NeuroImage xxx (2018) 1–5

Contents lists available at ScienceDirect

NeuroImage

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

Solved problems for Granger causality in neuroscience: A response to Stokes and Purdon

Lionel Barnett, Adam B. Barrett, Anil K. Seth *

Sackler Centre for Consciousness Science and Department of Informatics, University of Sussex, Brighton, BN1 9QJ, UK

ARTICLE INFO

Keywords
Granger causality
Functional connectivity
Effective connectivity
Statistical inference

ABSTRACT

Granger-Geweke causality (GGC) is a powerful and popular method for identifying directed functional ('causal') connectivity in neuroscience. In a recent paper, Stokes and Purdon (2017b) raise several concerns about its use. They make two primary claims: (1) that GGC estimates may be severely biased or of high variance, and (2) that GGC fails to reveal the full structural/causal mechanisms of a system. However, these claims rest, respectively, on an incomplete evaluation of the literature, and a misconception about what GGC can be said to measure. Here we explain how existing approaches resolve the first issue, and discuss the frequently-misunderstood distinction between *functional* and *effective* neural connectivity which underlies Stokes and Purdon's second claim.

Granger-Geweke causality (GGC) is a powerful analysis method for inferring directed functional ('causal') connectivity from time-series data, which has become increasingly popular in a variety of neuroimaging contexts (Hesse et al., 2003; Roebroeck et al., 2005; Ding et al., 2006; Dhamala et al., 2008a; Bressler and Seth, 2011; Valdes-Sosa et al., 2011; Barrett et al., 2012; Seth et al., 2015). GGC operationalises a statistical, predictive notion of causality in which causes precede, and help predict their effects. When implemented using autoregressive modelling, GGC can be computed in both time and frequency domains, in both bivariate and multivariate (conditional) formulations. Despite its popularity and power, the use of GGC in neuroscience and neuroimaging has remained controversial. In a recent paper, Stokes and Purdon (2017b) raise two primary concerns: (1) that GGC estimates may be severely biased or of high variance, and (2) that GGC fails to reveal the full structural/causal mechanisms of a system. We explain why these concerns are misplaced.

We note that Stokes and Purdon (2017a) have since responded to critiques of their claims by Barnett et al. (2017),[1] and Faes et al. (2017). Here, we expand on the points made in those articles [see also Dhamala et al. (2018)], and reply in detail to Stokes and Purdon (2017a).

Regarding the first claim, Stokes and Purdon (2017b) describe how bias and variance in GGC estimation arise from the use of separate, independent, full and reduced regressions. While true, this problem has long been recognised (Chen et al., 2006; Barnett and Seth, 2014), and has already been solved by methods which derive GGC from a single full regression.[2] These methods essentially extract reduced model parameters from the full model via factorisation of the spectral density matrix. Well-documented approaches include Wilson's frequency-domain algorithm (Wilson, 1972; Dhamala et al., 2008b, 2018), Whittle's time-domain algorithm (Whittle, 1963; Barnett and Seth, 2014), and a state-space approach which devolves to solution of a discrete-time algebraic Riccati equation (Lancaster and Rodman, 1995; Barnett and Seth, 2015; Solo, 2016). Thus, the source of bias and variance discussed in Stokes and Purdon (2017b) has already been addressed and resolved by previously published methods.

In their reply, Stokes and Purdon (2017a) acknowledge some of this work by saying: "We also described the state space solution to these problems in Dr. Stokes' Ph.D. thesis [Stokes (2015)] in January 2015, but felt it was important to first characterize and describe the problem, before laying out a solution to that problem." It is however worth noting that, at that time, the problem itself was already long-acknowledged (Chen et al., 2006) and, even prior to publication of the state-space method, the distinct and equally effective methods of Dhamala et al. (2008b) and Barnett and Seth (2014) were already in the public domain.

To further illustrate the issue of bias and variance highlighted by Stokes and Purdon (2017a), and its resolution by single-regression methods, in Fig. 1 we plot estimated frequency-domain GGC for the 3-node vector-autoregressive (VAR) model in Stokes and Purdon

* Corresponding author.
E-mail address: a.k.seth@sussex.ac.uk (A.K. Seth).

[1] Barnett et al. (2017) is a preprint of an earlier version of the current article.

[2] But note that the "partition matrix" solution proposed by Chen et al. (2006) is incorrect; see, e.g., Solo (2016).

https://doi.org/10.1016/j.neuroimage.2018.05.067
Received 5 February 2018; Received in revised form 1 May 2018; Accepted 27 May 2018
Available online xxxx

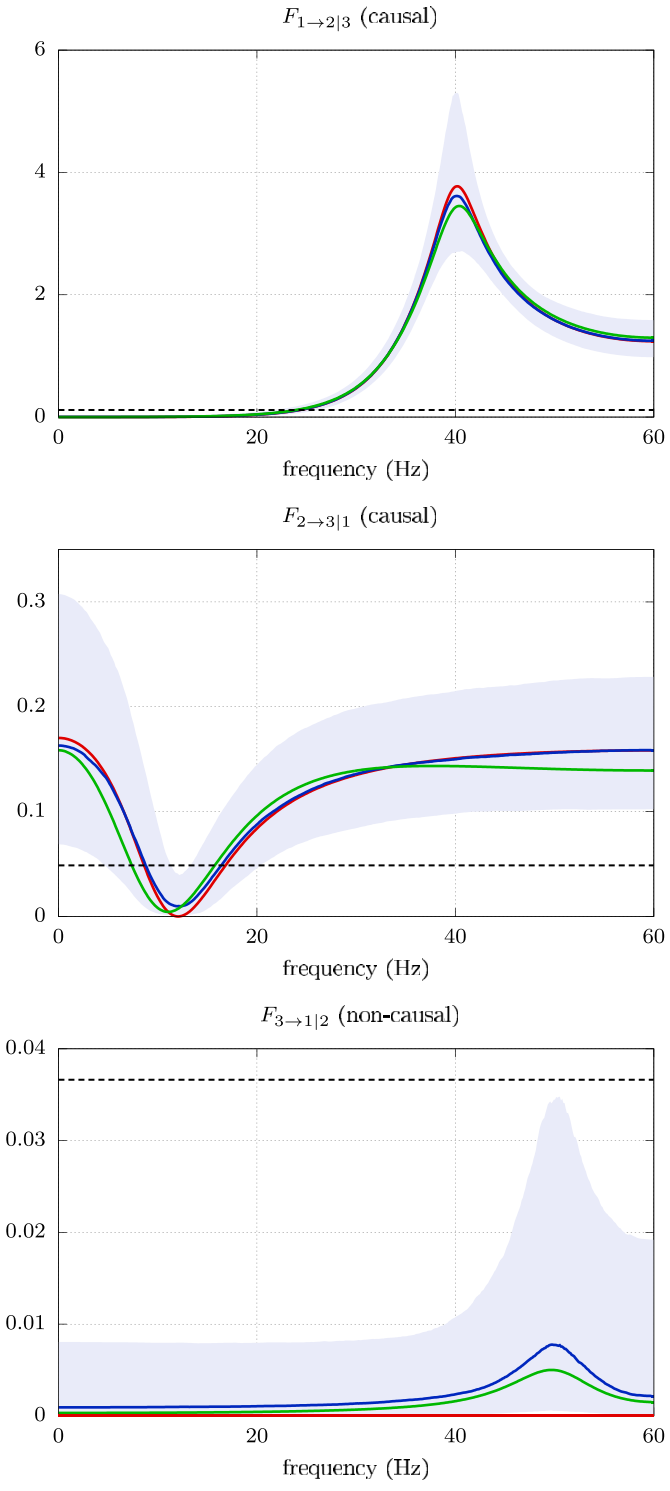

Fig. 1. Granger-Geweke frequency-domain causalities estimated by the single-regression state-space method (Barnett and Seth, 2015; Solo, 2016) for the 3-node VAR model in Stokes and Purdon (2017b), (Example 1, cf. Fig. 2). The true model order of 3 was used for the (single, full-model) VAR estimates. Plots are based on 10,000 time series realisations of 500 observations: red lines plot the exact causality for the model and blue lines sample estimate medians. The shaded areas indicate 90% central confidence intervals, while the green lines plot representative sample estimates. The dashed horizontal lines indicate critical thresholds over all frequencies [see Stokes and Purdon (2017b), Supporting Information, S9] at 95% significance, derived from simulation of the corresponding null model.

(2017b), Example 1, using the single-regression state-space method (Barnett and Seth, 2015; Solo, 2016); see also Faes et al. (2017), Fig. 1 and Dhamala et al. (2018), Fig. 1. We remark that identical results are obtained using the time-domain spectral factorisation method of Barnett and Seth (2014), as implemented in the current (v1.0, 2012) release of the associated MVGC Matlab© software package (Barnett and Seth, 2012). Our Fig. 1 may be directly compared with Fig. 2 in Stokes and Purdon (2017b); we see clearly that all estimates are strictly non-negative, and that exaggerated bias and variance associated with the dual-regression approach are absent. Therefore, Stokes and Purdon (2017b) are in error when they state that "Barnett and Seth […] have proposed fitting the reduced model and using it to directly compute the spectral components …". This is important to note because our MVGC toolbox has been widely adopted within the community, with $> 3,500$ downloads and a significant number of high-impact research publications using the method (e.g., Yellin et al., 2015; Bruneau et al., 2015; Place et al., 2016; Schmitt et al., 2017; Wilber et al., 2017). Thus, we can reassure users of the toolbox that problems of bias and variance as described by Stokes and Purdon (2017b) do not apply.

Sample variance is, of course, still evident, as is bias due to non-negativity of the GGC sample statistic (which may be countered by standard surrogate data methods), but both remain well below their minimum values across all model orders for the dual-regression case (as evidenced by Stokes and Purdon, 2017b, Fig. 2). Fig. 2 further compares bias and variance of time-domain GGC for the example system for single and dual regressions, at model order 3, across a wide-range of time-series lengths. A single regression consistently leads to substantially less bias and variance, except at high time-series lengths where there is a drop-off of bias and variance for both methods.

Stokes and Purdon (2017b) do correctly identify a fundamental cause of the problem with dual-regression GGC estimation: even if the full process is a finite-order autoregression, the reduced process will generally *not* be finite-order autoregressive; rather, it will be vector-autoregressive moving-average (VARMA), or equivalently, a finite-order state-space process (Hannan and Deistler, 2012) – which may be poorly modelled as a finite-order VAR (Barnett and Seth, 2014). The problem is in fact more pervasive than this: the full process *itself* may have a strong moving-average (MA) component and be poorly-modelled as a finite-order VAR. This is because common features of neurophysiological data acquisition, sampling and preprocessing procedures such as subsampling and other temporal aggregation, filtering, measurement noise and sub-process extraction will all, in general, induce an MA component (Barnett and Seth, 2011; Seth et al., 2013; Solo, 2016). This is particularly pertinent to fMRI data, where the haemodynamic response acts as a slow, MA filter. Fortunately, the state-space and non-parametric approaches handle VARMA data parsimoniously, hence avoiding this problem.

The second claim of Stokes and Purdon (2017b) is that GGC fails to reveal the full structural/causal mechanisms of a system. In their reply to our previous commentary, Stokes and Purdon (2017a) ask: "[Barnett et al. (2017)] emphasize that Granger causality reflects a 'directed information flow.' But how does one meaningfully interpret that information flow?" We address that question via a brief recap of the history, definition and interpretation of Granger causality. Wiener and Granger,[3] in their original conception, considered a notion of causality which, in Granger's words (Granger, 2004) comprises two components: (1) The cause occurs before the effect, and (2) The cause contains information about the effect that is unique, and is in no other variable. Granger refined these premises to a statement about dependencies between stochastic processes, essentially as follows: given two jointly-distributed stochastic processes X_t, Y_t in the context of a "universe of information" $\mathscr{U}_t$ (excluding X_t and Y_t) at each time stamp t, then Y does *not* Granger-cause X at time t iff

[3] Granger explicitly attributes the original premise to Norbert Wiener (1956).

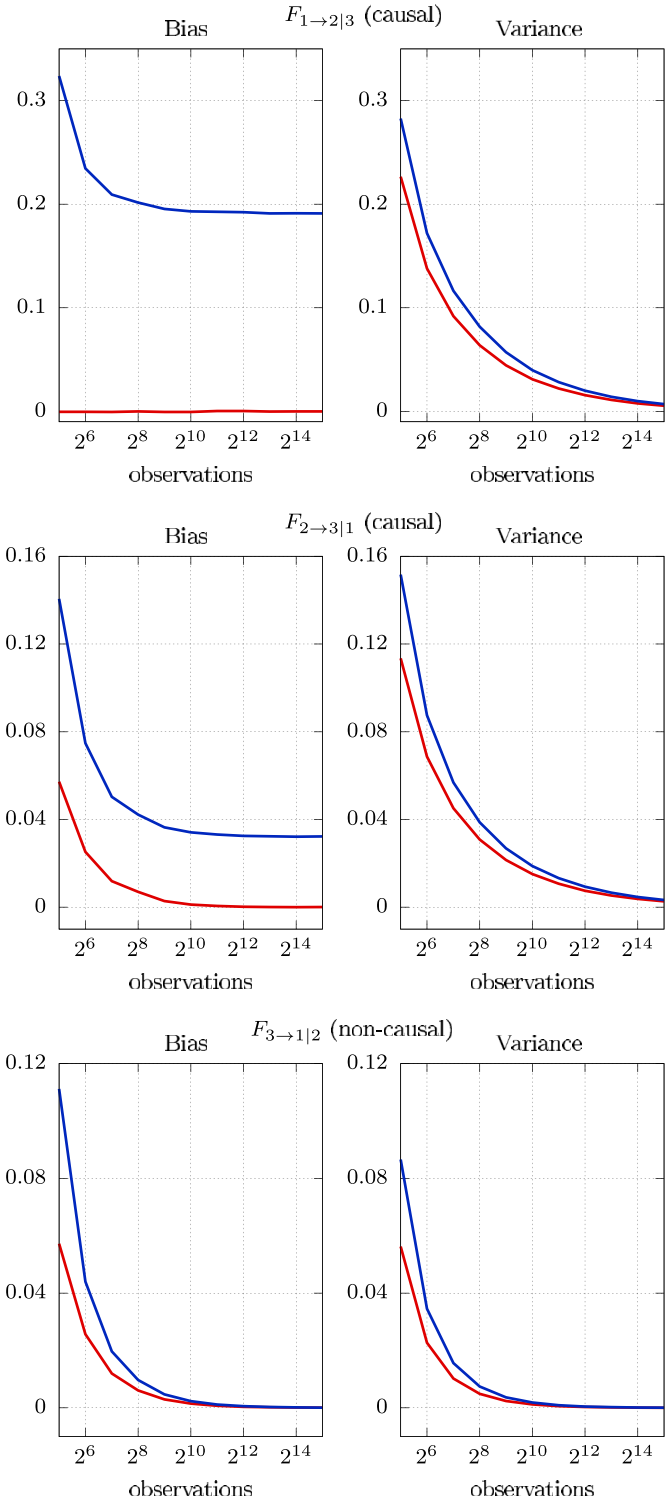

Fig. 2. Granger-Geweke time-domain causality bias (left column) and variance (right column) for estimation by the single-regression state-space method (red lines) and dual-regression method (blue lines), plotted aganst time series length, for the example 3-node VAR model in Stokes and Purdon (2017b). Bias is measured as the difference between the sample median and true causality, while variance is measured as the mean absolute deviation of the sample causality (we use non-parametric measures, as the GGC sample estimators are non-negative, non-Gaussian, and potentially highly skewed). The true model order of 3 was used for all VAR estimates. Plots are based on 10,000 time series realisations for each number of observations.

$$P\left(X_t \middle| X_t^-, Y_t^-, \mathscr{U}_t^-\right) = P\left(X_t \middle| X_t^-, \mathscr{U}_t^-\right) \tag{1}$$

where $P(\cdot|\cdot)$ denotes conditional distribution and superscript '-' denotes history up to (but not including) time t. That is, the distribution of X at time t, contingent on the full historical information set, is unchanged by exclusion of the history of Y. Granger (1963, 1969, 1981) then went on to devise statistical tests for the null hypothesis of non-causality. To do so, he considered (1) for linear VAR processes. The parametric VAR operationalisation was subsequently formalised by Geweke in two seminal papers (Geweke, 1982, 1984), where he defines (Geweke-)Granger causality as a log-likelihood ratio test statistic for the null hypothesis of non-causality, written $F_{Y\to X\,|\,\mathscr{U}}$, and also introduces a spectral decomposition of the statistic.

An explicit information-theoretic expression of Wiener-Granger causality surfaced nearly two decades later in the guise of *transfer entropy* (TE; Schreiber, 2000; Paluš et al., 2001), a form of conditional *mutual information* (MI):

$$T_{Y\to X|\mathscr{U}} \equiv I\left(X_t : Y_t^- \middle| X_t^-, \mathscr{U}_t^-\right) \tag{2}$$

where $I(\cdot : \cdot \mid \cdot)$ denotes conditional MI. Noting that variables A, B are independent conditional on variable C iff $I(A : B \mid C) = 0$, the connection with (1) is immediately clear: Y does not Granger-cause X at time t iff $T_{Y\to X\,|\,\mathscr{U}}$ vanishes at t. In this sense, TE is arguably a "pure-r"—nonparametric—expression of Wiener and Granger's notion than the parametric GGC form. Its interpretation as a metric for information transfer between stochastic variables rests on the appealing intuition of information as reduction of uncertainty. The precise quantitative relationship between GGC and TE was established by Barnett et al. (2009) and Barnett and Bossomaier (2013) which demonstrate, respectively, that if all stochastic processes are jointly Gaussian then there is an exact equivalence between GGC and TE, and that, more generally, if they are jointly Markovian (possibly nonlinear and/or non-Gaussian) then the corresponding Granger-Geweke log-likelihood ratio statistic is asymptotically equivalent to the TE. GGC and TE, in short, are parametric and nonparametric cousins, which instantiate the same formal concept.[4] Interpretation and intuition accordingly transfer between the notions. The relationship between parametric GGC and information-theoretic TE in fact closely mirrors that between *correlation* and MI, which, again, are equivalent for jointly Gaussian variables. Indeed, correlation statistics are widely regarded as standard—and uncontentious—measures of (undirected) functional connectivity.

Stokes and Purdon (2017b) note that GGC reflects a combination of 'transmitter' and 'channel' dynamics, and is independent of 'receiver' dynamics. Again, this independence has been previously identified, as a direct consequence of the invariance of GGC under certain affine transformations (Barrett et al., 2010; Barnett and Seth, 2011). But why should this independence matter? They suggest that it runs "counter to intuitive notions of causality intended to explain observed effects" since, according to them, "neuroscientists seek to determine the mechanisms that produce 'effects' within a neural system or circuit as a function of inputs or 'causes' observed at other locations". In fact, this view resonates more strongly with approaches such as Dynamic Causal Modelling (DCM; Friston et al., 2003)—usually characterised as *effective connectivity*—which attempt to find the optimal mechanistic (circuit-level) description that explains observed data. GGC, on the other hand, models statistical dependencies among observed responses and is therefore an example of (directed) *functional connectivity* (see Seth et al., 2015; Friston et al.,

[4] We concur with Stokes and Purdon (2017a) that other related, but non-equivalent, measures have been misleadingly conflated with Granger causality, in particular the "directed transfer function" (DTF; Kaminski and Blinowska, 1991) and "partial directed coherence" (PDC; Baccala and Sameshima, 2001). While these are valid (spectral) measures in their own right, they cannot be said to explicitly reflect the Wiener-Granger notion of causality.

2013, for in-depth comparison). Essentially, the distinction is between making inferences about an underlying *physical causal mechanism* (DCM; Valdes-Sosa et al., 2011) and—as explained above—making inferences about *directed information flow* (GGC; Barnett et al., 2009). DCM is able to deliver evidence for circuit-level descriptions of neural mechanism from a limited repertoire of tightly-framed hypotheses, which must be independently motivated and validated (Stephan et al., 2010); it is, in particular, unsuited to *exploratory* analyses. GGC inference, on the other hand, is data-driven and "data-agnostic" (it makes few assumptions about the generative process, beyond that it be reasonably parsimoniously modelled as a linear stochastic system), and as such is well-suited to exploratory analyses. It delivers an information-theoretic interpretation of the neural process which is both amenable to statistical inference, and which also stands as an *effect size* for directed information flow between components of the system (Barrett and Barnett, 2013). Our view is that both approaches address valid questions of interest for neuroscientific analyses, and indeed, that this is reflected in the burgeoning literature in both effective and functional connectivity analysis.

Stokes and Purdon (2017a), in their reply, go on to say: "While GG-causality is decipherable in reference to the selected model and its component dynamics, it is not understandable without these details." This statement appears to be based on an (unfortunately common) misunderstanding. The VAR (or, more recently, state-space) models that underlie the most common GGC inference methods fulfil an entirely different function from the circuit-level models that underpin, e.g., DCM. They are, in a sense "generic" [this statement can be made more precise; see, e.g., Geweke (1982)] and do not pretend to represent physical mechanism. In the sense of the (asymptotic) equivalence with TE, GGC may be considered an approximation to the nonparametric TE, and the underlying VAR model a mathematical construct to operationalise this approximation, in the same way that the linear regression model underlying a parametric correlation statistic might be deployed to approximate the information-theoretic MI. Indeed, the power of the information-theoretic approach is precisely that it furnishes intuitive, *model-free* accounts of dynamical processes.

Concluding, the primary claims in Stokes and Purdon (2017b) are invalid. Currently available implementations deal appropriately with issues of bias and variance associated with dual-regression methods, and invariance to receiver dynamics does not undermine GGC's ability to characterize information flow. Altogether, when used with appropriate care, GGC represents a conceptually satisfying and statistically powerful method for directed functional connectivity analysis in neuroscience and neuroimaging. However, a range of additional challenges remain in further developing this useful technique. These include issues of stationarity, linearity and exogenous influences, as noted by Stokes and Purdon (2017b), and in addition the influences of noise, sampling rates and temporal/spatial aggregation engendered by neural data acquisition (Solo, 2016; Barnett and Seth, 2017).

Acknowledgements

ABB is funded by EPSRC grant EP/L005131/1. All authors are grateful to the Dr. Mortimer and Theresa Sackler Foundation, which supports the Sackler Centre for Consciousness Science.

References

Baccala, L.A., Sameshima, K., 2001. Partial directed coherence: a new conception in neural structure determination. Biol. Cybern. 84, 463–474.
Barnett, L., Barrett, A.B., Seth, A.K., 2009. Granger causality and transfer entropy are equivalent for Gaussian variables. Phys. Rev. Lett. 103, 0238701.
Barnett, L., Barrett, A.B., Seth, A.K., 2017. Solved Problems and Remaining Challenges for Granger Causality Analysis in Neuroscience: a Response to Stokes and Purdon. ArXiv e-prints, 1708.08001.
Barnett, L., Bossomaier, T., 2013. Transfer entropy as a log-likelihood ratio. Phys. Rev. Lett. 109, 0138105.
Barnett, L., Seth, A.K., 2011. Behaviour of Granger causality under filtering: theoretical invariance and practical application. J. Neurosci. Meth. 201, 404–419.
Barnett, L., Seth, A.K., 2012. The MVGC multivariate granger causality Matlab© toolbox. http://users.sussex.ac.uk/~lionelb/MVGC/.
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.
Barnett, L., Seth, A.K., 2015. Granger causality for state-space models. Phys. Rev. 91, 040101(R).
Barnett, L., Seth, A.K., 2017. Detectability of Granger causality for subsampled continuous-time neurophysiological processes. J. Neurosci. Meth. 275, 93–121.
Barrett, A.B., Barnett, L., 2013. Granger causality is designed to measure effect, not mechanism. Front. Neuroinf. 7, 6.
Barrett, A.B., Barnett, L., Seth, A.K., 2010. Multivariate Granger causality and generalized variance. Phys. Rev. E 81, 041907.
Barrett, A.B., Murphy, M., Bruno, M.A., Noirhomme, Q., Boly, M., Laureys, S., Seth, A.K., 2012. Granger causality analysis of steady-state electroencephalographic signals during propofol-induced anaesthesia. PLoS One 7, 1–12.
Bressler, S.L., Seth, A.K., 2011. Wiener-Granger causality: a well established methodology. Neuroimage 58, 323–329.
Bruneau, E.G., Jacoby, N., Saxe, R., 2015. Empathic control through coordinated interaction of amygdala, theory of mind and extended pain matrix brain regions. Neuroimage 114, 105–119.
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.
Dhamala, M., Liang, H., Bressler, S.L., Ding, M., 2018. Granger-Geweke causality: estimation and interpretation. Neuroimage (in press).
Dhamala, M., Rangarajan, G., Ding, M., 2008a Jun. Analyzing information flow in brain networks with nonparametric Granger causality. Neuroimage 41 (2), 354–362.
Dhamala, M., Rangarajan, G., Ding, M., 2008b. Estimating Granger causality from Fourier and wavelet transforms of time series data. Phys. Rev. Lett. 100, 018701.
Ding, M., Chen, Y., Bressler, S.L., 2006. Granger causality: basic theory and application to neuroscience. In: Schelter, B., Winterhalder, M., Timmer, J. (Eds.), Handbook of Time Series Analysis: Recent Theoretical Developments and Applications. Wiley-VCH Verlag GmbH & Co. KGaA, pp. 437–460.
Faes, L., Stramaglia, S., Marinazzo, D., 2017. On the Interpretability and Computational Reliability of Frequency-domain Granger Causality. F1000Research 6. Version 1; Referees: 2 approved.
Friston, K., Moran, R., Seth, A.K., 2013. Analyzing connectivity with granger causality and dynamic causal modelling. Curr. Opin. Neurobiol. 23, 172–178.
Friston, K.J., Harrison, L., Penny, W., 2003. Dynamic causal modelling. Neuroimage 19, 1273–1302.
Geweke, J., 1982. Measurement of linear dependence and feedback between multiple time series. J. Am. Stat. Assoc. 77, 304–313.
Geweke, J., 1984. Measures of conditional linear dependence and feedback between time series. J. Am. Stat. Assoc. 79, 907–915.
Granger, C.W.J., 1963. Economic processes involving feedback. Inf. Control 6, 28–48.
Granger, C.W.J., 1969. Investigating causal relations by econometric models and cross-spectral methods. Econometrica 37, 424–438.
Granger, C.W.J., 1981. Some properties of time series data and their use in econometric model specification. J. Econom. 16, 121–130.
Granger, C.W.J., 2004. Time series analysis, cointegration, and applications. Am. Econ. Rev. 94, 421–425.
Hannan, E.J., Deistler, M., 2012. The Statistical Theory of Linear Systems. SIAM, Philadelphia, PA, USA.
Hesse, W., Möller, E., Arnold, M., Schack, B., 2003. The use of time-variant EEG Granger causality for inspecting directed interdependencies of neural assemblies. J. Neurosci. Meth. 124, 27–44.
Kaminski, M., Blinowska, K.J., 1991. A new method of the description of the information flow in brain structures. Biol. Cybern. 65, 203–210.
Lancaster, P., Rodman, L., 1995. Algebraic Riccati Equations. Oxford University Press, Oxford, UK.
Paluš, M., Komárek, V., Hrnčíř, Z., Štěrbová, K., 2001. Synchronization as adjustment of information rates: detection from bivariate time series. Phys. Rev. E 63, 046211.
Place, R., Farovik, A., Brockmann, M., Eichenbaum, H., 2016. Bidirectional prefrontal-hippocampal interactions support context-guided memory. Nat. Neurosci. 19, 992–994.
Roebroeck, A., Formisano, E., Goebel, R., 2005. Mapping directed influence over the brain using Granger causality and fMRI. Neuroimage 25, 230–242.
Schmitt, L.I., Wimmer, R.D., Nakajima, M., Happ, M., Mofakham, S., Halassa, M.M., 2017. Thalamic amplification of cortical connectivity sustains attentional control. Nature 545, 219–223.
Schreiber, T., 2000. Measuring information transfer. Phys. Rev. Lett. 85, 461–464.
Seth, A.K., Barrett, A.B., Barnett, L.C., 2015. Granger causality analysis in neuroscience and neuroimaging. J. Neurosci. 35, 3293–3297.
Seth, A.K., Chorley, P., Barnett, L., 2013. Granger causality analysis of fMRI BOLD signals is invariant to hemodynamic convolution but not downsampling. Neuroimage 65, 540–555.
Solo, V., 2016. State-space analysis of Granger-Geweke causality measures with application to fMRI. Neural Comput. 28, 914–949.
Stephan, K.E., Penny, W.D., Moran, R.J., den Ouden, H.E.M., Daunizeau, J., Friston, K.J., 2010. Ten simple rules for dynamic causal modeling. Neuroimage 49, 3099–3109.
Stokes, P.A., 2015. Fundamental Problems in Granger Causality Analysis of Neuroscience Data. Ph.D. thesis. Massachusetts Institute of Technology.
Stokes, P.A., Purdon, P.L., 2017a. In reply to Faes, et al. and Barnett, et al. regarding "A study of problems encountered in Granger causality analysis from a neuroscience perspective". ArXiv e-prints 1709.10248.

Stokes, P.A., Purdon, P.L., 2017b. A study of problems encountered in Granger causality analysis from a neuroscience perspective. Proc. Natl. Acad. Sci. U.S.A. 114, 7063–7072.

Valdes-Sosa, P.A., Roebroeck, A., Daunizeau, J., Friston, K., 2011. Effective connectivity: influence, causality and biophysical modeling. Neuroimage 58, 339–361.

Whittle, P., 1963. On the fitting of multivariate autoregressions, and the approximate canonical factorization of a spectral density matrix. Biometrika 50, 129–134.

Wiener, N., 1956. The theory of prediction. In: Beckenbach, E.F. (Ed.), Modern Mathematics for Engineers. McGraw Hill, New York, pp. 165–190.

Wilber, A.A., Skelin, I., Wei, W., McNaughton, B.L., 2017. Organization of encoding and memory reactivation in the parietal cortex. Neuron 95, 1406–1419 e5.

Wilson, G.T., 1972. The factorization of matricial spectral densities. SIAM J. Appl. Math. 23, 420–426.

Yellin, D., Berkovich-Ohana, A., Malach, R., 2015. Coupling between pupil fluctuations and resting-state fMRI uncovers a slow build-up of antagonistic responses in the human cortex. Neuroimage 106, 414–427.