



Benchmarking nonparametric Granger causality: Robustness against downsampling and influence of spectral decomposition parameters



Mattia F. Pagnotta^{a,*}, Mukesh Dhamala^{b,c}, Gijs Plomp^a

^a *Perceptual Networks Group, Department of Psychology, University of Fribourg, Fribourg, CH-1701, Switzerland*

^b *Department of Physics and Astronomy, Georgia State University, Atlanta, GA 30303, USA*

^c *Neuroscience Institute, Georgia State University, Atlanta, GA 30303, USA*

ARTICLE INFO

Keywords:

Barrel cortex
Brain connectivity
Conditional Granger causality
EEG
Nonparametric Granger causality
Multitaper method
Spectral factorization
Wavelet transform

ABSTRACT

Brain function arises from networks of distributed brain areas whose directed interactions vary at subsecond time scales. To investigate such interactions, functional directed connectivity methods based on nonparametric spectral factorization are promising tools, because they can be straightforwardly extended to the nonstationary case using wavelet transforms or multitapers on sliding time window, and allow estimating time-varying spectral measures of Granger–Geweke causality (GGC) from multivariate data. Here we systematically assess the performance of various nonparametric GGC methods in real EEG data recorded over rat cortex during unilateral whisker stimulations, where somatosensory evoked potentials (SEPs) propagate over known areas at known latencies and therefore allow defining fixed criteria to measure the performance of time-varying directed connectivity measures. In doing so, we provide a comprehensive benchmark evaluation of the spectral decomposition parameters that might influence the performance of wavelet and multitaper approaches. Our results show that, under the majority of parameter settings, nonparametric methods can correctly identify the contralateral primary sensory cortex (cS1) as the principal driver of the cortical network. Furthermore, we observe that, when properly optimized, the approach based on Morlet wavelet provided the best detection of the preferential functional targets of cS1; while, the best temporal characterization of whisker-evoked interactions was obtained with a sliding-window multitaper. In addition, we find that nonparametric methods provide GGC estimates that are robust against signal downsampling. Taken together our results provide a range of plausible application values for the spectral decomposition parameters of nonparametric methods, and show that they are well suited to characterize time-varying directed causal influences between neural systems with good temporal resolution.

1. Introduction

Human brain function emerges at large-scale from networks of dynamic interactions between distributed cortical areas (Bressler, 1995; Horowitz, 2003; Sporns, 2014). These interactions are inherently directed and reflect how neuronal activity in one area changes the excitability in the areas that it targets through synaptic connections (Felleman and Van Essen, 1991; Markov et al., 2014). The strength of these interactions, furthermore, varies on subsecond time scales both in stimulus-evoked data and in recordings obtained at rest, i.e. without stimulation or explicit task instructions (Britz et al., 2010; Bullier, 2001). In order to understand how these large-scale networks support sensory and cognitive processes it is essential to accurately characterize the underlying pattern of directed interactions and how this evolves in time.

The interactions between brain areas can be modeled from statistical regularities between signals simultaneously recorded from those areas, by using measures of directed connectivity (Baccalà and Sameshima, 2001; Geweke, 1982; Granger, 1969; Kaminski and Blinowska, 1991). We can distinguish at least three important goals of such directed connectivity analyses. The first goal is to accurately identify the most influential areas within the network; this can help identify functionally relevant areas in normal processing (Astolfi et al., 2007) but also the spread of pathological activity, for example in epileptic patients (Ding et al., 2007). A second goal of directed connectivity analyses is to accurately reflect which areas are preferentially targeted by any given brain area; this is a prerequisite for correctly inferring network properties through topological analyses of the resulting weighted directed graphs (Bassett and Sporns, 2017; Fallani et al., 2007). A third important goal of

* Corresponding author. Department of Psychology, University of Fribourg, Rue P.-A.-de-Faucigny 2, Fribourg, CH-1701, Switzerland.

E-mail address: mattia.pagnotta@unifr.ch (M.F. Pagnotta).

<https://doi.org/10.1016/j.neuroimage.2018.07.046>

Received 21 June 2018; Received in revised form 14 July 2018; Accepted 18 July 2018

Available online 20 July 2018

1053-8119/© 2018 Elsevier Inc. All rights reserved.

directed connectivity analyses is to show how interactions change over time, so as to accurately characterize transient interactions and network states (Baker et al., 2014; Deco et al., 2011; Vidaurre et al., 2017). These three goals are not exhaustive, but any directed connectivity measure that meets these goals would make for a powerful neuroimaging tool.

One class of directed connectivity measures that aims for these three goals is based on a nonparametric implementation of Granger causality (Dhamala et al., 2008a). Granger causality analyses make use of inferences based on the notion of temporal precedence and statistical predictability among simultaneously recorded neural time series. Nonparametric methods come from a long tradition that began with the notion of causality introduced by Wiener (1956), and continued with the statistical definitions proposed in time domain (Granger, 1969) and in frequency domain (Geweke, 1982). This spectral variant of causality is here referred to as Granger–Geweke Causality (GGC). Measures closely related to GGC have been defined on the basis of spectral quantities obtained from a multivariate autoregressive (MVAR) model of the time series, such as the spectral transfer matrix or the coefficients matrix (Baccalá and Sameshima, 2001; Kamiński et al., 2001).

Causality measures have been widely used to study directional influences in different neurophysiological data, such as electroencephalography (EEG) and magnetoencephalography (MEG) (Astolfi et al., 2007; Chand and Dhamala, 2016; Gow et al., 2008; Schoffelen and Gross, 2009), functional magnetic resonance imaging (fMRI) (Bajaj et al., 2015; Bressler et al., 2008; Roebroek et al., 2005; Wen et al., 2012), and local field potentials (LFPs) (Brovelli et al., 2004).

In their classical formulation, spectral GGC estimates are derived from MVAR models of the recorded time series (Geweke, 1982). This requires a priori choice of the model order, which is the parameter that determines how many past time samples are taken into account for predicting activities at present time. The MVAR-based methods, commonly referred to as parametric methods, have been successfully used to investigate directional influences between visual areas (Bernasconi et al., 2000; Bernasconi and König, 1999), sensorimotor processing (Brovelli et al., 2004; Zhang et al., 2008), cognitive tasks (Ding et al., 2006; Roebroek et al., 2005) and interictal epileptic activity (Coito et al., 2015; Lin et al., 2009).

Alternative methods allow deriving GGC estimates from a spectral factorization of the time series, and are referred to as nonparametric methods (Dhamala et al., 2008a, 2008b). Nonparametric GGC has been successfully used to investigate stationary causal influences in visual processing and selective attention (Bastos et al., 2015; Bosman et al., 2012; Roberts et al., 2013; Saalman et al., 2012), information processing in auditory cortex (Fontolan et al., 2014), and to evaluate directional influences between spike trains (Cao et al., 2012; Chen et al., 2014; Nedungadi et al., 2009).

Notably, nonparametric methods allow for a straightforward implementation of time-varying GGC, either by using wavelet transforms (Daubechies, 1990; Torrence and Compo, 1998) or by using the multitaper method (Thomson, 1982) on overlapping short time windows. While time-varying nonparametric approaches have been applied to neural data (Dhamala et al., 2008a), their use is less widespread and their validity is not fully established across various data types.

Another important aspect is that, even if nonparametric methods do not require model order selection, because the explicit MVAR-modeling is completely bypassed, they still need some initial choices of parameters, depending on the spectral decomposition approach employed. These parameters regulate the trade-offs between temporal and spectral resolutions, and between variance and bias of the spectral estimates, and therefore setting their values properly is critical to obtain interpretable causal measures. A systematic evaluation of the effects associated with varying spectral decomposition parameters in nonparametric GGC is currently missing.

Here, we aim to provide a systematic and objective assessment of the performance of time-varying nonparametric GGC methods in real data, by quantifying how well they fulfill the three main goals outlined

above, and to evaluate the effects of varying their spectral decomposition parameters. In order to do that, we used benchmark somatosensory evoked potentials (SEPs) obtained from Wistar rats by unilateral whisker stimulations (Plomp et al., 2014b; Quairiaux et al., 2011). Whisker-evoked SEPs show a highly dynamic activation pattern across rat's cortex that follows relatively well known structural pathways so that three objective performance criteria can be defined in terms of i) the ability to detect the major driver of evoked cortical activity as contralateral primary sensory cortex (cS1), ii) the ability to identify the main targets of cS1 in line with known synaptic connections, and iii) the ability to correctly represent the dynamics of the underlying process (Plomp et al., 2014b). We used these criteria to systematically assess the performance of nonparametric GGC, obtained using either wavelet transforms or multitaper method for spectral decomposition, under different analysis conditions.

Lowering sampling rate (downsampling) is a common step in EEG/MEG preprocessing; however, downsampling has been shown to potentially degrade Granger causality detectability depending on the interactions between sampling rate and generative time scales (Barnett and Seth, 2017). For this reason, we first investigated how time-varying nonparametric GGC performance depends on the sampling rate.

We then provided a benchmark evaluation of the spectral decomposition parameters of wavelet-based and multitaper-based nonparametric GGC. We critically assessed the effects on wavelet transforms performance of the wavelet parameters, which regulate time-frequency resolutions, and of window size and time-bandwidth parameter for the sliding-window multitaper (refer to sections 2.2.3–2.2.4).

Lastly, when investigating interactions between multiple brain areas we can either perform causality estimation by considering a bivariate model for each pair of variables independently, namely pairwise GGC, or consider a multivariate extension, also called conditional GGC (Chen et al., 2006; Dhamala et al., 2008a; Geweke, 1984). To explore possible confounds due to indirect paths of connections we compared performance of the spectral decomposition methods for both pairwise GGC and conditional GGC formulations.

2. Materials and methods

2.1. Benchmark EEG data

We considered previously recorded whisker-evoked SEPs obtained from ten p21 Wistar rats (Plomp et al., 2014b; Quairiaux et al., 2011). The SEPs were acquired with a grid of stainless steel electrodes, which were positioned in contact with the skull of the rats while they were under light isoflurane anesthesia. Fig. 1A provides a schematic representation of the recording setup. The signals were referenced to an electrode placed above the cerebellum, which is labeled as 'R' in the illustration (Fig. 1A). The sampling rate for signals acquisition was 2000 Hz and a bandpass filter (1–500 Hz) was applied online to the signals. The dataset has been made freely available.¹

Rejection of epochs contaminated with artifacts was performed using a semi-automatic approach, for further details refer to (Pagnotta and Plomp, 2018). Artifact-free epochs were then filtered using a zero-phase antialiasing filter with cutoff frequency of 125 Hz, which is the Nyquist frequency at the lowest sampling rate considered in the successive analyses ($F_s = 250$ Hz). We then computed grand-average whisker-evoked SEPs and grand-average amplitude spectra (Fig. 1B–C). The latter were derived using fast Fourier transform (FFT) algorithm on each single rat and then averaging amplitude spectra across rats. More in details, the single-rat spectrum of every node was computed with FFT algorithm at original sampling rate using the time interval [0, 200 ms] after stimulus onset for each trial; FFT estimates were then averaged across trials and normalized with respect to the maximum amplitude across nodes within

¹ <https://doi.org/10.6084/m9.figshare.5909122.v1>.

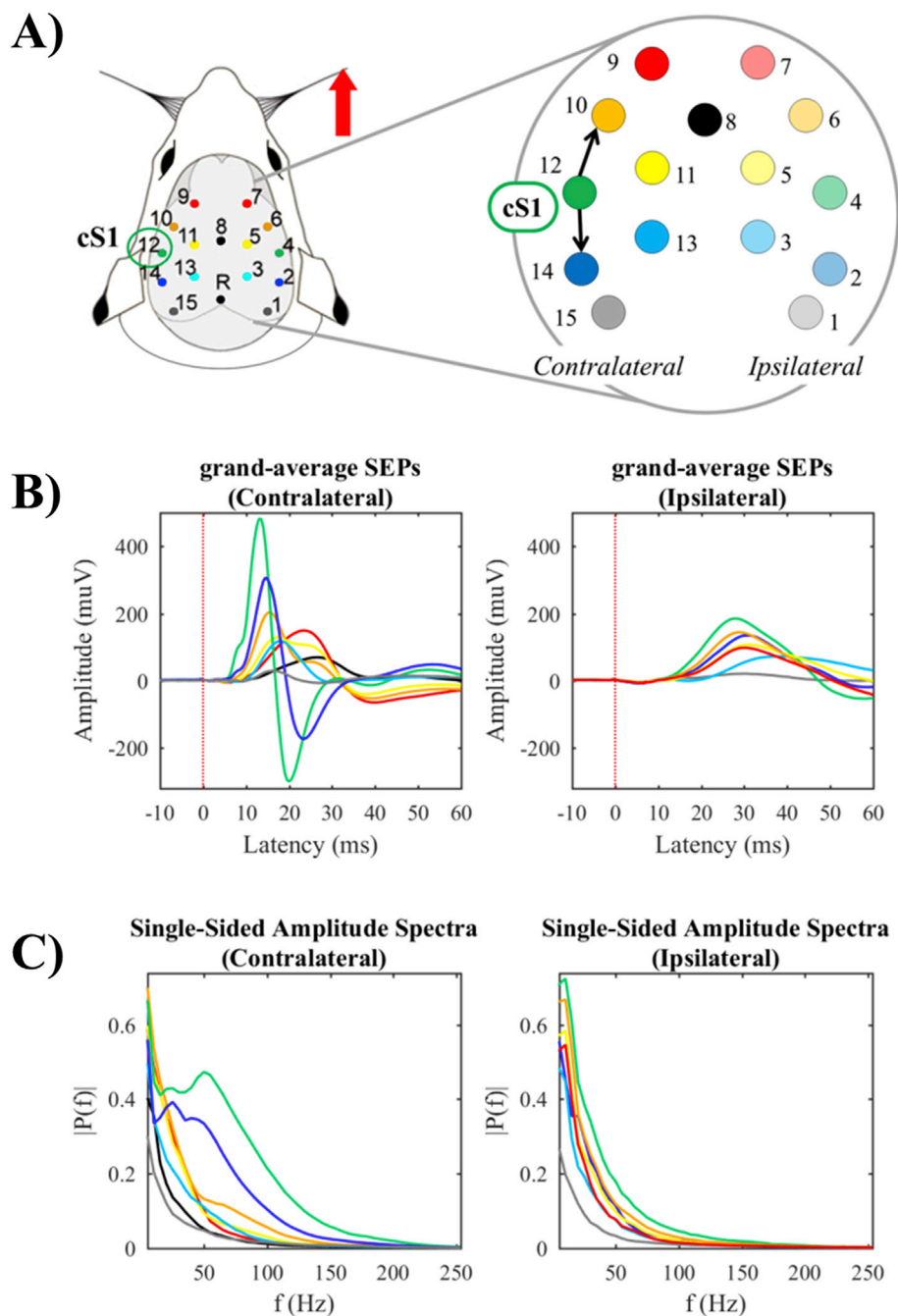


Fig. 1. A) Shows the grid of electrodes placed on rat's skull bone for recording; the red arrow schematically represents the unilateral whisker stimulation. Electrodes from 1 to 7 are on the ipsilateral hemisphere with respect to the whisker stimulation; while electrodes from 9 to 15 are on the contralateral hemisphere. The schematic diagram shows the physiologically plausible targets (nodes 10 and 14) of the early driving from cS1 (node 12). B) Plots grand-average whisker-evoked SEPs, separately for nodes on contralateral hemisphere and nodes on ipsilateral hemisphere. C) Plots normalized grand-average single-sided spectra of nodes on contralateral hemisphere and of nodes on ipsilateral hemisphere. The color-coding here introduced for each node of the network will be used in the entire manuscript to report the results.

each rat; single-rat FFT estimates were finally averaged across rats for each node of the network.

After preprocessing the whisker-evoked benchmark data allowed performance assessment based on three previously proposed criteria (Plomp et al., 2014b), which consider three different characteristics expected in the cortical network that comprises 15 nodes (Fig. 1A). More specifically, criterion I determines whether the method is able to distinguish the main driver of the cortical network as cS1, which is identified by node 12. Criterion II determines whether the driving from cS1 is preferentially directed towards contralateral parietal and frontal sensory-motor cortex, respectively defined by node 14 and node 10 (Fig. 1A). Criterion III evaluates whether the peak of cS1's driving is at physiologically plausible latency, i.e. between 5 and 20 ms after stimulus onset, which is the range around the 11 ms latency observed with

single-unit responses in layer V of cS1 (Shuler et al., 2001).

On criterion I the driving from cS1 was compared to that of the second largest driver to evaluate the identifiability of cS1 as the main driver. To assess the main targets of cS1 (criterion II), the connections from cS1 towards node 14 and node 10 were compared with those towards the equidistant medial nodes 13 and 11, respectively (Fig. 1A). Statistical comparisons between drivers or targets were performed at the latency of peak-driving from cS1, by using a bootstrap approach across rats (Plomp et al., 2014b). Such approach consisted in resampling with replacement the observed differences between drivers (criterion I) or targets (criterion II), in order to create a bootstrap distribution of differences ($n = 10,000$); here each resample was of the same size as the original dataset (Efron, 1987; Efron and Tibshirani, 1993). Afterwards, the bias-corrected and accelerated (BCa) method was used to compute the 95% confidence

interval (CI) of the bootstrap distribution. The BCa method corrects for both bias and skewness in the bootstrap distribution (Efron, 1987; Efron and Tibshirani, 1993). The comparisons assumed null hypotheses of no difference between drivers (criterion I) and no difference between targets (criterion II); hence, for a specific comparison a significant difference was obtained when the lower bound of the BCa 95% CI exceeded zero. For each comparison we computed effect sizes using Cohen's d (Cohen, 1992).

2.2. Nonparametric Granger causality

2.2.1. Frequency domain Granger causality

The statistical definition introduced by Granger (1969) relies on the concept of temporal precedence and refers to causality as improvement in the predictability of one time series x_1 by incorporating past samples of a second simultaneously recorded time series x_2 , with respect to the predictability purely based on past samples of x_1 itself; vice versa for the causal influence from x_1 to x_2 . The time domain Granger causality provides a measure that can be seen as amount of variance of one time series explained by the history of another (Brovelli et al., 2004).

This concept has been successively extended to the frequency domain (Geweke, 1982). If we consider multivariate data X at time t and with the number of channels given by d , the MVAR model of order p can be expressed as:

$$\sum_{k=0}^p A_k X_{t-k} = E_t \quad (1)$$

where A_k are d -by- d coefficient matrices and E_t is a temporally uncorrelated residual error with covariance matrix Σ , which is often referred to as noise covariance matrix. In the frequency domain the cross-spectral density matrix S can be computed as follows:

$$S(f) = \begin{bmatrix} S_{11}(f) & \dots & S_{1d}(f) \\ \vdots & \ddots & \vdots \\ S_{d1}(f) & \dots & S_{dd}(f) \end{bmatrix} = H(f)\Sigma H^*(f) \quad (2)$$

where $H(f) = (\sum_{k=0}^p A_k e^{-i2\pi f k})^{-1}$ is the spectral transfer matrix of the system and the asterisk denotes matrix transposition and complex conjugation (i.e., Hermitian transpose).

The GGC from channel j to channel i at frequency f can then be expressed as a function of elements of the three matrices from equation (2), i.e. S , H and Σ , in the following way:

$$GGC_{j \rightarrow i}(f) = \ln \frac{S_{ii}(f)}{S_{ii}(f) - \left(\sum_{j \neq i} \frac{\Sigma_{jj} - \Sigma_{ij}^2}{\Sigma_{ii}} \right) |H_{ij}(f)|^2} \quad (3)$$

This measure is basically the ratio between the total spectral density of i -th channel and its intrinsic part, which is obtained as the difference between total spectral density and causal part.

When more than two simultaneous time series are analyzed, we can either perform a pairwise analysis, which consists in using the definition in equation (3) for each pair of channels in each direction, or consider a conditional definition of GGC (Chen et al., 2006; Geweke, 1984). If we consider a multivariate stochastic process, the direct causal influence from channel j to channel i conditional to the remaining w channels can be calculated as:

$$GGC_{j \rightarrow i|w}(f) = \ln \left(\frac{\Omega_{ij}}{Q_{ii}(f)\Sigma_{ii}Q_{ii}^*(f)} \right) \quad (4)$$

where Σ and Ω are the noise covariance matrices of the full system and of the subsystem in which channel j is excluded, respectively; while, Q_{ii} is obtained from:

$$Q(f) = \begin{bmatrix} Q_{ii}(f) & Q_{ij}(f) & Q_{iw}(f) \\ Q_{ji}(f) & Q_{jj}(f) & Q_{jw}(f) \\ Q_{wi}(f) & Q_{wj}(f) & Q_{ww}(f) \end{bmatrix} = \begin{bmatrix} G_{ii}(f) & 0 & G_{iw}(f) \\ 0 & 1 & 0 \\ G_{wi}(f) & 0 & G_{ww}(f) \end{bmatrix}^{-1} \begin{bmatrix} H_{ii}(f) & H_{ij}(f) & H_{iw}(f) \\ H_{ji}(f) & H_{jj}(f) & H_{jw}(f) \\ H_{wi}(f) & H_{wj}(f) & H_{ww}(f) \end{bmatrix} \quad (5)$$

where H and G are the transfer function matrices of the full system and of the subsystem in which channel j is excluded, respectively. These two spectral transfer matrices are obtained using Geweke's normalization method, which consist of multiplying transfer functions and covariance matrices by transformation matrices to make the noise terms independent (Ding et al., 2006; Wen et al., 2013a).

2.2.2. Time reversal testing

A known problem of causality analyses is that spurious estimates can be produced by instantaneous mixing due to volume conduction, or by signal-to-noise ratio (SNR) differences between channels or sources (Bastos and Schoffelen, 2016). Haufe and colleagues introduced the concepts of “strong asymmetries” in contraposition to “weak asymmetries” in multivariate time series (Haufe et al., 2012). Strong asymmetries are due to true time-lagged causal relationships, while weak asymmetries may also reflect non-causal relationships. A framework based on the use of time-reversed data as surrogates for statistical testing of causality measures has been proposed to reduce the influence of weak asymmetries (Haufe et al., 2013, 2012). This approach has been shown to be valuable to detect spurious GGC directional differences purely caused by differences in SNR between channels (Bastos and Schoffelen, 2016), and also to potentially reduce the detrimental effect of linearly mixed noise (Vinck et al., 2015), which is known to strongly influence the amount of false positives (spurious estimates) in standard measures of causality directionality. The downside of this approach is that it may also produce a reduction in true positives (Vinck et al., 2015; Winkler et al., 2016).

In this study we used time reversal testing by exploiting the “difference-based” definition of time-reversed (tr -) Granger causality (Winkler et al., 2016), according to which the net causal influence from channel j to channel i is inferred if:

$$trGGC_{j \rightarrow i}^{(net)} := GGC_{j \rightarrow i}^{(net)} - \widetilde{GGC}_{j \rightarrow i}^{(net)} = (GGC_{j \rightarrow i} - GGC_{i \rightarrow j}) - (\widetilde{GGC}_{j \rightarrow i} - \widetilde{GGC}_{i \rightarrow j}) > 0 \quad (6)$$

where the tilde denotes GGC estimates computed on time-reversed data.

2.2.3. Nonparametric framework

To estimate GGC three quantities are needed: the cross-spectral density matrix S , the spectral transfer matrix H , and the noise covariance matrix Σ . Nonparametric methods start from estimating the cross-spectral density matrix and then spectral transfer and noise covariance matrices are derived from it, such that the equality in (2) holds (Dhamala et al., 2008a, 2008b). More specifically, nonparametric methods are based on a two-steps process that consists in: (i) computing the cross-spectral density matrix with nonparametric spectral density estimation techniques; (ii) applying a spectral matrix factorization to the cross-spectral density matrix, which allows for computing spectral transfer matrix and noise covariance matrix.

There are several nonparametric spectral decomposition approaches available to compute the cross-spectral density matrix; here we focus on multitaper method and wavelet transforms. Spectral estimation based on multitaper method (Thomson, 1982) involves the use of different data tapers that are orthogonal to each other (Mittra and Pesaran, 1999; Percival and Walden, 1993). These orthogonal tapers are given by discrete prolate spheroidal sequences (DPSS), also known as Slepian sequences (Slepian, 1978; Slepian and Pollak, 1961). If we consider time series simultaneously recorded from multiple channels $\{x_r\}(r = 1, \dots, d; t = 1,$

..., N), where r is the channel index and t the discrete time index, the multitaper cross-spectrum estimator between channels l and m at frequency f for a single trial is given by:

$$S_{lm}(f) = X_l(f)X_m^*(f) = \frac{\Delta}{K} \sum_{k=1}^K \left\{ \sum_{s=1}^N w_s(k)x_{ls}e^{-i2\pi f s \Delta} \right\} \left\{ \sum_{t=1}^N w_t(k)x_{mt}e^{i2\pi f t \Delta} \right\} \quad (7)$$

where $w(k)$ ($k = 1, 2, \dots, K$) are K orthogonal tapers of length N given by DPSS and Δ is the sampling interval. Beside the length N , the tapers are also characterized by a bandwidth parameter W . The time-bandwidth product NW regulates the trade-off between variance and bias of the spectral estimates, and it is also related to the number of tapers K because the typical approach is to take the leading $K = 2NW - 1$ DPSS as data tapers (Mitra and Pesaran, 1999). Increasing NW produces a decrease in variance, but also an increase in spectral leakage and bias of the spectral estimate; vice versa when NW is decreased. When $l = m$ from equation (7) we obtain the auto-spectrum estimator for that channel. Finally, the full cross-spectral density matrix is obtained by averaging over multiple trials auto-spectrum estimators and cross-spectrum estimators for all pairs of channels. In order to obtain a time-varying estimation of GGC the multitaper approach can be used on sliding time window.

The wavelet transform allows decomposing a time series directly into time-frequency space, by computing the convolution of the time series with a scaled and translated version of a prototype wavelet function $\psi(\eta)$ that satisfies zero mean and unity square-norm conditions, which is also known as “mother wavelet”, being η a dimensionless variable (Daubechies, 1990; Farge, 1992; Percival and Walden, 2000; Torrence and Compo, 1998). For the time series $x_l(t)$ recorded from channel l , the continuous wavelet transform at time t and scale s can be calculated as:

$$W_{x_l}(t, s) = \frac{1}{\sqrt{s}} \int_{-\infty}^{+\infty} \psi^*\left(\frac{\eta - t}{s}\right) x_l(\eta) d\eta \quad (8)$$

Similarly to the multitaper-based approach (equation (7)), the wavelet cross-spectrum estimator between channels l and m (or auto-spectrum estimator when $l = m$) at time t and scale s is:

$$WS_{lm}(t, s) = W_{x_l}(t, s)W_{x_m}^*(t, s) \quad (9)$$

where the angle brackets denote averaging over multiple trials. We can then obtain the full cross-spectral density matrix $WS(t, f)$ using the relationship between Fourier frequency f and scale s for the specific mother wavelet used (Torrence and Compo, 1998).

In this study we evaluated the performance of two different mother wavelets, the Morlet wavelet and the Paul wavelet, which both are complex-valued. The Morlet wavelet consists of a complex exponential with a Gaussian envelop: $\psi(\eta) = \pi^{-1/4} e^{i\omega_0 \eta} e^{-\eta^2/2}$, where ω_0 is a parameter that regulates the trade-off between frequency and temporal resolution (Morlet et al., 1982). The Paul wavelet is defined by the prototype function $\psi(\eta) = 2^m i^m m! (\pi(2m)!)^{-1/2} (1 - i\eta)^{-(m+1)}$, where m is the wavelet parameter that affects the resolutions of this wavelet transform (Torrence and Compo, 1998).

One problem associated with the use of wavelet transforms is the presence of edge effects due to the “cone of influence” of the wavelet used (Torrence and Compo, 1998). To solve this issue we padded with zeros the end of the time series before applying the wavelet transform (Meyers et al., 1993; Percival and Walden, 2000; Torrence and Compo, 1998).

Once the cross-spectral density matrix has been computed, either with the multitaper method or with wavelet transforms, we can apply spectral matrix factorization to obtain the spectral transfer matrix and the noise covariance matrix (Dhamala et al., 2008b; Wiener and Masani, 1957). Here, we employed the Wilson's algorithm for spectral factorization (Wilson, 1972). A review of alternative spectral factorization methods can be found in (Sayed and Kailath, 2001).

2.2.4. Benchmark analyses of nonparametric GGC

We performed a series of analyses to benchmark nonparametric GGC. If not stated otherwise the analyses were performed at sampling rate of 1000 Hz and using the conditional implementation of GGC in combination with time reversal testing, according to equation (6). For simplicity, in the description of the results we will often refer to these conditional time-reversed estimates just as GGC. In all the analyses with the multitaper we employed sliding windows with maximum overlap in time.

In the first analysis we aimed to evaluate the robustness of nonparametric methods against downsampling; we applied multitaper and Morlet wavelet approaches at original sampling rate of 2000 Hz and after downsampling the dataset to 1000 Hz, 500 Hz, and 250 Hz, using decimation factors of 2, 4, and 8, respectively. Here we used $NW = 7.5$ and a sliding-window of 20 ms for the multitaper, and $\omega_0 = 1$ for the Morlet wavelet.

In analyses two and three we benchmarked the two wavelet-based approaches. In the second analysis we evaluated the influence of the central frequency parameter ω_0 on the performance of the Morlet wavelet, by considering ω_0 values of 3, 5, 6, 9, and 12. In the third analysis we assessed the effects of varying the wavelet parameter m in the Paul wavelet, for values of 1, 3, 6, 12, and 20.

We then critically assessed the performance of the multitaper by varying time-bandwidth parameter NW and window size, in the fourth and fifth analysis respectively. In the fourth analysis we tested NW of 1.5, 2, 4, and 7.5, while the window size was set to 20 ms. In the fifth analysis we employed window sizes of 10 ms, 30 ms, and 40 ms, in addition to the 20 ms window, while considering $NW = 4$.

Finally, we assessed the performance of pairwise GGC in both multitaper and wavelet methods. In this complementary analysis we selected $\omega_0 = 6$ for the Morlet wavelet and $m = 12$ for the Paul wavelet, while for the multitaper we set $NW = 4$ combined with window size of 20 ms.

In every condition evaluated, the total driving from each node of the network was estimated as sum of the outgoing GGC estimates from that specific node, and then averaged across the gamma-band [40–90 Hz], which is predominant over cS1 and surrounding nodes (Cardin et al., 2009; Gerasimova et al., 2014; Minlebaev et al., 2011). Data processing and analyses were performed using in-house codes implemented in MATLAB (The MathWorks, Inc.). For each analysis the results obtained using the traditional GGC, i.e. without time reversal testing, are provided in the Supplementary material.

3. Results

3.1. The effect of downsampling

We evaluated the effects of varying sampling rate on the performance of nonparametric methods because downsampling may potentially corrupt causality estimates, as previously shown via analytic solutions (Barnett and Seth, 2017).

Results at original sampling rate of 2000 Hz showed good performance in cS1 identification (criterion I) for both Morlet wavelet (Fig. 2A) and multitaper approach (Fig. 2E). Analogous results were also obtained after downsampling with both approaches. We observed in fact an overall robustness in effect sizes across the four sampling rates considered, with d constantly above 1 for the Morlet wavelet (Fig. 2A–D) and around 0.8 for the multitaper (Fig. 2E–H).

Stable results across sampling rates of 2000, 1000 and 500 Hz were also obtained on criterion II, where the contralateral parietal region (node 14) was consistently identified as the main target of cS1, with stable effect sizes (Table 1). At these three sampling rates, only the multitaper was able to correctly identify the frontal sensory-motor region (node 10) as cS1's target. Downsampling to 250 Hz produced a detrimental effect on the multitaper, which was no longer able to correctly identify the two targets; differently, a slight improvement on criterion II performance was obtained with the Morlet wavelet at the lowest sampling rate.

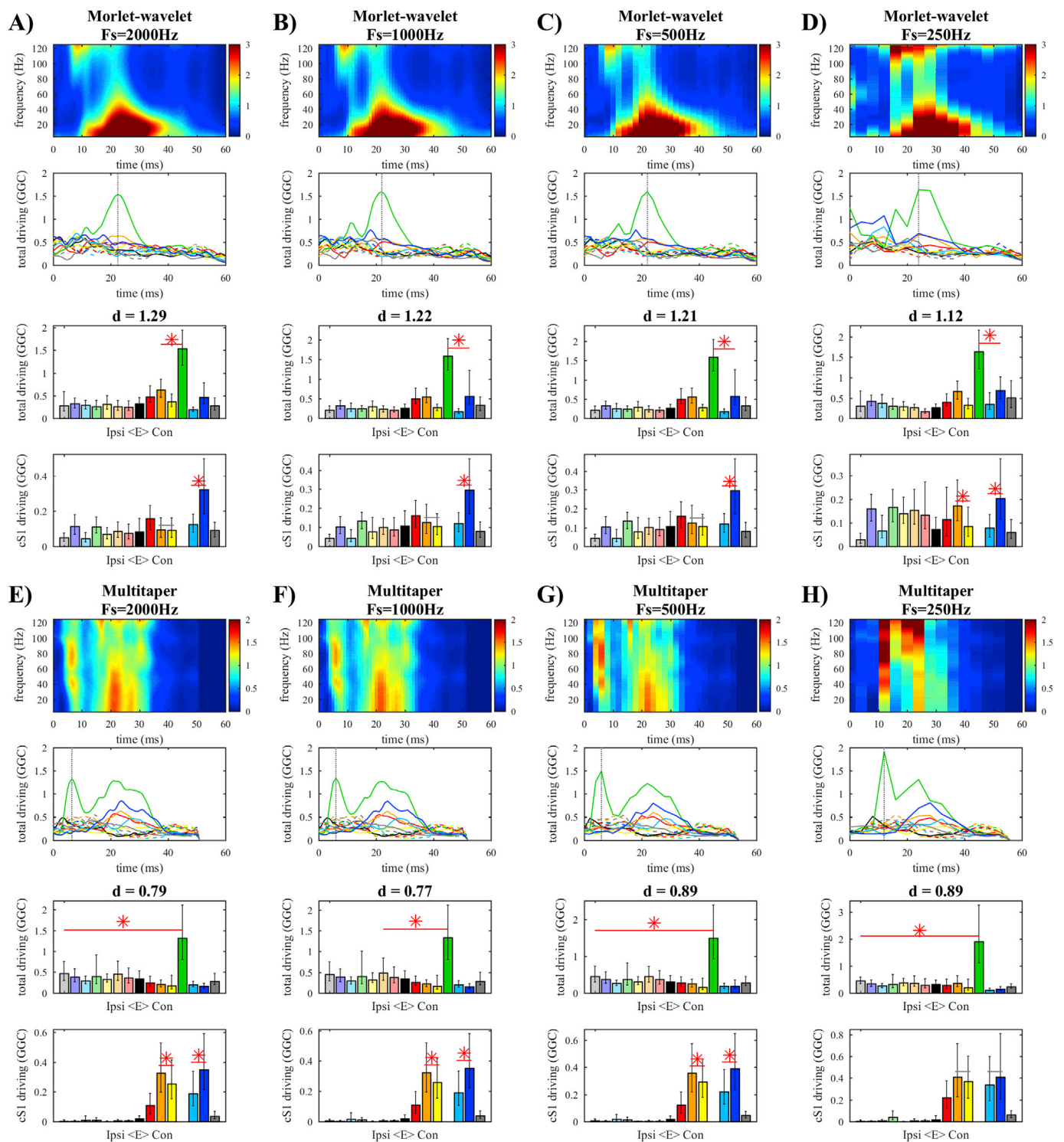


Fig. 2. The results of the analyses varying sampling rate are shown for the following methods: wavelet-based at A) 2000 Hz, B) 1000 Hz, C) 500 Hz, and D) 250 Hz; multitaper-based at E) 2000 Hz, F) 1000 Hz, G) 500 Hz, and H) 250 Hz. First row: time-frequency representation of the total driving from cS1. Second row: temporal dynamics of total driving in the gamma-band for each node and discriminability of cS1 peak-driving. Third row: barplots show average total driving from each ipsilateral (Ipsi) and contralateral (Con) node at the latency of peak-driving from cS1 (gray dotted line in the second row picture), with corresponding 95% CIs across rats and red asterisks denoting statistically significant differences with respect to the second largest driver (criterion I). Fourth row: barplots show the driving from cS1 to all the other channels, at the latency of peak-driving from cS1 (criterion II), and red asterisks denoting statistically significant differences between targets. Color coding of all bars is consistent with Fig. 1.

While the representations of the temporal dynamics of gamma-driving from cS1 seem to suggest that the multitaper may provide a clear advantage in temporal characterization (Fig. 2E–H), being cS1's peak-latency better in line with the latency of maximal activity over cS1

(Fig. 1B), we actually observe higher variability in latencies across rats (criterion III) for the multitaper compared to the Morlet wavelet (Table 2). The explanation to this apparent discrepancy is quite simple: the representations in Fig. 2 are the result of a grand-average across rats;

Table 1
Effect sizes (d) for distinguishing cS1's main targets (criterion II).

	Fs = 2000 Hz		Fs = 1000 Hz		Fs = 500 Hz		Fs = 250 Hz	
	<i>c. parietal</i>	<i>c. frontal</i>	<i>c. parietal</i>	<i>c. frontal</i>	<i>c. parietal</i>	<i>c. frontal</i>	<i>c. parietal</i>	<i>c. frontal</i>
Morlet-wavelet	0.736*	0.017	0.713*	0.140	0.713*	0.137	0.609*	0.541*
Multitaper	0.491*	0.228*	0.496*	0.199*	0.450*	0.192*	0.130	0.082

Note. *c. parietal* refers to the contralateral parietal region and *c. frontal* to contralateral frontal sensory-motor region; asterisks denote significant differences between targets (bootstrap lower 95% CI above zero).

Table 2
Latencies (ms) of peak-driving from cS1 (criterion III) and corresponding 95% CIs across rats.

	Fs = 2000 Hz	Fs = 1000 Hz	Fs = 500 Hz	Fs = 250 Hz
Morlet-wavelet	17.7 (14.9–20.4)	18.2 (15.3–20.9)	18.3 (15.5–21.1)	19.4 (14.5–23.2)
Multitaper	20.2 (15.3–24.5)	19.7 (14.9–24.0)	13.3 (9.2–18.1)	21.5 (17.3–25.7)

differently, criterion III considers the latency of maximum cS1 gamma-driving at single-rat level, and this latency varies from rat to rat. As a matter of fact, even in the representations of Fig. 2 we did not obtain a single isolated and well-defined peak for the multitaper, but rather two components, with the first one a bit more sharp and showing a peak at latency of 5–10 ms, and the second one more flattened and distributed over time (around 20 ms after stimulus onset), which are the results of estimates variability between animals. Nonetheless, an important aspect is that the results on criterion III confirm an overall robustness of nonparametric methods against signal downsampling.

In sum, our results from the three performance criteria suggest that both Morlet wavelet and multitaper approaches are quite robust to signal downsampling in the range of sampling rates here considered. Similar results are expected when using the Paul wavelet because the robustness against downsampling in terms of causal influences derives from robustness against downsampling of the spectral density estimates. Hence, as long as i) we properly avoid aliasing distortions/artifacts, ii) we use the same zero-phase antialiasing filter across sampling rates, and iii) the interactions time-scales are still observable at the considered sampling rates, we would observe robustness against downsampling of the spectral causal estimates, not only for the Paul wavelet, but also for any other nonparametric spectral decomposition approach.

3.2. Morlet wavelet (ω_0)

The Morlet wavelet used here (see section 2.2.3) is a truncated version of the complete Morlet wavelet, which is defined by the complete prototype function: $\psi(\eta) = \pi^{-1/4}(e^{i\omega_0\eta} - e^{-\omega_0^2/2})e^{-\eta^2/2}$. In this prototype function the second term inside the brackets corrects for non-zero mean of the first term inside the brackets, and is determined by the admissibility criterion. The central frequency parameter ω_0 controls the trade-off between temporal and spectral resolution in the Morlet. Reducing ω_0 produces an increase in temporal resolution together with a reduction in spectral resolution; vice versa, when increasing ω_0 . In the truncated version of the prototype, which is commonly used in practice, the correction term is always approximated to zero, but its value depends on ω_0 . The bigger ω_0 is, the more the correction term is actually close to zero; for example, when $\omega_0 \geq 5$ the correction term is lower than 10^{-5} . For this reason, previous studies have suggested to use ω_0 equal to 6 or bigger (De Moortel et al., 2004; Farge, 1992; Torrence and Compo, 1998). In the previous analysis we used $\omega_0 = 1$; here we extended the evaluation of the central frequency parameter ω_0 considering values of 3, 5, 6, 9, and 12.

The use of $\omega_0 = 3$ in the Morlet guaranteed a correct identification of cS1 as main driver (Fig. 3A). On criterion II, we observed better performance with $\omega_0 = 3$, characterized by improved identification of node 10

as functional target (Table 3), with respect to the case $\omega_0 = 1$ (Table 1); however, we also obtained a temporal bias for cS1's peak-driving towards longer latencies (Table 4).

Similar results were obtained across ω_0 values of 5 and 6; these were characterized by a correct identification of cS1 as main driver of the network with $d = 1.45$ and $d = 1.37$, respectively (Fig. 3B–C), which are higher than those obtained using smaller values of ω_0 (Figs. 2B and 3A). Compared to the case $\omega_0 = 1$, setting $\omega_0 = 5$ or $\omega_0 = 6$ allowed also to improve performance on criterion II, by correctly distinguishing both cS1's main targets with large effect sizes (Table 3). Differently, on criterion III we observed a further increased temporal bias towards longer latencies for cS1's peak-driving, with average latency across rats around 35 ms (Table 4), which is significantly later with respect to the physiologically plausible range; this confirms that increasing ω_0 reduces the temporal resolution of the Morlet. The total driving from cS1 showed a distinct pattern in the time-frequency space for $\omega_0 = 5$ and $\omega_0 = 6$ (Fig. 3B–C, first row), characterized by an early driving before 10 ms, which was however biased towards high-gamma frequencies (above 90 Hz); we also observed a successive 20 Hz-driving at latencies around 20 ms, followed by gamma-driving at later latencies (after 30–40 ms).

A qualitatively similar pattern of cS1's driving was also obtained using $\omega_0 = 9$, but with remarkably degraded temporal characterization. In terms of performance criteria, we observed an overall drop in performances when ω_0 was equal to 9 or 12 (Fig. 3D–E). This can be attributed to the further reduction in temporal resolution due to increasing ω_0 itself. Slightly better performances were obtained with these big values of ω_0 without using time reversal testing on GGC (see Supplementary material), but since the whisker-evoked activity is characterized by fast interactions, this temporal inaccuracy poses significant problems for the interpretability of the results.

3.3. Paul wavelet (m)

Compared to the Morlet wavelet, the Paul wavelet potentially provides better temporal resolution, but at the expense of reduced frequency resolution (De Moortel et al., 2004; Torrence and Compo, 1998). Here we investigated the effects of varying the wavelet parameter m in the Paul wavelet, which regulates the time-frequency resolutions trade-off.

On criterion I, we obtained a correct identification of cS1 across m levels (Fig. 4). The Paul wavelet was however able to correctly distinguish both cS1's functional targets only when $m = 3$ and when $m = 20$ (criterion II); while, with the remaining values of m only the contralateral parietal region (node 14) was reliably identified (Table 5).

On criterion III, we observed a general improvement in performance considering the Paul wavelet compared to the Morlet wavelet. The latencies of cS1's peak-driving were estimated earlier using the Paul (Table 6), with respect to those previously obtained with the Morlet (Table 4). The estimated latencies were close to the physiologically plausible range when m was equal to 1 and 3, while by increasing m we observed a consequent increase in temporal bias towards later latencies.

In terms of time-frequency components of total driving from cS1, we observed a very early driving at high-gamma frequencies only using small values of m , especially when $m = 1$ (Fig. 4A, first row), where we found in fact two peaks in cS1's gamma-driving (Fig. 4A, second row). Differently, increasing the parameter m made a later gamma-driving more distinguishable, in particular using $m = 20$ (Fig. 4E, first row).

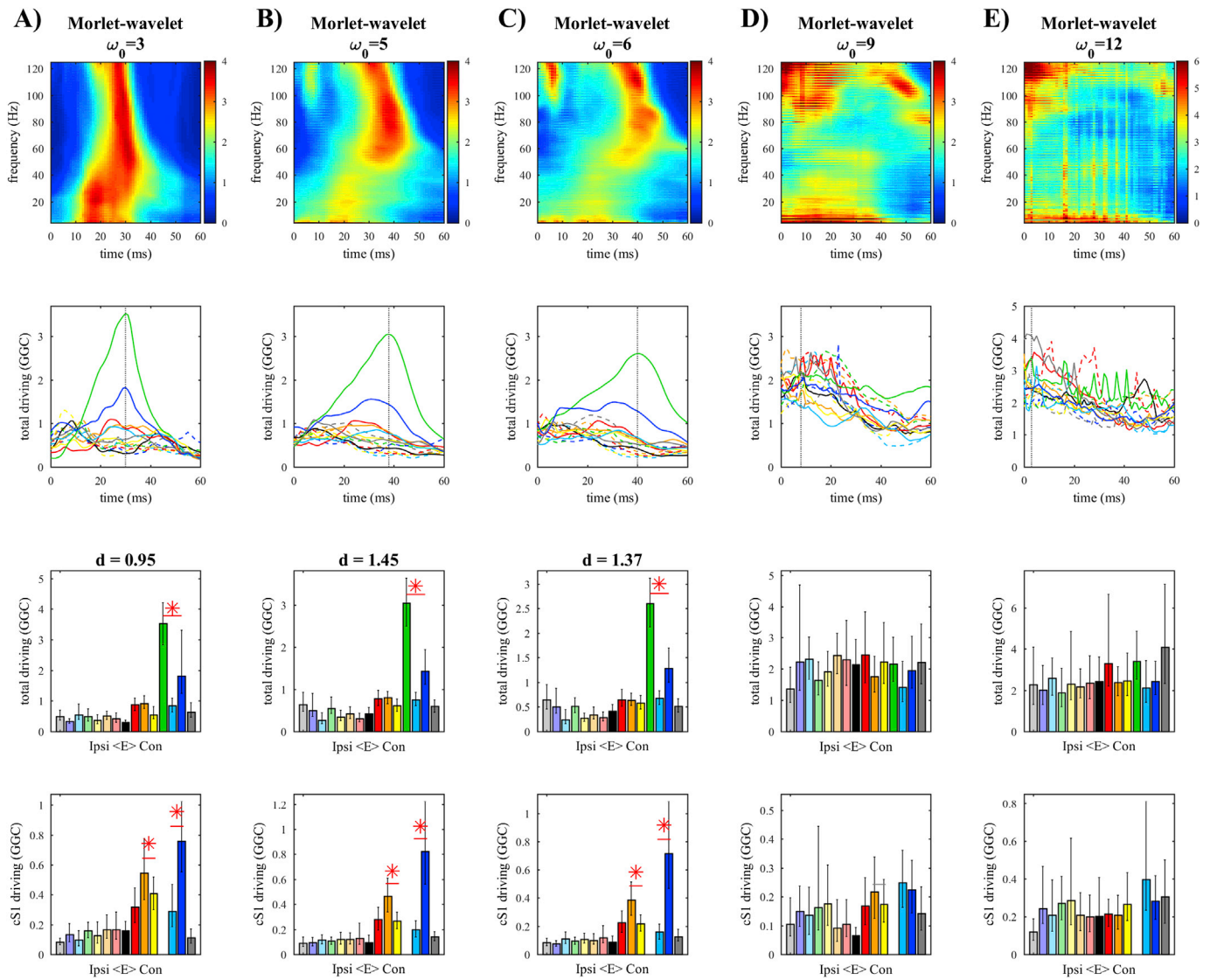


Fig. 3. Shows the results of the analyses varying the central frequency parameter (ω_0) in the Morlet-wavelet approach. Results are plotted for the following levels of the parameter: A) $\omega_0 = 3$; B) $\omega_0 = 5$; C) $\omega_0 = 6$; D) $\omega_0 = 9$; E) $\omega_0 = 12$. First row: time-frequency representation of the total driving from cS1. Second row: temporal dynamics of total driving in the gamma-band for each node and discriminability of cS1 peak-driving. Third row: barplots show average total driving from each ipsilateral (Ipsi) and contralateral (Con) node at the latency of peak-driving from cS1 (gray dotted line in the second row picture), with corresponding 95% CIs across rats and red asterisks denoting statistically significant differences with respect to the second largest driver (criterion I). Fourth row: barplots show the driving from cS1 to all the other channels, at the latency of peak-driving from cS1 (criterion II), and red asterisks denoting statistically significant differences between targets. Color-coding of all bars is consistent with Fig. 1.

Table 3
Effect sizes (d) for distinguishing cS1's main targets (criterion II).

ω_0	Morlet-wavelet	
	<i>c. parietal</i>	<i>c. frontal</i>
3	1.065*	0.361*
5	1.180*	0.819*
6	1.125*	0.785*
9	n.a.	0.206
12	n.a.	n.a.

Note. *c. parietal* refers to the contralateral parietal region and *c. frontal* to contralateral frontal sensory-motor region; asterisks denote significant differences between targets (bootstrap lower 95% CI above zero).

3.4. The time-bandwidth product (NW) in the multitaper

The time-bandwidth product NW controls the bias-variance trade-off in the multitaper method (Mitra and Pesaran, 1999). The results of the

Table 4
Latencies (ms) of peak-driving from cS1 (criterion III) and corresponding 95% CIs across rats.

ω_0	Morlet-wavelet
	Peak-latencies
3	27.0 (23.0–29.4)
5	35.1 (30.6–38.4)
6	36.5 (28.7–40.8)
9	31.6 (21.9–41.5)
12	20.2 (11.2–31.8)

analysis varying this parameter showed similar time-frequency patterns of cS1's total driving across NW levels (Fig. 5). We found an early driving from cS1 at high-gamma frequencies, which is similar to that previously observed in Morlet and Paul wavelets using small values for their respective parameters (Figs. 3–4), but this seemed more correctly localized in the expected frequency range [40–90 Hz] by using the multitaper.

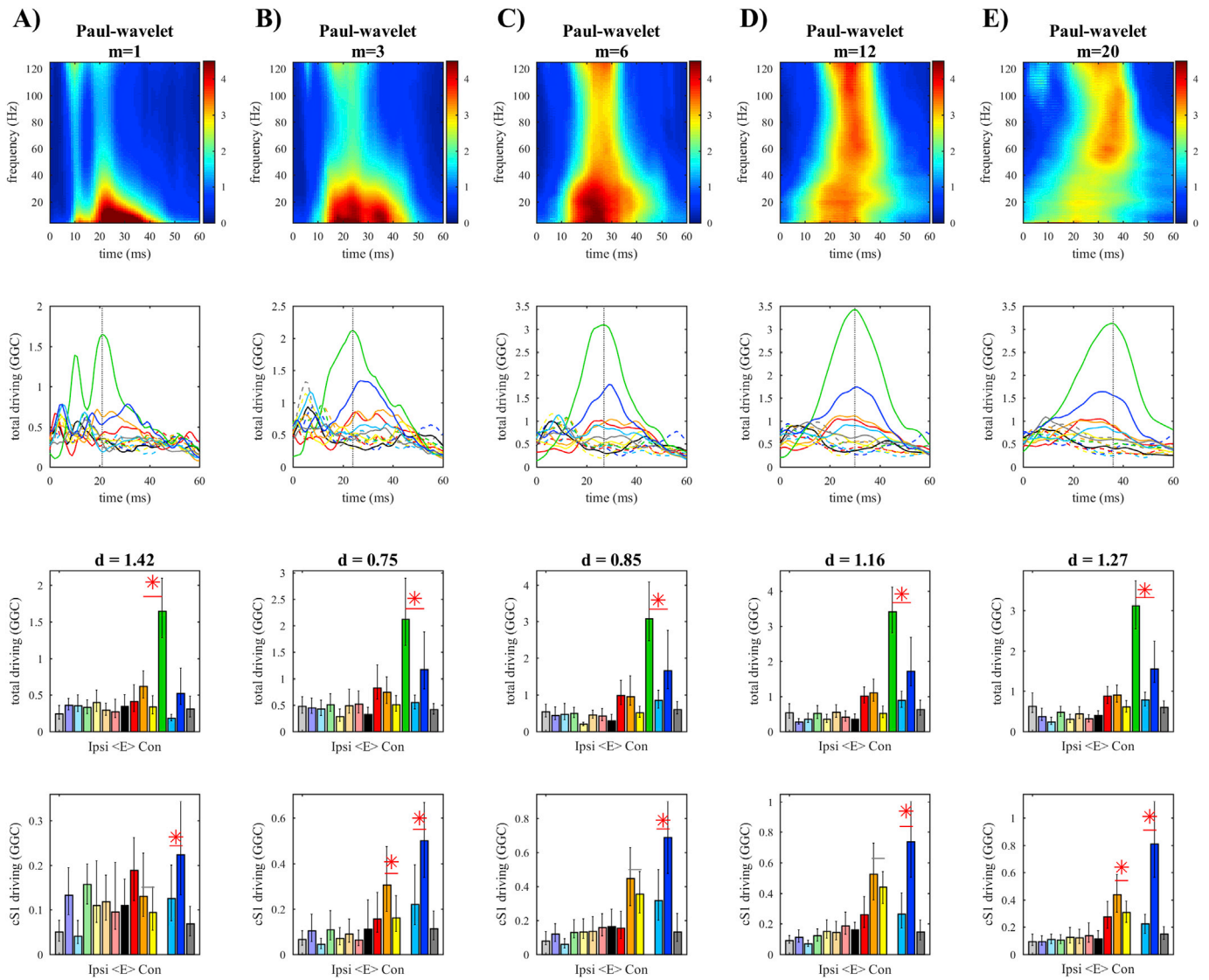


Fig. 4. Shows the results of the analyses varying the wavelet parameter (m) in the Paul-wavelet approach. Results are plotted for the following levels of the parameter: A) $m = 1$; B) $m = 3$; C) $m = 6$; D) $m = 12$; E) $m = 20$. First row: time-frequency representation of the total driving from cS1. Second row: temporal dynamics of total driving in the gamma-band for each node and discriminability of cS1 peak-driving. Third row: barplots show average total driving from each ipsilateral (Ipsi) and contralateral (Con) node at the latency of peak-driving from cS1 (gray dotted line in the second row picture), with corresponding 95% CIs across rats and red asterisks denoting statistically significant differences with respect to the second largest driver (criterion I). Fourth row: barplots show the driving from cS1 to all the other channels, at the latency of peak-driving from cS1 (criterion II), and red asterisks denoting statistically significant differences between targets. Color-coding of all bars is consistent with Fig. 1.

Table 5
Effect sizes (d) for distinguishing cS1's main targets (criterion II).

m	Paul-wavelet	
	<i>c. parietal</i>	<i>c. frontal</i>
1	0.508*	0.280
3	0.842*	0.572*
6	0.896*	0.285
12	1.060*	0.248
20	1.261*	0.511*

Note. *c. parietal* refers to the contralateral parietal region and *c. frontal* to contralateral frontal sensory-motor region; asterisks denote significant differences between targets (bootstrap lower 95% CI above zero).

The successive low-frequency driving at around 20 ms was identified as a double-peaked component across NW levels, and it was more prominent when considering a value of at least 4 for NW (Fig. 5C–D, first row). Differently, the late gamma-driving component seemed more prominent

Table 6
Latencies (ms) of peak-driving from cS1 (criterion III) and corresponding 95% CIs across rats.

m	Paul-wavelet
	Peak-latencies
1	19.8 (15.3–26.7)
3	20.9 (16.9–25.4)
6	24.8 (21.6–27.0)
12	30.6 (28.0–33.2)
20	33.6 (29.8–36.8)

with small NW (Fig. 5A–B, first row).

In all cases we obtained two peaks in the grand-average gamma-driving from cS1, the first one in the range 5–10 ms after stimulus onset and the second one at latency between 25 and 30 ms. For NW equal to 1.5 and 2 the highest peak was the second one, while for NW equal to 4 and

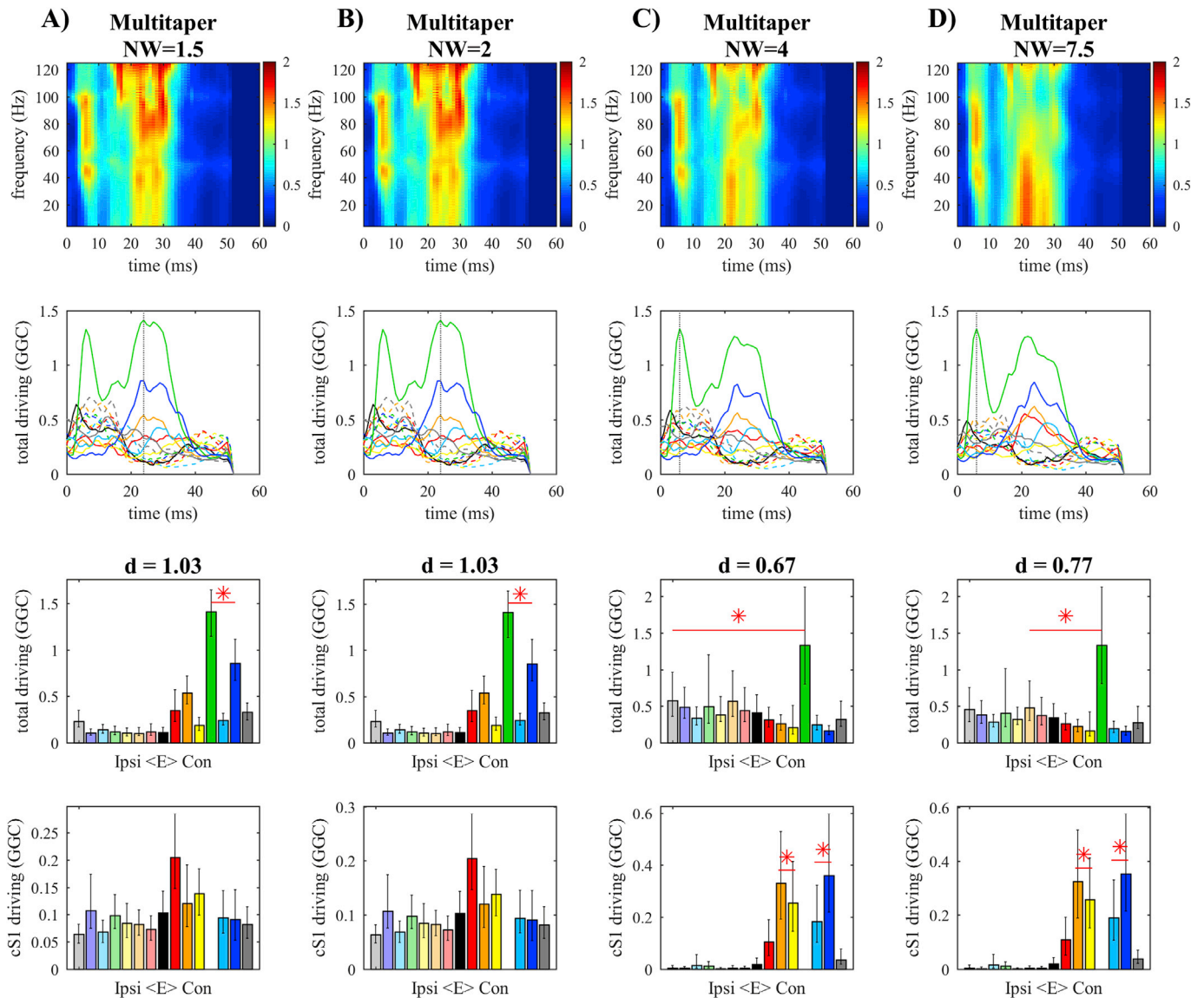


Fig. 5. Shows the results of the analyses varying the time-bandwidth parameter (NW) in the multitaper approach. Results are plotted for the following levels of the parameter: A) $NW = 1.5$; B) $NW = 2$; C) $NW = 4$; D) $NW = 7.5$. First row: time-frequency representation of the total driving from cS1. Second row: temporal dynamics of total driving in the gamma-band for each node and discriminability of cS1 peak-driving. Third row: barplots show average total driving from each ipsilateral (Ipsi) and contralateral (Con) node at the latency of peak-driving from cS1 (gray dotted line in the second row picture), with corresponding 95% CIs across rats and red asterisks denoting statistically significant differences with respect to the second largest driver (criterion I). Fourth row: barplots show the driving from cS1 to all the other channels, at the latency of peak-driving from cS1 (criterion II), and red asterisks denoting statistically significant differences between targets. Color-coding of all bars is consistent with Fig. 1.

7.5 it was the first one. This influenced the performance on criteria I and II. On the one hand, we observed basically identical results for NW equal to 1.5 and 2 (Fig. 5A–B): the multitaper was able to correctly distinguish cS1 as main driver ($d = 1.03$), but it was not able to identify cS1's targets. On the other hand, the multitaper showed good performances when using $NW = 4$ and $NW = 7.5$ (Fig. 5C–D): on criterion I we obtained correct identification of cS1 with effect sizes $d = 0.67$ for $NW = 4$ and $d = 0.77$ for $NW = 7.5$; both these NW values provided also correct identification of the targets of cS1 (Table 7).

Despite the differences on criteria I and II between the two small NW values and the two big ones, we observed similarities in results across all NW values on criterion III. As a matter of fact, we obtained robust average peak-driving latencies and 95% CIs across values of NW (Table 8), and overall the estimated latencies overlapped with upper bound of the physiologically plausible range.

For each NW value, two peaks of cS1 driving were observed. When

Table 7

Effect sizes (d) for distinguishing cS1's main targets (criterion II).

NW	Multitaper	
	<i>c. parietal</i>	<i>c. frontal</i>
1.5	n.a	n.a.
2	n.a.	n.a.
4	0.531*	0.223*
7.5	0.496*	0.199*

Note. *c. parietal* refers to the contralateral parietal region and *c. frontal* to contralateral frontal sensory-motor region; asterisks denote significant differences between targets (bootstrap lower 95% CI above zero).

analyzing results at the first peak latency, the two lower NW values showed correct identification of contralateral parietal ($d \approx 0.55$) and sensory-motor ($d \approx 0.23$) regions on criterion II; while, on criterion I the multitaper was still unable to correctly detect cS1.

Table 8
Latencies (ms) of peak-driving from cS1 (criterion III) and corresponding 95% CIs across rats.

NW	Multitaper
	Peak-latencies
1.5	20.0 (15.1–24.1)
2	20.0 (15.2–24.2)
4	20.1 (15.1–24.3)
7.5	19.7 (14.9–24.0)

3.5. Window size in the multitaper-based method

Since multitapers require a-priori choice of window size, we evaluated the effects of varying window size on the connectivity performance; we here set $NW = 4$. Looking at the grand-average driving from cS1 in the time-frequency space, we observed a good detection of the early gamma-

band component using the 20 ms window (Fig. 6, first row). The use of the shortest window (10 ms) qualitatively produced a detrimental effect in terms of spectral characterization; while on the other side, negative effects on temporal characterization were obtained using window sizes of 30–40 ms. More specifically, when we used the 40 ms window the dominant component of driving was observed at low frequencies (below 20 Hz), and this component had a broad temporal extension approximately in the time interval 20–35 ms after stimulus onset. These qualitative results confirmed the temporal-spectral resolutions trade-off depending on window size considered.

Regardless of window size, the multitaper was able to distinguish cS1 as main driver of the network (criterion I), with effect sizes $d = 0.88$ for the 10 ms window, $d \approx 0.67$ – 0.68 for the 20 ms and 30 ms windows, and $d = 1.07$ for the 40 ms window (Fig. 6).

On criterion II, the method correctly identified the two main targets of cS1, but only for windows of 20 ms or larger, with similar effect sizes across these windows for both cS1's targets (Table 9).

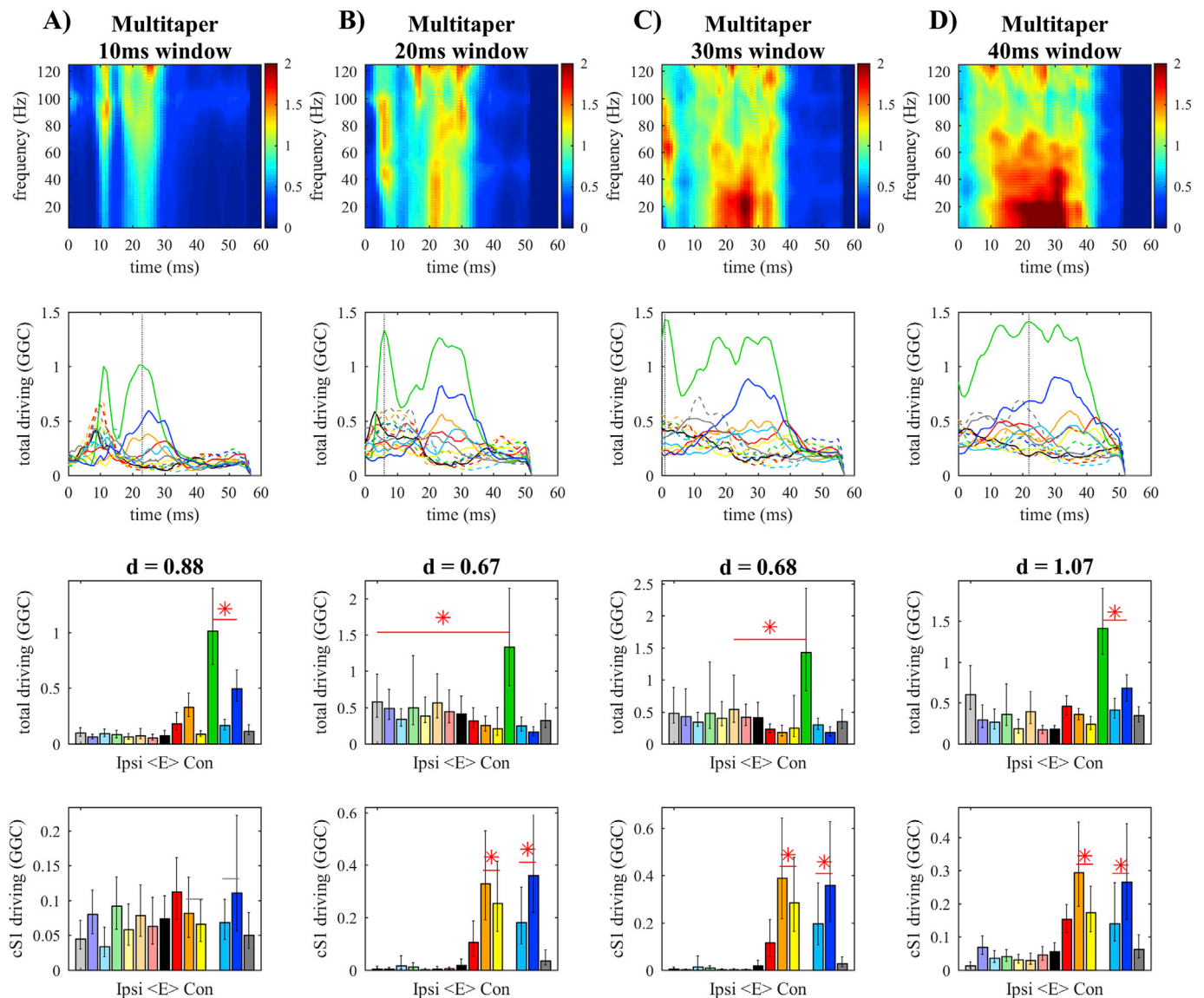


Fig. 6. Shows the results of the analyses varying window size in the multitaper approach. Results are plotted for the following window sizes: A) 10 ms; B) 20 ms; C) 30 ms; D) 40 ms. First row: time-frequency representation of the total driving from cS1. Second row: temporal dynamics of total driving in the gamma-band for each node and discriminability of cS1 peak-driving. Third row: barplots show average total driving from each ipsilateral (Ipsi) and contralateral (Con) node at the latency of peak-driving from cS1 (gray dotted line in the second row picture), with corresponding 95% CIs across rats and red asterisks denoting statistically significant differences with respect to the second largest driver (criterion I). Fourth row: barplots show the driving from cS1 to all the other channels, at the latency of peak-driving from cS1 (criterion II), and red asterisks denoting statistically significant differences between targets. Color-coding of all bars is consistent with Fig. 1.

On criterion III, however, window size had a clear effect: the shorter the window, the earlier the estimated latencies of cS1's peak-driving; reducing window size resulted in peak driving latencies closer to the physiologically plausible range (Table 10).

This shows that the dynamics of interactions strongly depend on window size and that overall the 20 ms window guarantees results most in line with the expected activity pattern.

3.6. Pairwise GGC

We next assessed the performance of pairwise GGC using three spectral decomposition approaches with optimal parameter settings obtained from the multivariate results above: Morlet-wavelet with $\omega_0 = 6$, Paul-wavelet with $m = 12$, and multitaper method with $NW = 4$ and 20 ms sliding-window. Historically, pairwise approaches have been often used for the easiness of their implementations (i.e., repeating simple bivariate analyses), and because they require the fitting of less parameters compared to a full-multivariate approach. However, pairwise approaches may not clearly distinguish between direct and indirect causal influences (Chen et al., 2006; Dhamala et al., 2008a; Ding et al., 2006).

The use of pairwise GGC produced here an overall reduction in performance for all three approaches, compared to the respective results obtained using conditional GGC under the same conditions. On criterion I, bivariate implementations of both Morlet and Paul wavelets failed in the identification of main driver; in fact, while cS1's gamma-driving on average reached the highest peak among the nodes of the network, in both wavelet-based approaches this was not significantly different from the driving from node 14, which was the second largest driver (Fig. 7A–B). Differently, the multitaper was able to correctly identify cS1 (Fig. 7C), but with effect size $d = 0.34$ lower than that obtained with conditional GGC, i.e. $d = 0.67$ (Fig. 6B).

The limitations of pairwise GGC became most clear on criterion II. The three nonparametric approaches performed in fact poorly in identifying the preferential functional targets of cS1 (Table 11). Moreover, we found the predominant receivers of cS1 on the ipsilateral hemisphere in most of the cases (Fig. 7, fourth row), which is clearly not physiologically plausible.

Finally, within each method the estimated range of cS1's peak-driving latencies did not differ much from that obtained with conditional GGC (Table 12), with only the multitaper showing peak latencies convergent with the activity profile of cS1.

4. Discussion

In this study, we systematically compared the performance of several nonparametric GGC methods based on time-varying spectral factorization (Dhamala et al., 2008a, 2008b), by using highly dynamic benchmark EEG data recorded from rats during unilateral whisker stimulations (Plomp et al., 2014b; Quairiaux et al., 2011). Nonparametric methods require an initial choice of spectral decomposition parameters and our results demonstrate that such selection is of crucial importance; nonetheless, an informed choice of these parameters can provide good and meaningful results in practice.

Table 9
Effect sizes (d) for distinguishing cS1's main targets (criterion II).

Window size	Multitaper	
	<i>c. parietal</i>	<i>c. frontal</i>
10 ms	0.330	0.189
20 ms	0.531*	0.224*
30 ms	0.422*	0.248*
40 ms	0.493*	0.540*

Note. *c. parietal* refers to the contralateral parietal region and *c. frontal* to contralateral frontal sensory-motor region; asterisks denote significant differences between targets (bootstrap lower 95% CI above zero).

Table 10

Latencies (ms) of peak-driving from cS1 (criterion III) and corresponding 95% CIs across rats.

Window size	Multitaper
	Peak-latencies
10 ms	17.5 (14.8–20.2)
20 ms	20.1 (15.1–24.3)
30 ms	20.9 (14.3–26.7)
40 ms	26.1 (19.8–31.1)

In order to be a reliable and valuable tool for real data applications, any time-varying directed connectivity analysis should be able to identify functionally relevant drivers and preferential connections in the network, and furthermore correctly characterize how such connections evolve over time. By evaluating main driver detection in the benchmark dataset, our results show that nonparametric GGC methods are quite robust in correctly identifying the dominant functional driver, which, in general, is extremely important for the identification of areas functionally involved in normal processing or pathological activity spread. We observed in fact that all nonparametric methods successfully detected cS1 as the main driver of the whisker-evoked cortical network, irrespective of parameter settings considered, when using conditional GGC. Only the Morlet wavelet failed in the identification of cS1, when its central frequency parameter ω_0 was equal to 9 or above. This notwithstanding, when ω_0 was in the range between 3 and 6, the Morlet wavelet guaranteed not only a good detection of cS1, but also the best identification of its two preferential functional targets. The Paul wavelet was able to distinguish only the contralateral parietal region (node 14, Fig. 1A), regardless of m value; while, we observed inconsistent identification of the other cS1's functional target (node 10, Fig. 1A) across m values. For the sliding-window multitaper good performance in target identification was obtained only with time-bandwidth product $NW \geq 4$ and window size of at least 20 ms. These findings suggest that care should be taken in selecting spectral decomposition parameters in nonparametric methods, when we are interested in correctly characterizing the functional connections in the network, which is crucial for inferring network properties in topological and graph-based analyses. In terms of temporal characterization, the wavelet transforms seem suboptimal, the Morlet wavelet especially, while the multitaper method appears more promising for investigating fast-evolving interactions. In the benchmark dataset, for the Morlet wavelet we observed an overall temporal bias towards longer latencies for cS1's peak-driving, which increased by increasing ω_0 . The use of the Paul wavelet with $m \leq 3$ and of the multitaper with small window guaranteed a better temporal characterization of the whisker-evoked network dynamics, with estimated latencies of cS1's peak-driving overlapping with the upper bound of the physiologically plausible range.

Altogether our findings show that connectivity performance of nonparametric methods can strongly depend on parameters choice; hence, a spectral decomposition method with certain parameter settings may be optimal for a specific scenario/dataset, but not for others. In practical applications, the choice of approach for GGC estimation and parameter settings should be guided by the main purpose of the analysis and by any prior information about the investigated data, when available.

If we are interested in identifying the main drivers and targets within the network, the Morlet wavelet is a promising solution, choosing ω_0 between 5 and 6. With these settings the Morlet wavelet is suitable for studying causal influences with relatively slow temporal evolution, but it may be ineffective for the investigation of interactions whose temporal evolution is fast. Since ω_0 controls the trade-off between temporal and spectral resolution, and better temporal resolution can be obtained when ω_0 is small (Percival and Walden, 2000; Torrence and Compo, 1998), some may be tempted to significantly lower ω_0 to improve temporal characterization. However, selecting ω_0 too small introduces a deformation in the scale-space, and this may produce significant negative

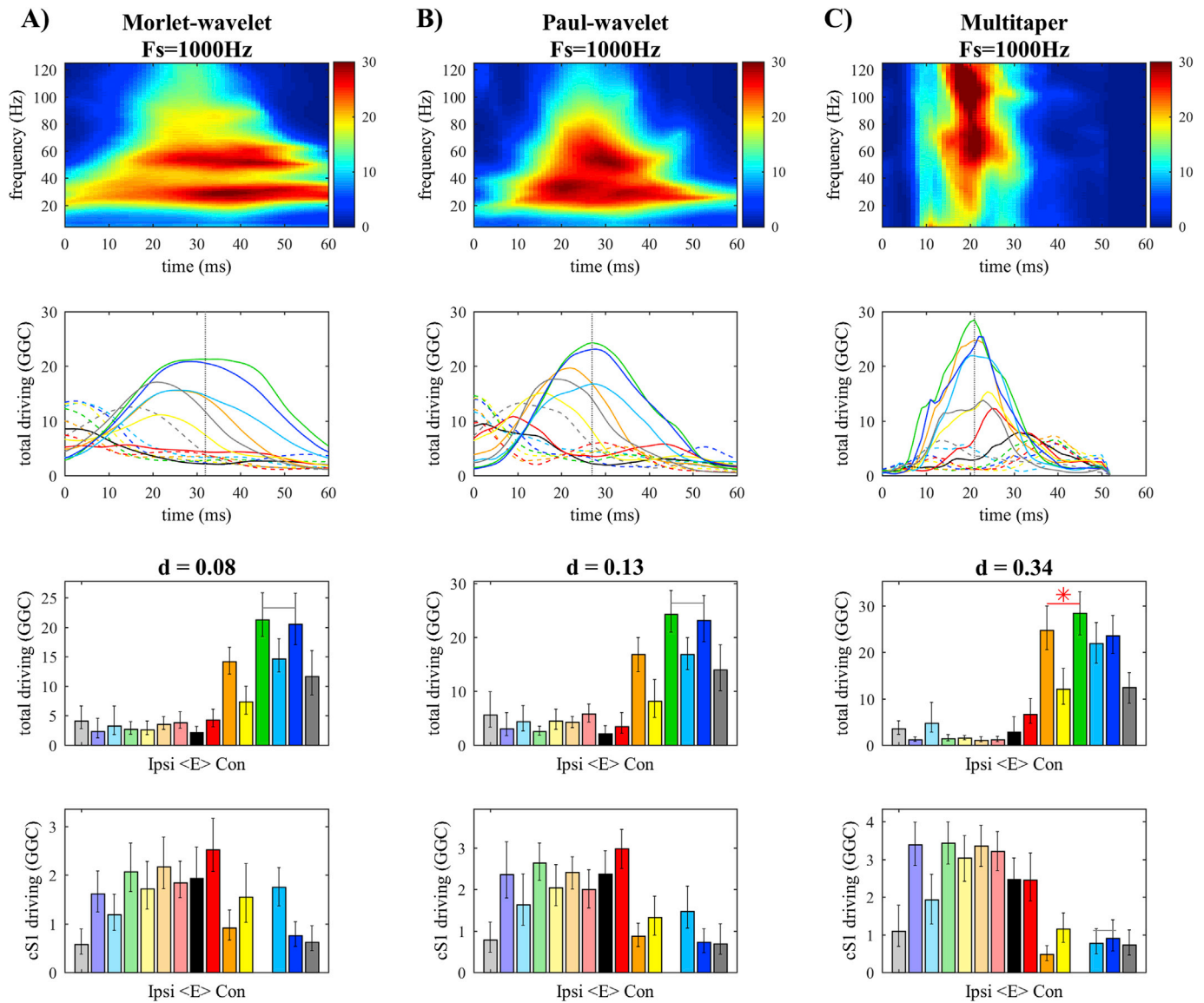


Fig. 7. The results of the analyses using pairwise GGC are shown for: A) Morlet-wavelet with $\omega_0 = 6$; B) Paul-wavelet with $m = 12$; C) sliding-window multitaper with $NW = 4$ and window size of 20 ms. First row: time-frequency representation of the total driving from cS1. Second row: temporal dynamics of total driving in the gamma-band for each node and discriminability of cS1 peak-driving. Third row: barplots show average total driving from each ipsilateral (Ipsi) and contralateral (Con) node at the latency of peak-driving from cS1 (gray dotted line in the second row picture), with corresponding 95% CIs across rats and red asterisks denoting statistically significant differences with respect to the second largest driver (criterion I). Fourth row: barplots show the driving from cS1 to all the other channels, at the latency of peak-driving from cS1 (criterion II), and red asterisks denoting statistically significant differences between targets. Color-coding of all bars is consistent with Fig. 1.

effects; while, $\omega_0 \geq 5$ allows to confidently approximate the complete Morlet prototype to the truncated version described in section 2.2.3 (De Moortel et al., 2004; Farge, 1992; Torrence and Compo, 1998). From a

Table 11

Effect sizes (d) for distinguishing cS1's main targets (criterion II).

	pairwise GGC		conditional GGC	
	c.	c.	c.	c. frontal
	parietal	frontal	parietal	
Morlet-wavelet ($\omega_0=6$)	n.a.	n.a.	1.125*	0.785*
Paul-wavelet ($m=12$)	n.a.	n.a.	1.060*	0.248
Multitaper ($NW=4$; 20 ms window)	0.165	n.a.	0.531*	0.224*

Note. *c. parietal* refers to the contralateral parietal region and *c. frontal* to contralateral frontal sensory-motor region; asterisks denote significant differences between targets (bootstrap lower 95% CI above zero); n.a. is used for comparisons not available.

practical standpoint, on the opposite extreme, increasing ω_0 produces a progressive temporal flattening of the causal influences between cortical areas and, consequently, problems in identifying expected drivers and targets. Our results confirmed in fact that the time localization of the Morlet transform is almost completely lost when $\omega_0 = 12$ (De Moortel et al., 2004).

The Paul wavelet provides a valid alternative when we aim to better characterize interactions timing, because this prototype is more localized

Table 12

Latencies (ms) of peak-driving from cS1 (criterion III) and corresponding 95% CIs across rats.

	pairwise GGC	conditional GGC
Morlet-wavelet ($\omega_0=6$)	32.6 (29.2–36.4)	36.5 (28.7–40.8)
Paul-wavelet ($m=12$)	28.6 (26.2–31.8)	30.6 (28.0–33.2)
Multitaper ($NW=4$; 20 ms window)	18.7 (16.9–20.5)	20.1 (15.1–24.3)

in time and is characterized by less oscillations, and then it is expected to have better temporal resolution but worse spectral resolution compared to the Morlet wavelet, for equal wavelet parameters (De Moortel et al., 2004; Torrence and Compo, 1998). Moreover, since the Paul is admissible for any value of the parameter m , we can here lower more the parameter to improve temporal resolution, but at the expenses of spectral resolution. Our results reproduced these expected behaviors.

The sliding-window multitaper method may be the best choice for investigating fast-evolving interactions. Starting from Ding and colleagues' early work (Ding et al., 2000), several studies have exploited a short-window adaptive MVAR method to investigate neural interactions, e.g. (Brovelli et al., 2004; Zhang et al., 2008). This approach is based on the assumption of signals stationarity in short time intervals. In general, sufficiently short windows may guarantee the “local stationarity” of some data (Ding et al., 2000), which is usually considered in the weak-sense, i.e. when first (mean) and second order (variance) ensemble statistics are constant in time (Liégeois et al., 2017). Despite the possible violation of statistical assumptions in highly nonstationary data, the size of the window influences the time-frequency resolutions trade-off: reducing window size lowers spectral resolution, which may produce significant detrimental effects on interpreting the functional roles of different brain areas, but it also helps to improve temporal resolution and detect interactions that appear and disappear quickly in time; vice versa when increasing window size. We here confirmed these aspects for the multitaper-based nonparametric GGC. As rule of thumb, in order to avoid a complete loss of spectral characterization, it seems reasonable to select window size based on the frequency of interest, such that at least one period fits the length of the window. Moreover, values of NW around 4 allow in general obtaining a good compromise in the spectral estimates trade-off between variance and bias, making such values a reasonable choice for practical applications with the multitaper.

Nonparametric GGC has been initially applied to LFPs recorded from monkeys while performing a GO/NOGO visual discrimination task (Dhamala et al., 2008a) and provided a pattern of interactions between S1 and parietal areas in line with well-established functional hierarchy and anatomical connections (Felleman and Van Essen, 1991). Our current work extends this initial application substantially, by providing critical benchmark assessment of the effects of varying spectral decomposition parameters and quantifying the performance of nonparametric GGC on fixed criteria, and by introducing the Paul as an alternative wavelet prototype for time-varying spectral decomposition. In addition, we observed that nonparametric methods are quite robust against signal downsampling. However, this may not always be the case, and in general downsampling may significantly reduce detectability of causality estimates (Barnett and Seth, 2017; Seth et al., 2013; Wen et al., 2013b). Robustness against downsampling is an appealing feature of nonparametric methods because it allows obtaining reliable estimates even when signals are downsampled to lower the computational time of processing. Nonetheless, reducing sampling rate goes together with decreasing temporal resolution, and an excessive reduction in temporal resolution may produce detrimental effects on connectivity estimation performance and results interpretability also for the nonparametric methods. In our dataset, for example, a further downsampling to 125 Hz would produce negative effects, because a temporal resolution of 8 ms is too low to investigate interactions that happen in the order of about 10 ms, i.e. the time series would not be resolved enough to observe properly the phenomenon. In practice, for computational reasons, there may be scenarios in which we want to use sampling rates around 100 Hz. The reliability of the results obtained at these sampling rates depends on the temporal scale of the phenomenon under investigation, and in principle a sampling rate of 100 Hz may be suitable to investigate causal influences that evolve more slowly compared to those considered here. While downsampling seems unproblematic for nonparametric methods, temporal aspect of the signal of interest put a lower bound on the sampling rate. In this respect, by

using analytic solutions Barnett and Seth (2017) compared GGC in continuous-time at finite prediction horizons and GGC based on discrete-time subsampling of the underlying continuous-time stochastic process, and showed that both continuous-time GGC and discrete-time GGC decay exponentially as the sampling interval increases beyond the longest internal time scale of circuit-level feedback delay. Moreover, they showed that for sampling intervals smaller than the longest feedback delay, there may be “black spots” where the discrete-time (downsampled) GGC becomes smaller than the “ground truth” continuous-time GGC, and hence harder to detect in finite sample. However, the actual detectability rate will depend on the sampling distribution of the particular discrete GGC estimator used, thus some estimation methods may potentially alleviate the degradation, while others may aggravate it. This may account for the results observed in our study for nonparametric methods.

Our results provide a clear demonstration that pairwise GGC can lead to physiologically misleading results in real data. Using pairwise GGC we obtained erroneous early preferential driving from cS1 towards areas in the ipsilateral hemisphere, while in the rat model of unilateral whisker stimulation cS1's targets at early latencies are expected to be in the contralateral hemisphere. These findings can be explained in the context of the unobserved common inputs, a scenario where dependencies between two nodes are detected only because they receive inputs from other possibly unobserved nodes, called common sources (Bastos and Schoffelen, 2016). The common input is a problem in the application of causality measures because it produces ambiguities between paths of connections that are direct and those that are indirect, i.e. mediated by other nodes. In general, if common source signals are not recorded the problem is fundamentally unsolvable. Pairwise GGC is affected by the ambiguities due to indirect paths even when the common sources are recorded, as previously shown through simulated models (Chen et al., 2006; Dhamala et al., 2008a; Ding et al., 2006; Kus et al., 2004), because for each bivariate model all the other nodes are basically unobserved. Differently, the conditional GGC is able to overcome the common input problem by taking into account all available information in the recorded dataset, and consequently it can determine whether a causal influence between any pair of nodes is direct or indirect (Ding et al., 2006; Geweke, 1984; Wen et al., 2013a). Our results confirm that conditional GGC should be preferred over pairwise GGC in order to unambiguously distinguish direct connections from indirect ones.

While the findings of this study are important for better understanding the influence of spectral decomposition parameters on connectivity performance, and we show practical ways on how to apply time-varying nonparametric GGC, there are certain issues we do not explicitly address. Besides the common input problem, other practical issues are in fact associated with the application of causality analyses to real data, and these should be taken into account to further optimize results from nonparametric GGC, as well parametric GGC.

Well-known practical issues are for instance the presence on recorded signals of additive noise, also referred to as measurement/observational noise (Vinck et al., 2015; Winkler et al., 2016), and SNR differences between channels (Bastos and Schoffelen, 2016). While both these problems can be partly mitigated at experimental level before recording, time reversal testing (Haufe et al., 2013, 2012) has been proved to reduce most of the negative effects associated with them (Bastos and Schoffelen, 2016; Vinck et al., 2015; Winkler et al., 2016). This has motivated the use of time-reversed GGC in the current work. In practice, interpretability can be further improved employing analytical denoising algorithms, which help reducing the negative effects of noise on causality analyses, by explicitly modeling the additive noise with state-space model representations of the MVAR process (Nalatore et al., 2014, 2007; Sommerlade et al., 2015), or by detecting linear mixing effects from structural MVAR models (Vinck et al., 2015).

The use of a common reference may lead to spurious causality estimates when the reference signal is not electrically silent. Re-referencing

using bipolar derivations has been shown to remove adverse effects of the common reference in electrocorticography (ECoG) and laminar recordings (Bastos et al., 2015; Bosman et al., 2012; Trongnetpunya et al., 2016). We used the recording reference for analysis here, because SNR was high and common noise not apparent in the data, and most importantly our goal was to compare performance of various algorithms.

Some studies have demonstrated beneficial effects of time reversal on parametric GGC, e.g. (Winkler et al., 2016); these results can be easily extended to the nonparametric implementations considered in the current work using a simulation framework (see the accompanied data article). This framework can also be used to demonstrate that, when using a common reference, nonparametric time-reversed GGC allows obtaining unambiguous detection of the dominant spectral component of the imposed causal influence (strong asymmetry), irrespective of the type of common reference signal considered; although spurious estimates are not completely eliminated outside the frequency band of the dominant interaction.

Interpretational issues for connectivity analyses can also derive from volume conduction in EEG or from field spread in MEG, and generally connectivity analyses on sensor space do not permit any meaningful interpretation in terms of interactions between functionally connected neuronal populations on source space (Brunner et al., 2016; Van de Steen et al., 2016). In the rat model the negative effects associated with volume conduction are less severe compared to traditional human EEG, because even throughout adulthood rat brains remain lissencephalic, i.e. they lack surface convolutions; this aspect combined with the recording setup of the benchmark dataset allowed obtaining a good location-specific characterization of each sensor (Plomp et al., 2014b), matching intracortical recordings (Plomp et al., 2014a; Quairiaux et al., 2011).

One downside for nonparametric methods is that they usually require more computational time with respect to parametric methods. In the benchmark dataset at original sampling rate of 2000 Hz, nonparametric methods required a computational time for each animal between 8 and 11 h,² while for parametric recursive algorithms computational time ranged between less than a minute and 3 h, depending on the algorithm. Our finding that nonparametric GGC is robust against downsampling suggests however that reducing sampling rate may safely cut computational time.

The nonparametric methods here considered allow modeling dynamic nonstationary signals through short-time multitaper or wavelet transforms, which estimate patterns of interactions whose evolution in time is relatively smooth. Similarly, smoothly evolving patterns can be obtained when applying recursive algorithms for parametric time-varying MVAR-modeling (Astolfi et al., 2008; Milde et al., 2010; Möller et al., 2001), which can also provide meaningful results in the whisker-evoked benchmark dataset, as shown in a previous study (Pagnotta and Plomp, 2018). In addition, for parametric connectivity measures a weighting by the spectral power of the source region has been found to enhance their interpretability and performance in the benchmark dataset (Plomp et al., 2014b).

We conclude by addressing a recent claim of pitfall of conditional GGC (Stokes and Purdon, 2017), which was previously observed and appropriately addressed (Barnett and Seth, 2014; Chen et al., 2006). The problem arises from considering model subset for the conditional definition of GGC. The fitting of separate full and reduced MVAR models introduces in fact a bias-variance trade-off in the estimation of GGC, depending on the selected model order. Alternative methods to overcome such problem already exist and the spectral matrix factorization-based implementation such as nonparametric method that we used here is among them (Barnett et al., 2018; Barnett and Seth, 2015, 2014; Dhamala et al., 2018; Faes et al., 2017), as can be demonstrated using our simulation framework (see the accompanied data article).

² All computations were performed using an Intel® Core™ i5-4590 Processor (3.3 GHz).

Acknowledgements

This work was supported by the Swiss National Science Foundation (grant PP00P1_157420 to GP). We thank Dr. Lionel Barnett for valuable comments on the effects of downsampling on Granger causality detectability. The authors declare that this study has been conducted in the absence of any conflict of interest.

Appendix A. Supplementary data

Supplementary data related to this article can be found at <https://doi.org/10.1016/j.neuroimage.2018.07.046>.

References

- Astolfi, L., Cincotti, F., Mattia, D., Fallani, F.D.V., Tocci, A., Colosimo, A., Salinari, S., Marciani, M.G., Hesse, W., Witte, H., others, 2008. Tracking the time-varying cortical connectivity patterns by adaptive multivariate estimators. *IEEE Trans. Biomed. Eng.* 55, 902–913. <https://doi.org/10.1109/TBME.2007.905419>.
- Astolfi, L., Cincotti, F., Mattia, D., Marciani, M.G., Baccala, L.A., de Vico Fallani, F., Salinari, S., Ursino, M., Zavanaglia, M., Ding, L., others, 2007. Comparison of different cortical connectivity estimators for high-resolution EEG recordings. *Hum. Brain Mapp.* 28, 143–157. <https://doi.org/10.1002/hbm.20263>.
- Baccalá, L.A., Sameshima, K., 2001. Partial directed coherence: a new concept in neural structure determination. *Biol. Cybern.* 84, 463–474. <https://doi.org/10.1007/PL00007990>.
- Bajaj, S., Butler, A.J., Drake, D., Dhamala, M., 2015. Functional organization and restoration of the brain motor-execution network after stroke and rehabilitation. *Front. Hum. Neurosci.* 9. <https://doi.org/10.3389/fnhum.2015.00173>.
- Baker, A.P., Brookes, M.J., Rezek, I.A., Smith, S.M., Behrens, T., Probert Smith, P.J., Woolrich, M., 2014. Fast transient networks in spontaneous human brain activity. *eLife* 3. <https://doi.org/10.7554/eLife.01867>.
- Barnett, L., Barrett, A.B., Seth, A.K., 2018. Solved problems for Granger causality in neuroscience: a response to Stokes and Purdon. *Neuroimage*. <https://doi.org/10.1016/j.neuroimage.2018.05.067>.
- Barnett, L., Seth, A.K., 2017. Detectability of Granger causality for subsampled continuous-time neurophysiological processes. *J. Neurosci. Meth.* 275, 93–121. <https://doi.org/10.1016/j.jneumeth.2016.10.016>.
- Barnett, L., Seth, A.K., 2015. Granger causality for state-space models. *Phys. Rev. E* 91. <https://doi.org/10.1103/PhysRevE.91.040101>.
- Barnett, L., Seth, A.K., 2014. The MVGC multivariate Granger causality toolbox: a new approach to Granger-causal inference. *J. Neurosci. Meth.* 223, 50–68. <https://doi.org/10.1016/j.jneumeth.2013.10.018>.
- Bassett, D.S., Sporns, O., 2017. Network neuroscience. *Nat. Neurosci.* 20, 353–364. <https://doi.org/10.1038/nn.4502>.
- Bastos, A.M., Schoffelen, J.-M., 2016. A tutorial review of functional connectivity analysis methods and their interpretational pitfalls. *Front. Syst. Neurosci.* 9. <https://doi.org/10.3389/fnsys.2015.00175>.
- Bastos, A.M., Vezoli, J., Bosman, C.A., Schoffelen, J.-M., Oostenveld, R., Dowdall, J.R., De Weerd, P., Kennedy, H., Fries, P., 2015. Visual areas exert feedforward and feedback influences through distinct frequency channels. *Neuron* 85, 390–401. <https://doi.org/10.1016/j.neuron.2014.12.018>.
- Bernasconi, C., König, P., 1999. On the directionality of cortical interactions studied by structural analysis of electrophysiological recordings. *Biol. Cybern.* 81, 199–210. <https://doi.org/10.1007/s004220050556>.
- Bernasconi, C., Von Stein, A., Chiang, C., König, P., 2000. Bi-directional interactions between visual areas in the awake behaving cat. *Neuroreport* 11, 689–692. <https://doi.org/10.1097/00001756-200003200-00007>.
- Bosman, C.A., Schoffelen, J.-M., Brunet, N., Oostenveld, R., Bastos, A.M., Womelsdorf, T., Rubehn, B., Stieglitz, T., De Weerd, P., Fries, P., 2012. Attentional stimulus selection through selective synchronization between monkey visual areas. *Neuron* 75, 875–888. <https://doi.org/10.1016/j.neuron.2012.06.037>.
- Bressler, S.L., 1995. Large-scale cortical networks and cognition. *Brain Res. Rev.* 20, 288–304. [https://doi.org/10.1016/0165-0173\(94\)00016-1](https://doi.org/10.1016/0165-0173(94)00016-1).
- Bressler, S.L., Tang, W., Sylvester, C.M., Shulman, G.L., Corbetta, M., 2008. Top-down control of human visual cortex by frontal and parietal cortex in anticipatory visual spatial attention. *J. Neurosci.* 28, 10056–10061. <https://doi.org/10.1523/JNEUROSCI.1776-08.2008>.
- Britz, J., Van De Ville, D., Michel, C.M., 2010. BOLD correlates of EEG topography reveal rapid resting-state network dynamics. *Neuroimage* 52, 1162–1170. <https://doi.org/10.1016/j.neuroimage.2010.02.052>.
- Brovelli, A., Ding, M., Ledberg, A., Chen, Y., Nakamura, R., Bressler, S.L., 2004. Beta oscillations in a large-scale sensorimotor cortical network: directional influences revealed by Granger causality. *Proc. Natl. Acad. Sci. U. S. A.* 101, 9849–9854. <https://doi.org/10.1073/pnas.0308538101>.
- Brunner, C., Billinger, M., Seeber, M., Mullen, T.R., Makeig, S., 2016. Volume conduction influences scalp-based connectivity estimates. *Front. Comput. Neurosci.* 10. <https://doi.org/10.3389/fncom.2016.00121>.
- Bullier, J., 2001. Integrated model of visual processing. *Brain Res. Rev., the brain in health and disease - from molecules to man. Swiss National Foundation Symposium NRP 38* (36), 96–107. [https://doi.org/10.1016/S0165-0173\(01\)00085-6](https://doi.org/10.1016/S0165-0173(01)00085-6).

- Cao, Y., Maran, S.K., Dhamala, M., Jaeger, D., Heck, D.H., 2012. Behavior-related pauses in simple-spike activity of mouse purkinje cells are linked to spike rate modulation. *J. Neurosci.* 32, 8678–8685. <https://doi.org/10.1523/JNEUROSCI.4969-11.2012>.
- Cardin, J.A., Carlén, M., Meletis, K., Knoblich, U., Zhang, F., Deisseroth, K., Tsai, L.-H., Moore, C.I., 2009. Driving fast-spiking cells induces gamma rhythm and controls sensory responses. *Nature* 459, 663–667. <https://doi.org/10.1038/nature08002>.
- Chand, G.B., Dhamala, M., 2016. The salience network dynamics in perceptual decision-making. *Neuroimage* 134, 85–93. <https://doi.org/10.1016/j.neuroimage.2016.04.018>.
- Chen, M., Yan, Y., Gong, X., Gilbert, C.D., Liang, H., Li, W., 2014. Incremental integration of global contours through interplay between visual cortical areas. *Neuron* 82, 682–694. <https://doi.org/10.1016/j.neuron.2014.03.023>.
- Chen, Y., Bressler, S.L., Ding, M., 2006. Frequency decomposition of conditional Granger causality and application to multivariate neural field potential data. *J. Neurosci. Meth.* 150, 228–237. <https://doi.org/10.1016/j.jneumeth.2005.06.011>.
- Cohen, J., 1992. A power primer. *Psychol. Bull.* 112, 155–159. <https://doi.org/10.1037/0033-2909.112.1.155>.
- Coito, A., Plomp, G., Genetti, M., Abela, E., Wiest, R., Seeck, M., Michel, C.M., Vuilleumoz, S., 2015. Dynamic directed interictal connectivity in left and right temporal lobe epilepsy. *Epilepsia* 56, 207–217. <https://doi.org/10.1111/epi.12904>.
- Daubechies, I., 1990. The wavelet transform, time-frequency localization and signal analysis. *IEEE Trans. Inf. Theor.* 36, 961–1005. <https://doi.org/10.1109/18.57199>.
- De Moor, T., Munday, S.A., Hood, A.W., 2004. Wavelet Analysis: the effect of varying basic wavelet parameters. *Sol. Phys.* 222, 203–228. <https://doi.org/10.1023/B:SOLA.0000043578.01201.2d>.
- Deco, G., Jirsa, V.K., McIntosh, A.R., 2011. Emerging concepts for the dynamical organization of resting-state activity in the brain. *Nat. Rev. Neurosci.* 12, 43–56. <https://doi.org/10.1038/nrn2961>.
- Dhamala, M., Liang, H., Bressler, S.L., Ding, M., 2018. Granger-Geweke causality: estimation and interpretation. *Neuroimage* 175, 460–463. <https://doi.org/10.1016/j.neuroimage.2018.04.043>.
- Dhamala, M., Rangarajan, G., Ding, M., 2008a. Analyzing information flow in brain networks with nonparametric Granger causality. *Neuroimage* 41, 354–362. <https://doi.org/10.1016/j.neuroimage.2008.02.020>.
- Dhamala, M., Rangarajan, G., Ding, M., 2008b. Estimating granger causality from fourier and wavelet transforms of time series data. *Phys. Rev. Lett.* 100. <https://doi.org/10.1103/PhysRevLett.100.018701>.
- Ding, L., Worrell, G.A., Lagerlund, T.D., He, B., 2007. Ictal source analysis: localization and imaging of causal interactions in humans. *Neuroimage* 34, 575–586. <https://doi.org/10.1016/j.neuroimage.2006.09.042>.
- Ding, M., Bressler, S.L., Yang, W., Liang, H., 2000. Short-window spectral analysis of cortical event-related potentials by adaptive multivariate autoregressive modeling: data preprocessing, model validation, and variability assessment. *Biol. Cybern.* 83, 35–45. <https://doi.org/10.1007/s004229900137>.
- Ding, M., Chen, Y., Bressler, S.L., 2006. 17 Granger causality: basic theory and application to neuroscience. *Handb. Time Ser. Anal. Recent Theor. Dev. Appl.* 437. <https://doi.org/10.1002/9783527609970.ch17>.
- Efron, B., 1987. Better bootstrap confidence intervals. *J. Am. Stat. Assoc.* 82, 171–185. <https://doi.org/10.1080/01621459.1987.10478410>.
- Efron, B., Tibshirani, R.J., 1993. *An Introduction to the Bootstrap*. Springer US, Boston, MA. <https://doi.org/10.1007/978-1-4899-4541-9>.
- Faes, L., Stramaglia, S., Marinazzo, D., 2017. On the interpretability and computational reliability of frequency-domain Granger causality. *F1000Research* 6 (1710). <https://doi.org/10.12688/f1000research.12694.1>.
- Fallani, F.D.V., Astolfi, L., Cincotti, F., Mattia, D., Marciani, M.G., Salinari, S., Kurths, J., Gao, S., Cichocki, A., Colosimo, A., Babiloni, F., 2007. Cortical functional connectivity networks in normal and spinal cord injured patients: evaluation by graph analysis. *Hum. Brain Mapp.* 28, 1334–1346. <https://doi.org/10.1002/hbm.20353>.
- Farge, M., 1992. Wavelet transforms and their applications to turbulence. *Annu. Rev. Fluid Mech.* 24, 395–458. <https://doi.org/10.1146/annurev.fl.24.010192.002143>.
- Felleman, D.J., Van Essen, D.C., 1991. Distributed hierarchical processing in the primate cerebral cortex. *Cerebr. Cortex* 1, 1–47. <https://doi.org/10.1093/cercor/1.1.1>.
- Fontolan, L., Morillon, B., Liegeois-Chauvel, C., Giraud, A.-L., 2014. The contribution of frequency-specific activity to hierarchical information processing in the human auditory cortex. *Nat. Commun.* 5 (4694). <https://doi.org/10.1038/ncomms5694>.
- Gerasimova, E.V., Zakharov, A.V., Lebedeva, Y.A., Inacio, A.R., Minlebaev, M.G., Sitdikova, G.F., Khazipov, R.N., 2014. Gamma oscillations in the somatosensory cortex of newborn rats. *Bull. Exp. Biol. Med.* 156, 295–298. <https://doi.org/10.1007/s10517-014-2333-3>.
- Geweke, J., 1982. Measurement of linear dependence and feedback between multiple time series. *J. Am. Stat. Assoc.* 77, 304–313. <https://doi.org/10.2307/2287238>.
- Geweke, J.F., 1984. Measures of conditional linear dependence and feedback between time series. *J. Am. Stat. Assoc.* 79, 907–915. <https://doi.org/10.2307/2288723>.
- Gow, D.W., Segawa, J.A., Ahlfors, S.P., Lin, F.-H., 2008. Lexical influences on speech perception: a Granger causality analysis of MEG and EEG source estimates. *Neuroimage* 43, 614–623. <https://doi.org/10.1016/j.neuroimage.2008.07.027>.
- Granger, C.W.J., 1969. Investigating causal relations by econometric models and cross-spectral methods. *Econometrica* 37, 424–438. <https://doi.org/10.2307/1912791>.
- Haufe, S., Nikulin, V.V., Müller, K.-R., Nolte, G., 2013. A critical assessment of connectivity measures for EEG data: a simulation study. *Neuroimage* 64, 120–133. <https://doi.org/10.1016/j.neuroimage.2012.09.036>.
- Haufe, S., Nikulin, V.V., Nolte, G., 2012. Alleviating the influence of weak data asymmetries on granger-causal analyses. In: Theis, F., Cichocki, A., Yeredor, A., Zibulevsky, M. (Eds.), *Latent Variable Analysis and Signal Separation*. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 25–33. https://doi.org/10.1007/978-3-642-28551-6_4.
- Horwitz, B., 2003. The elusive concept of brain connectivity. *Neuroimage* 19, 466–470. [https://doi.org/10.1016/S1053-8119\(03\)00112-5](https://doi.org/10.1016/S1053-8119(03)00112-5).
- Kamiński, M., Ding, M., Truccolo, W.A., Bressler, S.L., 2001. Evaluating causal relations in neural systems: granger causality, directed transfer function and statistical assessment of significance. *Biol. Cybern.* 85, 145–157. <https://doi.org/10.1007/s004220000235>.
- Kaminski, M.J., Blinowska, K.J., 1991. A new method of the description of the information flow in the brain structures. *Biol. Cybern.* 65, 203–210. <https://doi.org/10.1007/BF00198091>.
- Kus, R., Kaminski, M., Blinowska, K.J., 2004. Determination of EEG activity propagation: pair-wise versus multichannel estimate. *IEEE Trans. Biomed. Eng.* 51, 1501–1510. <https://doi.org/10.1109/TBME.2004.827929>.
- Liégeois, R., Laumann, T.O., Snyder, A.Z., Zhou, J., Yeo, B.T.T., 2017. Interpreting temporal fluctuations in resting-state functional connectivity MRI. *Neuroimage* 163, 437–455. <https://doi.org/10.1016/j.neuroimage.2017.09.012>.
- Lin, F.-H., Hara, K., Solo, V., Vangel, M., Belliveau, J.W., Stufflebeam, S.M., Hämäläinen, M.S., 2009. Dynamic Granger-Geweke causality modeling with application to interictal spike propagation. *Hum. Brain Mapp.* 30, 1877–1886. <https://doi.org/10.1002/hbm.20772>.
- 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. Comp. Neurol.* 522, 225–259. <https://doi.org/10.1002/cne.23458>.
- Meyers, S.D., Kelly, B.G., O'Brien, J.J., 1993. An introduction to wavelet analysis in oceanography and meteorology: with application to the dispersion of yanai waves. *Mon. Weather Rev.* 121, 2858–2866. [https://doi.org/10.1175/1520-0493\(1993\)121<2858:AITWAI>2.0.CO;2](https://doi.org/10.1175/1520-0493(1993)121<2858:AITWAI>2.0.CO;2).
- Milde, T., Leistriz, L., Astolfi, L., Miltner, W.H., Weiss, T., Babiloni, F., Witte, H., 2010. A new Kalman filter approach for the estimation of high-dimensional time-variant multivariate AR models and its application in analysis of laser-evoked brain potentials. *Neuroimage* 50, 960–969. <https://doi.org/10.1016/j.neuroimage.2009.12.110>.
- Minlebaev, M., Colonnese, M., Tsintsadze, T., Sirota, A., Khazipov, R., 2011. Early gamma oscillations synchronize developing thalamus and cortex. *Science* 334, 226–229. <https://doi.org/10.1126/science.1210574>.
- Mitra, P.P., Pesaran, B., 1999. Analysis of dynamic brain imaging data. *Biophys. J.* 76, 691–708. [https://doi.org/10.1016/S0006-3495\(99\)77236-X](https://doi.org/10.1016/S0006-3495(99)77236-X).
- Möller, E., Schack, B., Arnold, M., Witte, H., 2001. Instantaneous multivariate EEG coherence analysis by means of adaptive high-dimensional autoregressive models. *J. Neurosci. Meth.* 105, 143–158. [https://doi.org/10.1016/S0165-0270\(00\)00350-2](https://doi.org/10.1016/S0165-0270(00)00350-2).
- Morlet, J., Arens, G., Fourgeau, E., Giard, D., 1982. Wave propagation and sampling theory—Part II: sampling theory and complex waves. *Geophysics* 47, 222–236. <https://doi.org/10.1190/1.1441329>.
- Nalatore, H., Ding, M., Rangarajan, G., 2007. Mitigating the effects of measurement noise on Granger causality. *Phys. Rev. E* 75, 031123. <https://doi.org/10.1103/PhysRevE.75.031123>.
- Nalatore, H., Sasikumar, N., Rangarajan, G., 2014. Effect of measurement noise on Granger causality. *Phys. Rev. E* 90, 062127. <https://doi.org/10.1103/PhysRevE.90.062127>.
- Nedungadi, A.G., Rangarajan, G., Jain, N., Ding, M., 2009. Analyzing multiple spike trains with nonparametric granger causality. *J. Comput. Neurosci.* 27, 55–64. <https://doi.org/10.1007/s10827-008-0126-2>.
- Pagnotta, M.F., Plomp, G., 2018. Time-varying MVAR algorithms for directed connectivity analysis: critical comparison in simulations and benchmark EEG data. *PLoS One* 13, e0198846. <https://doi.org/10.1371/journal.pone.0198846>.
- Percival, D.B., Walden, A.T., 2000. *Wavelet Methods for Time Series Analysis*. Cambridge University Press, Cambridge. <https://doi.org/10.1017/CBO9780511841040>.
- Percival, D.B., Walden, A.T., 1993. *Spectral Analysis for Physical Applications: Multitaper and Conventional Univariate Techniques*. Cambridge University Press, Cambridge. <https://doi.org/10.1017/CBO9780511622762>.
- Plomp, G., Quairiaux, C., Kiss, J.Z., Astolfi, L., Michel, C.M., 2014a. Dynamic connectivity among cortical layers in local and large-scale sensory processing. *Eur. J. Neurosci.* 40, 3215–3223. <https://doi.org/10.1111/ejn.12687>.
- Plomp, G., Quairiaux, C., Michel, C.M., Astolfi, L., 2014b. The physiological plausibility of time-varying Granger-causal modeling: normalization and weighting by spectral power. *Neuroimage* 97, 206–216. <https://doi.org/10.1016/j.neuroimage.2014.04.016>.
- Quairiaux, C., Mégevand, P., Kiss, J.Z., Michel, C.M., 2011. Functional development of large-scale sensorimotor cortical networks in the brain. *J. Neurosci.* 31, 9574–9584. <https://doi.org/10.1523/JNEUROSCI.5995-10.2011>.
- Roberts, M.J., Lowet, E., Brunet, N.M., Ter Wal, M., Tiesinga, P., Fries, P., De Weerd, P., 2013. Robust gamma coherence between macaque V1 and V2 by dynamic frequency matching. *Neuron* 78, 523–536. <https://doi.org/10.1016/j.neuron.2013.03.003>.
- Roebroeck, A., Formisano, E., Goebel, R., 2005. Mapping directed influence over the brain using Granger causality and fMRI. *Neuroimage* 25, 230–242. <https://doi.org/10.1016/j.neuroimage.2004.11.017>.
- Saalmann, Y.B., Pinsk, M.A., Wang, L., Li, X., Kastner, S., 2012. The pulvinar regulates information transmission between cortical areas based on attention demands. *Science* 337, 753–756. <https://doi.org/10.1126/science.1223082>.
- Sayed, A.H., Kailath, T., 2001. A survey of spectral factorization methods. *Numer. Lin. Algebra Appl.* 8, 467–496. <https://doi.org/10.1002/nla.250>.
- Schoffelen, J.-M., Gross, J., 2009. Source connectivity analysis with MEG and EEG. *Hum. Brain Mapp.* 30, 1857–1865. <https://doi.org/10.1002/hbm.20745>.

- Seth, A.K., Chorley, P., Barnett, L.C., 2013. Granger causality analysis of fMRI BOLD signals is invariant to hemodynamic convolution but not downsampling. *Neuroimage* 65, 540–555. <https://doi.org/10.1016/j.neuroimage.2012.09.049>.
- Shuler, M.G., Krupa, D.J., Nicolelis, M.A., 2001. Bilateral integration of whisker information in the primary somatosensory cortex of rats. *J. Neurosci.* 21, 5251–5261. <https://doi.org/10.1523/JNEUROSCI.21-14-05251.2001>.
- Slepian, D., 1978. Prolate spheroidal wave functions, fourier analysis, and Uncertainty-V: the discrete case. *Bell Syst. Tech. J.* 57, 1371–1430. <https://doi.org/10.1002/j.1538-7305.1978.tb02104.x>.
- Slepian, D., Pollak, H.O., 1961. Prolate spheroidal wave functions, fourier analysis and uncertainty - I. *Bell Syst. Tech. J.* 40, 43–63. <https://doi.org/10.1002/j.1538-7305.1961.tb03976.x>.
- Sommerlade, L., Thiel, M., Mader, M., Mader, W., Timmer, J., Platt, B., Schelter, B., 2015. Assessing the strength of directed influences among neural signals: an approach to noisy data. *J. Neurosci. Meth.* 239, 47–64. <https://doi.org/10.1016/j.jneumeth.2014.09.007>.
- Sporns, O., 2014. Contributions and challenges for network models in cognitive neuroscience. *Nat. Neurosci.* 17, 652–660. <https://doi.org/10.1038/nn.3690>.
- Stokes, P.A., Purdon, P.L., 2017. A study of problems encountered in Granger causality analysis from a neuroscience perspective. *Proc. Natl. Acad. Sci. Unit. States Am.* 114, E7063–E7072. <https://doi.org/10.1073/pnas.1704663114>.
- Thomson, D.J., 1982. Spectrum estimation and harmonic analysis. *Proc. IEEE* 70, 1055–1096. In: <https://doi.org/10.1109/PROC.1982.12433>.
- Torrence, C., Compo, G.P., 1998. A practical guide to wavelet analysis. *Bull. Am. Meteorol. Soc.* 79, 61–78 [https://doi.org/10.1175/1520-0477\(1998\)079<0061:APGTWA>2.0.CO;2](https://doi.org/10.1175/1520-0477(1998)079<0061:APGTWA>2.0.CO;2).
- Trongnetrpunya, A., Nandi, B., Kang, D., Kocsis, B., Schroeder, C.E., Ding, M., 2016. Assessing granger causality in electrophysiological data: Removing the adverse effects of common signals via bipolar derivations. *Front. Syst. Neurosci.* 9. <https://doi.org/10.3389/fnsys.2015.00189>.
- Van de Steen, F., Faes, L., Karahan, E., Songsiri, J., Valdes-Sosa, P.A., Marinazzo, D., 2016. Critical comments on EEG sensor space dynamical connectivity analysis. *Brain Topogr.* <https://doi.org/10.1007/s10548-016-0538-7>.
- Vidaurre, D., Abeyurriya, R., Becker, R., Quinn, A.J., Alfaro-Almagro, F., Smith, S.M., Woolrich, M.W., 2017. Discovering dynamic brain networks from big data in rest and task. *Neuroimage.* <https://doi.org/10.1016/j.neuroimage.2017.06.077>.
- 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. <https://doi.org/10.1016/j.neuroimage.2014.12.017>.
- Wen, X., Rangarajan, G., Ding, M., 2013a. Multivariate Granger causality: an estimation framework based on factorization of the spectral density matrix. *Phil Trans R Soc A* 371, 20110610. <https://doi.org/10.1098/rsta.2011.0610>.
- Wen, X., Rangarajan, G., Ding, M., 2013b. Is Granger causality a viable technique for analyzing fMRI data? *PLoS One* 8, e67428. <https://doi.org/10.1371/journal.pone.0067428>.
- Wen, X., Yao, L., Liu, Y., Ding, M., 2012. Causal interactions in attention networks predict behavioral performance. *J. Neurosci.* 32, 1284–1292. <https://doi.org/10.1523/JNEUROSCI.2817-11.2012>.
- Wiener, N., 1956. *The theory of prediction.* *Mod. Math. Eng* 1, 125–139.
- Wiener, N., Masani, P., 1957. The prediction theory of multivariate stochastic processes: I. The regularity condition. *Acta Math.* 98, 111–150. <https://doi.org/10.1007/BF02404472>.
- Wilson, G.T., 1972. The factorization of matricial spectral densities. *SIAM J. Appl. Math.* 23, 420–426. <https://doi.org/10.1137/0123044>.
- Winkler, I., Panknin, D., Bartz, D., Müller, K.-R., Haufe, S., 2016. Validity of time reversal for testing Granger causality. *IEEE Trans. Signal Process.* 64, 2746–2760. <https://doi.org/10.1109/TSP.2016.2531628>.
- Zhang, Y., Chen, Y., Bressler, S.L., Ding, M., 2008. Response preparation and inhibition: the role of the cortical sensorimotor beta rhythm. *Neuroscience* 156, 238–246. <https://doi.org/10.1016/j.neuroscience.2008.06.061>.