/AgCl electrodes positioned according to the 10-20 system with additional electrode positions AF7,AF3, AF4, AF8, F5, F1, F2, F6, F10, FT9, FT7, FC3, FC4, FT8, FT10, C5, C1,C2, C6, TP7, CPz, TP8, P5, P1, P2, P6, PO3, POz and PO4 mounted on an EEG cap (ActiCaps, Brain Products, Munich, Germany), Impedance was kept below 5 k throughout the experiments and EEG data was recorded using the Brain Vision Recorder software version 1.10 (Brain Products, Munich, Germany). Postprocessing involved ICA decomposition, artefact-removal by visual inspection and application of a 0.1-70 Hz bandpass filter. Furthermore, the data was down-sampled to 256Hz and re-referenced to the common average reference.

3.2. Illustrative results for empirical and simulated data

3.2.1. EEG data

In Fig.1 we show typical bicoherence patterns from selected single EEG channels. The most dominant patterns of bicoherence (upper row)) are sharp peaks at a base frequency (around 10 Hz) and higher harmonics. Note, that the peak at f_1=11 Hz and f_2=11 Hz corresponds to a coupling of 11 Hz and $f_1 + f_2$=22 Hz, and represents a coupling between the base frequency and its first higher harmonic. This peak can be observed in essentially all subjects, while even higher harmonics (e.g. atf_1=11 Hz and f_2=22 Hz) are more prominent in motor alpha as shown in this example. These bicoherence patterns are clearly caused by non-sinusoidal properties of alpha rhythms in visual and sensorimotor areas, as we will show in more detail below.

The most relevant confounder for bicoherence are heart artifacts, as shown in the second row which can easily be confused with theta-gamma or theta-beta coupling. The heart beat itself can be seen in the raw data shown in the left panel. These are sensor data, which do not isolate the effects, and this channel also contains alpha rhythm, but this alpha rhythm does not dominate the bicoherence pattern.

Also, muscles produce bicoherence patterns as shown in the third row which can typically be observed in electrodes near the neck. The fourth row shows a structered bicoherence pattern, which we believe to originate also from muscles. It is typically observable in temporal and central channels in around 20% of the subjects. We also found such a signal in EEG measurements using different EEG devices and, to a lower extent, also in MEG data. We therefore do not believe that this is a technical artifact. However, we were not able to produce this signal by teeth clenching, and addressing this signal to muscle artifacts might be premature.

3.2.2. LFP snd simulated data

To our opinion, the only neuronal origins of non-vanishing bicoherence observable in EEG resting state data are higher harmonics of the alpha rhythm. Such a pattern can be produced qualitatively using an artifical non-sinusoidal wave shape. As en example we used a sawtooth signal with a period od 50 ms, i.e. a linear decreasing signal which is repeated every 50 ms, with additional white noise such that the sawtooth and the noise have equal variances. We simulated 10 minutes of data with a sampling rate of 500 Hz. For this case we used a high frequency resolution of 0.5 Hz. An illustrative portion of the original time series and the resulting absolute value of bicoherence are shown in the left panels of Fig.2. In the bicoherence plots we can observe the higher harmonics as sharp peaks at 20 Hz and multiples of it.

To illustrate phase amplitude coupling, we simulated a coupling between the phase of a theta oscillation and the amplitude of a broad band gamma activity. To get the low-frequency signal, we narrowly filtered white noise at 7 Hz with a band width of 1 Hz. From this signal, $x(t)$, we calculated its phase, $\Phi(t)$, using the Hilbert transform. The high frequency signal, $y(t)$, was constructed as white noise filtered between 40 Hz and 80 Hz. The phase-modulated high frequency signal was constructed as

$$z(t) = (1 - \cos(\Phi(t))y(t) \tag{20}$$

and the final signal was constructed as

$$u(t) = 2\frac{x(t)}{\sigma_x} + \frac{z(t)}{\sigma_z} + \eta(t) \tag{21}$$

with $\eta(t)$ being white noise with unit standard deviation and σ_x and σ_z being the standard deviations of x and z, respectively. White noise was added to avoid leakage artifacts at frequencies containing

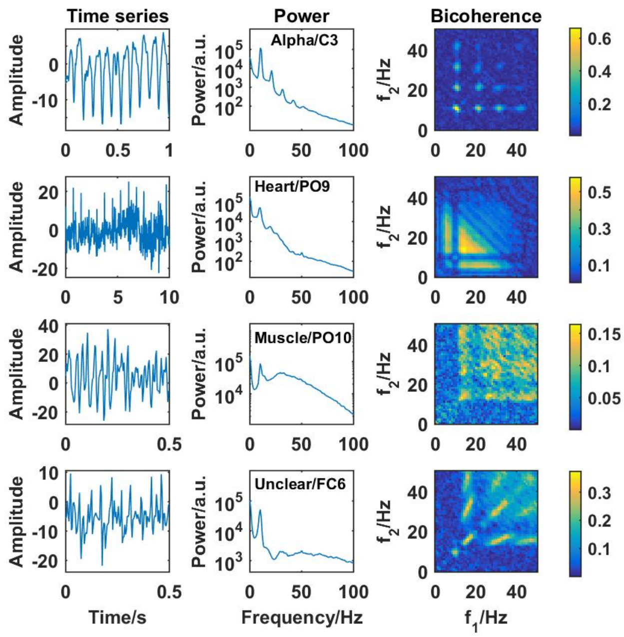

Figure 1: Typical pattern of selected channels and subjects of resting state data. We show raw data of selected time segments (left column), power (middle column) and absolute value of bicoherence (right column). Only the upper row panel, showing a bicoherence pattern of an alpha rhythm, is clearly of neural origin. The origin of the structure shown in the fourth row is not entirely clear to us but we assume that it is a muscle artifact. The name of the EEG sensor and the origin, to which we address the patterns, are noted in the power plots.

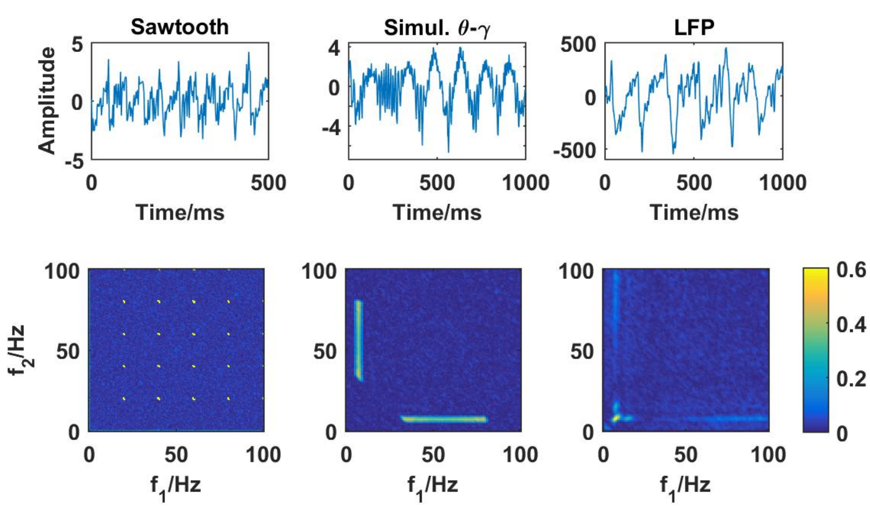

Figure 2: Raw data (top row) and absolute value of bicoherence (bottom row) of simulated and empirical LFP data. Left column: artifical sawtooth signal with a period of 50 ms and additive white noise. Middle column: Simulated Theta-Gamma phase amplitude coupling. Right column. LFP data of a rat.

no signal, and the factor 2 in front of the low frequency signal was included only to better visualize the raw data. Results are shown in the middle panels of Fig.2. For bicoherence we observe a narrow strip corresponding to the coupling between the narrow band theta rhythm and the wide band gamma rhythm.

We finally present results for empirical data. Scheffer-Teixeira and Tort [2016b] reported a coupling between the amplitude of the gamma rhythm and the phase of the theta rhythm in LFP data of the hippocampus of rats during maze exploration and REM sleep (and also analyzed phase-phase coupling). The data for REM sleep is publicily available [Scheffer-Teixeira and Tort, 2016a], and we reanalyzed these data with bicoherence. An illustrative example is shown in the right column of Fig.2. The theta-gamma coupling can be recovered with bicoherence as a thin line at around f_1=7 Hz going from around f_2=50 Hz up to 100 Hz (and vice versa). We only show results up to 100 Hz, but in fact, this line goes up to around 150 Hz. A similar structure can be observed for around half of the sensors of all rats. In contrast to the previous simulation we can also observe separated peaks at low frequencies corrsponding to higher harmonics of the empirical theta rhythm.

3.3. Power Maximization versus Bicoherence Maximization

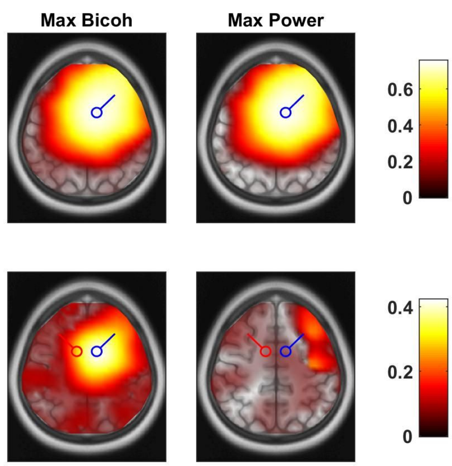

Figure 3: Top row: Estimate of univariate bicoherence for a simulated single source with strong bicoherence plus additive noise found by maximizing bicoherence (left) and by maximizing power (right). Bottom row: an additional strong source having vanishing bicoherence was placed in the vicinity of the first source with orthogonal dipole direction.

Before applying the method to the empirical EEG data we evaluated the performance on simulated data. The goal of this simulation is to illustrate the idea of the proposed method by constructing a case where optimization of source orientation with respect to power misses the bicoherence in source space. For this, a single source with strong bicoherence between 10 Hz and 20 Hz is simulated and projected to the sensors. The source activity was constructed from white noise narrow band filtered at 10 Hz. The filtered noise, $f(t)$, was squared with mean subtracted, $g(t) = f^2(t) - < f^2(t) >$, and then the bicoherent source was taken as

$$h(t) = \frac{f(t)}{std(f(t))} + \frac{g(t)}{std(g(t))} \tag{22}$$

with $std(\cdot)$ denoting standard deviation. This construction is rather trivial and serves only the purpose to illustrate the reasoning behind the proposed method. For a source without bicoherence we simply omitted the second term on the right hand side of the last equation.

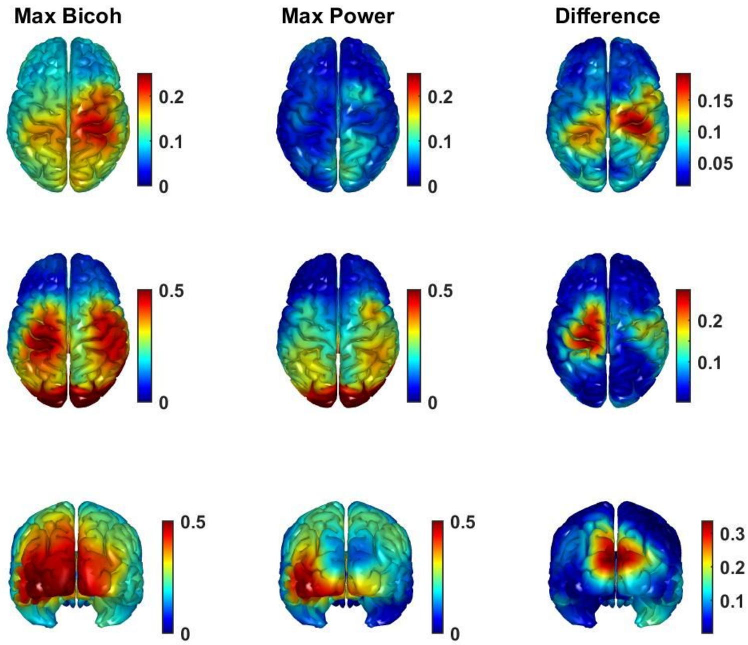

Figure 4: Each row is an example from empirical data where univariate bicoherence was estimated with dipole direction chosen to maximize bicoherence (left column) and power (middle column). The difference is shown in the right column. In all these cases we set $f_1 = f_2$ and f_1 was chosen from the alpha band.

The forward model estimation is based on the method described in [Nolte and Dassios, 2005]. The sampling frequency is 200 Hz and the total duration was 300 seconds. Small independent and equally distributed random noise was simulated on all voxels and mapped into sensor space. Bicoherence is estimated once by the dipole directions set to maximize power and once to maximize bicoherence. To avoid confusion, we emphasize that in both cases bicoherence is calculated. The difference is only the criterion to find the source direction. The results are shown in Fig.3. In the first row we simulated two sources each with large bicoherence. In the second row a second source with vanishing bicoherence but strong power at 10 Hz is added to the simulations and the bicoherence is again estimated using the two methods. Apparently, the maximization of bicoherence gives a proper estimation of the source with strong bicoherence. However, the localization using the maximization of power fails to localize the bicoherent source. The reason can be attributed to the non-uniqueness of the inverse solution. The activity of the strong source has a large impact at the location of the weak nonlinear source. Optimization with respect to power then leads to a source orientation corresponding to the strong source also at the location of the weak source.

To illustrate differences for real data, we have compared the case in which the dipole directions, α in Eq.11, are estimated by the maximization of the local power in each grid point instead of the maximization of bicoherence. Bicoherence is then estimated in each grid point and plotted in Fig.4. The column on the left shows the magnitude of bicoherence for dipole directions obtained from the maximization of bicoherence and the figures in the middle are the bicoherence values obtained from the dipole directions maximizing the power. The differences between these two are plotted in the right column. Each row is an example calculated from empirical data from a single subject where f_1=f_2=11 Hz. We observe a large difference between the two columns. These results suggest that the dipole directions corresponding to the largest power do not necessarily correspond to the directions of the maximum bicoherence. We emphasize that we picked these examples for illustration. In most cases, we found similar results for the two optimization strategies when applied to alpha-beta coupling.

3.4. Comparison of norms

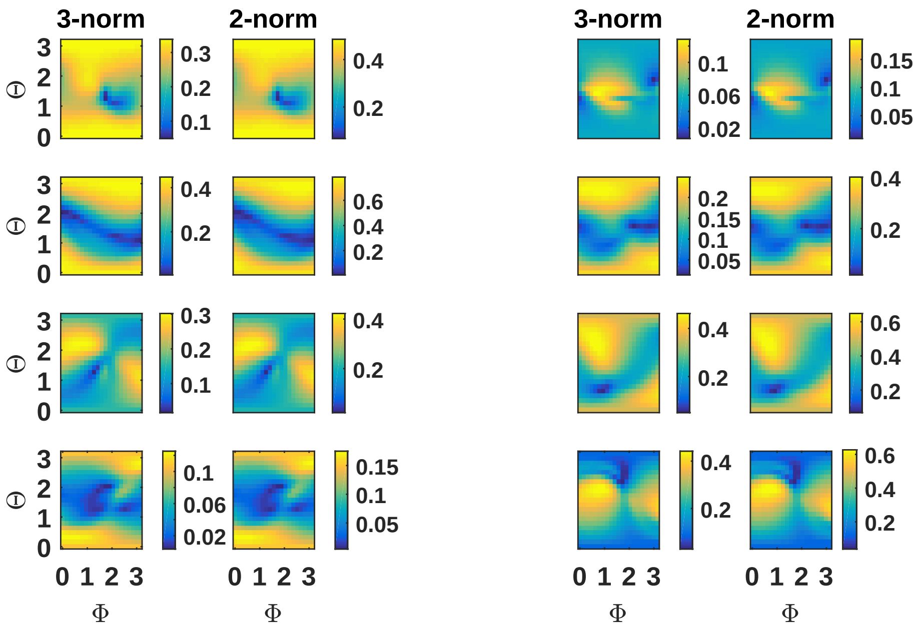

Figure 5: Absolute value of univariate Bicoherence as a function of source direction for illustrative cases using 2-norm and 3-norm normalization for 8 random voxels of one subject. In all examples we set $f_1 = f_2 = 10$ Hz. The first and the third column show results for bicoherence using the 3-norm, and the second and fourth column show the corresponding results calculated with the 2-norm. As we see, the patterns are almost identical apart from a global scale.

As was mentioned before, the optimization of bicoherence was computationally too costly using the 3-norm for normalization and therefore we replaced it by a 2-norm. To see how this modification affects the maximum values of bicoherence we estimated this value for 8 randomly chosen voxels both for the 2-norm and 3-norm for f_1=f_2=10 Hz on the empirical data set from a single subject. For each voxel, we discretized the angles in an orientation vector $\alpha = (\sin(\Theta)\cos(\Phi), \sin(\Theta)\sin(\Phi), \cos(\Theta))^T$ with 21 values between 0 and π for each polar angle. For each orientation we calculated bicoherence with either normalization. Results are plotted in Figure 5. There is a strong similarity between patterns for the same voxel and maxima occur at similar angles even though the amplitudes are different for the two nomalizations. That illustrates that the 2-norm does not change optimization results for bicoherence beyond a negligible amount.

To analyze this in a statistically systematic way, we calculated bicoherence at 20 random voxels for each of the 24 healthy controls at the frequency pair corresponding to dominant peak of alpha-beta coupling. First, for each voxel we calculated and maximized the absolute value of bicoherence using 3-norm normalization across the 441 orientations. Second, we maximized the absolute value of bicoherence with 2-norm normalization across the same 441 orientations and calculated the absolute value of bicoherence with 3-norm normalization for that orientation. Third, we calculated analytically the orientation which maximized source power and calculated the absolute value of bicoherence with

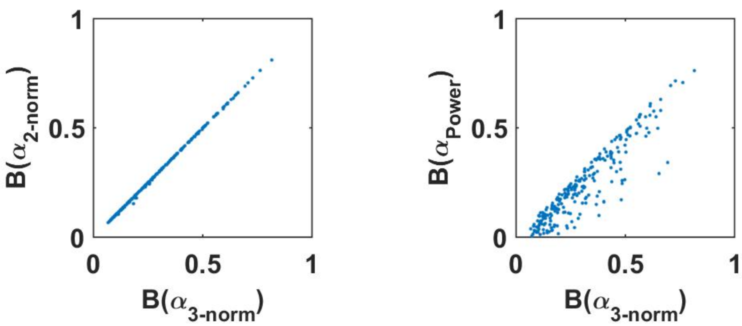

Figure 6: Absolute value of univariate bicoherence with 3-norm normalization for random voxels using three different choices for source orientation. Left: orientation chosen to maximize absolute value of bicoherence with 3-norm normalization versus orientation chosen to maximize absolute value of bicoherence with 2-norm normalization. Right: orientation chosen to maximize absolute value of bicoherence with 3-norm normalization versus orientation chosen to maximize power.

3-norm normalization for this orientation. Note, that the final quantity is always the absolute of bicoherence with 3-norm normalization, and it is only the chosen orientation which varied across these three approaches. Results are shown in Fig.6. We see that optimization with respect to 2-norm normalization results in almost identical values for the absolute value of bicoherence whereas maximizing power can result in substantially lower values.

3.5. Bicoherence of the alpha rhythm across all subjects

In the following analysis, all 24 healthy subjects were combined and bicoherence between a frequency f and $2f$ and power at frequencies f and $2f$ were estimated for all frequencies between 8 and 13 Hz. As an inverse method we chose eLORETA. In Fig.7, the logarithm of power in the source space after estimating the dipole directions at each voxel by the maximization of power is averaged over 24 subjects and plotted in the alpha band for 6 different frequencies between 8 Hz and 13 Hz. We see a strong activity in the occipital region which is almost identical in all frequencies. In Fig.8 the same results are shown in the beta band. In the beta band the major activity appears to be in the frontal region and fairly consistent across all frequencies similar to the alpha band. However, Fig.9 shows the absolute value of bicoherence estimated by the maximization of local bicoherence at each voxel after the eLORETA step and averaged across 24 subjects for the same frequency band. This figure shows varying activity regions for the 6 frequencies. While the activities at 9 Hz and 10 Hz are larger in the occipital region, at 11 Hz the inter-frequency interaction between 11 Hz and 22 Hz in the motor region has increased and at 12 Hz the right and left hemispheres in the motor region are dominant while the occipital region has a smaller activity. This effect is also noticeable at 13 Hz where the occipital region is even less active than in 12 Hz. We recall that here the activity at frequency f reflects the bicoherence at $f_1 = f_2 = f$ and $f_3 = 2f$.

3.6. Difference between patients with schizophrenia and healthy controls

As was seen in the previous section, bicoherence results are qualitatively different in high and low alpha bands. To compare between patients and controls we performed separate averages in the low

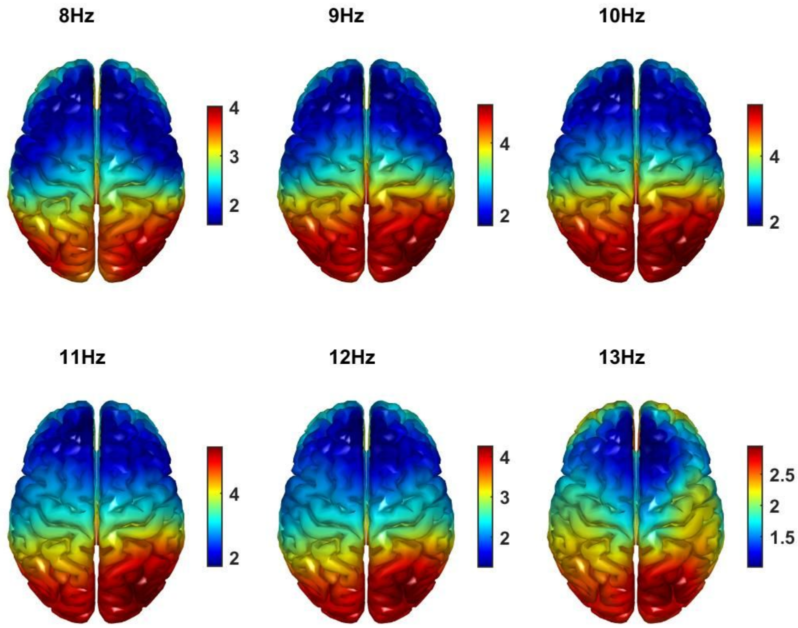

Figure 7: Logarithm of the power averaged across 24 healthy subjects for all frequencies in the alpha band.

alpha band from 8 Hz to 10 Hz and in the high alpha band between 11 Hz and 13 Hz. We found significant differences between the groups (healthy controls minus patients) only in the high alpha band and in the motor areas as shown in Figure 10. Significance was estimated using a permutation test: bicoherence for all 46 subjects were randomly assigned to 24 controls and 22 patients, and the difference of respective group averages was calculated. For $N = 10.000$ permutations for each voxel the number M of permutations was calculated for which the randomized data have a larger difference of the group averaged absolute values of bicoherence than the original data, and the p-value was set to $p = (M+1)/N$. We used the false discovery rate (FDR) to correct for multiple comparison at a q-level of .05. The result with non-significant results set to 0 are shown in Fig.11. We observe significant differences mostly in left and right motor areas. We also corrected for multiple comparisons using the maximum statistics: for each permuted data set we calculated the maximum of the differences across all voxels. A threshold for bicoherence was then set such that 95% of these maxima were lower than this threshold. Not shown as a figure, we found that only a focal spot in the left motor area, coinciding with the apparent local maximum of the FDR-corrected result, survived this correction.

Also not shown are corresponding results for power differences. None of the power differences reached significance and the non-significant differences were not maximal in motor areas but rather in parietal areas.

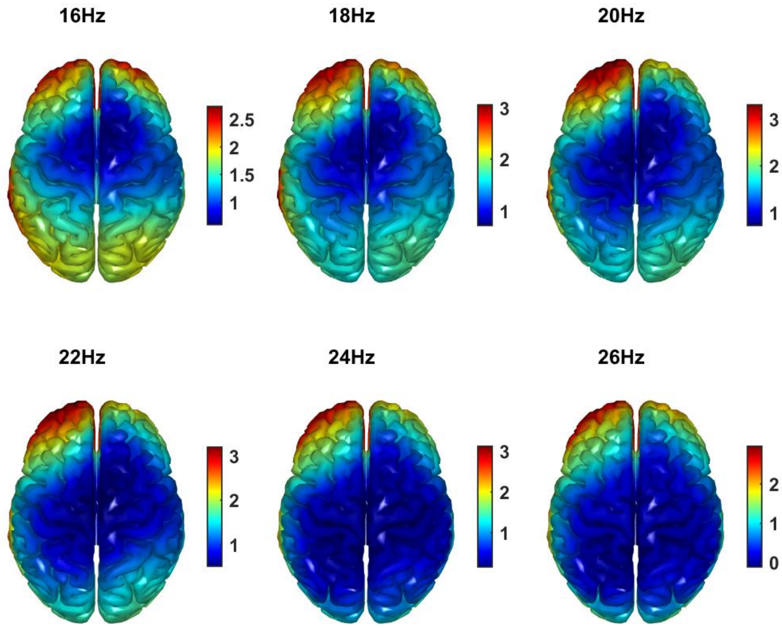

Figure 8: Logarithm of the power averaged across 24 healthy subjects for all frequencies in the beta band.

4. Conclusion and Discussion

We have proposed to estimate univariate bicoherence from EEG/MEG data in source space by optimizing the source direction to maximize the absolute value of bicoherence rather than commonly used power. First, bicoherence is a nonlinear measure of brain dynamics, thus it can better capture the weak nonlinear effects in EEG and MEG data. Additionally, inverse methods to estimate activity in source space have a poor spatial resolution: the estimated activity in a specified voxel in general originates from sources from a relatively large area centered at the specified voxel. Hence, the estimated activity in one voxel in general represents activity from sources with many different orientations. Choosing the orientation according to power bares the risk to miss the nonlinearity. As shown in our experiments, bicoherence is not affected by this problem. Also, power maximization is conceptually ambiguous in this context: for cross-frequency coupling it is unclear at which frequency power should be maximized or how various frequencies should be combined. Whether or not specifying the orientation by maximizing power could still be reasonable, depends on the case. We presented cases from empirical data where the results from power maximization missed major bicoherence of the alpha rhythms and its higher harmonics in source space. However, we found that those were exceptions when studying alpha-beta coupling in resting state. We also identified clearer signals from heart artifacts very deep in the brain using our method. While that can be considered a drawback, it also shows that in general bicoherence *is* missed when maximizing power even though in this particular case that source is uninteresting. We leave it as an open question whether for the detection of other sources of cross-frequency coupling (probably not measured during rest with eyes

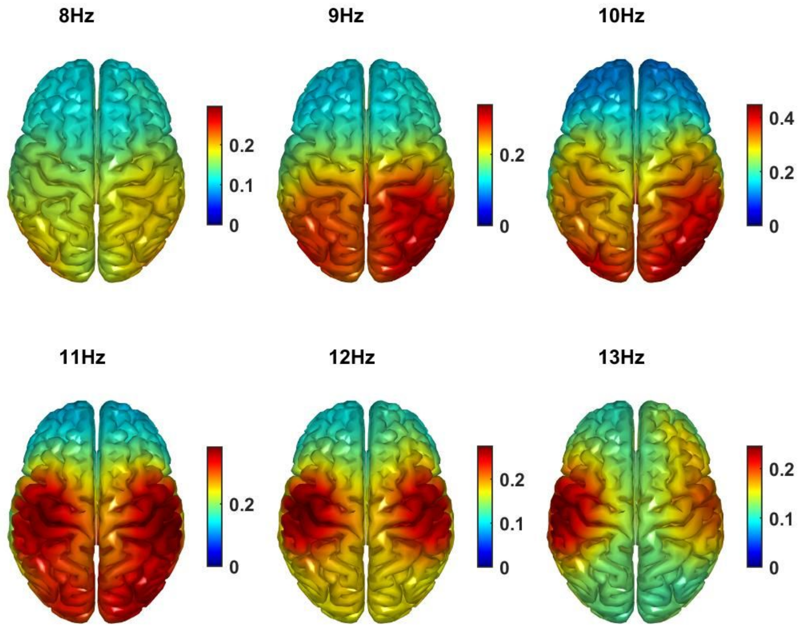

Figure 9: Absolute value of univariate bicoherence averaged across 24 healthy subjects for all frequencies f_1 in the alpha band with $f_1 = f_2$.

closed) the optimization with respect to bicoherence, which we believe to be conceptually convincing, is also necessary practically.

The proposed technique exploits that cross-bispectra are low order statistical moments such that one can first calculate those moments in sensor space and, using a linear inverse method, map them into source space with reasonable computational effort. Although counterintuitive, bicoherence, which is a measure of phase-phase coupling, is strongly related to phase-amplitude coupling which was demonstrated by Hyafil [2015] in examples. The question then is to what extent those findings generalize to arbitrary signals. If they do not, phase amplitude coupling, which became a major focus of research in this area, can be missed when using bicoherence. To find exact and generally valid relations between PAC and cross-bispectra for hypothetical infinite averages we slightly modified the definition of PAC by weighting the phases of the low frequency signal and squaring the amplitude. The former is actually also done in the definition of PAC given by Osipova et al. [2008]. With the limitation of this modification, which to our opinion is essentially irrelevant in practice, we come to the conclusion that bicoherence covers phase amplitude coupling completely: if bicoherence vanishes so does PAC. We also showed for simulated data and empirical LFP data of rats, containing coupling between the phase of the theta rhythm and the amplitude of gamma activity, that this coupling is well detectable with bicoherence. However, it is an open question which of the methods is preferable in a statistical sense for which case.

We analyzed EEG of 24 healthy subjects and 22 subjects with schizophrenia. We observed that the most relevant signals containing bicoherence originate in the presence of higher harmonics of motor

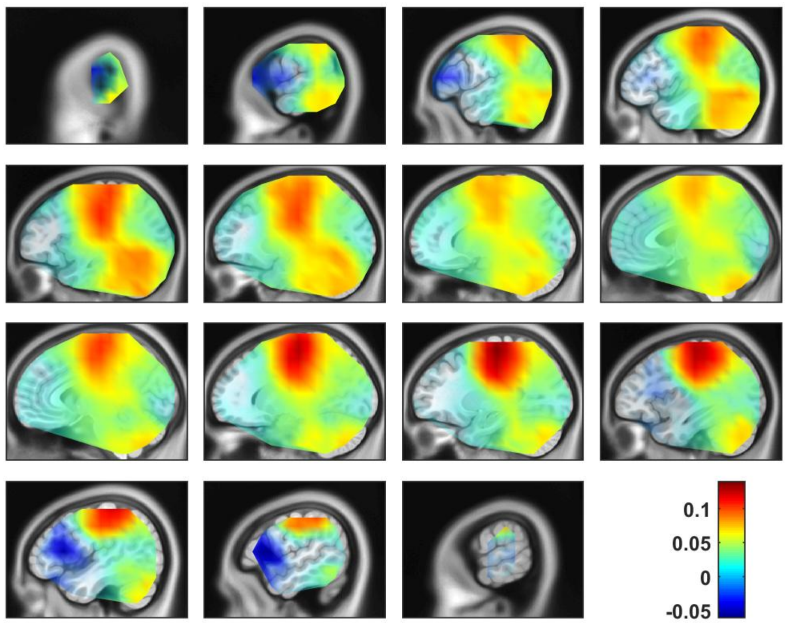

Figure 10: Difference of absolute values of univariate bicoherence with f_1 in the high alpha band and with $f_1 = f_2$ averaged across 24 subjects and 22 patients, respectively.

and occipital alpha rhythms. The motor alpha is barely visible in power analysis which is obscured by the dominant occipital alpha rhythm. These higher harmonics are clearly visible as such because of the high frequency resolution of bicoherence and could potentially be misinterpreted using PAC after the inherent smearing in the frequency domain. In resting state, we do not see other relevant neural sources of bicoherence. However, when applying our method we observed (partly strange) apparent muscle artifacts and heart artifacts. We emphasize that these data were carefully cleaned of artifacts, but the cleaning just may not be perfect, and bicoherence is very sensitive to nonlinear activities produced by the heart and muscles. In particular the heart signal shows a very complicated structure in the bicoherence plots, which could easily be confused with neural cross-frequency coupling.

Localization of the alpha-beta coupling, i.e. the first two harmonics of the alpha rhythms, showed that in the high alpha range between 11 Hz and 13 Hz, bicoherence becomes increasingly strong in left and right motor areas whereas bicoherence of the low alpha range is mainly located in parietal and occipital regions. We compared patients and healthy controls in both bands, and found a significant difference of bicoherence in motor areas in the high alpha band. These results might reflect the well-known phenomenon of genuine (i.e., medication-independent) motor abnormalities in patients with schizophrenia, thought to be due to dysfunctions across a cortico-subcortical circuit involving the motor cortex [Hirjak et al., 2015]. We found no significant power differences in either of the bands and no significant difference of bicoherence in the low alpha band. We conclude that bicoherence is not only conceptually a useful measure to study cross-frequency coupling, but also differentiates between subjects groups and might be useful as a biomarker or as feature to study brain diseases or effects of

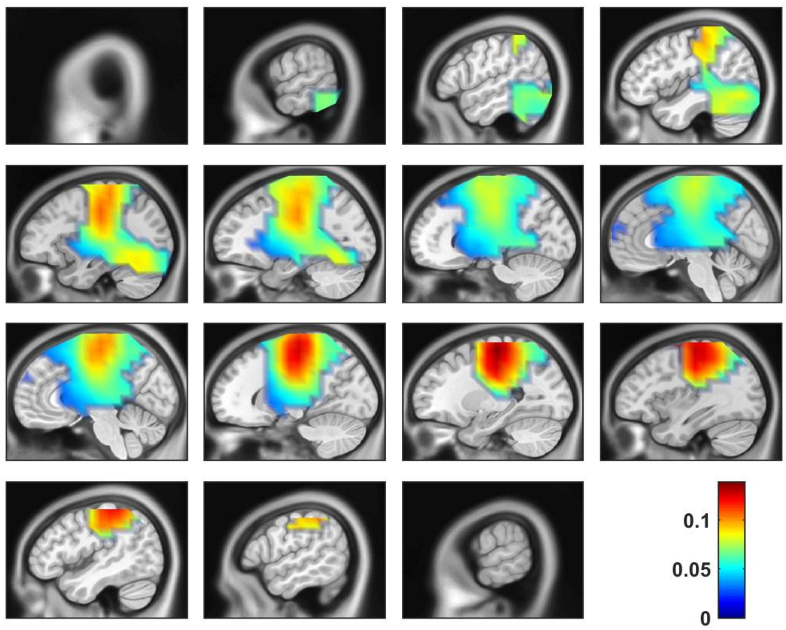

Figure 11: Same as Figure 10 with non-significant values set to zero.

medical treatment.

In our statistical testing of differences in bicoherence between patients and healthy controls we made corrections for multiple comparisons across voxels but not across pairs of frequencies. Rather, the frequencies were fixed to reflect alpha-beta coupling. This was based on the general observation that this phenomenon is extremely stable across all subjects. It was not picked to find a large difference. In fact, with a multiple comparison across all frequency pairs it would be difficult to find significant results also with the rather tolerant false discovery rate because the phenomenon is sparse as a function of two frequencies. It is tempting to select frequencies according to power within bands, but that could also be misleading. E.g., power in the alpha band is dominated by occipital alpha, but motor alpha has typically larger bicoherence values. We leave it is an open question, how to generally address this problem.

Acknowledgement

This research was partially funded by the German Research Foundation (DFG, SFB936/A3/C6/Z3 and TRR169/B1/B4), from the Landesforschungsförderung Hamburg (CROSS, FV25), and by the Fraunhofer Society of Germany.

5. Appendix

In this appendix we derive the gradient of the cost function, i.e. the squared absolute value of coherence, given in Eq. 19. The numerator reads

$$|D|^2 = \left|\sum \alpha_i \alpha_j \alpha_k E_{ijk}\right|^2 = \sum_{ijklmn} \alpha_i \alpha_j \alpha_k \alpha_l \alpha_m \alpha_n E_{ijk} E^*_{lmn} \tag{23}$$

where α is the dipole direction to be estimated and E is the cross-bispectrum in Eq. 14 in source space for a given voxel. Then

$$\begin{aligned}\frac{\partial |D|^2}{\partial \alpha_p} = &\underbrace{\sum_{jklmn} \alpha_j \alpha_k \alpha_l \alpha_m \alpha_n E_{pjk} E^*_{lmn}}_{=:A_1} + \underbrace{\sum_{ijkmn} \alpha_i \alpha_j \alpha_k \alpha_m \alpha_n E_{ijk} E^*_{pmn}}_{=:A_2} + \\ &\underbrace{\sum_{iklmn} \alpha_i \alpha_k \alpha_l \alpha_m \alpha_n E_{ipk} E^*_{lmn}}_{=:B_1} + \underbrace{\sum_{ijkln} \alpha_i \alpha_j \alpha_k \alpha_l \alpha_n E_{ijk} E^*_{lpn}}_{=:B_2} + \\ &\underbrace{\sum_{ijlmn} \alpha_i \alpha_j \alpha_l \alpha_m \alpha_n E_{ijp} E^*_{lmn}}_{=:C_1} + \underbrace{\sum_{ijklmp} \alpha_i \alpha_j \alpha_k \alpha_l \alpha_m E_{ijk} E^*_{lmp}}_{=:C_2}\end{aligned} \tag{24}$$

After renaming the indices we have

$$\begin{aligned} A_2 &= A_1^* \\ B_2 &= B_1^* \\ C_2 &= C_1^* \end{aligned} \tag{25}$$

On the other hand

$$\begin{aligned} A_1 &= \sum_{jk} \alpha_j \alpha_k E_{pjk} D^* \\ B_1 &= \sum_{ik} \alpha_i \alpha_k E_{ipk} D^* \\ C_1 &= \sum_{ij} \alpha_i \alpha_j E_{ijp} D^* \end{aligned} \tag{26}$$

$$\tag{27}$$

Let's define vectors u_p, v_p and w_p as

$$\begin{aligned} u_p &= \sum_{jk} \alpha_j \alpha_k E_{pjk} D^* \\ v_p &= \sum_{jk} \alpha_j \alpha_k E_{jpk} D^* \\ w_p &= \sum_{jk} \alpha_j \alpha_k E_{jkp} D^* \end{aligned}$$

Then

$$\frac{\partial |D|^2}{\partial \alpha_p} = 2Re((u_p + v_p + w_p) D^*) \tag{28}$$

where $Re(x)$ stands for the real part of the complex value of x.

The derivative of the squared denominator in Eq.16, after some simplification steps reads

$$\begin{aligned}\frac{\partial N^2}{\partial \alpha_p} &= N^2\left(\frac{1}{P_1}\frac{\partial P_1}{\partial \alpha_p} + \frac{1}{P_2}\frac{\partial P_2}{\partial \alpha_p} + \frac{1}{P_3}\frac{\partial P_3}{\partial \alpha_p}\right) \\ &= 2N^2 \,\mathrm{Re}\left(\frac{C_1}{P_1} + \frac{C_2}{P_2} + \frac{C_3}{P_3}\right)\alpha\end{aligned} \tag{29}$$

with and $C_i \equiv C(f_i)$ and $P_i = \alpha^T C_i \alpha$.

Putting the derivatives together we get with $|B|^2 = |D|^2/N^2$

$$\frac{\partial |B|^2}{\partial \alpha} = \frac{2}{N^2}\,\mathrm{Re}\left((u+v+w)\,D^* - |D|^2\left(\frac{C_1}{P_1} + \frac{C_2}{P_2} + \frac{C_3}{P_3}\right)\alpha\right) \tag{30}$$

We use the formula in the compter code to minimize bicoherence by steepest descent approach.

References

Andreou, C., Leicht, G., Nolte, G., Polomac, N., Moritz, S., Karow, A., Hanganu-Opatz, I. L., Engel, A. K., and Mulert, C. (2015a). Resting-state theta-band connectivity and verbal memory in schizophrenia and in the high-risk state. *Schizophrenia Research*, 161(2-3):299–307.

Andreou, C., Nolte, G., Leicht, G., Polomac, N., Hanganu-Opatz, I. L., Lambert, M., Engel, A. K., and Mulert, C. (2015b). Increased Resting-State Gamma-Band Connectivity in First-Episode Schizophrenia. *SCHIZOPHRENIA BULLETIN*, 41(4):930–939.

Ansari-Asl, K., Senhadji, L., Bellanger, J.-J., and Wendling, F. (2006). Quantitative evaluation of linear and nonlinear methods characterizing interdependencies between brain signals. *Physical Review E*, 74(3):031916.

Aru, J., Aru, J., Priesemann, V., Wibral, M., Lana, L., Pipa, G., Singer, W., and Vicente, R. (2015). Untangling cross-frequency coupling in neuroscience. *Current Opinion in Neurobiology*, 31:51–61.

Bullock, T., Achimowicz, J., Duckrow, R., Spencer, S., and Iragui-Madoz, V. (1997). Bicoherence of intracranial {EEG} in sleep, wakefulness and seizures. *Electroencephalography and Clinical Neurophysiology*, 103(6):661 – 678.

Canolty, R. T., Edwards, E., Dalal, S. S., Soltani, M., Nagarajan, S. S., Kirsch, H. E., Berger, M. S., Barbaro, N. M., and Knight, R. T. (2006). High gamma power is phase-locked to theta oscillations in human neocortex. *Science*, 313(5793):1626–1628.

Chella, F., Marzetti, L., Pizzella, V., Zappasodi, F., and Nolte, G. (2014). Third order spectral analysis robust to mixing artifacts for mapping cross-frequency interactions in eeg/meg. *Neuroimage*, 91:146–161.

Cole, S. R., van der Meij, R., Peterson, E. J., de Hemptinne, C., Starr, P. A., and Voytek, B. (2017). Nonsinusoidal Beta Oscillations Reflect Cortical Pathophysiology in Parkinson's Disease. *Journal of Neuroscience*, 37(18):4830–4840.

Cole, S. R. and Voytek, B. (2017). Brain Oscillations and the Importance of Waveform Shape. *Trends in Cognitive Sciences*, 21(2):137–149.

Hirjak, D., Thomann, P. A., Kubera, K. M., Wolf, N. D., Sambataro, F., and Wolf, R. C. (2015). Motor dysfunction within the schizophrenia-spectrum: A dimensional step towards an underappreciated domain. *SCHIZOPHRENIA RESEARCH*, 169(1-3):217–233.

Hyafil, A. (2015). Misidentifications of specific forms of cross-frequency coupling: three warnings. *Frontiers in Neuroscience*, 9.

Jensen, O., Spaak, E., and Park, H. (2016). Discriminating Valid from Spurious Indices of Phase-Amplitude Coupling. *ENEURO*, 3(6).

Leicht, G., Andreou, C., Polomac, N., Lanig, C., Schoettle, D., Lambert, M., and Mulert, C. (2015). Reduced auditory evoked gamma band response and cognitive processing deficits in first episode schizophrenia. *World Journal of Biological Psychiatry*, 16(6):387–397.

Nolte, G., Bai, O., Wheaton, L., Mari, Z., Vorbach, S., and Hallett, M. (2004). Identifying true brain interaction from {EEG} data using the imaginary part of coherency. *Clinical Neurophysiology*, 115(10):2292 – 2307.

Nolte, G. and Dassios, G. (2005). Analytic expansion of the eeg lead field for realistic volume conductors. *Physics in Medicine and Biology*, 50(16):3807.

Nunez, P., Srinivasan, R., Westdorp, A., Wijesinghe, R., Tucker, D., Silberstein, R., and Cadusch, P. (1997). EEG coherency .1. Statistics, reference electrode, volume conduction, Laplacians, cortical imaging, and interpretation at multiple scales. *Electroencephalography and Clinical Neurophysiology*, 103(5):499–515.

Osipova, D., Hermes, D., and Jensen, O. (2008). Gamma Power Is Phase-Locked to Posterior Alpha Activity. *Plos One*, 3(12).

Sarnthein, J., Morel, A., von Stein, A., and Jeanmonod, D. (2003). Thalamic theta field potentials and eeg: high thalamocortical coherence in patients with neurogenic pain, epilepsy and movement disorders. *Thalamus and Related Systems*, 2(3):231 – 238.

Scheffer-Teixeira, R. and Tort, A. B. L. (2016a). Data from: On cross-frequency phase-phase coupling between theta and gamma oscillations in the hippocampus. . *Dryad Digital Repository*.

Scheffer-Teixeira, R. and Tort, A. B. L. (2016b). On cross-frequency phase-phase coupling between theta and gamma oscillations in the hippocampus. *ELIFE*, 5.

Shahbazi, F., Ewald, A., and Nolte, G. (2014). Univariate normalization of bispectrum using hlder's inequality. *Journal of Neuroscience Methods*, 233:177 – 186.

Tort, A. B. L., Komorowski, R., Eichenbaum, H., and Kopell, N. (2010). Measuring Phase-Amplitude Coupling Between Neuronal Oscillations of Different Frequencies. *Journal of Neurophysiology*, 104(2):1195–1210.

Watt, R. C., Sisemore, C., Kanemoto, A., and Polson, J. S. (1996). Bicoherence of eeg can be used to differentiate anesthetic levels. In *Proceedings of 18th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, volume 5, pages 2035–2036 vol.5.