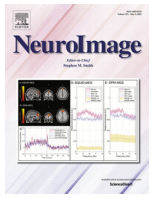
NeuroImage 271 (2023) 119998

Contents lists available at ScienceDirect

NeuroImage

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



Coherence fails to reliably capture inter-areal interactions in bidirectional neural systems with transmission delays

Jarrod Robert Dowdall [a,b,*], Martin Vinck [a,c,*]

[a] Ernst Strüngmann Institute (ESI) for Neuroscience in Cooperation with Max Planck Society, Frankfurt am Main, Germany
[b] Robarts Research Institute, Western University, London, Ontario, Canada
[c] Donders Centre for Neuroscience, Department of Neuroinformatics, Radboud University, Nijmegen, The Netherlands

ARTICLE INFO

Keywords:
Inter-areal coherence
Granger causality
Functional connectivity
Source mixing
Communication-Through-Coherence (CTC)

ABSTRACT

Accurately measuring and quantifying the underlying interactions between brain areas is crucial for understanding the flow of information in the brain. Of particular interest in the field of electrophysiology is the analysis and characterization of the spectral properties of these interactions. Coherence and Granger-Geweke causality are well-established, commonly used methods for quantifying inter-areal interactions, and are thought to reflect the strength of inter-areal interactions. Here we show that the application of both methods to bidirectional systems with transmission delays is problematic, especially for coherence. Under certain circumstances, coherence can be completely abolished despite there being a true underlying interaction. This problem occurs due to interference caused in the computation of coherence, and is an artifact of the method. We motivate an understanding of the problem through computational modelling and numerical simulations. In addition, we have developed two methods that can recover the true bidirectional interactions in the presence of transmission delays.

1. Introduction

A primary interest in neuroscience, particularly in the fields of fMRI and electrophysiology, is characterizing the neural correlates of functional connectivity and information flow between neuronal populations. In practice, researchers often rely on methods that quantify the statistical dependencies between time-series data (Pesaran et al., 2018). These statistical dependencies are often quantified with either coherence, an undirected measure of the dependencies, or Granger-causality, a directed measure of the dependencies. The understanding is that the statistical dependencies quantified by coherence and Granger-causality arise as a direct result of the functional interactions and propagation of signals between anatomically connected brain areas (Pesaran et al., 2018; Schneider et al., 2021). It follows then that differences in coherence and Granger-causality are typically interpreted as reflecting true differences in either the strength of inter-areal (network) interactions, or properties of the transmitted signals (Bastos et al., 2015a; Vezoli et al., 2021). Indeed, we note that both coherence and Granger-causality can be informative for analyzing the statistical dependencies in relatively noise-free unidirectional systems. However, neural data is not noise free and brain areas are reciprocally connected.

For instance, anatomical studies in non-human primates have found that the large majority inter-areal connections are reciprocal in nature (Markov et al., 2014a). Thus, with regard to the methods, this system is expected to be bidirectional with variable transmission delays (i.e., causal). Furthermore, transmission delays will not only depend on factors such as white matter distance, and conduction delays (due to the fiber diameter and myelination), but also factors such as membrane time constants (which also depend on conductance states), neuromodulators and receptor mechanisms (e.g., NMDA vs. AMPA) (Bucher and Goaillard, 2011; Destexhe et al., 2003; Tsodyks and Markram, 1997). Although it is assumed that coherence is invariant to transmission delays (i.e., a constant phase delay), we show that this does not hold for bidirectional systems with transmission delays. Surprisingly, we found that coherence and Granger-causality are highly dependent on transmission delays even when the inter-areal interactions were otherwise identical.

The reason bidirectional communication with delays is problematic for coherence can be understood by recognizing that the cross-spectral density (CSD) is the Fourier transform of the cross-covariance function. Because the cross-covariance that arises with bidirectional interactions contains two peaks separated by the total transmission delay, the Fourier transform will completely cancel for frequencies where the peak sepa-

Abbreviations: AR, Autoregressive; CSD, Cross-spectral density; E-I, Excitatory-Inhibitory; FB, Feedback; FF, Feedforward; FIR, Finite impulse response; LFP, Local field potential; GGC, Granger-Geweke causality; PLV, Phase locking value; PSD, Power spectral density; PUC, Proportion of unidirectional coherence; TCC, Truncated cross-covariance; VAR, Vector autoregressive.

* Corresponding authors.
E-mail addresses: jarrod.dowdall@gmail.com (J.R. Dowdall), martin.vinck@esi-frankfurt.de (M. Vinck).

https://doi.org/10.1016/j.neuroimage.2023.119998.
Received 20 August 2022; Received in revised form 5 February 2023; Accepted 27 February 2023
Available online 28 February 2023.

ration is a half-integer multiple of the cycle length. In fact, the only frequencies that do not exhibit some amount of interference are those where the cycle length is an exact integer multiple of the total transmission delay. It is important to emphasize here that this is a failure of the method (coherence) to capture the true underlying inter-areal interactions, rather than interference at the level of the communicating areas. That is, this interference is purely a consequence of the method, and is an artifact.

The implication is that differences in coherence and Granger-causality cannot be interpreted to be strictly the result of differences in inter-areal communication. However, given certain assumptions on the activity of each area and the transmitted signals, we show that it is possible to recover the underlying bidirectional interactions despite the transmission delays. To that end, we developed two methods for quantifying inter-areal in bidirectional systems with transmission delays.

2. Methods

2.1. Bidirectional source mixing model

We modelled the activity of each area, x_1 and x_2, as the superposition of the intrinsic activity, additive $1/f$, and the activity received from the other area weighted by the inter-areal connectivity strength, w_{21} and w_{12}, and delayed by τ to reflect the transmission delay,

$$x_1^{(obs)}[n] = x_1^{(int)}[n] + w_{21}\, x_2^{(int)}[n-\tau] + \eta_1[n]\,, \tag{1}$$

$$x_2^{(obs)}[n] = x_2^{(int)}[n] + w_{12}\, x_1^{(int)}[n-\tau] + \eta_2[n]\,, \tag{2}$$

where η is a term containing $1/f$ fluctuations.

Note, in all simulations shown here, $w_{21} = w_{12}$ and τ was the same in both directions.

The intrinsic activity in each area was described by a pseudo-periodic second order autoregressive model (AR(2)),

$$x_1^{(int)}[n] = a_{11} x_1^{(int)}[n-1] + a_{12} x_1^{(int)}[n-2] + \epsilon_1[n]\,, \tag{3}$$

$$x_2^{(int)}[n] = a_{21} x_2^{(int)}[n-1] + a_{22} x_2^{(int)}[n-2] + \epsilon_2[n]\,. \tag{4}$$

Here, $\epsilon[t]$ is the stochastic drive of the system, and $\epsilon[t] = \mathcal{N}(\mu, \sigma^2)$. We chose to model the intrinsic activity in each area using pseudo-periodic AR(2) models because they have been shown to reproduce the statistical properties of stationary gamma oscillations in primary visual cortex, and provide mean-field approximations of E-I circuits driven by stochastic input (Spyropoulos et al., 2022). Note, only the oscillatory part of each area was transmitted (i.e., not the background $1/f$), which follows from the empirical observations that coherence spectra are peaked, suggesting that only the oscillatory activity in the sender is coherent with the receiver (Schneider et al., 2021).

2.2. Parameters of the auto-regressive models

The following equations were used to determine the coefficients of the AR(2) models given the desired peak frequency and bandwidth:

$$a_1 = \frac{4\, a_2 \cos(\omega_{max})}{a_2 - 1}\,, \tag{5}$$

$$a_2 = -R^2\,, \tag{6}$$

where $\omega_{\max}$ is the peak frequency in radians, and R is the moduli of the models complex roots and controls the bandwidth of the pseudo-periodic oscillations.

The following equation was used to determine the variance of the white-noise process for the desired peak power $S_{\max}$ in the power spectral density (PSD) at the peak frequency $\omega_{\max}$:

$$\sigma_\epsilon^2 = S_{max}\, \frac{\left(R^2-1\right)^2 \left(R^4 - 2\cos(2\omega_{max})R^2 + 1\right)}{\left(R^2+1\right)^2}\,, \tag{7}$$

which can be derived from substituting Eq. (5) and Eq. (6) into Eq. (24).

2.3. Analytical description of the interference pattern

In the Source Mixing model the observed PSDs of x_1 and x_2 are:

$$S_{11}^{(obs)}(f) = S_{11}^{(int)}(f) + w^2 S_{22}^{(proj)}(f)\,, \tag{8}$$

$$S_{22}^{(obs)}(f) = S_{22}^{(int)}(f) + w^2 S_{11}^{(proj)}(f)\,, \tag{9}$$

omitting background $1/f$ and letting $w = w_{12} = w_{21}$ for brevity.

In the case of instantaneous transmission (i.e., $\tau = 0$), the CSD is given by:

$$S_{12}(f) = w\, S_{11}^{(proj)}(f) + w\, S_{22}^{(proj)}(f)\,. \tag{10}$$

Note, this assumes the intrinsic signals in each area are uncorrelated.

The transmission delay leads to a phase shift in the projected signal relative to the intrinsic signal such that the CSD becomes:

$$S_{12}(f) = w\, S_{11}^{(proj)}(f) + w\, S_{22}^{(proj)}(f)\, e^{-if2\pi\,\Delta\tau}\,, \tag{11}$$

where $\Delta\tau$ is the total transmission delay in seconds.

Note that

$$|S_{12}(f)|^2 = \Re(S_{12}(f))^2 + \Im(S_{12}(f))^2\,. \tag{12}$$

Given that $w^2 S_{11}^{(proj)}(f)$ is strictly real-valued, it follows that

$$\begin{aligned} |S_{12}(f)|^2 = {} & \left(wS_{11}^{(proj)}(f) + wS_{22}^{(proj)}(f)\,\cos(2\pi f\Delta\tau)\right)^2 \cdots \\ & + \left(wS_{22}^{(proj)}(f)\,\sin(2\pi f\Delta\tau)\right)^2\,, \end{aligned} \tag{13}$$

which becomes

$$\begin{aligned} |S_{12}(f)|^2 = {} & \left(wS_{11}^{(proj)}(f)\right)^2 + \left(wS_{22}^{(proj)}(f)\right)^2 \cdots \\ & + 2w^2 S_{11}^{(proj)}(f) S_{22}^{(proj)}(f)\,\cos(2\pi f\Delta\tau)\,. \end{aligned} \tag{14}$$

If $S^{(proj)}(f) = S_{11}^{(proj)}(f) = S_{22}^{(proj)}(f)$, then Eq. (14) reduces to

$$|S_{12}(f)|^2 = 2w^2 S^{(proj)}(f)^2(1 + \cos(2\pi f\Delta\tau))\,. \tag{15}$$

Thus, the total transmission delay introduces a complex term in the CSD that reflects the phase difference between the two transmitted signals per frequency as a function of the total transmission delay.

The net result is the sum of the transmitted signals depends on the magnitude of the real part of $e^{-if2\pi\,\Delta\tau}$ as seen in Eq. (15), which explains why the pattern of interference across frequencies manifests as a cosine function. Thus, there is no interference at frequency f where $cos(f2\pi\,\Delta\tau) = 1$, and maximal interference when $cos(f2\pi\,\Delta\tau) = -1$.

2.4. Method 1: Truncated cross-covariance

The truncated cross-covariance (TCC) method is a naive approach that assumes the signals projected from one area to the other will cause peaks in the cross-covariance function at opposite lags.

In this case, it is possible to partially isolate the influence of one area on another by multiplying the cross-covariance with a Heaviside step function Θ before applying the Fourier transform to derive the cross spectral density for computing directional coherence measure.

Thus, the directed cross-spectral densities are computed as follows:

$$S_{x_1\to x_2} = \mathcal{F}\left(s_{12}(\tau)\,\Theta_{x_1\to x_2}(\tau)\right) \tag{16}$$

and

$$S_{x_2\to x_1} = \mathcal{F}\Big(s_{12}(\tau)\,\Theta_{x_2\to x_1}(\tau)\Big), \tag{17}$$

where $s_{12}(\tau)$ is the cross-covariance function

$$s_{12}(\tau) = \mathbb{E}\big[x_1[\tau]\,x_2[t-\tau]\big] \tag{18}$$

and

$$\Theta_{x_1\to x_2}(\tau) := \begin{cases} 1, & \tau > 0 \\ 0, & \tau \le 0 \end{cases} \tag{19}$$

and

$$\Theta_{x_2\to x_1}(\tau) := \begin{cases} 0, & \tau > 0 \\ 1, & \tau \le 0\,. \end{cases} \tag{20}$$

Then the TCC estimate of the directed coherence is

$$C_{x_1\to x_2}(f) = \frac{|S_{x_1\to x_2}|^2}{S_{11}(f)\,S_{22}(f)}\,, \tag{21}$$

$$C_{x_2\to x_1}(f) = \frac{|S_{x_2\to x_1}|^2}{S_{11}(f)\,S_{22}(f)}\,. \tag{22}$$

2.5. Method 2: Model fit

The following additive model (sum of a spectral peak, AR(2) model, and background $1/f$) was fit to the observed power spectral density using non-linear least squares regression (using `lsqnonlin()`) in Matlab):

$$S_{obs}(\omega) = S_{peak}(\omega) + \eta(\omega)\,, \tag{23}$$

where $S_{peak}(\omega)$ is the power spectrum of the AR(2) and $\eta(\omega)$ is the power spectrum of the $1/f$ component.

The power spectral density of the AR(2) was given by the equation

$$S_{peak}(\omega) = \frac{\sigma^2}{1 + a_1^2 + a_2^2 - 2a_1(1-a_2)\cos(\omega) - 2a_2\cos(2\omega)}\,, \tag{24}$$

where a_1 and a_2 are the coefficients of the auto-regressive model, σ^2 is a scaling factor, and $\omega \in [0, \pi]$ is the frequency in radians.

And the power spectral density of $1/f$ component given by

$$\eta(\omega) = g^{\omega\,\alpha}\,, \tag{25}$$

where $g \in (0, \infty)$ is a gain factor that scales the $1/f$ spectrum, $\alpha \in [-2, 0]$ reflects the slope, and $\omega \in [0, \pi]$ is the frequency in radians per sample.

To fit the auto-covariance functions to the observed cross-covariance we followed these steps:

1. Each area's auto-covariance function, $s_{11}(\tau)$ and $s_{22}(\tau)$, were estimated from the parameters of the AR(2) coefficients of the model fit.
2. The individual auto-covariance functions were lag-shifted in their respective direction according to biophysically reasonable τ (e.g., 0 to 10 ms) (Ferraina et al., 2002; Miller, 1975; Swadlow et al., 1978).
3. The individual auto-covariance functions were then differently weighted by δ, and summed to produce an estimate cross-covariance function, $\hat{s}_{12}(\tau)$, as follows

$$\hat{s}_{12}(\tau) = \delta\,s_{11}(\tau) + (1-\delta)\,s_{22}(\tau)\,, \tag{26}$$

where $\delta \in [0, 1]$. In practice this step is necessary when $w_{12} \neq w_{21}$. However, for these simulations it was not strictly necessary as the connectivity weights were always equal in both directions (i.e., $w = 0.5$).
4. For each (τ, δ), we correlated the estimate cross-covariance with the observed cross-covariance to determine the best fit as follows:

$$\rho = \frac{\mathbb{E}\big[\hat{s}_{12}(\tau)\,s_{12}(\tau)\big]^2}{\mathbb{E}\big[\hat{s}_{12}(\tau)\big]^2\,\mathbb{E}\big[s_{12}(\tau)\big]^2}\,. \tag{27}$$

The best fit (determined by the maximum correlation coefficient ρ over all τ and δ) was taken as reflecting the underlying transmission delay. For simplicity, we assumed that τ was approximately equal in each direction (note, fitting unequal delays would require the additional step of taking the max cross-correlation across all lags and τ).
5. To estimate the connectivity weights w_{12} and w_{21}, we first estimated the total combined connectivity weight as follows

$$\hat{w}_{total} = \sqrt{\rho\,\frac{\mathbb{E}\Big[\big(s_{12}(\tau)\big)^2\Big]}{\mathbb{E}\Big[\big(\hat{s}_{12}(\tau)\big)^2\Big]}}\,. \tag{28}$$

Note, this is simply a re-scaling through a ratio of energies, and ρ forces the weight to the linearly correlated portion of the ratio.
6. Lastly, the total connectivity weight was divided into the connectivity weights for each direction according to the δ of the best fit as follows:

$$w_{12} = \delta\,\hat{w}_{total}\,, \tag{29}$$

$$w_{21} = (1-\delta)\,\hat{w}_{total}\,. \tag{30}$$

2.5.1. Approach to fitting the PSD

We noted two strategies that assisted in fitting the power spectral densities. One caveat of power spectral analysis is that power of each additive component is unevenly distributed across frequencies. This leads to difficulties fitting each component of the PSD as it implies that the information for each component is not uniformly distributed across frequencies (i.e., each frequency contains a different ratio of power from each component). For instance, when the spectral peaks are narrow band, the frequency range that contains the most information useful for fitting that peak is also limited to a narrow frequency range. In this case, fitting the model over a wide frequency range may fit the $1/f$ component well, but miss the spectral peak altogether. Whereas, when the spectral peak is broadband it is necessary to fit the model over a much wider frequency range in order to accurately estimate both the $1/f$ and spectral peak.

In practice, it is possible to overcome this problem by visually inspecting the PSD to pre-determine an appropriate frequency range that balances the information used to fit both components. However, because we aimed to fit the spectrum across a wide range of possible spectral peaks (from very broadband to very narrow), we needed a more general approach that would perform well regardless of the bandwidth of each component. We found that fitting the PSD over multiple (predetermined but fixed) frequency ranges, and selecting the best overall fit (minimum L1 norm), performed nearly equally well across spectra with very different bandwidths. We further refined this fit, by using it to set the initial conditions of a fit over a very wide frequency range, [5, 600] Hz. In all cases, the fit that minimized the L1 norm over the frequency range [5, 600] Hz was taken as the best fit.

The second strategy we employed was repeating the multiple bandwidth fitting procedure each time with random initial conditions (we found 35 random initializations were sufficient). As with the multiple bandwidth fit, the best fit (as determined by the L1 norm) across all initial conditions was retained. Thus, the final fit was the fit that minimized L1 norm over the frequency interval [5, 600] Hz. We did not perform any additional checks on the quality of the fit as these two approaches combined were sufficient to accurately fit both the spectral peak and $1/f$ components across a wide range of spectral bandwidths.

2.6. Proportion of unidirectional coherence (PUC)

We assessed the ability of each method to recover the true unidirectional coherence (defined analytically) by quantifying the proportion of unidirectional coherence (PUC) each method was able to recover.

For the source mixing model, the expected unidirectional coherence is:

$$C_{x_1 \to x_2}(f) = \frac{w^2 \left| S_{11}^{(int)}(f) \right|^2}{S_{11}^{(obs)}(f)\, S_{22}^{(obs)}(f)}, \tag{31}$$

$$C_{x_1 \leftarrow x_2}(f) = \frac{w^2 \left| S_{22}^{(int)}(f) \right|^2}{S_{11}^{(obs)}(f)\, S_{22}^{(obs)}(f)}. \tag{32}$$

The proportion of unidirectional coherence (PUC) was defined as

$$PUC_{x_1 \to x_2} := 1 - \sqrt{\frac{\sum_{f=k}^{N} \left(C_{x_1 \to x_2}(f) - \hat{C}_{x_1 \to x_2}(f) \right)^2}{\sum_{f=k}^{N} C_{x_1 \to x_2}^2(f)}}, \tag{33}$$

where $\hat{C}_{x_1 \to x_2}(f)$ is the estimated unidirectional coherence from either method. For both methods PUC was computed over the frequency range of 5 to 120 Hz.

2.7. Bivariate vector autoregressive model

To demonstrate that the interference in coherence due to transmission delays was not dependent on the source mixing model, we simulated two bidirectionally coupled areas with bivariate vector autoregressive (VAR) models:

$$x_{1,t} = a_{11}\, x_{1,t-1} + a_{12}\, x_{1,t-2} + w_{21}\, x_{2,t-\tau} + \epsilon_{1,t} \tag{34}$$

$$x_{2,t} = a_{21}\, x_{1,t-1} + a_{22}\, x_{1,t-2} + w_{12}\, x_{1,t-\tau} + \epsilon_{2,t}. \tag{35}$$

The parameters of the VAR simulations were otherwise identical to the source mixing simulations shown in Fig. 1, which included the model coefficients for the AR(2) models as well as the $1/f$ background fluctuations. The only difference was the connectivity weight for both areas was adjusted to make the VAR model stable ($w = 0.0015$).

2.8. Background $1/f$ fluctuations and numerical simulations

Note that in the presence of $1/f$ fluctuations we can describe

$$x^{(obs)}[t] = x^{(int)}[t] + \eta[t], \tag{36}$$

where η is a background term containing $1/f$ fluctuations. We assumed that these $1/f$ fluctuations were not projected, and uncorrelated with each area's intrinsic activity.

The power spectrum of $1/f$ is approximately equal to the inverse of the frequency

$$S(f) \propto \frac{1}{f}. \tag{37}$$

We simulated the $1/f$ spectrum according to

$$S(f) = \frac{f_0}{f} P, \tag{38}$$

then inverse Fourier transformed the spectrum to arrive at the time-domain coefficients of an N order FIR filter, where N was equal to the number of samples in each simulated time series epoch. For all simulations that included background $1/f$, f_0 was set to the receiver's peak frequency (60 Hz) and $P = 3^{-1}$.

The sampling rate for all simulations was $fs = 2000$. The epoch length was 1001 samples, 2500 independent epochs were generated for each run of a simulation, and 15 independent runs of each simulation were generated and then averaged to get a central estimate of the power, coherence and Granger-causality spectra.

3. Results

Our primary question was how inter-areal transmission delays lead to changes in the coherence between two areas with bidirectional communication. We modelled inter-areal interactions with the Source Mixing model (see Section 2.1) (Schneider et al., 2021). In the Source Mixing model, the observed signal in each area is composed of its intrinsic signal plus the transmitted signal received from the other area. The primary advantage of the Source Mixing model is there is no recurrent coupling that could potentially lead to intrinsic interference.

The Source Mixing model has been previously shown to account for changes in inter-areal coherence as a result of a complementary change in the oscillatory power of the source projecting area (Schneider et al., 2021). However, these results focused on the specific case of unidirectional communication between two brain areas. In the case of unidirectional communication, the magnitude of coherence is unaffected by transmission delays. However, the anatomical connectivity between brain areas is typically reciprocal (Chaudhuri et al., 2015; Markov et al., 2014b; Vezoli et al., 2021).

3.1. Coherence depends on inter-areal transmission delays

To investigate the influence of transmission delays on the magnitude of inter-areal coherence, we simulated the local field potential (LFP) of two bidirectionally connected brain areas with quasi-periodic causal autoregressive (AR) models. Note, the main result of these simulations, namely that coherence depends on transmission delays, is not dependent on AR models. In fact, coherence also shows a dependence on transmission delays for white-noise signals (see Fig. 2a). Thus, the results presented here generalize to all systems with bidirectional communication and transmission delays. We further note that this problem can also occur at the level of spiking signals when computing spike-field and spike-spike correlations.

We simulated two bidirectionally connected areas (x_1 and x_2) with equal connectivity weights ($w = 0.15$), and increased the transmission delay, τ, from 0 to 5 ms (Fig. 1a). Coherence and Granger-Geweke causality were computed between the observed signals of each area, which were the superposition of the intrinsic activity of that area, the input from the other area and additive $1/f$ background fluctuation (see Section 2.1). Note the intrinsic oscillations and background $1/f$ in both areas were all independent and uncorrelated (i.e., there was no synchronization or coupling between oscillators, thus the only source of coherence were the transmitted signals). For all simulations, we set $\tau_{x_1 \to x_2} = \tau_{x_1 \leftarrow x_2}$ under the assumption that the transmission delay between the two bidirectionally coupled areas would be approximately equal in both directions. However, it should be noted that the relevant metric for the effect on coherence is the total transmission delay between the two areas, i.e., $\Delta\tau = \tau_{x_1 \leftarrow x_2} + \tau_{x_1 \to x_2}$. The power spectral densities of x_1 and x_2 overlapped, but with a slight shift in the peak frequency (60 and 65 Hz, respectively, see Fig. 1b). Note, the peak frequency shift serves to illustrate that it is not necessary that the transmitted signals have exactly the same peak frequency, however some spectral overlap is necessary.

We observed that increasing the transmission delay (τ) led to a decrease in the magnitude of coherence (Fig. 1c). A similar pattern of interference was also observed for phase locking value (PLV) (Fig. 1d). The maximum decrease in coherence (and PLV) occurred for $\tau = 4$ ms, and resulted in a visible dip at 62.5 Hz. Importantly, due to the nature of the simulated model, this change in coherence cannot be explained by a change in power, connectivity weight, nor a shift in the peak frequencies, and therefore does not reflect a true change in the strength of the interaction between these areas.

We considered the possibility that Granger-Geweke causality (GGC) analysis may be able to recover the inter-areal interactions independent of transmission delays (i.e., the spectral influence of $x_1 \to x_2$ and

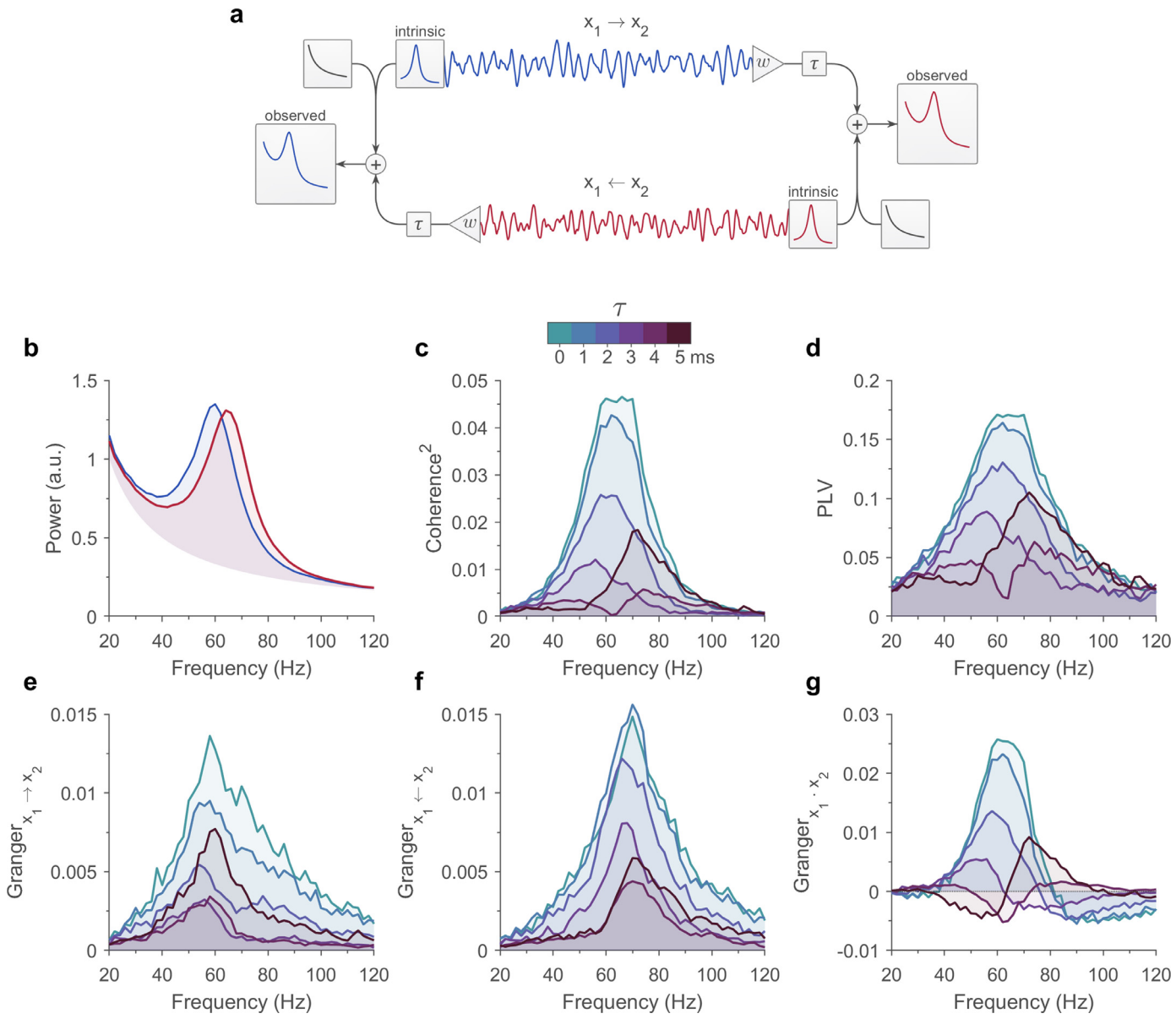


Fig. 1. Bidirectional communication with inter-areal delays. Comparison between two reciprocally connected areas with different transmission delays (τ) using the Source Mixing model. (a) Source Mixing model circuit diagram. For these simulations, both areas were reciprocally connected with the same connectivity weight was $w = 0.15$. Only the intrinsic gamma-band oscillations in each area (peak frequency at 60 Hz for x_1 and 65 Hz for x_2) were projected. These simulations included additive $1/f$ background fluctuations, which were uncorrelated with the gamma-band oscillations. (b) The observed power spectra for x_1 and x_2 (blue and red lines, respectively) contains both the intrinsic gamma-band activity in each area, additive $1/f$ background fluctuations, and the projected signal from the other area (weighted by the connectivity weight). The power of the intrinsic gamma-band activity and $1/f$ background fluctuations were identical for both areas. Note, both the intrinsic gamma-band activity and $1/f$ background fluctuations were uncorrelated between each area. That is, the only part of the activity in each area that was correlated with the other area was the projected signal. (c) The coherence spectra between x_1 and x_2 for τ between 0 and 5 ms. Note, an interference pattern emerges with increasing transmission delay, which appears maximal for a $\tau = 4$ ms with a sharp dip at 62.5 Hz. (d) Phase locking value (PLV) between x_1 and x_2 for τ between 0 and 5 ms. Note, an interference pattern emerges with increasing transmission delay, similar to coherence, and also appears maximal at 4 ms delay and ≈ 62 Hz. (e) The non-parametric Granger-Geweke causality (GGC) spectra for x_1 to x_2. (f) The non-parametric GGC spectra for x_2 to x_1. (g) The instantaneous GGC spectra between x_1 and x_2. Notice that the instantaneous GGC is negative for some frequencies. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

$x_1 \leftarrow x_2$) (Dhamala et al., 2018). Although GGC did not show the dip at 62.5 Hz observed in coherence, GGC did decrease with increasing transmission delay (Fig. 1e-g). However, we note that when no $1/f$ background fluctuations were added to the observed time series, Granger-causality did not depend on the transmission delay (see Fig. S1e-g). Nonetheless, these results indicate that both coherence and GGC are dependent on transmission delays in systems with bidirectional communication and additive noise.

3.2. The source of the interference

The pattern of coherence that results from varying transmission delays is due to interference in the cross-covariance function. To understand why this effect occurs, we note that the cross spectral density (CSD) can be computed by taking the Fourier transform of the cross-covariance function. In Fig. 2, we have illustrated the procedure of computing the Fourier transform on the cross-covariance function for two

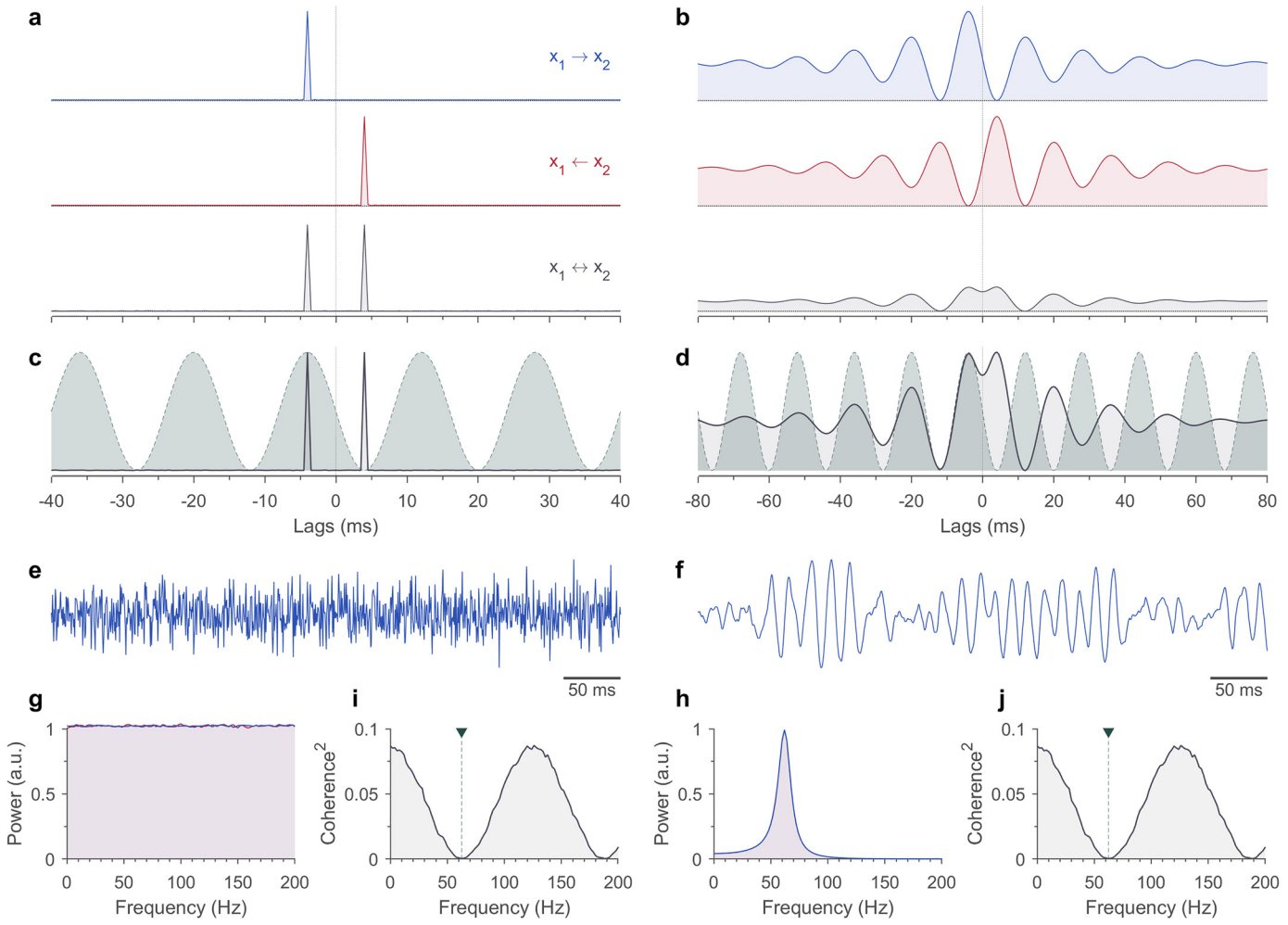


Fig. 2. Comparison of weak versus strong auto-correlation of the transmitted signals. (a) Auto-covariance functions for x_1 (blue) and x_2 (red) and cross-covariance (black) function for a bidirectional system with intrinsic signals modelled as zero-mean white-noise process (i.e., no auto-correlation). The connectivity weights, $w = 0.15$, and transmission delays $\tau = 4$ ms, were equal for both directions. (b) The same as in panel a, but with strongly auto-correlated signals. The intrinsic signals in this example were modelled as AR(2) processes, both with the same peak frequency at 62 Hz (ω_{max}), and $R = 0.98$. The auto-covariance and cross-covariance functions are plotted normalized to the overall max. Note, unlike the white-noise process shown in panel a, the magnitude of the cross-covariance for strongly auto-correlated signals with an anti-phase relationship is significantly reduced due to interference. (c-d) Respective cross-covariance functions (black) showing a cosine (dashed line) function superimposed ($cos(2\pi f[t-\tau])$, where $f = 62.5$ Hz and $\tau = 4$ ms.). The cross-covariance functions are plotted according to their max, and the cosine function illustrates the effect of the basis functions of the Fourier transform. (e-f) Example simulated time series of the intrinsic signals of area x_1 for the white-noise process and AR(2), respectively. (g-h) Observed power spectra (note, x_1 and x_2 overlap in both examples). (i-j) Coherence spectra for the white-noise process and AR(2), respectively. The vertical dashed line indicates the first frequency of maximal interference (i.e., 62.5 Hz). Note, despite distinctly different power spectra between the two examples shown here, their coherence spectra are identical, and show a cosine modulation across frequencies that depends on the transmission delay (4 ms). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

pairs of areas with contrasting power spectra, i.e., flat versus peaked power spectra. For this illustration, we simulated two pairs of areas with bidirectional communication, equal connectivity weight $w = 0.15$, and a fixed transmission delay in both directions $\tau = 4$ ms. For the purposes of illustrating the cause of the effect we omitted the additive $1/f$ background fluctuations in these simulations. The only difference between these two pairs of areas was the auto-correlation of their intrinsic signals. Thus, each pair reflects one extreme of autocorrelation that the transmitted signals between two area could have (that is, anywhere from weak or no auto-correlation up to strong auto-correlation, Fig. 2a vs. b respectively).

In the case where the signals transmitted between each area are serially uncorrelated (i.e., have white noise statistics), the cross-covariance function shows two delta peaks at $\pm\tau$ lags reflecting the transmission of two serially uncorrelated signals in each direction (Fig. 2a). Because the transmitted signals have equal power at all frequencies, coherence should be a constant across all frequencies. However, the coherence shows an interference pattern that is maximal at frequencies where $2 \cdot f\Delta\tau = 1$, where $\Delta\tau$ is the total transmission delay in seconds (see Section 2.3). Thus, in the absence of additive noise, the coherence spectra appear as a cosine modulation across frequencies whose frequency depends on the total transmission delay (Fig. 2i).

One way to understand this interference pattern is to envision the sum of two cosine waves each centered on one of the two peaks in the cross-covariance function (see Fig. 2c). It can be seen that the superposition of two cosines that are each shifted by $\tau = 0.25 \cdot$ (cycle duration) will maximally interfere as they are in anti-phase with respect to each another. Thus, for weakly auto-correlated signals, the interference pattern occurs, because the basis functions underlying the Fourier transform are sines and cosines. Importantly, this particular example exemplifies the nature of the problem, namely that the interference is the result of the method rather than a true modulation in the strength of the inter-areal communication.

In the second example, we considered the case where the transmitted signals were strongly auto-correlated. For these simulations we modelled the intrinsic activity of each area with AR(2) models that produced relatively strong pseudo-periodic oscillations with a center frequency of 62 Hz. The same AR(2) coefficients were used for both areas, and as in the previous example, they had equal connectivity weights $w = 0.15$, and a fixed transmission delay in both directions $\tau = 4$ ms. Note that although both areas were spectrally similar (i.e., identical AR(2) model coefficients), the white-noise processes used to drive each model were independent and uncorrelated, and thus so were the intrinsic signals of each area. In this second example, it is apparent that there is an additional source for the interference, which in this case occurs at the level of the cross-covariance function itself (see Fig. 2b). Because the cross-covariance function is the sum of the individual unidirectional cross-covariance functions, it follows that the superposition of two unidirectional cross-covariance functions that are in anti-phase will result in destructive interference. Thus, the stronger the auto-correlation, the more the individual unidirectional cross-covariance functions overlap in the cross-covariance function. Taken to its extreme, this result implies that the cross-covariance function between two areas with bidirectional transmission of equal magnitude cosine waves with a frequency of 62.5 Hz and a transmission delay of 4 ms will be entirely flat. One may suspect, especially in this extreme example, that the modulation in coherence is somehow the result of interference between the intrinsic signal in one area and the transmitted signal it receives from the other. However, it important to remember that in these simulations the intrinsic signals, and likewise the signals each area transmitted to the other, were strictly uncorrelated between the two areas. More plainly, there was no systematic phase relationship between the intrinsic activity of one area and the signal received from the other, and as such there was no interference at the level of the communication between areas.

These two examples clearly illustrate that this dependence of coherence on transmission delays is strictly methodological. In the case of weak auto-correlation, the interference occurs primarily at the level of the Fourier transform, and in the case of strong auto-correlation the interference can occur both in the cross-covariance function and Fourier transform.

3.3. The truncated cross-covariance method

We aimed to develop a method that could recover the underlying inter-areal interaction independent of the transmission delays. As a first naive approach, we reasoned that it should be possible to isolate the directed influences by appropriately truncating (or windowing) the cross-covariance function. The observation that weakly auto-correlated signals do not, or at least minimally, interfere with each other in the cross-covariance function given sufficient transmission delays, suggests that windowing the cross-covariance function based on the sign of lags may be sufficient to recover the directional influences in most cases. We refer to this approach as the truncated cross-covariance (TCC) method.

To that end, we multiplied the cross-covariance with a step function, a Heaviside conditional on the sign of the lags, and then applied the Fourier transform to estimate the cross-spectral density of the directional influences. Coherence can then be computed using this windowed estimate of the cross-spectral density (see Section 2.4). Note, TCC provides two estimates of directed coherence, one for each direction, based on the sign of the lags for which the Heaviside was equal to 1.

Given the transmitted signals show minimal overlap in the cross-covariance, the TCC method should approximate the coherence that would have occurred if the communication was unidirectional in each direction. Therefore, the effectiveness of the TCC method is expected to be highly dependent on auto-correlation structure of the transmitted signals, and to a lesser extent the transmission delay. To gain an understanding of the range signals over which the TCC is effective, we repeated the simulation depicted in Fig. 1a, but for fixed the transmission delay, and instead varied the auto-correlation of the transmitted signals (Fig. 3a). Similar to the previous simulation, the peak frequencies of each area were 60 and 65 Hz, with equal connectivity weight ($w = 0.15$), the transmission delay was 4 ms, and the observed activity in both areas included additive 1/f background fluctuations.

The auto-correlation of the transmitted signals was controlled by varying the modulus of the eigenvalue of the AR(2) models from 0.01 to 0.99 used to simulate the intrinsic activity of each area (Fig. 3c). Another way to understand the effect of changing the modulus of the eigenvalue of the AR(2) is that it controls the bandwidth (or peakiness) of the oscillations as seen in the power spectral density, and with that the auto-correlation structure. The modulus of the eigenvalues of the AR(2) models were always equal for both areas. The performance of TCC was evaluated by computing the proportion of unidirectional coherence (PUC), which was defined as the proportion of unidirectional coherence explained by TCC relative to the expected unidirectional coherence given the analytical model (see Section 2.6). As anticipated, we found that the effectiveness of the method decreased as the auto-correlations became stronger (i.e., eigenvalue moduli closer to 1.0). However, the TCC method performed relatively well for signals with eigenvalues up to ≈ 0.8 (Fig. 3c). It is worth noting that the performance of the TCC method improved slightly for eigenvalues close to 1.0. The likely explanation for this effect is the frequency difference between each area. That is, as the transmitted signals become more auto-correlated, their bandwidths become more narrow and the PSDs overlap less.

Although the TCC method was able to estimate the true unidirectional coherence reasonably well over a range of signals, the method quickly failed for signals with strong auto-correlation. Thus, the TCC method may be sufficient for certain signals, but does not generally solve the problem of destructive interference.

3.4. Fitting the power spectrum to recover the unidirectional coherence

Because the TCC method failed to recover the true directed coherence for strongly auto-correlated signals, this method may be of little utility for those interested in quantifying oscillatory inter-areal interactions. Thus, we were motivated to develop a method more suited for quantifying inter-areal interactions given strongly auto-correlated (oscillatory) signals.

The TCC method failed to isolate the directed influences, because the strong auto-correlations of the transmitted signals causes more overlap in the cross-covariance function, which leads to destructive interference. This implies that it may be infeasible to decompose the observed cross-covariance into the individual directional cross-covariance functions when the transmitted signals are strongly auto-correlated. However, we reasoned that it may be possible to infer the unidirectional cross-covariance functions by finding a re-composition that best explains the observed cross-covariance function.

In order to reconstruct the observed cross-covariance it is necessary to estimate the auto-covariance functions of the transmitted signals. In the simplest case, that is when there are no additive background fluctuations in either area, the observed power spectra accurately reflect the power spectra of the transmitted signals themselves. In this case, estimating the auto-covariance function of the transmitted signals is straightforward. However, in practice it is likely that the data being analyzed (e.g., EEG and LFPs) contains auto-correlations that are not transmitted (e.g., additive $1/f$). Therefore, it is necessary to decompose the observed power spectra and isolate the component that reflects the transmitted signal.

To that end, we fit an additive model (sum of a spectral peak, AR(2) model, and background $1/f$) to the power spectrum of each area using a non-linear least squares fit (see Section 2.5). Each area's auto-covariance function could then be estimated from the parameters of the AR(2) coefficients. The individual auto-covariance functions were shifted in their respective directions according to biophysically reasonable τ (e.g., 0 to 10 ms), and the weighted sum was correlated with the observed cross-covariance to determine the best fit.

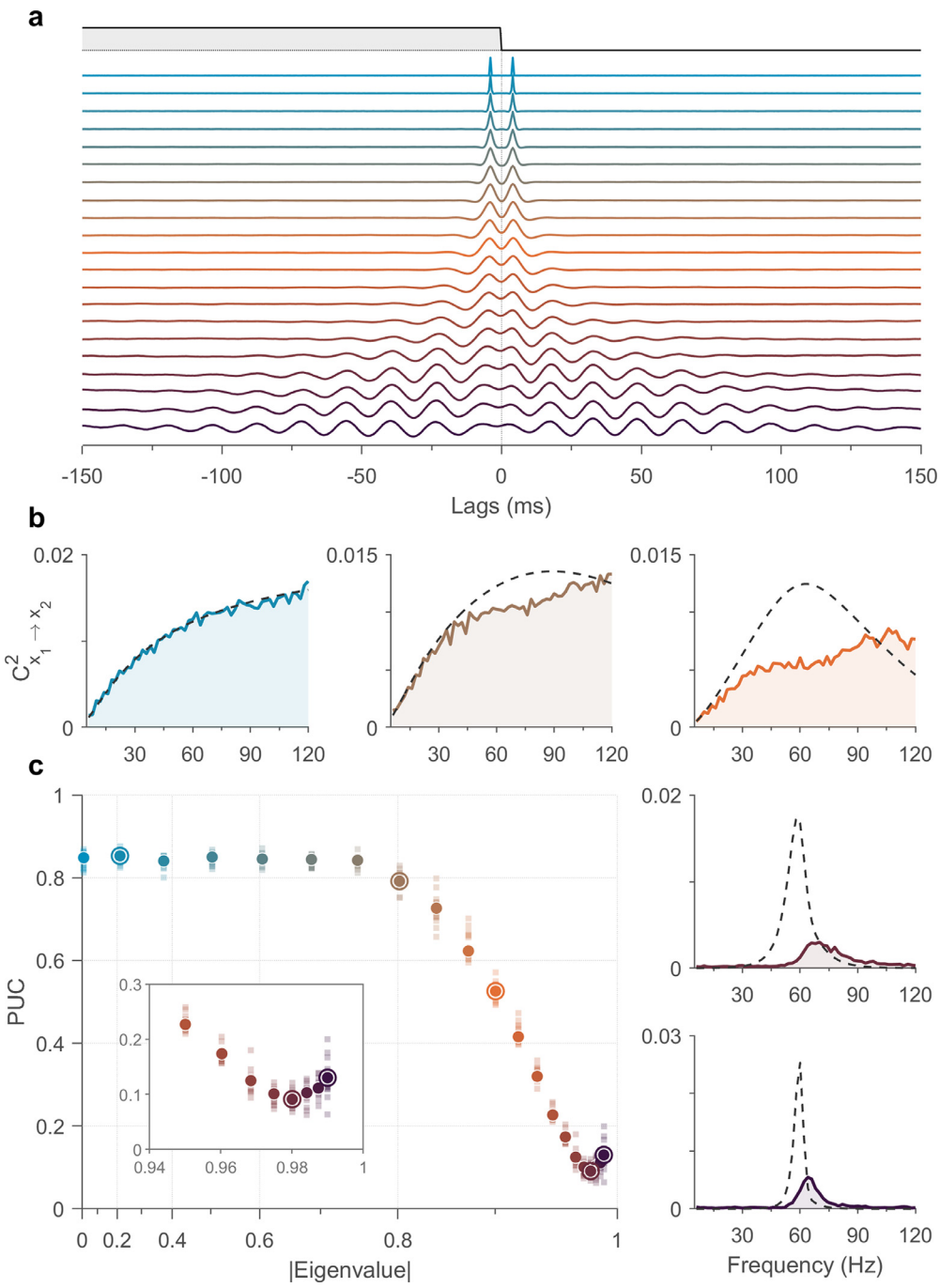


Fig. 3. Method 1: The truncated cross-covariance (TCC) approach. For these simulations, both areas were reciprocally connected with the same connectivity weight, $w = 0.15$ and $\tau = 4$ ms in both directions. Only the intrinsic gamma-band activity in each area (peak frequency at 60 hz for x_1 and 65 Hz for x_2) was projected. These simulations included uncorrelated additive $1/f$ background fluctuations. (a) Cross-covariance functions for different eigenvalue moduli. The Heaviside applied to the cross-covariance functions for these results is plotted at the top. (b) Five examples of the directional coherence $x_1 \to x_2$ compared to the expected derived analytically (dashed line). Colors correspond to results shown in panel c. (c) Proportion of unidirectional coherence (PUC) for the directional coherence $x_1 \to x_2$. Examples shown in panel b are circled.

We fit this model to the same data derived from the simulations used for the TCC method. We found that model fit approach performed much better than the TCC method across the entire range of signals tested (see Fig. 4c). Surprisingly, the model fit approach also performed better than the TCC method for weakly auto-correlated signals.

3.5. Bivariate vector autoregressive model simulations

We motivated using the source mixing model as the primary model for the simulations because the Source Mixing model has been shown to capture interactions between brain areas as measured in the LFP signal (Schneider et al., 2021). However, we do not exclude the possibility that the interactions between brain areas may be more accurately described by a different model, e.g., vector autoregressive model (VAR) (Ding et al., 2006). In the VAR model, the signal that is sent from one area passes through the transfer function of the receiver, and at least to some extent is projected back to its original source. Therefore, in the VAR model it possible that there can be some amount of self-interference, because of this recursion. For the purpose of demonstrating the interference effect, the source mixing model has the advantage that we can exclude the possibility of any interference at the level of the signals/system.

However, we predicted that the problem of interference would still occur in VAR model. To that end, we constructed an otherwise similar set of simulations as those show in Fig. 1, i.e., with the source mixing model, but now using the VAR(2) model. We found that Granger-causality, phase locking value and coherence behaved similarly in the VAR model as in the Source Mixing model (Fig. 5). As in the Source Mixing model, coherence and PLV were more strongly affected by interference than Granger-causality, yet magnitude of Granger-causality

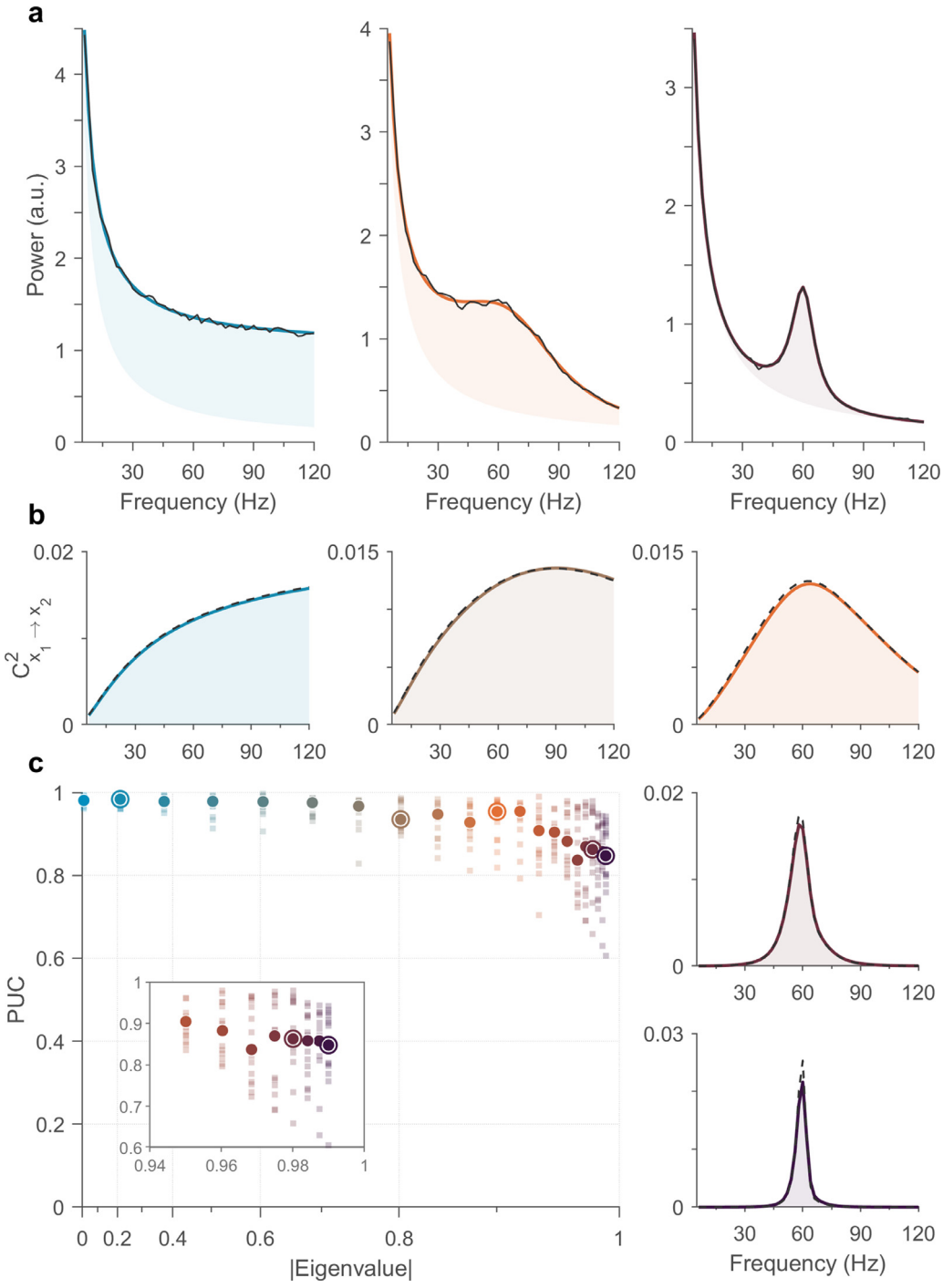


Fig. 4. Method 2: The model fit approach. The simulation data used for these analyses is the same data that was analyzed using the TCC method (see Fig. 3). (a) Example model fits to the observed power spectra for x_1. The modelodel fit is colored according to the eigenvalue moduli shown in panel c. (b) Five examples of the directional coherence $x_1 \to x_2$ compared to the expected derived analytically (dashed line). (c) Proportion of unidirectional coherence (PUC) for the directional coherence $x_1 \to x_2$. Examples shown in panel b are circled.

still showed a dependence on transmission delays. Similarly, we found that Granger-causality was able to recover the true interactions in the absence of $1/f$ background fluctuations (see Fig. S2e-f). This is expected as the extent to which a signal predicts the future values of another signal does not depend on the delay between the signals. However, in either case the interference problem remained for coherence, which in this case, reflects interference at the level of computing coherence.

3.6. Interference and other connectivity measures

For systems with bidirectional interactions, we have shown that coherence and Granger-Geweke causality (in the presence of additive noise) depend on the total transmission delay. However, the question is whether other commonly used connectivity measures also depend on transmission delays. In fact, any measure that is derived from cross-correlogram will depend on transmission delays.

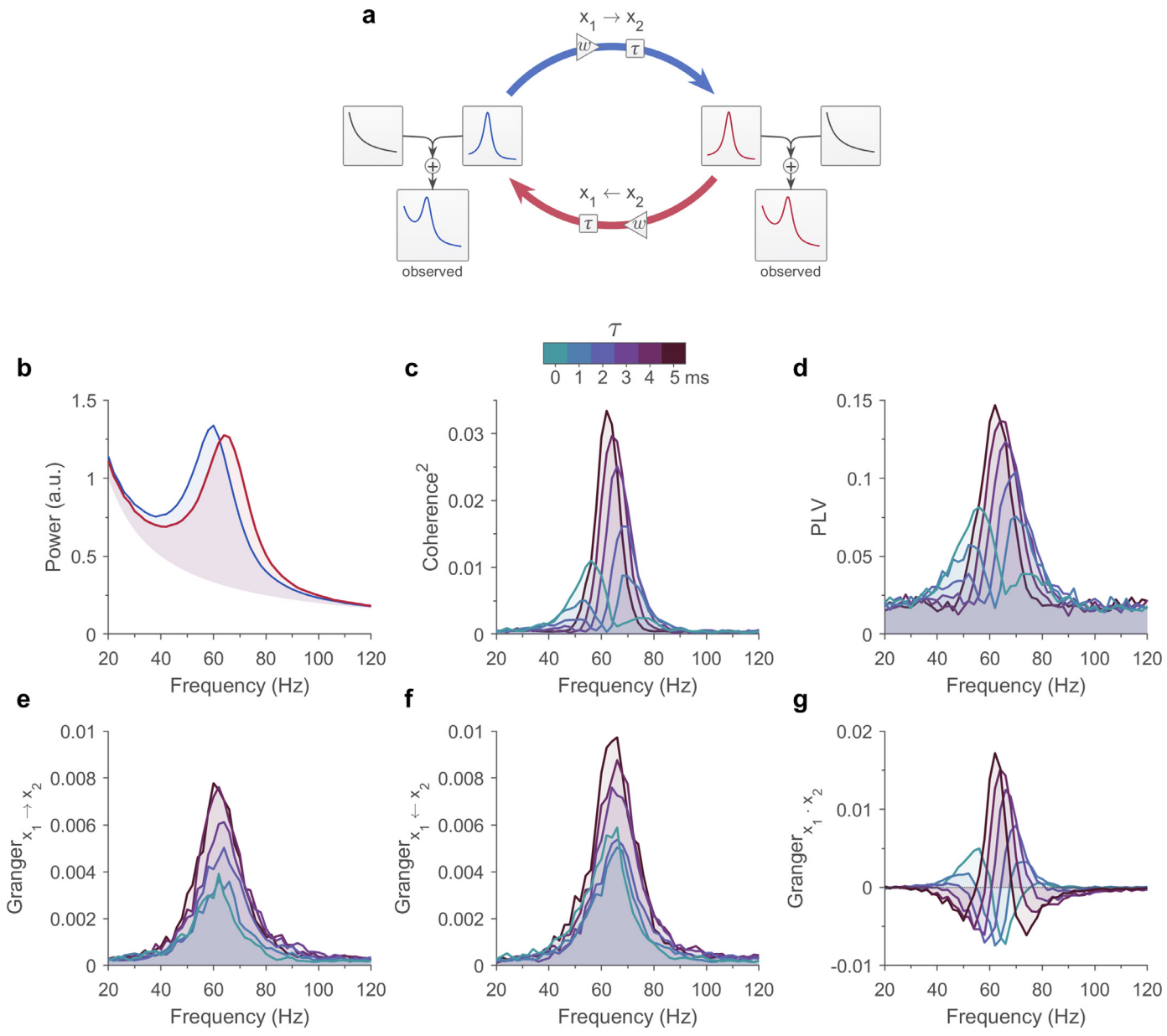


Fig. 5. Comparison between two reciprocally connected areas with different transmission delays (τ) using a VAR model. (a) Simulation circuit diagram for the bivariate VAR model. For these simulations, both areas were reciprocally connected with the same connectivity weight was $w = 0.0015$. The peak frequency intrinsic to each area was 60 Hz for x_1 and 65 Hz for x_2. These simulations included uncorrelated additive $1/f$ background fluctuations. (b) The observed power spectra for x_1 and x_2 (blue and red lines, respectively) contained both the intrinsic gamma-band activity (including the inputs form the other area) and additive $1/f$ background fluctuations. The power and slope of the $1/f$ background fluctuations were identical for both areas. Note, both the intrinsic gamma-band activity and $1/f$ background fluctuations were uncorrelated between each area. (c) The coherence spectra between x_1 and x_2 for τ over the range of 0 to 5 ms. Note, there is already interference at $\tau = 0$ ms. This occurs because each area filters its inputs and thus shifts the phase accordingly. Therefore, in addition to the transmission delay there is an additional filter delay. (d) Phase locking value (PLV) between x_1 and x_2 for τ between 0 and 5 ms. Note, an interference pattern is similar to coherence. (e) The non-parametric Granger-Geweke causality (GGC) spectra for x_1 to x_2. (f) The non-parametric GGC spectra for x_2 to x_1. (g) The instantaneous GGC spectra between x_1 and x_2. Notice that the instantaneous GGC is negative for some frequencies. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Phase locking value (PLV) is typically defined as

$$PLV(f) := \left|\mathbb{E}\left[e^{i\Delta\phi_f}\right]\right|, \tag{39}$$

where $\Delta\phi$ is the phase difference between observed signals in each area (Lachaux et al., 1999).

However, PLV can also be written as the expected per trial, or per epoch, normalized cross-spectral density (CSD)

$$PLV(f) := \left|\mathbb{E}\left[\frac{\hat{S}_{12}(f)}{\sqrt{\hat{S}_{11}(f)\,\hat{S}_{22}(f)}}\right]\right|, \tag{40}$$

where $\widehat{S}_{12}(f)$ is the cross-spectral density estimate in a given trial, and $\widehat{S}_{11}(f)$ and $\widehat{S}_{22}(f)$ are the respective estimates of the power spectral densities for that trial.

Thus, both PLV and coherence depend on the transmission delays given bidirectional interactions, because both measures are derived from the cross-covariance function.

We confirmed that PLV does indeed depend on transmission delays as does coherence through numerical simulations. PLV depended on transmission delays similar to coherence for both the Source Mixing model (Fig. 1d) and bivariate autoregressive model with additive $1/f$ (Fig. 5d), as well as without $1/f$ (see Figs. S1 and S2, respectively). Furthermore, the relationship between PLV and coherence from the numerical simulations with varying transmission delays followed the relationship given by their analytical relationship (see Fig. S3) (Aydore et al., 2013).

4. Discussion

An active area of study in systems neuroscience involves characterizing and understanding the temporal and spectral properties of the signals communicated between brain areas. Coherence, and more recently Granger-causality, have become standard methods that are routine in analyzing LFPs, EEG/MEG, spike-field and spiking data. These methods are so influential that they underlie prominent theories in systems neuroscience on attention (Bosman et al., 2012; Ferro et al., 2021; Grothe et al., 2012), inter-areal communication (Bastos et al., 2015a; 2015b; Buschman and Miller, 2007; Fries, 2005; 2015; Michalareas et al., 2016), inter-areal networks (Vezoli et al., 2021) and predictive coding (Bastos et al., 2020; Chao et al., 2018). The crucial assumption for these theories is that differences in coherence and Granger-causality directly correspond to differences in inter-areal communication, stimulus processing, or prediction/prediction error. However, we show that differences in coherence and Granger-causality can occur simply because of the transmission delays between two areas. The implication of these results is that previously reported differences in coherence are not unequivocal evidence of differences in information flow, inter-areal communication, stimulus processing, or prediction/prediction error.

4.1. Understanding the cause of the interference

We have shown that the magnitude and spectral characteristics of coherence are dependent on the transmission delays between two areas with bidirectional communication. Furthermore, this dependence on transmission delay is not a property of the inter-areal communication itself, but rather a consequence of how coherence is computed. Therefore, the modulation of coherence due to transmission delays shown here is an artifact of the method, and does not reflect a true modulation of the strength of inter-areal communication.

In Fig. 2, we illustrated that the source of the interference can appear in two forms. When the transmitted signals are weakly auto-correlated, the interference is predominantly due to the sine and cosine basis functions underlying the Fourier transform. Conversely, when the transmitted signals are strongly auto-correlated, there is an additional source of interference that arises at the level of the cross-covariance function itself.

It is worth recognizing that this interference effect extends to computing spike-field and spike-spike correlations. The reality is, this interference effect will occur for any signal/system where there are two peaks in the cross-covariance function that are separated by some non-zero lag. In fact, this interference effect will also occur for unidirectional systems when both areas also receive common input. Note that it is sufficient for this common input to be the result of volume conduction (Vinck et al., 2015). Furthermore, the interference pattern will change if there is common input in a bidirectional system. In such a situation there will be three peaks in the cross-covariance function, which will lead to a more complex pattern of interference dependent on the relative delays between each pair of the three peaks.

An intuition behind the source of this interference can be gained by recognizing that the superposition of two anti-phase cosine waves will result in destructive interference. The cosine waves themselves can be a component of the transmitted signals, or underlie the method applied to the signals (e.g., the Fourier transform) (see Fig. 2). The effect of transmission delays on coherence is that it shifts the cosine waves into an anti-phase relationship, which then manifests as an interference pattern across frequencies. This interference pattern appears as a cosine modulation across frequencies, because the CSD is modulated by the real part of a complex term that reflects the phase difference between the transmitted signals as a function of the total transmission delay.

Although the Granger-causality spectra did not show the same interference pattern that was observed for coherence, Granger-causality is still dependent on transmission delays in the presence of additive $1/f$ background fluctuations. We believe this problem is the result of the spectral factorization method in estimating the minimum phase transfer functions given additive noise. It is worth noting, that although it has been previously reported that instantaneous Granger-causality spectra can be negative for some frequencies, to our knowledge the circumstances under which this occurs has not been previously described (Ding et al., 2006). Interestingly, although the directional Granger-causality spectra did not show interference in the absence of additive noise, the instantaneous Granger-causality spectra did (see Figs. S1 and S2). Note that the total interdependence (F_{x_1,x_2}) between two time series, x_1 and x_2, is the sum of the directional and instantaneous influences, i.e., $F_{x_1,x_2} = F_{x_1 \to x_2} + F_{x_1 \leftarrow x_2} + F_{x_1 \cdot x_2}$, (Ding et al., 2006). Thus, the interference pattern is not evident in the directional Granger-causality spectra because it is counteracted by the negative instantaneous Granger-causality values. Therefore, interference due to transmission delays is one case under which instantaneous Granger-causality can be negative, and further exemplifies the difficulty in interpreting instantaneous Granger-causality.

Nonetheless, it is the case that coherence depended on transmission delays in all our simulations (i.e., for both the Source Mixing and VAR models, with and without $1/f$). Therefore, coherence as method for analyzing and characterizing the inter-areal interactions in bidirectional systems is severely limited (i.e., to cases where the power spectra of the transmitted are completely non-overlapping).

4.2. Implications for previous studies

Previous studies have interpreted differences in coherence and Granger-causality as reflecting differences in the strength or spectral characteristics of inter-areal communication, network interactions, or predictive processing (Bastos et al., 2015a; 2020; Bosman et al., 2012; Buschman and Miller, 2007; Ferro et al., 2021; Grothe et al., 2012; Michalareas et al., 2016; Vezoli et al., 2021). We show however, that these differences may in fact be explained by differences in transmission delays between bidirectionally coupled populations of neurons. The interference pattern that arises as a result of the transmission delays is in fact an artifact of the coherence method.

Several previous studies have shown that the strength of Granger-causality and coherence correlate with measures of anatomical connectivity (Bastos et al., 2015a; Vezoli et al., 2021). For example, Vezoli et al. (2021) showed a correlation between the strength of anatomical connectivity (i.e., the relative number of neurons projecting from area A to B, and from B to A) with the coherence. They found that the anatomical connectivity strength could be positively predicted from the coherence in different bands. The interpretation of these findings is that areas that are more strongly connected also exhibit stronger functional interactions. However, it should be noted that conduction delays and the anatomical connectivity strength have the opposite relationship with white-matter distance. That is, areas that are physically closer tend to be more strongly connected, as measured by the proportion of extrinsic projections between the areas (often referred to as FLN) (Markov et al., 2013; 2014a; 2011), and their white-matter distance

is also shorter. Consequently, shorter white-matter distance implies shorter transmission delays, and thus less interference in the coherence spectra. Furthermore, given a fixed transmission delay across all frequencies, lower frequencies should be less affected by the interference problem than higher frequencies. And indeed, Vezoli et al. (2021) reported that higher frequencies were more predictive of the anatomical connectivity. Therefore, it is possible that the correlation between coherence and anatomical connectivity strength is at least partially explained by the interference effect we have described here.

Several studies have observed changes in coherence with cognitive factors such as attention (Bosman et al., 2012; Ferro et al., 2021; Grothe et al., 2012). These results have been interpreted as empirical evidence for the hypothesis that coherence facilitates inter-areal communication (Bosman et al., 2012; Fries, 2005; 2015). In addition to an increase in inter-areal coherence (between areas V1 and V4), Bosman et al. (2012) reported a shift in the peak frequency of the V1 gamma-oscillations with attention (see also Ferro et al., 2021). Taking into consideration the interference pattern across frequencies in coherence spectra, which depends on the transmission delay, it is clear how a shift in the peak frequency alone could result in a change the magnitude of coherence. That is, two V1 populations sending the same signal to V4, although with a different frequency, can show different coherence spectra with V4 simply because their peak frequencies land at different phases of the interference pattern across frequencies.

It is clear that the interference effect can lead to the appearance of differences in coherence that do not correspond to differences in the magnitude nor strength of the transmitted signals. However, there may be true differences in the strength of feedforward (FF) and feedback (FB) signals as a result of differences in anatomical connectivity (Vezoli et al., 2021), or cognitive factors such as attention and prediction (Bastos et al., 2020; Bosman et al., 2012; Ferro et al., 2021; Grothe et al., 2012; Rao and Ballard, 1999). Nonetheless, the interference effect presents a problem for interpreting any coherence differences that may result. The problem becomes apparent when we consider two areas that transmit the same signal to a third area, but receive differential feedback. If it is the case that one area receives feedback while the other does not, then this means in one area interference can arise but in the other not. Paradoxically, this leads to the possibility that an increase in the strength of FF or FB communication may result in reduced inter-areal coherence - despite an increase in the strength of the communication. Thus, the difficulty in interpreting true differences in FF or FB communication is that the interference also depends on the relative strengths of FF and FB signals.

It is evident that coherence depends on factors such as transmission delay and peak frequency, and as a result the true meaningful differences in FF or FB communication are rendered uninterpretable by the interference that results from computing coherence. Therefore, strictly interpreting coherence differences as the result of differences in the strength of communication is problematic. This means that in most circumstances, coherence is an ill-suited measure for analyzing and characterizing interactions in bidirectional systems.

4.3. Practical recommendations and future work

We have developed two methods that can recover the directed coherence independent of transmission delays. These methods may be useful, but are also didactic as they illustrate that the interference that arises is the result of the method (namely computing coherence), and that it is possible to recover the true directional influences. However, it is important to recognize that these methods have limitations. For instance, the truncated cross-covariance method (method 1) is highly dependent on the auto-correlation structure of the transmitted signals. That is, strongly auto-correlated signals are difficult to separate because their auto-correlation functions overlap in the cross-covariance. Thus, method 1 is most effective when the auto-correlation of the transmitted signals is weak, and the individual auto-covariance functions can be well separated with appropriate windowing of the cross-covariance function. The model fit method (method 2) is less dependent on the auto-correlation of the transmitted signals than method 1, however, method 2 requires model specification and is much more computationally expensive.

Between the two methods, method 2 performed much better across the range of power spectra in our simulations. That said, our simulated power spectra were composed of a single oscillatory peak with additive $1/f$, whereas empirical power spectra typically contain multiple oscillatory peaks including $1/f$. One approach here would be to utilize existing toolboxes developed for fitting power spectra, such as FOOOF (Donoghue et al., 2020). Regardless of the algorithm used for fitting the power spectra, it is important to specify the appropriate model that accurately reflects the underlying statistics and features of the data being analyzed. We chose to model the LFP oscillations with pseudo-periodic AR(2) models, which have been shown to reproduce the statistical properties of stationary gamma oscillations in primary visual cortex, and provide mean-field approximations of E-I circuits driven by stochastic input (Spyropoulos et al., 2022). Therefore, we recommend fitting the power spectral peaks with autoregressive models, rather than arbitrary functions such as a Gaussian.

Unlike coherence, in the absence of additive $1/f$, Granger-causality can accurately estimate the bidirectional interactions, and is therefore insensitive to transmission delays (see Figs. S1 and S2). However, in the presence of additive $1/f$, which is expected in empirical data, Granger-causality is sensitive to transmission delays (see Figs. 1e-g and 5e-g). Therefore, in practice it should be assumed that Granger-causality will also depend on transmission delays. Importantly, the cause of this dependence appears to be more related to estimating the transfer function in the presence of noise. The problem of estimating Granger-causality in the presence of additive noise has been previously discussed (Vinck et al., 2015).

It is important to note that the interference pattern we have described here depends on the presence of bidirectional interactions, and overlapping power spectra of the transmitted signals (which may or may not be consistent with the peaks in the observed power spectra). This means of course, that in systems where the interactions are always unidirectional, or in systems with non-overlapping power spectra, coherence may be a useful measure of their interaction.

However, in systems with overlapping power spectra and bidirectional interactions, transmission delays can influence the shape and magnitude of the coherence spectra. Thus, caution should be taken when interpreting coherence spectra unless it is known that the interaction is entirely unidirectional, or it can be reasonably assumed that the power spectra of transmitted signals do not overlap.

We have shown that the interference pattern appears as a cosine modulation across frequencies, which depends on the total transmission delay (see Eq. (15)). The maximal interference for a given frequency occurs when the transmission delay shifts the directed cross-correlations such they are in anti-phase (i.e., they are shifted by half of their cycle length relative to each other). For example, for two areas interacting at alpha (≈10 Hz), the transmission delay in each direction would need to be $\tau = 25$ ms to cause maximal interference. Therefore, it is reasonable to assume then that low frequencies in the coherence spectra would be less sensitive to interference.

However, it is important to note that conduction delays have been shown to range from 2–4 ms for cortico-cortico connections (within a hemisphere) (Ferraina et al., 2002), to 2–18 ms for inter-hemispheric connections (Miller, 1975; Swadlow et al., 1978). Thus, whether or not there is a possibility for interference at a given frequency is highly dependent on the areas being studied.

To that end, we have presented two possible solutions to the problem of estimating bidirectional interactions when there are transmission delays. Although these methods have limitations, they may prove useful for recovering the true bidirectional interactions in many cases. However, as this work shows, there is still a need for further development of methods suitable for analyzing and characterizing interac-

tions in bidirectional systems with transmission delays and additive noise.

Code availability statement

All code was written in Matlab and run with Matlab 2021a, The MathWorks, Inc., Natick, Massachusetts, United States.

The code for the numerical simulations and methods was written by the corresponding author (J.R.D) and is available upon request.

External toolboxes

For the Granger-Geweke causality (GGC) analyses, spectral factorization was computed using the Wilson-Burg algorithm written in Matlab by Dhamala et al. (2018), available here: https://scholarworks.gsu.edu/phy_astr_facupub/13/. Pairwise Granger-Geweke causality was computed using the modified code from the FieldTrip toolbox (Oostenveld et al., 2011), available here: https://github.com/fieldtrip/fieldtrip/blob/master/connectivity/ft_connectivity_granger.m. The Gaussian hypergeometric functions were computed in Matlab using the external function hypergeometric2f1.m available here https://www.mathworks.com/matlabcentral/fileexchange/1844-gaussian-hypergeometric-function. Figures were generated in Matlab and exported using a combination of custom code, and the following external toolboxes: arrow.m available here https://www.mathworks.com/matlabcentral/fileexchange/278-arrow, and export_fig available here https://github.com/altmany/export_fig. Colormaps were created using David Johnstone's colour gradient and cubehelix picker, available here https://davidjohnstone.net/lch-lab-colour-gradient-picker, and https://davidjohnstone.net/cubehelix-gradient-picker

Data availability

No data was used for the research described in the article.

Credit authorship contribution statement

Jarrod Robert Dowdall: Conceptualization, Investigation, Formal analysis, Methodology, Validation, Software, Visualization, Writing – original draft, Writing – review & editing. **Martin Vinck:** Methodology, Validation, Writing – review & editing, Supervision.

Acknowledgements

This project was supported by an ERC Starting Grant (SPATEMP, EU), a BMBF (Germany) Grant to M.V. (Computational Life Sciences, project BINDA, 031L0167).

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.neuroimage.2023.119998.

References

Aydore, S., Pantazis, D., Leahy, R.M., 2013. A note on the phase locking value and its properties. Neuroimage 74, 231–244. doi:10.1016/j.neuroimage.2013.02.008.

Bastos, A., Vezoli, J., Bosman, C., Schoffelen, J.M., Oostenveld, R., Dowdall, J., DeWeerd, P., Kennedy, H., Fries, P., 2015. Visual areas exert feedforward and feedback influences through distinct frequency channels. Neuron 85. doi:10.1016/j.neuron.2014.12.018.

Bastos, A.M., Lundqvist, M., Waite, A.S., Kopell, N., Miller, E.K., 2020. Layer and rhythm specificity for predictive routing. Proc. Natl. Acad. Sci. 117, 31459–31469. doi:10.1073/pnas.2014868117.

Bastos, A.M., Vezoli, J., Fries, P., 2015. Communication through coherence with inter-areal delays. Curr. Opin. Neurobiol. 31, 173–180. doi:10.1016/j.conb.2014.11.001.

Bosman, C.A., Schoffelen, J.M., Brunet, N., Oostenveld, R., Bastos, A.M., Womelsdorf, T., Rubehn, B., Stieglitz, T., Weerd, P.D., Fries, P., 2012. Attentional stimulus selection through selective synchronization between monkey visual areas. Neuron 75, 875–888. doi:10.1016/j.neuron.2012.06.037.

Bucher, D., Goaillard, J.M., 2011. Beyond faithful conduction: short-term dynamics, neuromodulation, and long-term regulation of spike propagation in the axon. Prog. Neurobiol. 94, 307–346. doi:10.1016/j.pneurobio.2011.06.001.

Buschman, T.J., Miller, E.K., 2007. Top-down versus bottom-up control of attention in the prefrontal and posterior parietal cortices. Science 315, 1860–1862. doi:10.1126/science.1138071.

Chao, Z.C., Takaura, K., Wang, L., Fujii, N., Dehaene, S., 2018. Large-scale cortical networks for hierarchical prediction and prediction error in the primate brain. Neuron 100, 1252–1266.e3. doi:10.1016/j.neuron.2018.10.004.

Chaudhuri, R., Knoblauch, K., Gariel, M.A., Kennedy, H., Wang, X.J., 2015. A large-scale circuit mechanism for hierarchical dynamical processing in the primate cortex. Neuron 88, 419–431. doi:10.1016/j.neuron.2015.09.008.

Destexhe, A., Rudolph, M., Paré, D., 2003. The high-conductance state of neocortical neurons in vivo. Nat. Rev. Neurosci. 4, 739–751. doi:10.1038/nrn1198.

Dhamala, M., Liang, H., Bressler, S.L., Ding, M., 2018. Granger-geweke causality: estimation and interpretation. Neuroimage 175, 460–463. doi:10.1016/j.neuroimage.2018.04.043.

Ding, M., Chen, Y., Bressler, S.L., 2006. Granger Causality: Basic Theory and Application to Neuroscience. Wiley-VCH Verlag GmbH & Co. KGaA, pp. 437–460.

Donoghue, T., Haller, M., Peterson, E.J., Varma, P., Sebastian, P., Gao, R., Noto, T., Lara, A.H., Wallis, J.D., Knight, R.T., Shestyuk, A., Voytek, B., 2020. Parameterizing neural power spectra into periodic and aperiodic components. Nat. Neurosci. 23, 1655–1665. doi:10.1038/s41593-020-00744-x.

Ferraina, S., Paré, M., Wurtz, R.H., 2002. Comparison of cortico-cortical and cortico-collicular signals for the generation of saccadic eye movements. J. Neurophysiol. 87, 845–858. doi:10.1152/jn.00317.2001.

Ferro, D., van Kempen, J., Boyd, M., Panzeri, S., Thiele, A., 2021. Directed information exchange between cortical layers in macaque v1 and v4 and its modulation by selective attention. Proc. Natl. Acad. Sci. 118, e2022097118. doi:10.1073/pnas.2022097118.

Fries, P., 2005. A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. Trends Cogn. Sci. (Regul. Ed.) 9, 474–480. doi:10.1016/j.tics.2005.08.011.

Fries, P., 2015. Rhythms for cognition: communication through coherence. Neuron 88, 220–235. doi:10.1016/j.neuron.2015.09.034.

Grothe, I., Neitzel, S.D., Mandon, S., Kreiter, A.K., 2012. Switching neuronal inputs by differential modulations of gamma-band phase-coherence. J. Neurosci. 32, 16172–16180. doi:10.1523/JNEUROSCI.0890-12.2012.

Lachaux, J.P., Rodriguez, E., Martinerie, J., Varela, F.J., 1999. Measuring phase synchrony in brain signals. Hum. Brain Mapp. 8, 194–208. doi:10.1002/(SICI)1097-0193(1999)8:4<194::AID-HBM4>3.0.CO;2-C.

Markov, N.T., Ercsey-Ravasz, M., Lamy, C., Gomes, A.R.R., Magrou, L., Misery, P., Giroud, P., Barone, P., Dehay, C., Toroczkai, Z., Knoblauch, K., Essen, D.C.V., Kennedy, H., 2013. The role of long-range connections on the specificity of the macaque interareal cortical network. Proc. Natl. Acad. Sci. 110, 5187–5192. doi:10.1073/pnas.1218972110.

Markov, N.T., Ercsey-Ravasz, M.M., Gomes, A.R.R., Lamy, C., Magrou, L., Vezoli, J., Misery, P., Falchier, A., Quilodran, R., Gariel, M.A., Sallet, J., Gamanut, R., Huissoud, C., Clavagnier, S., Giroud, P., Sappey-Marinier, D., Barone, P., Dehay, C., Toroczkai, Z., Knoblauch, K., Essen, D.C.V., Kennedy, H., 2014. A weighted and directed interareal connectivity matrix for macaque cerebral cortex. Cerebral Cortex 24, 17–36. doi:10.1093/cercor/bhs270.

Markov, N.T., Misery, P., Falchier, A., Lamy, C., Vezoli, J., Quilodran, R., Gariel, M.A., Giroud, P., Ercsey-Ravasz, M., Pilaz, L.J., Huissoud, C., Barone, P., Dehay, C., Toroczkai, Z., Essen, D.C.V., Kennedy, H., Knoblauch, K., 2011. Weight consistency specifies regularities of macaque cortical networks. Cerebral Cortex 21, 1254–1272. doi:10.1093/cercor/bhq201.

Markov, N.T., Vezoli, J., Chameau, P., Falchier, A., Quilodran, R., Huissoud, C., Lamy, C., Misery, P., Giroud, P., Ullman, S., Barone, P., Dehay, C., Knoblauch, K., Kennedy, H., 2014. Anatomy of hierarchy: feedforward and feedback pathways in macaque visual cortex. J. Comparat. Neurol. 522, 225–259. doi:10.1002/cne.23458.

Michalareas, G., Vezoli, J., van Pelt, S., Schoffelen, J.M., Kennedy, H., Fries, P., 2016. Alpha-beta and gamma rhythms subserve feedback and feedforward influences among human visual cortical areas. Neuron 89, 384–397. doi:10.1016/j.neuron.2015.12.018.

Miller, R., 1975. Distribution and properties of commissural and other neurons in cat sensorimotor cortex. J. Comp. Neurol. 164, 361–373. doi:10.1002/cne.901640307.

Oostenveld, R., Fries, P., Maris, E., Schoffelen, J.M., 2011. Fieldtrip: open source software for advanced analysis of meg, eeg, and invasive electrophysiological data. Comput. Intell. Neurosci. 2011, 1–9. doi:10.1155/2011/156869.

Pesaran, B., Vinck, M., Einevoll, G.T., Sirota, A., Fries, P., Siegel, M., Truccolo, W., Schroeder, C.E., Srinivasan, R., 2018. Investigating large-scale brain dynamics using field potential recordings: analysis and interpretation. Nat. Neurosci. 21, 903–919. doi:10.1038/s41593-018-0171-8.

Rao, R.P.N., Ballard, D.H., 1999. Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. Nat. Neurosci. 2, 79–87. doi:10.1038/4580.

Schneider, M., Broggini, A.C., Dann, B., Tzanou, A., Uran, C., Sheshadri, S., Scherberger, H., Vinck, M., 2021. A mechanism for inter-areal coherence through communication based on connectivity and oscillatory power. Neuron doi:10.1016/j.neuron.2021.09.037.

Spyropoulos, G., Saponati, M., Dowdall, J.R., Schölvinck, M.L., Bosman, C.A., Lima, B., Peter, A., Onorato, I., Klon-Lipok, J., Roese, R., Neuenschwander, S., Fries, P., Vinck, M., 2022. Spontaneous variability in gamma dynamics described by a damped harmonic oscillator driven by noise. Nat. Commun. 13, 2019. doi:10.1038/s41467-022-29674-x.

Swadlow, H., Rosene, D., Waxman, S., 1978. Characteristics of interhemispheric impulse conduction between prelunate gyri of the rhesus monkey. Exp. Brain Res. 33. doi:10.1007/BF00235567.

Tsodyks, M., Markram, H., 1997. The neural code between neocortical pyramidal neurons depends on neurotransmitter release-probability. Proc. Natl. Acad. Sci. 94, 719–723. doi:10.1073/pnas.94.2.719.

Vezoli, J., Vinck, M., Bosman, C.A., Bastos, A.M., Lewis, C.M., Kennedy, H., Fries, P., 2021. Brain rhythms define distinct interaction networks with differential dependence on anatomy. Neuron 109, 3862–3878.e5. doi:10.1016/j.neuron.2021.09.052.

Vinck, M., Huurdeman, L., Bosman, C.A., Fries, P., Battaglia, F.P., Pennartz, C.M., Tiesinga, P.H., 2015. How to detect the granger-causal flow direction in the presence of additive noise? Neuroimage 108, 301–318. doi:10.1016/j.neuroimage.2014.12.017.