Published in final edited form as:
J Neurosci Methods. 2013 October 30; 220(1): . doi:10.1016/j.jneumeth.2013.08.006.

Assessment of cross-frequency coupling with confidence using generalized linear models

M. A. Kramer and **U. T. Eden**
Department of Mathematics and Statistics, Boston University, 111 Cummington Mall, Boston, MA, 02215

Abstract

Background—Brain voltage activity displays distinct neuronal rhythms spanning a wide frequency range. How rhythms of different frequency interact – and the function of these interactions – remains an active area of research. Many methods have been proposed to assess the interactions between different frequency rhythms, in particular measures that characterize the relationship between the phase of a low frequency rhythm and the amplitude envelope of a high frequency rhythm. However, an optimal analysis method to assess this cross-frequency coupling (CFC) does not yet exist.

New Method—Here we describe a new procedure to assess CFC that utilizes the generalized linear modeling (GLM) framework.

Results—We illustrate the utility of this procedure in three synthetic examples. The proposed GLM-CFC procedure allows a rapid and principled assessment of CFC with confidence bounds, scales with the intensity of the CFC, and accurately detects biphasic coupling.

Comparison with Existing Methods—Compared to existing methods, the proposed GLM-CFC procedure is easily interpretable, possesses confidence intervals that are easy and efficient to compute, and accurately detects biphasic coupling.

Conclusions—The GLM-CFC statistic provides a method for accurate and statistically rigorous assessment of CFC.

Keywords

Phase-amplitude coupling; oscillations; theta; gamma

1. Introduction

Oscillations are assumed to play a critical role in coordinating separate brain locations by providing an effective means to control the timing of neuronal firing (Engel et al., 2001; Varela et al., 2001; Buzsaki and Draguhn, 2004; Buzsáki, 2006). These rhythms have been categorized into different frequency bands (e.g., theta [4–8 Hz], beta [13–30 Hz] and gamma [30–90 Hz]) and associated with different functions: the theta band with memory and coordination (Buzsáki, 2006), the beta band with functional coupling across large distances (Kopell et al., 2000; Bibbig et al., 2002), “readiness”, and maintenance of the *status quo*

(Engel and Fries, 2010); the gamma band with local coupling (Kopell et al., 2000), competition (Börgers et al., 2008), and binding of cell assemblies (Engel and Singer, 2001). In general, lower frequency rhythms engage larger brain areas and modulate spatially localized fast oscillations (Bragin et al., 1995; Chrobak and Buzsaki, 1998; von Stein and Sarnthein, 2000; Lakatos et al., 2005, 2008). This cross-frequency coupling (CFC) between the power (or amplitude) of high frequency rhythms and the phase of low frequency rhythms has been observed in many brain regions including the hippocampus (Bragin et al., 1995; Csicsvari et al., 2003), entorhinal cortex (Chrobak and Buzsaki, 1998; Mormann et al., 2005), and the temporoparietal lobe of human patients (Canolty et al., 2006b), and has been shown to change with task demands (Tort et al., 2008; Voytek et al., 2013). CFC might play a functional role in working memory (Jones and Wilson, 2005; Lisman, 2005; Sirota et al., 2008; Lisman and Jensen, 2013), neuronal computation, communication and learning (Jensen and Colgin, 2007; Tort et al., 2009; Canolty and Knight, 2010). Although the cellular and dynamical mechanisms of specific rhythms associated with CFC have been extensively studies (e.g., gamma (Whittington et al., 2000 e.g., gamma (Whittington et al., 2011; Mann and Mody, 2010), beta (Roopun et al., 2006; Kramer et al., 2008a), and theta (Hutcheon and Yarom, 2000; Bao and Wu, 2003; Whittington and Traub, 2003)), the mechanisms governing interactions between different frequency rhythms, and the appropriate techniques for measuring CFC, remain active research areas (Kramer et al., 2008b; Penny et al., 2008; Wulff et al., 2009; Tort et al., 2010b).

Many quantitative measures have been proposed to characterize cross-frequency coupling. These include: the mean vector length or modulation index (Canolty et al., 2006b; Tort et al., 2010b) the phase-locking value (Lachaux et al., 1999; Vanhatalo et al., 2004; Mormann et al., 2005), the envelope-to-signal correlation (Bruns and Eckhorn, 2004), the analysis of amplitude spectra (Cohen, 2008), the coherence between amplitude and signal (Colgin et al., 2009), and the coherence between the time course of power and signal (Osipova et al., 2008). Overall, these different measures have been developed from different principles and made suitable for different purposes, as shown in comparative studies (Penny et al., 2008; Cohen, 2008; Tort et al., 2010b; Onslow et al., 2011). Recent approaches have considered the application of the generalized linear model (GLM) framework to assess CFC (Penny et al., 2008; Voytek et al., 2013). Although the advantages of the GLM approach have been described, specific disadvantages of previously developed GLM methods have been reported as well, including the failure to detect biphasic coupling when the model consists of three covariates, representing the amplitude as a function of a constant and the sine and cosine of the phase (Penny et al., 2008; Tort et al., 2010b; Özkurt and Schnitzler, 2011). Here, to address these disadvantages, we extend the GLM approach to include a more flexible representation of phase association, as described below.

Although successfully employed, these measures typically exhibit three common limitations. First, most measures lack a direct procedure to assess the confidence level of the CFC. Instead, existing methods assess the significance of CFC through surrogate control analysis (Lachaux et al., 1999; Canolty et al., 2006a; Tort et al., 2010b). This can involve shuffling trials (i.e., the amplitude time series of one trial is associated with the phase time series of another trial (Tort et al., 2010b)), or shifting of the signals (e.g., the high frequency amplitude envelope signal is shifted in time (Onslow et al., 2011)). Although powerful, such surrogate approaches are computationally expensive. A recent approach to determine an analytic procedure assessing the significance of CFC was presented in (Ozkurt, 2012). Second, many existing measures accurately detect only monophasic relationships (i.e., the phase-amplitude probability density has at most one peak (Penny et al., 2008; Özkurt and Schnitzler, 2011)), although entropy-based measures try to overcome this (Canolty et al., 2006b; Tort et al., 2010b). Third, some sophisticated CFC measures – although useful in practice – lack an immediate, intuitive understanding (Tort et al., 2010b).

Here we propose a new procedure that addresses these limitations. This procedure is based on the GLM framework, in which we model the high-frequency amplitude envelope as a function of the low-frequency phase using a spline basis. We show in simulation that the method accurately detects both monophasic and biphasic coupling, and provides an easily interpretable assessment of CFC with a principled procedure to establish confidence intervals that performs well when compared to two existing methods. The procedure is relatively easily implemented in existing software, and the analysis code is provided. We propose that this GLM-CFC procedure provides a simple and robust statistic to assess CFC.

2. Methods

In this section we describe the steps required to compute the GLM-CFC statistic. We first describe the procedure for extracting the phase and amplitude envelope from the observed data. We then describe the GLM procedure, the associated CFC statistic with confidence intervals, and a technique to visualize the results. Finally, we describe the synthetic time series utilized to illustrate the statistic, and two additional measures of CFC that are examined in the Results. The associated MATLAB code to compute the GLM-CFC procedure is provided in Appendix A.

2.1. Estimation of the phase and amplitude envelopes

The first step in assessing CFC is to extract the phase of the low frequency signal and the amplitude envelope of the high frequency signal. There exist numerous techniques to do so; we utilize here a typical approach, although the analysis framework applies for other approaches as well. First, we band-pass filter the data to isolate two frequency bands of interest: a low frequency band and a high frequency band (least square linear-phase FIR filter of order 375 for the low frequency band, and order 50 for the high frequency band, with zero-phase digital filtering). For the synthetic data, we consider only two frequency bands: low (4–7 Hz) and high (100–140 Hz). We choose to study only these two frequency bands; the analysis can be easily extended to explore many frequency bands (i.e., to compute the comodulogram (Tort et al., 2010a)). We then construct the analytic signal by applying the Hilbert transform, and extract the phase and amplitude envelope (Canolty et al., 2006b; Tort et al., 2010b). Artifacts resulting from bandpass filtering (Kramer et al., 2008b) and inaccurate phase estimation (Lepage et al., 2013) may potentially impact the CFC results; for the synthetic time series analyzed here, these associated issues are not considered in detail.

2.2. The GLM-CFC framework

Generalized linear models (GLMs) have been utilized in many neuroscience contexts, especially in the analysis and characterization of spike train data (Brown et al., 2004; Truccolo et al., 2005; Czanner et al., 2008; MacDonald et al., 2011; Eden et al., 2012). The purpose of GLMs is to relate response variables to linear combinations of predictor variables (McCullagh and Nelder, 1989). The GLM approach extends the normal linear regression model in two important ways: 1) The response variables may follow non-normal distributions, and 2) A nonlinear link function can be used to relate the responses and predictors. The latter feature allows flexibility in choosing the function that relates the expectation of the response variable to a linear combination of the predictors; common choices are the log, inverse, and inverse-square functions (McCullagh and Nelder, 1989). Here, for the GLM-CFC statistic, the response variable is the amplitude (A) of the high frequency band and the predictor variable is a function of the phase (ϕ) of the low frequency band. We choose a gamma distribution for the response variables because the amplitudes are always real and positive. The gamma model is common for data for which the standard deviation increases linearly with the mean (that is, when larger amplitude data has more

variability). For the case where the mean is large compared to the standard deviation, the gamma distribution approximates a normal distribution. The gamma distribution is flexible, and can capture both normal distributions and distributions with large positive tails. For the link function, we use the log link,

$$\log(\mu)=\beta X \quad (1)$$

where μ is the expected value of A, the design matrix X is a function of the predictors ϕ, and β are the unknown coefficients to determine. The log link function is common for GLMs using a gamma distribution, and leads to models where predictors have multiplicative effects on the response.

We construct two GLMs to fit the amplitude A as a function of the phase ϕ. In the first, we assume that the amplitude does not depend on the phase; we label this the *null model* because this model represents the null hypothesis of no CFC. In this case, the design matrix is a constant $X = 1$, and there exists a single unknown coefficient β_0 to estimate. Conceptually, the null model estimates the average amplitude across all phases. In the second model, which we label the *spline model*, we use cardinal splines to fit a smooth function for the expected amplitude as a function of the phase. Cardinal splines are smooth, piece-wise connected third order polynomial functions that are defined by a set of control points. The advantage of the cardinal spline is that it is capable of approximating any continuous functional relationship between phase and amplitude with a small number of parameters (Hearn and Baker, 1996). These parameters, the control point values, are directly interpretable as the expected amplitude at a specific set of phase values. The spline fit then smoothly interpolates between the estimated control point values.

To fit a spline model, we generate a design matrix X by applying a set of cardinal spline basis functions to the observed phase values, ϕ, at each time step. We select a number of control points, n, and space these evenly between 0 and 2π. The value of the spline estimate at any phase is determined by the two nearest control points to the left and two nearest control points to the right, where the control point values below zero or above 2π are taken modulo 2π. In this manner, we define a spline function over the circular topology of phase values. This means that the expected amplitude at 0 is equal to that at 2π and that the fit will be exactly the same whether we define the range of phases to be $[0, 2\pi]$ or $[-\pi, \pi]$. We also select the tension parameter for the spline, which controls the curviness of the function at the control points, to be 0.5; this is a standard choice, with the advantage of providing a smooth fit with fidelity to the data. In this case, the design matrix consists of n independent variables, and therefore n unknown coefficients of the vector β_S to estimate. Exponentiating these estimates represents the multiplicative effect of the phase on the expected value of the amplitude envelope at each control point.

To determine the number of control points n in the spline model, we evaluate the Akaike information criterion (AIC), defined as

$$\mathrm{AIC}=\Delta+2n \quad (2)$$

where Δ is the deviance of the spline model. For the simulation scenarios, we compute the AIC for different choices of n, and choose the n that minimizes the AIC, as described in the Results.

Upon estimating the unknown coefficients β_0 and β_S of the two models, we then compute the predicted values for the amplitude using the spline model (A_S) and the null model (A_0)

at 100 phase values evenly spaced between $-\pi$ and π. As a scalar statistic to characterize the CFC we compute,

$$r=\mathbf{max}[\mathbf{abs}[1-A_S/A_0]] \quad (3)$$

which is the maximum absolute fractional change between the spline and null models. This statistic is therefore simply interpreted as the largest proportional (or percentage) change between the null and spline models. As described in the Results, a large value of r is indicative of CFC. When r is large, the amplitude at some phase in the spline model differs from the constant amplitude of the null model.

To establish a confidence interval for the statistic r we utilize a parametric bootstrap procedure. To do so, we generate a surrogate distribution of the statistic r directly from the GLM; we use the observed coefficient estimates β_S and the estimated covariance of these estimates to generate 10000 normally distributed samples of the coefficients (β_S^j, where j = {0, 1, 2, …10000}). For each j, we then compute the predicted values for the amplitude using the spline model (A_S^j) and re-estimate the amplitude for the null model (A_θ^j) as the mean of A_S^j. Finally, we compute the measure r_S^j for each j. From the resulting distribution of r_S^j, we determine the 0.025 and 0.975 quantiles. In this way we use the surrogate distribution to define the 95% confidence interval for the statistic r.

The statistic r provides a single scalar value representative of the CFC between the phase and amplitude time series, and the associated confidence interval provides a range of certainty in the statistic. To visualize the CFC, we plot the predicted amplitude as a function of phase for the null and spline GLMs, and the point wise 95% confidence bounds of the predicted amplitude values, for both models. Doing so provides a graphical representation of the differences between the two models; strong CFC at some phase results in large differences between the two models.

To summarize, we measure CFC using the single quantity r in (3) with a corresponding 95% confidence interval determined from the GLM. This measure represents the largest deviation between the null model (which allows no variation in amplitude with phase) and the spline model (which permits variation of amplitude with phase). To visualize the CFC as a function of phase, we plot both models with corresponding point wise 95% confidence intervals. In the Results, the statistic and visualizations are illustrated for simulated examples, and compared with two existing measures of CFC in current use.

2.3. Synthetic time series

To generate the synthetic time series, we implement the following steps. First, we simulate 60 s of pink noise data (sampling rate 500 Hz), in which the power is proportional to the inverse of the frequency, consistent with brain field voltage recordings (e.g., (He et al., 2010)). Second, we filter these data into low (4–7 Hz) and high (100–140 Hz) frequency bands (Section 2.1). Third, we introduce coupling between the high and low frequency band signals. To do so, from the low pass filtered data we detect the times of the relative maxima and minima (i.e., the peaks and troughs in the low pass filtered signal). Then, at each extremum time, we multiply the high frequency signal by one plus a Hanning window of duration 42 ms centered at the extremum time. The result is to increase smoothly the amplitude of the high frequency signal for a short duration (42 ms) at a fixed phase of the low frequency signal. We scale the height of the Hanning window to adjust the intensity, I, of the CFC: setting I = 0.0 produces no modulation of the high frequency signal, while setting I = 1.0 increases the high frequency signal by a factor of 2 at the preferred phase of

the low frequency signal. In the simulations that follow, we consider the case of no modulation ($I = 0.0$, Section 3.1), the case of modulation only at the peaks of the low frequency signal (monophasic coupling, Section 3.2), and the case of modulation at the peaks and troughs of the low frequency signal (biphasic coupling, Section 3.3). To introduce noise disruptive to the CFC, we add additional time series to the low and high frequency bands. For each band, we first generate a new instance of pink noise. We then filter these data into the low (4–7 Hz) and high (100–140 Hz) frequency bands, as described above. Finally, we add these new instances of low and high frequency noise to the existing low and high frequency signals, respectively. For $I > 0$ the existing low and high frequency signals possess CFC, while the new signals provide additional, frequency matched noise without CFC.

2.4. Additional CFC measures

For comparison, we apply two additional measures of CFC. The first makes use of the GLM framework, but consists of only three covariates corresponding to the sine of the low frequency phase, the cosine of the low frequency phase, and a constant (Penny et al., 2008). The second measure is the modulation index (MI) (Canolty et al., 2006b), as defined in (Tort et al., 2010b). Briefly, to compute the MI, the average amplitude of the high frequency activity is determined for small intervals of the low frequency phase; here we divide the phase 0 to 2π into 18 intervals of equal length. The corresponding distribution of amplitude versus phase is then analyzed, and a scaled Kullback-Leibler distance between the amplitude distribution and the uniform distribution determined (Tort et al., 2010b).

3. Results

We first illustrate the utility of the GLM-CFC procedure through three simulation examples. In each case, we show that the statistic r provides a method to characterize CFC, increases with the intensity of CFC, and can distinguish even low intensity coupling with sufficient durations of data. We then examine how the number of control points, and large amplitude outliers, impacts the statistic r. We show that the GLM-CFC procedure can success-fully characterize both monophasic and biphasic coupling, and performs well when compared to existing measures in common use.

3.1. Simulation study: no CFC

We begin with an example that possesses no CFC. The synthetic data consist of 60 s of pink noise time series (sampling rate 500 Hz), filtered into low (4–7 Hz) and high (100–140 Hz) frequency bands (Figure 1A), from which the phase and amplitude envelope are extracted using the Hilbert transform (see *Methods*). To determine the number of control points in the spline model, we compute the AIC and find a minimum at $n = 5$ (Figure 1B). Plotting the amplitude versus phase for all time points reveals no obvious structure (gray points in Figure 1C), and the GLM fits for the spline and null models are similar (red and black, respectively, in Figure 1C). In this case, we find for the CFC measure $r = 0.037$ with 95% confidence interval [0.024, 0.050]. We conclude that the CFC is small, and consists of a modulation of amplitude versus phase between approximately 2.4% and 5.0%. Repeating this analysis for 1000 simulated pink noise time series (without modulation) we find the measure r typically remains small (median 3% change, maximum 8% change, Figure 1D). To summarize, the maximum difference between the null and spline models is small – and therefore the measure r is small – when time series lack CFC.

3.2. Simulation study: monophasic CFC

In the second simulation study, we consider the scenario in which CFC exists. We begin with the simplest case, in which the amplitude of the high frequency signal increases for a

single phase interval, so called “monophasic” coupling. To generate these synthetic data, we again simulate a pink noise time series (duration 60 s), and then low pass filter (4–7 Hz) and high pass filter (100–140 Hz) these data. From the low pass filtered data we detect the times of the relative maxima (i.e., the peaks in the low pass filtered signal), and introduce coupling between the amplitude of the high frequency activity and phase of the low frequency activity, as described in *Methods* (Section 2.3). We adjust a single parameter corresponding to the intensity, I, of the CFC; setting $I = 0.0$ produces no modulation of the high frequency signal by the low frequency phase, while setting $I = 1.0$ increases the high frequency signal by a factor of 2 at the preferred phase of the low frequency signal.

We show in Figure 2A example traces of the low (black) and high (blue) bandpass filtered signals for an intensity of $I = 0.5$. Visual inspection does not immediately suggest monophasic CFC. However, fixing the number of control points $n = 9$ in the spline model (as deduced from the AIC, Figure 2B), and performing the GLM-CFC analysis, we detect CFC between the two signals (Figure 2C). The statistic $r = 0.31$, with 95% confidence interval [0.28, 0.33], indicative of modulation of amplitude versus phase between 28% and 33%. Repeating this analysis for 1000 synthetic data sets with monophasic coupling, we find values of r between 0.22 and 0.46 (median 0.33), as shown in Figure 2D.

To compare the statistic r in the conditions of no CFC versus monophasic CFC, we plot a receiver operating characteristic (ROC) curve for three values of the intensity I. The ROC curve illustrates the performance of a binary classifier as a discrimination threshold applied to the statistic r is varied. Here we compare two simulation scenarios in which either monophasic CFC exists ($I > 0$) or does not ($I = 0$). For both scenarios, we consider 1000 simulations (duration 60 s, as above), resulting in 1000 values for r, and use the ROC curve to illustrate the effectiveness of the classification. When the intensity is low ($I = 0.05$), the monophasic CFC is weak, and detection of CFC is poor (Figure 2E, hollow circles). However, for increased intensity ($I = 0.10$ and $I = 0.15$), the statistic r becomes an accurate method to detect the CFC (Figure 2E, filled circles). We note that the statistic r scales with the modulation intensity (Figure 2F) – a desirable CFC measure property (Tort et al., 2010b) – and that the power of the statistic improves as the duration of the time series increases (Figure 2G). To compute each power curve, we first determine a threshold value equal to the 95% quantile of r from 1000 simulations for the $I = 0$ case (i.e., the case of no CFC). Then, at each value of the intensity I, we determine the proportion of r values exceeding this threshold from 1000 simulations; this proportion is plotted in Figure 2G. We conclude that the proposed GLM-CFC method accurately detects monophasic CFC, even for weak intensities when the duration of the data is sufficiently long.

3.3. Simulation study: biphasic CFC

As a second example of CFC, we consider the case in which the amplitude envelope of the high frequency activity increases at two different values of the phase. Such “biphasic” coupling may occur in neuronal systems, for example involving activation and refractory periods (Canolty et al., 2006b). To simulate biphasic coupling, we generate the synthetic data as for the monophasic CFC simulation in Section 3.2. We then select the relative maxima and minima of the low pass filtered signal (i.e., the peaks and troughs of the low pass filtered signal) and at the times of these extrema scale the high frequency signal by a value of $I = 0.5$, as for the monophasic CFC simulation (Figure 3A).

Fixing the number of control points $n = 10$ in the spline model (as deduced from the AIC, Figure 3B), and performing the GLM-CFC analysis, we detect CFC between the two signals ($r = 0.25$, with 95% confidence interval [0.23, 0.27]). Visual inspection of the model fits reveal two phase intervals – near 0 and $\pm\pi$ radians – at which the spline model deviates strongly from the null model (Figure 3C). Repeating this analysis for 1000 synthetic data

sets with biphasic coupling, we find values of r between 0.12 and 0.31 (median 0.20, Figure 3D). The statistic r provides good classification of CFC (ROC curve in Figure 3E), increases with the intensity of the biphasic CFC (Figure 3F), and exhibits increased power for longer durations of data (Figure 3G). We conclude that, for the simulated biphasic CFC, the proposed statistic accurately detects the coupling.

3.4. Impact of the number of control points on the GLM-CFC method

In computing the statistics r, we must specify the number of control points in the spline model. A principled method to do so is to choose the number of control points that minimizes the AIC (Eq. 2). An alternative method for selecting the number of control points is to employ prior physical or observational knowledge about the system. For example, if we believe that the amplitude increases at one or two broad phase intervals, then we may choose to utilize $n = 4$ or $n = 10$, respectively, in the spline model. However, if instead we believe that the amplitude increases in a sharp phase interval, then we may choose to utilize more control points (e.g., more than 10). By selecting too few control points, we may fail to detect amplitude increases restricted to narrow phase intervals, and by selecting too many control points, we may lose statistical power. In general, by selecting the number of control points, we impose a class of models in the GLM-CFC procedure, and we may do so using either quantitative techniques (e.g., AIC) or prior physical knowledge about the system.

To illustrate the impact of different choices for the number of control points on the statistic r, we consider the case of no CFC, and generate 60 s of data in the low and high frequency bands, as in the simulations of Section 3.1. We find that increasing the number of control points tends to increase the statistic r; the median value of r increases from $r = 0.03$ using $n = 5$, to $r = 0.06$ using $n = 30$ (Figure 4A). The reason for this increase is that, as n increases, the smoothness of the spline fit decreases. Therefore, localized increases in the amplitude – here, driven by noise – produce increased values of r. To illustrate this concept, we show in Figure 4B,C the same synthetic data with no CFC analyzed with 5 and 30 control points, respectively. Both spline models fluctuate around the mean amplitude (the null model indicated by the black horizontal line). However, when n is large (Figure 4C), more features in the amplitude versus phase curve are captured, resulting in larger fluctuations.

For the case of monophasic CFC (simulated as in Section 3.2) increasing the number of control points above 5 shifts the distribution of r values from a median of 0.22 to a median of 0.34 (Figure 4D). In this case, by employing too few control points (here, 5), the smoothing imposed by the spline curves underestimates the extent of the amplitude modulation. However, with a sufficient number of control points (here, ≥ 10, consistent with the AIC results in Section 3.2), we accurately capture the phase-amplitude relationship. To illustrate this effect, we show in Figure 4E,F the same synthetic data with monophasic CFC analyzed with $n = 5$ and $n = 30$, respectively. Although both figures capture the CFC, increasing n more accurately captures the sharp increase in phase. We note the increased variability of the spline model with $n = 30$, which in this case does not mask the true CFC effect.

These simulation results suggest that the proposed statistic r does depend on the number of control points chosen for the spline fit. In the case of no CFC, increasing the number of control points increases the statistic r. Although this effect is small, we recommend caution in interpreting small values of r when n is large. In the case of monophasic CFC, choosing too few control points may result in an underestimate of r.

3.5. Impact of outliers on the GLM-CFC method

Sudden, unexpected events often appear in brain voltage recordings, whether due to "noise" (e.g., non-neuronal activity) or true but unusual neuronal activity. To illustrate the impact of these outliers on the GLM-CFC procedure, we perform the following simulation study. We start by simulating 60 s of low and high frequency band data with no CFC, as in Section 3.1. Then, we introduce an outlier in the high frequency activity by increasing this time series by a factor of 10, 50, or 100 for a single, short interval (10 ms) chosen to occur at a random time (uniformly distributed) in the 60 s. We then perform the GLM-CFC analysis with $n = 5$, as in Section 3.1, and compute the statistic r. We find that a brief, factor of 10 increase in the high frequency activity does not impact the resulting r values (Figure 5); the median remains $r = 0.03$, identical to case of no outliers (Section 3.1). However, increasing the magnitude of this outlier increases the values of r observed; we find median values of $r = 0.05$ for a factor of 50 outlier, and $r = 0.09$ for a factor of 100 outlier. More importantly, as the outlier magnitude increases, the distribution of r values exhibits a longer tail that extends to high r values (i.e., $r > 0.1$, Figure 5).

We conclude that this type of outlier impacts the statistic r. However, the GLM framework provides a principled approach to detect and account for these types of outliers. In the GLM-CFC method, we assume that the amplitudes arise from a gamma distribution that depends on the associated phase values. Although the gamma distribution is flexible, and capable of representing both Gaussian and long-tailed distributions, rare outliers are not consistent with this distribution; therefore, for data with infrequent, large amplitude outliers, the model applied in the GLM framework is misspecified. To check for this misspecification, we may examine the residuals of the GLM procedure. Outliers in the amplitude time series will produce extreme values in the residuals. Multiple approaches exist to account for these outliers, including updating the model – for example, specifying a mixed model capable of capturing both the expected amplitude distribution and outliers – or detecting extreme values in the residuals and eliminating these time points from the analysis.

3.6. Comparison with existing CFC methods

To compare the performance of the proposed GLM-CFC procedure to existing methods, we implement two additional measures in common use: a GLM measure with only three covariates (as motivated by (Penny et al., 2008)), and the modulation index (Canolty et al., 2006b; Tort et al., 2010b); see *Methods*. To illustrate the performance of these measures, we focus on the simulation scenario consisting of biphasic coupling with increasing intensity, employed in Section 3.3. As reported in other studies (e.g., (Penny et al., 2008)), the GLM measure with two covariates performs poorly in this case (Figure 6A). The reason for this poor performance is that the three covariates do not provide enough flexibility to capture the biphasic coupling. However, applying the spline basis to the phase signal, and then computing the GLM fit, provides sufficient flexibility to detect the biphasic structure in the proposed statistic r, as described in Section 3.3. Both the proposed statistic r and the modulation index (MI) distinguish between different levels of coupling strength (Figure 6A,B). In addition, the proposed statistic r and the MI behave similarly when the CFC becomes obscured by noise. For these simulations, we first fix the intensity ($I = 0.5$) to produce CFC, and then add filtered pink noise to the low and high frequency bands (see *Methods*). As the noise level increases, the CFC becomes more difficult to detect, and both measures decrease (Figure 6C,D). We note that at high noise levels, both measures approach values consistent with no CFC (compare Figure 6C with large noise to Figure 6A with $I = 0$, and Figure 6D with large noise to Figure 6B with $I = 0$).

These results suggest that the MI and GLM-CFC procedures perform similarly. In fact, the two methods characterize similar features of the CFC. Like the proposed statistic r, the MI

assesses CFC by characterizing the deviation of the amplitude distribution from the uniform distribution in a phase-amplitude plot. The MI measure is a normalized distance; it is proportional to Kullback-Leibler distance between the amplitude distribution and the uniform distribution (Tort et al., 2010b). Although these two methods perform similarly for the case of biphasic coupling, the proposed statistic r offers two advantages. First, the proposed statistic r has a simple, intuitive interpretation as the percentage change from a flat phase-amplitude distribution. For example, from Figure 6A, we conclude that as the intensity increases, the amplitude modulation increases from a small modulation (less than 10%) about the average amplitude, to a 35% modulation ($r \approx 0.35$) about the average amplitude. The MI also increases with the intensity, from a value of 0.020 to 0.028; we interpret this change to indicate that the (normalized) distance between the empirical amplitude distribution and the uniform distribution increases by approximately 0.008. Although r and MI provide a consistent characterization of the CFC, the interpretation of the proposed statistic r is simpler. Second, the proposed measure, and the associated visualization of amplitude versus phase (e.g., Figure 3C), possesses well-defined confidence intervals. These confidence intervals allow an assessment of the range of r values consistent with the data; for example, in Section 3.3 we reported an r value representing a 23% to 27% modulation about the average amplitude. Using the GLM framework, these confidence intervals are efficient and simple to compute.

4. Discussion

Dynamic brain activity exhibits a multitude of rhythms (Buzsáki and Draguhn, 2004), and understanding how different frequency rhythms interact is an important research question in neuroscience (Canolty and Knight, 2010; Lisman and Jensen, 2013). Many methods exist for assessing CFC, each with specific merits (as reviewed in (Tort et al., 2010b)). Here, we propose a new method to assess CFC based on a generalized linear modeling (GLM) framework, and show how this method performs in simulation scenarios. Compared to existing methods, the proposed GLM-CFC procedure possesses three advantages. First, the statistic r is easily interpretable; for example, a 2% difference between the null and spline models indicates that the amplitude varies little with phase, while a 20% difference indicates large deviations in the amplitude versus phase. Second, the statistic r possesses confidence intervals that are easy and efficient to compute. Moreover, the visualization of amplitude versus phase also possesses well defined confidence intervals that are easily computed (e.g., Figure 3C). Third, the GLM-CFC procedure accurately detects biphasic coupling.

As part of computing the CFC for observational data, many analysis issues must be considered. These include the choice of filtering method and frequency ranges to isolate the low and high frequency oscillations, and the method to extract the phase and amplitude envelope. Here we implemented particular choices (see *Methods*) to address these issues. Different types of data may require different choices of filtering or phase estimation, or techniques to analyze short epochs of trial specific data (Voytek et al., 2013). The proposed GLM approach can apply in these different scenarios. Finally, we note that certain artifacts in the data (Kramer et al., 2008b), and inaccurate estimation of the phase (Lepage et al., 2013) have the potential to confound all CFC measures.

In practice, assessment of CFC typically occurs for a group of subjects. To assess the statistic r computed for a group of subjects, we propose two approaches. First, a value r may be computed for each subject, and compared to the null hypothesis of no CFC in any subject. Under the null hypothesis, the distribution of r can be estimated (for example, Figure 3.1D) or additional analysis may be undertaken to determine a corresponding theoretical distribution. Using this distribution, each statistic r may be assigned a p-value, and the results corrected for multiple comparisons. Second, a new model could be developed

that incorporates the CFC and possible subject specific effects. This model would utilize all of the data (i.e., the data from all subjects) and return a single r value. The covariates in the model – corresponding to indicator functions for individual subjects – could also be investigated to determine single subject effects.

We do not assess directly the significance of the CFC statistic r. Instead, we establish confidence bounds for r, and then interpret the results. We note that, because we choose r as a maximum, we do not expect these confidence bounds to include zero. Therefore, even in the case of no CFC, we find nonzero values of r; these values tend to be small (e.g., representative of a 5% change, Figure 1D). Because we expect that knowledge of the magnitude of the phase-amplitude relationship is important for understanding and characterizing CFC, we did not focus specifically on calculating a significance of the statistic r. Instead, we have described how the statistic r and the associated visualizations (e.g., Figure 3C) provide a direct and intuitive assessment of the CFC magnitude. If a significance value is required, standard techniques (e.g., a bootstrap resampling based on shuffled intervals of the data (Canolty et al., 2006b; Tort et al., 2010b)) can be employed. Although no gold standard for assessing CFC exists, we have shown that the proposed GLM-CFC procedure possesses advantageous properties, including detection of monophasic and biphasic coupling, characterization of CFC intensity, and principled methods to evaluate uncertainty.

References

Bao, Weili; Wu, Jian-Young. Propagating wave and irregular dynamics: spatiotemporal patterns of cholinergic theta oscillations in neocortex in vitro. JOURNAL OF NEUROPHYSIOLOGY. Jul; 2003 90(1):333–41. [PubMed: 12612003]

Bibbig, Andrea; Traub, Roger D.; Whittington, Miles A. Long-range synchronization of gamma and beta oscillations and the plasticity of excitatory and inhibitory synapses: a network model. JOURNAL OF NEUROPHYSIOLOGY. Oct; 2002 88(4):1634–54. [PubMed: 12364494]

Börgers, Christoph; Epstein, Steven; Kopell, Nancy J. Gamma oscillations mediate stimulus competition and attentional selection in a cortical network model. Proc Natl Acad Sci USA. Nov; 2008 105(46):18023–8. [PubMed: 19004759]

Bragin A, Jando G, Nadasdy Z, Hetke J, Wise K, Buzsaki G. Gamma (40–100-hz) oscillation in the hippocampus of the behaving rat. J Neurosci. Jan; 1995 15(1):47–60. [PubMed: 7823151]

Brown, Emery N.; Kass, Robert E.; Mitra, Partha P. Multiple neural spike train data analysis: state-of-the-art and future challenges. Nat Neurosci. May; 2004 7(5):456–61. [PubMed: 15114358]

Bruns, Andreas; Eckhorn, Reinhard. Task-related coupling from high-to low-frequency signals among visual cortical areas in human subdural recordings. International journal of psychophysiology: official journal of the International Organization of Psychophysiology. Jan; 2004 51(2):97–116. [PubMed: 14693360]

Buzsáki, G. Rhythms of the brain. Oxford University Press; 2006.

Buzsaki, Gyorgy; Draguhn, Andreas. Neuronal oscillations in cortical networks. Science. 2004; 304(5679):1926–1929. [PubMed: 15218136]

Buzsáki, György; Draguhn, Andreas. Neuronal oscillations in cortical networks. Science. Jun; 2004 304(5679):1926–9. [PubMed: 15218136]

Canolty RT, Edwards E, Dalal SS, Soltani M, Nagarajan SS, Kirsch HE, Berger MS, Barbaro NM, Knight RT. Supp info high gamma. Science. Sep; 2006a 313(5793):1626–8. [PubMed: 16973878]

Canolty RT, Edwards E, Dalal SS, Soltani M, Nagarajan SS, Kirsch HE, Berger MS, Barbaro NM, Knight RT. High gamma power is phase-locked to theta oscillations in human neocortex. Science. Sep; 2006b 313(5793):1626–8. [PubMed: 16973878]

Canolty, Ryan T.; Knight, Robert T. The functional role of cross-frequency coupling. Trends Cogn Sci (Regul Ed). Nov; 2010 14(11):506–15. [PubMed: 20932795]

Chrobak J, Buzsaki G. Gamma oscillations in the entorhinal cortex of the freely behaving rat. J Neurosci. Jan; 1998 18(1):388–398. [PubMed: 9412515]

Cohen, Michael X. Assessing transient cross-frequency coupling in eeg data. J Neurosci Methods. Mar; 2008 168(2):494–9. [PubMed: 18061683]

Colgin, Laura Lee; Denninger, Tobias; Fyhn, Marianne; Hafting, Torkel; Bonnevie, Tora; Jensen, Ole; Moser, May-Britt; Moser, Edvard I. Frequency of gamma oscillations routes flow of information in the hippocampus. Nature. Nov; 2009 462(7271):353–7. [PubMed: 19924214]

Csicsvari, Jozsef; Jamieson, Brian; Wise, Kensall D.; Buzśaki, Gÿorgy. Mechanisms of gamma oscillations in the hippocampus of the behaving rat. NEURON. Jan; 2003 37(2):311–22. [PubMed: 12546825]

Czanner, Gabriela; Eden, Uri T.; Wirth, Sylvia; Yanike, Marianna; Suzuki, Wendy A.; Brown, Emery N. Analysis of between-trial and within-trial neural spiking dynamics. JOURNAL OF NEUROPHYSIOLOGY. May; 2008 99(5): 2672–93. [PubMed: 18216233]

Eden, Uri T.; Gale, John T.; Amirnovin, Ramin; Eskandar, Emad N. Characterizing the spiking dynamics of subthalamic nucleus neurons in parkinson's disease using generalized linear models. Front Integr Neurosci. Jan.2012 6:28. [PubMed: 22723771]

Engel AK, Fries P, Singer W. Dynamic predictions: oscillations and synchrony in top-down processing. Nat Rev Neurosci. Oct; 2001 2(10):704–16. [PubMed: 11584308]

Engel AK, Singer W. Temporal binding and the neural correlates of sensory awareness. Trends Cogn Sci. Jan; 2001 5(1):16–25. [PubMed: 11164732]

Engel, Andreas K.; Fries, Pascal. Beta-band oscillations–signalling the status quo? Curr Opin Neurobiol. Apr; 2010 20(2):156–65. [PubMed: 20359884]

He, Biyu J.; Zempel, John M.; Snyder, Abraham Z.; Raichle, Marcus E. The temporal structures and functional significance of scale-free brain activity. NEURON. May; 2010 66(3):353–69. [PubMed: 20471349]

Hearn, D.; Baker, MP. Computer Graphics: C Version. Prentice Hall; 1996.

Hutcheon B, Yarom Y. Resonance, oscillation and the intrinsic frequency preferences of neurons. Trends Neurosci. May; 2000 23(5):216–22. [PubMed: 10782127]

Jensen, Ole; Colgin, Laura L. Cross-frequency coupling between neuronal oscillations. Trends Cogn Sci. Jul; 2007 11(7):267–9. [PubMed: 17548233]

Jones, Matthew W.; Wilson, Matthew A. Theta rhythms coordinate hippocampal-prefrontal interactions in a spatial memory task. PLoS Biol. Dec.2005 3(12):e402. [PubMed: 16279838]

Kopell N, Ermentrout GB, Whittington MA, Traub RD. Gamma rhythms and beta rhythms have different synchronization properties. Proc Natl Acad Sci USA. Feb; 2000 97(4):1867–72. [PubMed: 10677548]

Kramer, Mark A.; Roopun, Anita K.; Carracedo, Lucy M.; Traub, Roger D.; Whittington, Miles A.; Kopell, Nancy J. Rhythm generation through period concatenation in rat somatosensory cortex. PLoS Comput Biol. Jan.2008a 4(9):e1000169. [PubMed: 18773075]

Kramer, Mark A.; Tort, Adriano BL.; Kopell, Nancy J. Sharp edge artifacts and spurious coupling in eeg frequency comodulation measures. Journal of Neuroscience Methods. May; 2008b 170(2): 352–7. [PubMed: 18328571]

Lachaux, Jean-Philippe; Rodriguez, Eugenio; Martinerie, Jacques; Varela, Francisco. Measuring phase synchrony in brain signals. Human brain mapping. 1999; 8(4):194–208. [PubMed: 10619414]

Lakatos P, Shah A, Knuth K, Ulbert I, Karmos G, Schroeder C. An oscillatory hierarchy controlling neuronal excitability and stimulus processing in the auditory cortex. JOURNAL OF NEUROPHYSIOLOGY. Sep; 2005 94(3): 1904–1911. [PubMed: 15901760]

Lakatos, Peter; Karmos, George; Mehta, Ashesh D.; Ulbert, Istvan; Schroeder, Charles E. Entrainment of neuronal oscillations as a mechanism of attentional selection. Science. Apr; 2008 320(5872): 110–3. [PubMed: 18388295]

Lepage, Kyle; Kramer, Mark; Eden, Uri. Some sampling properties of common phase estimators. Neural Computation. Jan.2013 :1–21. [PubMed: 23020108]

Lisman, John. The theta/gamma discrete phase code occuring during the hippocampal phase precession may be a more general brain coding scheme. Hippocampus. Jan; 2005 15(7):913–22. [PubMed: 16161035]

Lisman, John E.; Jensen, Ole. The theta-gamma neural code. Neuron. Mar; 2013 77 (6):1002–1016. [PubMed: 23522038]

MacDonald, Christopher J.; Lepage, Kyle Q.; Eden, Uri T.; Eichenbaum, Howard. Hippocampal time cells bridge the gap in memory for discontiguous events. Neuron. Aug; 2011 71(4):737–749. [PubMed: 21867888]

Mann, Edward O.; Mody, Istvan. Control of hippocampal gamma oscillation frequency by tonic inhibition and excitation of interneurons. Nat Neurosci. Feb; 2010 13(2):205–12. [PubMed: 20023655]

McCullagh, P.; Nelder, JA. Generalized Linear Models. 2. Chapman and Hall/CRC; 1989.

Mormann, Florian; Fell, Juergen; Axmacher, Nikolai; Weber, Bernd; Lehnertz, Klaus; Elger, Christian E.; Ferńandez, GuilÍen. Phase/amplitude reset and theta-gamma interaction in the human medial temporal lobe during a continuous word recognition memory task. Hippocampus. Jan; 2005 15(7): 890–900. [PubMed: 16114010]

Onslow, Angela CE.; Bogacz, Rafal; Jones, Matthew W. Quantifying phase-amplitude coupling in neuronal network oscillations. Prog Biophys Mol Biol. Mar; 2011 105(1–2):49–57. [PubMed: 20869387]

Osipova, Daria; Hermes, Dora; Jensen, Ole. Gamma power is phase-locked to posterior alpha activity. PLoS ONE. Jan.2008 3(12):e3990. [PubMed: 19098986]

Ozkurt, Tolga Esat. Statistically reliable and fast direct estimation of phase-amplitude cross-frequency coupling. IEEE Trans Biomed Eng. Jul; 2012 59(7):1943–50. [PubMed: 22531738]

Özkurt, Tolga Esat; Schnitzler, Alfons. A critical note on the definition of phase-amplitude cross-frequency coupling. Journal of Neuroscience Methods. Oct; 2011 201(2):438–43. [PubMed: 21871489]

Penny WD, Duzel E, Miller KJ, Ojemann JG. Testing for nested oscillation. Journal of Neuroscience Methods. Sep; 2008 174(1):50–61. [PubMed: 18674562]

Roopun A, Middleton S, Cunningham M, LeBeau F, Bibbig A, Whittington M, Traub R. A beta2-frequency (20–30 hz) oscillation in nonsynaptic networks of somatosensory cortex. Proc Natl Acad Sci USA. Oct; 2006 103(42): 15646–15650. [PubMed: 17030821]

Sirota, Anton; Montgomery, Sean; Fujisawa, Shigeyoshi; Isomura, Yoshikazu; Zugaro, Michael; Buzśaki, Gÿorgy. Entrainment of neocortical neurons and gamma oscillations by the hippocampal theta rhythm. NEURON. Nov; 2008 60 (4):683–97. [PubMed: 19038224]

Tort, Adriano BL.; Kramer, Mark A.; Thorn, Catherine; Gibson, Daniel J.; Kubota, Yasuo; Graybiel, Ann M.; Kopell, Nancy J. Dynamic cross-frequency couplings of local field potential oscillations in rat striatum and hippocampus during performance of a t-maze task. PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA. Dec; 2008 105(51):20517–22. [PubMed: 19074268]

Tort, Adriano BL.; Komorowski, Robert W.; Manns, Joseph R.; Kopell, Nancy J.; Eichenbaum, Howard. Theta-gamma coupling increases during the learning of item-context associations. PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA. Dec; 2009 106(49):20942–7. [PubMed: 19934062]

Tort, Adriano BL.; Fontanini, Alfredo; Kramer, Mark A.; Jones-Lush, Lauren M.; Kopell, Nancy J.; Katz, Donald B. Cortical networks produce three distinct 7–12 hz rhythms during single sensory responses in the awake rat. J Neurosci. Mar; 2010a 30(12):4315–24. [PubMed: 20335467]

Tort, Adriano BL.; Komorowski, Robert; Eichenbaum, Howard; Kopell, Nancy. Measuring phase-amplitude coupling between neuronal oscillations of different frequencies. JOURNAL OF NEUROPHYSIOLOGY. Aug; 2010b 104(2): 1195–210. [PubMed: 20463205]

Truccolo, Wilson; Eden, Uri T.; Fellows, Matthew R.; Donoghue, John P.; Brown, Emery N. A point process framework for relating neural spiking activity to spiking history, neural ensemble, and extrinsic covariate effects. JOURNAL OF NEUROPHYSIOLOGY. Feb; 2005 93(2):1074–89. [PubMed: 15356183]

Vanhatalo S, Palva JM, Holmes MD, Miller JW, Voipio J, Kaila K. Infraslow oscillations modulate excitability and interictal epileptic activity in the human cortex during sleep. Proc Natl Acad Sci USA. Apr; 2004 101(14):5053–7. [PubMed: 15044698]

Varela F, Lachaux JP, Rodriguez E, Martinerie J. The brainweb: phase synchronization and large-scale integration. Nat Rev Neurosci. Apr; 2001 2(4):229–39. [PubMed: 11283746]

von Stein A, Sarnthein J. Different frequencies for different scales of cortical integration: from local gamma to long range alpha/theta synchronization. International journal of psychophysiology: official journal of the International Organization of Psychophysiology. Dec; 2000 38(3):301–13. [PubMed: 11102669]

Voytek, Bradley; D'Esposito, Mark; Crone, Nathan; Knight, Robert T. A method for event-related phase/amplitude coupling. NEUROIMAGE. Jan.2013 64: 416–24. [PubMed: 22986076]

Whittington M, Traub R, Kopell N, Ermentrout B, Buhl E. Inhibition-based rhythms: experimental and mathematical observations on network dynamics. Int J Psychophysiol. Dec; 2000 38(3):315–336. [PubMed: 11102670]

Whittington, Miles A.; Traub, Roger D. Interneuron diversity series: inhibitory interneurons and network oscillations in vitro. Trends Neurosci. Dec; 2003 26(12):676–82. [PubMed: 14624852]

Whittington, Miles A.; Cunningham, Mark O.; LeBeau, Fiona EN.; Racca, Claudia; Traub, Roger D. Multiple origins of the cortical rhythm. Dev Neurobiol. Jan; 2011 71(1):92–106. [PubMed: 21154913]

Wulff, Peer; Ponomarenko, Alexey A.; Bartos, Marlene; Korotkova, Tatiana M.; Fuchs, Elke C.; Bähner, Florian; Both, Martin; Tort, Adriano BL.; Kopell, Nancy J.; Wisden, William; Monyer, Hannah. Hippocampal theta rhythm and its coupling with gamma oscillations require fast inhibition onto parvalbumin-positive interneurons. PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA. Mar; 2009 106(9): 3561–6. [PubMed: 19204281]

Appendix A

In this Appendix we provide example MATLAB code for computing the GLM-CFC measure, and displaying the measure results.

```
%Function to apply GLM-CFC procedure.
%
%INPUTS.
% Vlo = low frequency band signal.
% Vhi = high frequency band signal.
% nCtlPts = the number of control points to use in the spline fitting of
phase.
%
%OPTIONAL INPUT.
% The 4th input is optional. Set the 4th input to:
% 'noplot' to prevent plotting of results.
% 'AIC' to compute the # control points using AIC. When this option is used,
% the number of control points that minimizes the AIC is used, and the
% 3rd input (nCtlPts) is ignored.
%
%OUTPUTS.
% r = the GLM-CFC measure.
% r_CI = the 95% confidence intervals for r.
% nCtlPts = the number of control points used.
%
% By default, this function plots the results.
function [r, r_CI, nCtlPts] = GLM_CFC_for_paper(Vlo, Vhi, nCtlPts, varargin)
%Compute phase and amplitude.
```

```
phi = angle(hilbert(Vlo));
amp = abs(hilbert(Vhi));
%Compute AIC to determine number of control points.
if ~isempty(varargin) & strcmp(varargin{1}, 'AIC')
fprintf('Running AIC … \n')
Y = amp';
CtlPts = (4:1:30);
%Compute the AIC.
AIC = zeros(size(CtlPts));
for k=1:length(CtlPts)
nCtlPts = CtlPts(k);
X = splinephase(phi', nCtlPts);
[b, dev, stats] = glmfit(X, Y, 'gamma', 'link', 'log', 'constant', 'off');
AIC(k) = dev + 2*nCtlPts;
fprintf([num2str(nCtlPts) ' ' num2str(AIC(k)) '\n'])
end
%Select the # control points from AIC, and plot the AIC.
[~, imn] = min(AIC);
nCtlPts = CtlPts(imn);
figure(1); clf();
plot(CtlPts, AIC, 'k', 'LineWidth', 2)
hold on
plot([nCtlPts, nCtlPts], [min(AIC) max(AIC)], 'r', 'LineWidth', 2)
hold off
axis tight
xlabel('# control points')
ylabel('AIC')
fprintf(['Suggested number of control points is ' num2str(nCtlPts) '\n'])
end
%Define variables for GLM procedure.
Y = amp';
X = spline_phase0(phi',nCtlPts);
XC = ones(size(Y));
%Perform GLM.
[b, ~, stats] = glmfit(X, Y, 'gamma', 'link', 'log', 'constant', 'off');
[bC, ~, statsC] = glmfit(XC, Y, 'gamma', 'link', 'log', 'constant', 'off');
%Define dense phase points for interpolation.
phi0 = linspace(-pi,pi,100);
X0 = spline_phase0(phi0',nCtlPts);
%Determine spline fit and CI.
[spline0, dylo, dyhi] = glmval(b,X0,'log',stats,'constant', 'off');
splineU = spline0+dyhi;
splineL = spline0-dylo;
%Determine null fit and CI.
[null0, dylo, dyhi] =
glmval(bC,ones(size(phi0)),'log',statsC,'constant', 'off');
nullU = null0+dyhi;
nullL = null0-dylo;
%Find the max absolute percentage change between the two models.
[r imx] = max(abs(1-spline0./null0))
```

```
%Determine CI for the measure r.
M = 10000;
bMC = b*ones(1,M) + sqrtm(stats.covb)*normrnd(0,1,nCtlPts,M);
splineMC = glmval(bMC,X0,'log',stats,'constant', 'off');
nullMC = mean(splineMC,1);
mx = zeros(M,1);
for k=1:M m
mx(k) = max(abs(1-splineMC(:,k)./nullMC(k)));
end
r_CI = quantile(mx, [0.025, 0.975]);
%Plot the results.
if isempty(varargin) || ~strcmp(varargin{1}, 'noplot')
figure(2)
plot(phi0, spline0, 'r', 'LineWidth', 2)
hold on
plot(phi0, splineU, ':r', 'LineWidth', 2)
plot(phi0, splineL, ':r', 'LineWidth', 2)
plot(phi0, null0, 'k', 'LineWidth', 2)
plot(phi0, nullL, 'k:', 'LineWidth', 2)
plot(phi0, nullU, 'k:', 'LineWidth', 2)
plot([phi0(imx) phi0(imx)], [null0(imx), spline0(imx)], 'LineWidth', 2)
hold off
ylabel('Amplitude')
xlabel('Phase')
axis tight
end
end
% Generate a design matrix X (n by nCtlPts) for a phase signal (n by 1)
function X = spline_phase0(phase,nCtlPts)
% Define Control Point Locations
c_pt_times_all = linspace(0,2*pi,nCtlPts+1);
s = 0.5; % Define Tension Parameter
% Construct spline regressors
X = zeros(length(phase),nCtlPts);
for i=1:length(phase)
nearest_c_pt_index = max(find(c_pt_times_all<=mod(phase(i),2*pi)));
nearest_c_pt_time = c_pt_times_all(nearest_c_pt_index);
next_c_pt_time = c_pt_times_all(nearest_c_pt_index+1);
u = (mod(phase(i),2*pi)-nearest_c_pt_time)/(next_c_pt_time-
nearest_c_pt_time);
p=[u^3 u^2 u 1]*[-s 2-s s-2 s;2*s s-3 3-2*s -s;-s 0 s 0;0 1 0 0];
X(i,mod(nearest_c_pt_index-2:nearest_c_pt_index+1,nCtlPts)+1) = p;
end
end
```

Highlights

- We use generalized linear models to assess cross-frequency coupling.
- The method allows direct computation of confidence in the resulting statistic.
- The method accurately detects biphasic cross-frequency coupling.
- The resulting statistic is easily interpretable.

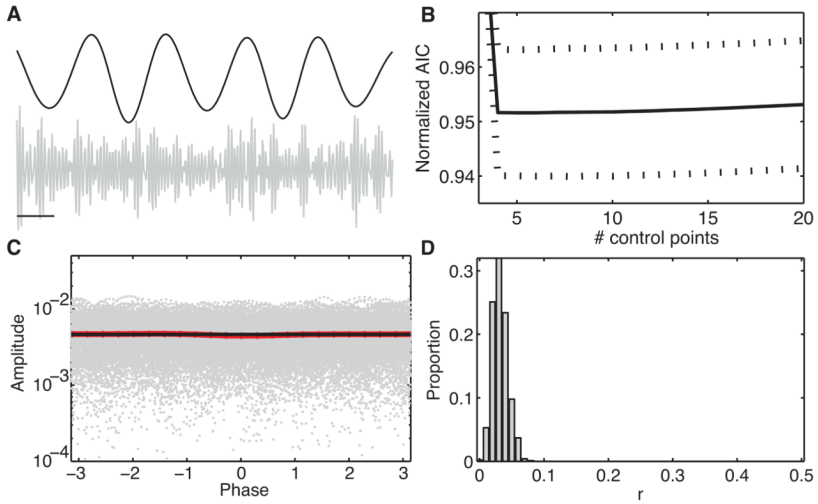

Figure 1. No phase-amplitude coupling results in small values of the CFC statistic *r*

(A) Example traces of the low (black) and high (blue) frequency activities. Scale bar indicates 0.1 s. (B) Normalized AIC computed for 100 simulations of pink noise time series. The minimum occurs at $n = 5$. (C) Example phase-amplitude values (gray points) and fits of the null model (black; mean is solid curve, and 95% confidence interval is dotted curve) and spline model (red; mean is solid curve, and 95% confidence interval is dotted curve); both confidence intervals are close to the mean curves. In this example, the largest difference between the two models occurs near phase 0. (D) Distribution of *r* values for 1000 simulations of pink noise data.

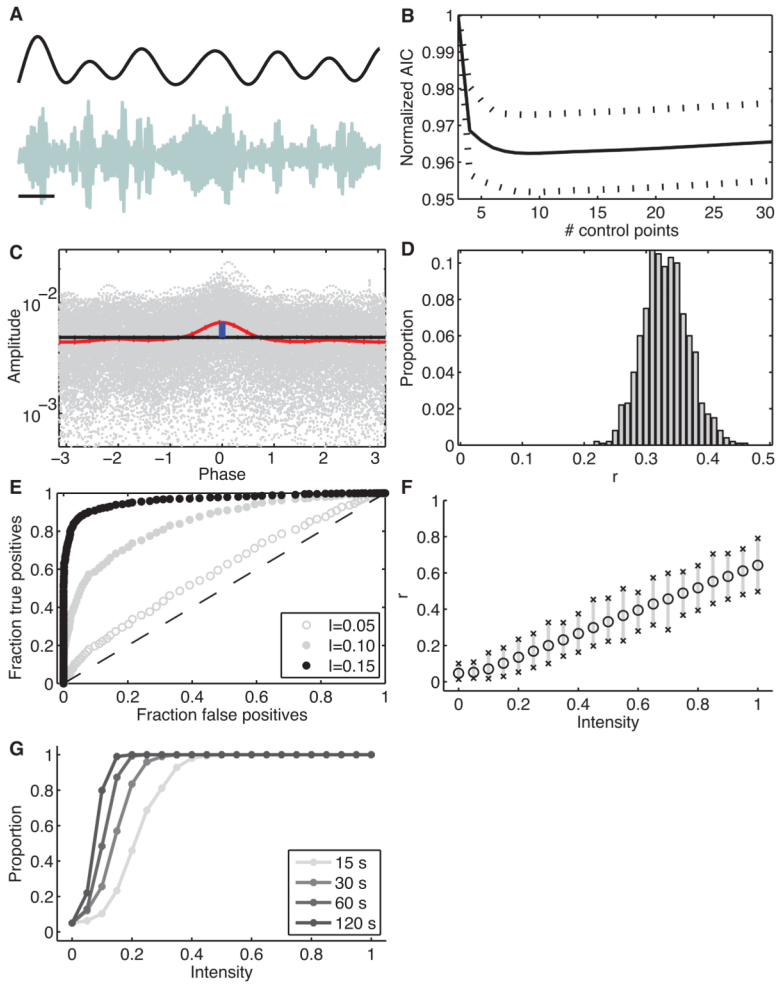

Figure 2. The GLM-CFC statistic detects monophasic coupling

(A) Example traces of the low (black) and high (blue) frequency activities, with $I = 2$. Scale bar indicates 0.1 s. (B) Normalized AIC computed for 100 simulations of monophasic CFC. The minimum occurs near $n = 8$. (C) Example phase-amplitude values (gray points), fits of the null (black curve) and spline (red curve) models, and the largest difference between the two models (blue vertical line). (D) Distribution of r values for 1000 simulations of monophasic CFC. (E) ROC for the statistic r for three different values of the intensity I (see legend). (F) The statistic r versus the modulation intensity I. Circles indicate the median, and x the maximum and minimum values at each intensity. (G) The power of the statistic versus intensity for four different durations of data (see legend). As the duration of data increases, lower intensity monophasic coupling is detectable.

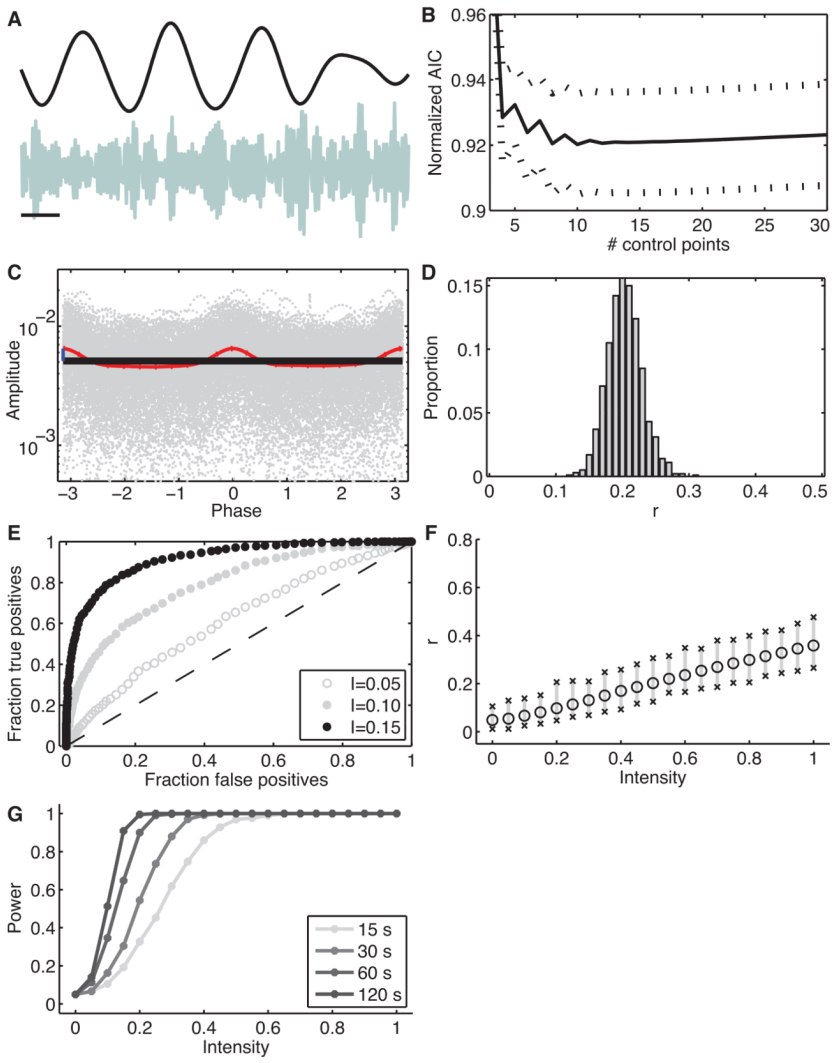

Figure 3. The GLM-CFC statistic detects biphasic coupling

(A) Example traces of the low (black) and high (blue) frequency activities, with $I = 0.5$. Scale bar indicates 0.1 s. (B) Normalized AIC computed for 100 simulations of biphasic CFC. The minimum occurs near $n = 10$. (C) Example phase-amplitude values (gray points), fits of the null (black) and spline (red) models, and the largest difference between the two models (blue vertical line). (D) Distribution of r values for 1000 simulations of biphasic CFC. (E) ROC for the statistic r for three different values of the intensity I (see legend). (F) The statistic r versus the modulation intensity I; see Figure 2F. (G) The power of the statistic versus intensity for four different durations of data (see legend).

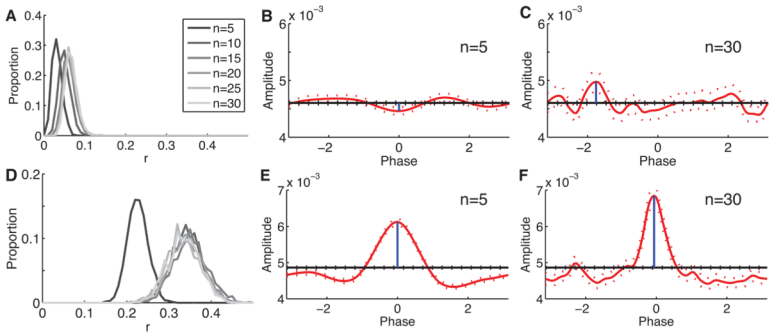

Figure 4. Impact of the number of control points on the statistic *r*

(A) Distribution of *r* values for the case of no CFC using different choices for *n* (see legend). (B,C) Example phase-amplitude GLM-CFC model fits for data with no CFC using $n = 5$ (B) and $n = 30$ (C). (D) Distribution of *r* values for the case of monophasic CFC using different choices for *n* (see legend in (A)). (E,F) Example phase-amplitude GLM-CFC model fits for data with monophasic CFC using $n = 5$ (B) and $n = 30$ (C).

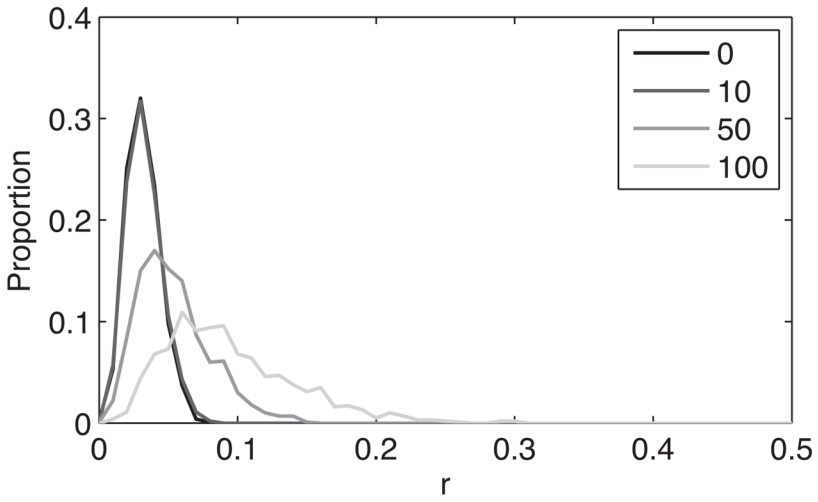

Figure 5. Impact of large amplitude outliers on the statistic *r*

Distribution of *r* values for outliers of different strength (indicated in legend). The distributions without an outlier (label 0) and with an outlier of strength 10 overlap.

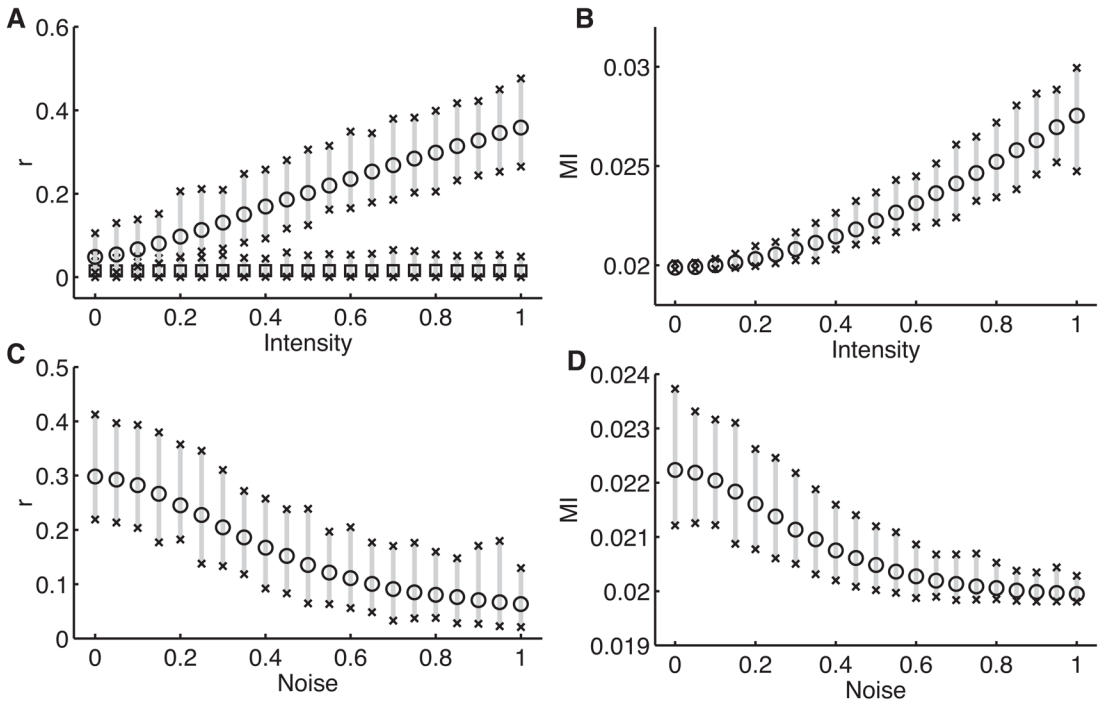

Figure 6. Comparison of the GLM-CFC method and two existing procedures

(A,B) Three CFC measures versus the intensity. In (A), the proposed GLM-CFC method (black circles), and the GLM method using only three covariates (black squares) are compared. In (B), the MI is shown. (C,D) The proposed statistic r (C) and MI (D) versus the amount of noise (i.e., independent low and high frequency activity) added to the signal. Noise degrades both CFC measures. Symbols in all subfigures follow the conventions in Figure 2F.