Directed neural interactions in fMRI: a comparison between Granger Causality and Effective Connectivity

Michele Allegra[1, 2], Matthieu Gilson[3, *], Andrea Brovelli[3, *]

1 Dipartimento di Fisica "Galileo Galilei", University of Padova, Padova, Italy
2 Padova Neuroscience Center, University of Padova, Padova, Italy
3 Institut de Neurosciences de la Timone, UMR 7289, CNRS and Aix-Marseille University, Marseille, France

Corresponding author: michele.allegra@unipd.it
* equal contribution

Abstract

Understanding how neural populations interact is crucial to understand brain function. Most common approaches to infer neural interactions are based on Granger causality (GC) analyses and effective connectivity (EC) models of neural time series. However, an in-depth investigation of the similarity and complementarity of these approaches is currently lacking. GC and EC are classically thought to provide complementary information about the interdependence between neural signals. Whereas GC quantifies the amount of predictability between time series and it is interpreted as a measure of information flow, EC quantifies the amount and sign of the interaction, and it is often interpreted as the causal influence that a neural unit exert over another. Here, we show that, in the context of functional magnetic resonance imaging (fMRI) data analysis and first-order autoregressive models, GC and EC share common assumptions and are mathematically related. More precisely, by defining a 'corrected' version of GC accounting for unequal noise variances affecting the source and target node, we show that the two measures are linked by an approximately quadratic relation, where positive or negative values of EC are associated with identical values of GC. While the relation is obtained in limit of infinite sampling time, we use simulations to show that it can be observed in finite data samples as classically observed in neuroimaging studies, provided sufficiently long sampling, multiple sessions or group averaging. Finally, we compare the GC and EC analyses on fMRI data from the Human Connectome Project, and obtain results consistent with simulation outcomes. While GC and EC analyses do not provide reliable estimates at the single subject or single connection level, they become stable at the group level (more than approximately 20 subjects), where the predicted relation between GC and EC can be clearly observed from the data. To conclude, our study provides a common mathematical framework to make grounded methodological choices in the reconstruction and analysis of directed brain networks from neuroimaging time series.

Introduction

The study of large-scale brain networks [1, 2] strongly relies on the analysis of functional magnetic resonance imaging (fMRI) data. The standard approach for reconstructing brain networks from ongoing fMRI activity [3], especially spontaneous activity at rest (rs-fMRI), is based on *undirected functional connectivity* (UFC), often defined as correlation

of the blood-oxygen-level-dependent (BOLD) signals of two brain areas. Despite its simplicity, UFC has yielded an extremely rich insight into the large-scale organization of spontaneous activity in health and pathology, the relations with the underlying anatomy, and neural signatures of individual behavioral traits [4–10].

While UFC-based methods are statistically reliable and give reproducible findings [11], they do not provide information about the directionality or asymmetry of interaction between brain areas (i.e., UFC is a symmetric measure) and they are sensitive to third-party effects. These limitations have prompted the development of methods that make use of the temporal structure of BOLD signals to characterize *direct* (non-mediated) and *directed* connections between brain areas [12]. Going further, it is possible to embed anatomical priors in the estimation of connections, constraining the topology of network models to be more biologically realistic than a naive fully connected one [13, 14]. All such methods assume, implicitly or explicitly, a generative model for the dynamics underlying the observed signals. Thus, each method can be ideally placed within a continuum line going from the weakest to the strongest model assumptions.

A first approach for the inference of asymmetric interactions between neural time series is based on the Wiener-Granger principle [15–17], which identifies directed relations from the ability to determine whether one time series is useful in forecasting another [18]. While information theoretic measures based on the Wiener-Granger principle, such as Transfer Entropy [19] and Directed Information [20], provide purely data-driven tools capturing directed interactions, most *Granger causality* (GC) methods used in resting-state fMRI [21–25] assume a linear multivariate autoregressive (MAR) model as the generative process of BOLD time series. In the Granger-Geweke formalism, the total interdependence between two time series is split into two components, the Granger causality proper (GC) and the so-called *instantaneous causality* (IC), a non-directional measure that captures instantaneous interactions (i.e., occurring faster than the sampling time).

A second approach is based on effective connectivity (EC) models [26], which assume that the data are generated by a continuous dynamical system. EC weights between pairs of brain regions quantify the strength and sign of directional interactions and they can be inferred by fitting the model to reproduce empirical fMRI time series. Most EC approaches for whole-brain fMRI are based on similar models of neural activity: linear systems of stochastic differential equations. A popular class of EC models are dynamical causal models (DCMs) [27]. Due to computational complexity of model inversion, EC models are rarely applied to large brain networks. However, recent work has shown that with appropriate simplifications, EC models can be applied to multivariate time series including hundreds of brain regions. The first approach consists in linearizing the dynamics of the DCM, including approximating the hemodynamic convolution with a linear filter), thus allowing the analysis of networks with $n \gtrsim 50$ areas and $n^2 \gtrsim 2500$ links [27–30]. A second approach is based on the Lyapunov optimization for multivariate Ornstein-Uhlenbeck processes (MOU) with linear dynamics [13, 30, 31]. DCM and MOU mainly differ in terms of inference methodology (e.g., Lyapunov optimization vs. variational Bayes), rather than the generative model underlying neural time series.

GC and EC approaches for fMRI analysis are not dichotomous. Indeed, EC and GC for resting-state fMRI (rs-fMRI) activity (as compared to evoked activity) often assume *very similar* models of fMRI activity and they both attribute a large relevance to second-order statistics in the data (i.e., cross-correlations without and with lag among time series, or its equivalent in the frequency domain) to infer directed relations. Whereas GC assumes a linear multivariate autoregressive (MAR) model, EC is generally based on a system of linear ordinary differential equations (ODEs) in continuous time. If the noise is assumed to be Gaussian, a formal link exists between ODE systems and MAR processes (via the integration of ODEs), so GC and EC are actually based on mutually

consistent modeling despite seemingly different premises and formulation.

Given the widespread use of GC and EC models in network neuroscience, an open question is whether these approaches are comparable, and whether they provide similar or complementary information about brain interactions. In this work, we systematically address this question, focusing on whether GC and EC yield consistent interpretations in terms of 1) *topological structure*: presence of a directional and causal interaction between two areas, and 2) *asymmetry*: presence of a stronger connection in one direction as compared to the opposite direction. For both GC and EC, we focus on estimation schemes that extract the spatio-temporal relationships between all signals in the network. GC is estimated using the covariance-based method [32–34], where the estimation of MAR model parameters is bypassed, and GC is directly inferred from observed cross-covariance matrices. EC is estimated with the Lyapunov optimization of the MOU, also based on lagged cross-covariances [13].

In Methods, we discuss an analytical relation between EC and GC, which relies on the assumption that the time series are generated by linear network dynamics with Gaussian-like inputs (MOU model). We first analytically derive quadratic relations between EC and GC measures, and identify for which conditions EC is best captured by instantaneous Granger causality (IC). In Results, we discuss how these relations emerge in finite-length data affected by (large) sampling noise. We present results from numerical simulations to investigate to what extent the two methodologies converge in estimating the *topology* and *asymmetry* of connectivity estimates, both theoretically, and in the case of finite-sampling data. We then test our predictions on rs-fMRI data from the Human Connectome Project [35, 36] involving short time series of about $500 - 2000$ frames. Finally, we discuss to what extent our theoretical predictions are met in the data, and which degree of consistency between the two methods is obtained, at the individual and group level.

Models and analytical derivations

Multivariate autoregressive (MAR) models and Granger Causality

A 1st order N-dimensional multivariate autoregressive (MAR) process is defined for discrete time t (with steps Δ) by

$$\mathbf{x}(t+\Delta) = A\mathbf{x}(t) + \boldsymbol{\epsilon}(t) \tag{1}$$

where $\boldsymbol{\epsilon}(t)$ is Gaussian noise with zero mean and covariance matrix S,

$$\boldsymbol{\epsilon}(t) \sim \mathcal{N}(0, S)$$

and $||A|| \leq 1$ for stability. The innovation, or noisy input, $\epsilon_j(t)$ in Eq. (1) corresponds to the residual of the linear regression

$$\epsilon_j(t) = x_j(t+\Delta) - \sum_{ji} A_{ji} x_i(t)$$

The variance of $\epsilon_j(t)$ (corresponding to a diagonal element of matrix S) reflects the "magnitude" of the residual, which measures how well past values of the time series can predict the next value at j, $x_j(t+\Delta)$, namely

$$S_{jj} = Var[\epsilon_j] = Var[x_j(t+\Delta) - \sum_{ji} A_{ji} x_i(t)]$$

.

To quantify the Granger causal effect of node i on node j, one can measure the relevance

of $x_i(t)$ in predicting $x_j(t+\Delta)$. To this aim, one defines a a "reduced" MAR process where the influence of node i is removed:

$$\mathbf{x}(t+\Delta) = A'\mathbf{x}_{\neg i}(t) + \boldsymbol{\epsilon}'(t) \tag{2}$$

where $\mathbf{x}_{\neg i}(t)$ is obtained from $\mathbf{x}(t)$ by removing its i-th component and $\boldsymbol{\epsilon}'(t)$ is Gaussian with zero mean and covariance matrix S',

$$\boldsymbol{\epsilon}'(t) \sim \mathcal{N}(0, S')$$

Again, the magnitude of the residual $\epsilon'_j(t)$ can be assessed via its variance

$$S'_{jj} = Var[\epsilon'_j] = Var[x_j(t+\Delta) - \sum_{ji} A'_{ji} x_i(t)]$$

The effect of node i on node j defined by *Granger causality* (GC) [21, 37] is given by the log-ratio of the variances

$$G_{ij} = \log\left(\frac{S'_{jj}}{S_{jj}}\right) \tag{3}$$

Note that we consider the conditional version of GC, meaning that the linear regression in Eq. (2) includes all remaining nodes in the network. Also note that Eq. (3) defines a "model" GC, which differs from its *estimates* obtained from finite data and noisy covariance matrices (see below). We can show an approximate relation between the Granger Causality G_{ij} and the MAR coefficient a_{ji}. Assuming that $A' \simeq A^{\neg i}$ (where $A^{\neg i}$ is simply A without the i-th column), one obtains

$$\begin{aligned} S'_{jj} &\simeq Var[x_j(t+\Delta) - \sum_{k\neq i} A_{jk} x_k(t)] = Var[A_{ji} x_i(t) + \epsilon_j] \\ &= A_{ji}^2 Var[x_i(t)] + S_{jj} = A_{ji}^2 Q_{ii}^0 + S_{jj} \end{aligned}$$

using the further assumption of the statistical independence between $x_i(t)$ and ϵ_j. Therefore we have

$$G_{ij} \simeq \log\left(\frac{A_{ji}^2 Q_{ii}^0 + S_{jj}}{S_{jj}}\right) = \log\left(1 + \frac{A_{ji}^2 Q_{ii}^0}{S_{jj}}\right) \tag{4}$$

hence G_{ij} is approximately a monotonic function of A_{ji}^2. For sufficiently small $A_{ji} \ll S_{jj}/Q_{ii}^0$, G_{ij} is approximately a quadratic function of A_{ji}^2.

In addition to the standard Granger causality (3), Geweke [37] defined the "instantaneous" Granger causality I_{ij}, which compares the magnitude of the innovations when considered jointly or separately. The innovations $\{\epsilon_i(t), \epsilon_j(t)\}$ jointly have a covariance matrix

$$S^{[ij]} = \begin{pmatrix} S_{ii} & S_{ij} \\ S_{ij} & S_{jj} \end{pmatrix}$$

The total magnitude of the joint innovations can be measured as $\log(\det S^{[ij]})$. If one considers the innovations independently, one obtains instead $\log S_{ii} + \log S_{jj}$. The instantaneous Granger causality is defined as

$$\begin{aligned} I_{ij} &= \log(S_{ii}) + \log(S_{jj}) - \log(\det S^{[ij]}) = \log(S_{ii}S_{jj}) - \log(S_{ii}S_{jj} - S_{ij}^2) \\ &= -\log\left(\frac{S_{ii}S_{jj} - S_{ij}^2}{S_{ii}S_{jj}}\right) = -\log\left(1 - \frac{S_{ij}^2}{S_{ii}S_{jj}}\right) \end{aligned}$$

While the instantaneous causality is often discarded in Granger causality analyses, it may capture a large part of the interdependence between two time series.

Covariance-based GC

A standard method to estimate G_{ij} from the data is by inferring the parameters of the MAR (1) and the reduced MAR (2). An alternative and computationally simple approach exploits a relation between the variance of the MAR residuals and covariance terms, assuming Gaussian innovations [32, 38, 39]. Indeed, it can be shown that

$$Var[\epsilon_j] \ = Var[x_j(t+\Delta)|\mathbf{x}(t)] = \frac{\det(Cov[x_j(t+\Delta), \mathbf{x}(t)])}{\det(Cov[\mathbf{x}(t)])} \tag{5}$$

$$Var[\epsilon'_j] \ = Var[x_j(t+\Delta)|\mathbf{x}_{\neg i}(t)] = \frac{\det(Cov[x_j(t+\Delta), \mathbf{x}_{\neg i}(t)])}{\det(Cov[\mathbf{x}_{\neg i}(t)])} \tag{6}$$

Furthermore, we have:

$$\det S^{[ij]} = \det(Cov[x_i(t+\Delta)x_j(t+\Delta)|\mathbf{x}(t)]) = \frac{\det(Cov[x_i(t+\Delta), x_j(t+\Delta), \mathbf{x}(t)])}{\det(Cov[\mathbf{x}(t)])}$$

$$\det S_{ii} = \det(Cov[x_i(t+\Delta)|\mathbf{x}(t)]) = \frac{\det(Cov[x_i(t+\Delta), \mathbf{x}(t)])}{\det(Cov[\mathbf{x}(t)])}$$

$$\det S_{jj} = \det(Cov[x_j(t+\Delta)|\mathbf{x}(t)]) = \frac{\det(Cov[x_j(t+\Delta), \mathbf{x}(t)])}{\det(Cov[\mathbf{x}(t)])}$$

Thus, exploting covariance relations, we can express G_{ij} and I_{ij} as follows:

$$G_{ij} = \log \frac{\det(Cov[\mathbf{x}(t)]) \det(Cov[x_j(t+\Delta), \mathbf{x}_{\neg i}(t)])}{\det(Cov[x_j(t+\Delta), \mathbf{x}(t)]) \det(Cov[\mathbf{x}_{\neg i}(t)])} \tag{7}$$

$$I_{ij} = \log \ \frac{\det(Cov[x_i(t+\Delta), \mathbf{x}(t)]) \det(Cov[x_j(t+\Delta), \mathbf{x}(t)])}{\det(Cov[\mathbf{x}(t)]) \det(Cov[x_i(t+\Delta)x_j(t+\Delta)|\mathbf{x}(t)])} \tag{8}$$

Thus, G_{ij} and I_{ij} can be both expressed in terms of elements of the 0-lagged and the Δ-lagged covariance matrices

$$Q^0 = E[\mathbf{x}(t)\mathbf{x}^T(t)], \qquad Q^\Delta = E[\mathbf{x}(t)\mathbf{x}^T(t+\Delta)]$$

Note that for Gaussian systems, there is a relation between the entropy and the covariance matrix,

$$H(x) = -\log(\det(Cov(x))$$

Using this relation, one can show that the covariance-based GC is equivalent to the transfer entropy [40].Granger causality measures can therefore be formulated in completely information-theoretical terms, based on entropy estimates [32, 34].
Estimates $\hat{G}_{ij}$ and $\hat{I}_{ij}$ can be obtained from finite-sampling estimates of the covariance matrix,

$$\hat{Q}^0_{ij} = \frac{1}{L}\sum_{t=1}^{L}(x_i(t)-\mu_i)(x_j(t)-\mu_j), \qquad \mu_i = \frac{1}{L}\sum_{t=1}^{L} x_i(t)$$

$$\hat{Q}^1_{ij} = \frac{1}{L}\sum_{t=1}^{L-1}(x_i(t)-\mu_i)(x_j(t+1)-\mu_j)$$

Before computing covariance-matrices, we applied the cop-norm transformation on the signal of each region, $x \to F_{\mathcal{N}}^{-1}(F_{emp}(x))$ where $F_{emp}(x)$ is the empical cumulative distribution function of x, and $F_{\mathcal{N}}(x)$ is the cumulative distribution function of a standard normal. This transformation corrects for possible non-Gaussianity of the signal distributions, assuming a Gaussian copula [41].

Multivariate Ornstein-Uhlenbeck (MOU) process and relation to MAR

Another approach to characterize the causal effect of node i on node j is to assume a generative dynamical model in continuous time, and assess the value of the i, j coupling that can be seen as *model effective connectivity*. Gilson et al. rely on the multivariate Ornstein-Uhlenbeck process (MOU) given by

$$d\mathbf{x} = J\mathbf{x}dt + dw \tag{9}$$

where

$$J = -\frac{1}{\tau}\mathbb{I} + C^T \tag{10}$$

with the process time constant $\tau > 0$ and the identity matrix $\mathbb{I}$. Here w is a Wiener process (akin to white noise) corresponding to a diagonal covariance matrix Σ with input variances σ_i^2 and zero input cross-covariance, such that $\int_t^{t+dt} Cov(dw) = \Sigma dt$. At equilibrium, the covariance matrix Q^0 for $\mathbf{x}(t)$ following Eq. (9) satisfies the (continuous) Lyapunov equation

$$JQ^0 + Q^0 J^T + \Sigma = 0 \tag{11}$$

Furthermore, the 0-lagged and Δ-lagged covariance matrix Q^0 and Q^Δ obey the relation

$$Q^\Delta = Q^0 e^{J^T \Delta}$$

with Δ a given time lag. Therefore, one can use the matrix logarithm to obtain

$$J = \frac{1}{\Delta}\log((Q^0)^{-1}Q^\Delta)^T \tag{12}$$

The "causal" effect of node i on node j is quantified here by the *model effective connectivity* C_{ji}. In practice, the matrices Q^0 and Q^Δ in Eq. (12) can be replaced by their empirical counterparts ($\hat{Q}^0$,$\hat{Q}^1$) calculated from the data to obtain estimates $\hat{J}$, $\hat{C}$ of hence J, C. In this work, we rely on a more elaborate estimation procedure based on a gradient descent (or Lyapunov optimization) to robustly estimate C for fMRI data that typically consist of a limited number of time points due to the sampling rate [13]. In essence, it uses a partial differentiation of Eq. (12) to iteratively optimize $\hat{J}$ and $\hat{\Sigma}$. Here an important point to note is that the matrix logarithm can yield complex values for $\hat{J}$, while the iterative optimization keeps the matrix elements real-valued.

For any given MOU process (9), we can build an equivalent 1st order MAR process (1). Indeed, by integrating Eq. (9) for a given lag Δ, one obtains

$$\mathbf{x}(t+\Delta) = e^{J\Delta}\mathbf{x}(t) + \mathbf{w} \tag{13}$$

where $\mathbf{w} \sim \mathcal{N}(0, \tilde{\Sigma})$ is a Gaussian noise with covariance matrix

$$\tilde{\Sigma} = \int_0^\Delta dt\ e^{Jt}\Sigma\ e^{J^T t} = \sigma^2 \int_0^\Delta dt\ e^{J^* t} = \sigma^2(2(J^*)^{-1})(e^{2J^* t} - \mathbb{I})$$

where $J^* = \log e^J e^{J^T} \sim J + J^T + \frac{1}{2}[J, J^T]$.

However, for a given 1st order MAR process, the equivalent MOU process corresponds to

$$S = \tilde{\Sigma}, \qquad A = e^{J\Delta} = e^{-\frac{\Delta}{\tau}} e^{C^T \Delta} \tag{14}$$

which implies the following constraints on S and A:

$$\log A \propto \kappa_1 + C^T$$

$$S = \kappa_2 \int_0^\Delta dt\ A^t\ (A^T)^t$$

where $\kappa_1 < 0$, $\kappa_2 > 0$ are constants, and C is a real matrix with null diagonal, $C_{ii} = 0$.

In sum, any MOU process (9) with real coefficients is associated with an equivalent MAR process (1) with real coefficients. The reciprocal, however, is not true, as an arbitrary matrix A will not yield in general a real-valued, but a complex-valued matrix $J = \log(A)$ because of the matrix logarithm. We will refer to a *MOU-compatible MAR* when the process can be associated to a real-valued J .

Theoretical relation between GC and EC

Assuming compatibilty, the multivariate dynamics underlying the estimation of GC and EC are fully consistent. We can therefore compare C_{ij} and G_{ij}, which both measure the effect of node i on node j. The following analysis relies on the approximation of the matrix exponential by a first-order expansion:

$$e^{C^T\Delta} \simeq \mathbb{I} + C^T\Delta \tag{15}$$

Eq. (15) holds under the following condition

$$||C|| \ll \frac{1}{\Delta} \tag{16}$$

where $||\cdot||$ denotes matrix 2-norm. Note that this condition can be formulated considering the dominant eigenvalue of C: $\mu_1 \ll \frac{1}{\Delta}$ with μ_i being the real parts of the eigenvalues of C in decreasing order ($\mu_1 \geq \mu_2 \cdots \geq \mu_N$). Also note that the eigenvalues of J are equal to those of C shifted by $-\frac{1}{\tau}$, so the stability of the resulting network dynamics requires that the real parts of μ_i are smaller than $\frac{1}{\tau}$ (again, it is sufficient to check only the dominant eigenvalue). We thus distinguish three regimes,

$\Delta \ll \tau$	sampling faster than process = fast sampling rate (FSR)
$\Delta \approx \tau$	sampling as fast as process = matched sampling rate (MSR)
$\Delta \gg \tau$	sampling slower than process = slow sampling rate (SSR)

In FSR, Eq. (16) is always satisfied since $|\mu_i| \leq \frac{1}{\tau} \ll \frac{1}{\Delta}$. In MSR, Eq. (16) is satisfied whenever

$$||C|| \ll \frac{1}{\tau} \simeq \frac{1}{\Delta} \qquad \text{weak coupling}$$

becoming more constraining in SSR

$$||C|| \ll \frac{1}{\Delta} \ll \frac{1}{\tau} \qquad \text{very weak coupling}$$

In practice, condition (16) can often be considered to be satisfied in both FSR and MSR, essentially because large values of C_{ij} are incompatible with stability, so $||C|| \ll 1/\tau$. However, SSR might break the assumption.

We can show that, provided Eq. (15) holds, there exist quadratic relations between IC, GC and EC. As detailed in Eqs. (22) and (24) in S1Text., we have

$$G_{ij} \simeq \Delta^2 \frac{1}{e^{\frac{2\Delta}{\tau}} - 1} \frac{\sigma_i^2}{\sigma_j^2} C_{ij}^2 \tag{17}$$

$$I_{ij} \simeq \tau^2 \big(1 - \frac{2\Delta}{\tau} \frac{e^{-\frac{2\Delta}{\tau}}}{1 - e^{-\frac{2\Delta}{\tau}}}\big)^2 \frac{\sigma_i^2 C_{ij} + \sigma_j^2 C_{ji}}{2\sigma_i\sigma_j} \tag{18}$$

with $\tilde{C}_{ij} = \frac{\sigma_i^2 C_{ij} + \sigma_j^2 C_{ji}}{2\sigma_i \sigma_j}$ being a mean of the reciprocal connection weights between i and j weighted by the input variances. If the input noise is homogeneous across all nodes, $\sigma_i = \sigma$, these relations further simplify:

$$G_{ij} \simeq \Delta^2 \frac{C_{ij}^2}{e^{\frac{2\Delta}{\tau}} - 1}$$
$$I_{ij} \simeq \tau^2 \big(1 - \frac{2\Delta}{\tau} \frac{e^{-\frac{2\Delta}{\tau}}}{1 - e^{-\frac{2\Delta}{\tau}}}\big)^2 \bar{C}_{ij}{}^2$$

with $\bar{C}_{ij} = (C_{ij} + C_{ji})/2$ being the symmetrized effective connectivity between i and j. So, assuming equal input noise on all nodes and neglecting asymmetries, there is an approximately quadratic relation between the symmetrized effective connectivity C_{ij} and I_{ij}. In particular, these relations mean that Granger causality provides an estimate of the quadratic interaction between nodes governed by continuous dynamics (i.e., having a MOU as a generative process). In the general case of inhomogeneous noise variance, we can nevertheless retrieve the approximately quadratic relation using a "corrected" versions of G and I:

$$cG_{ij} = \frac{Q_{jj}^0}{Q_{ii}^0} G_{ij} \simeq \Delta^2 \frac{C_{ij}^2}{e^{\frac{2\Delta}{\tau}} - 1} \tag{19}$$
$$cI_{ij} = I_{ij} \cdot \frac{4 Q_{ii}^0 Q_{jj}^0}{Q_{ii}^0 + Q_{jj}^0} \simeq \tau^2 \big(1 - \frac{2\Delta}{\tau} \frac{e^{-\frac{2\Delta}{\tau}}}{1 - e^{-\frac{2\Delta}{\tau}}}\big)^2 C_{ij}{}^2 \tag{20}$$

Here we have used the fact that the node variance Q_{ii}^0 is strongly related to the corresponding input variance σ_i^2. This holds in practice for weak coupling and can be seen via the Lyapunov equation (11) where J is then dominated by its diagonal elements $-\frac{1}{\tau}$, yielding $Q_{ii}^0 \simeq \frac{\tau}{2}\sigma_i^2$.

A last quantity of interest is the ratio between Granger causality and instantaneous causality:

$$\rho = \frac{cG_{ij}}{cI_{ij}} \approx \frac{(\Delta^2/\tau^2)}{\big(e^{\frac{2\Delta}{\tau}} - 1 - \frac{2\Delta}{\tau}\big)^2} \simeq \frac{(\Delta^2/\tau^2)}{\big(2\frac{\Delta^2}{\tau^2}\big)^2} = \frac{1}{2}\frac{\tau^2}{\Delta^2} \tag{21}$$

This means that in FSR we have $cG \gg cI$, whereas $cG \ll cI$ in SSR (note that this also true for the uncorrected versions G and I). Thus, G and I do not always reflect the underlying network connectivity. In other words, when a continuous MOU model is a valid generative model for the observed time series, the information about C is differentially captured by G and I, crucially depending on the sampling regime.

Materials and Methods

Network simulations

We generated time series using the MOU model (9) with process time scales $\tau \in [0.1, 10]$ and a fixed sampling period $\Delta = 1$. The model connectivity C was a random matrix with probability $p_1 = 30\%$ of connection between each pair of nodes. We considered networks of $N = 10$, $N = 40$ and $N = 100$ nodes, The weights were drawn from a Pareto distribution $p(w) = \frac{\alpha w_0^\alpha}{w^{\alpha+1}}$ with $\alpha = 5$ and $w_0 = 0.1/\tau$ (for $N = 5$); with $\alpha = 3$ and $w_0 = 0.1/\tau$ (for $N = 40$); with $\alpha = 3$ and $w_0 = 0.5/\tau$ (for $N = 100$). With a probability $p_2 = 30\%$, a pair of reciprocal connections (C_{ij},C_{ji}) was chosen as their sign is flipped. Thus, roughly 30% of connections are negative. Finally, for all pairs of non-zero reciprocal links, a random number $0 < r < 0.2$ was extracted, and one of the two connections was multiplied by $\frac{1+r}{1-r}$, so as to generate asymmetries in reciprocal

connections.

In all simulations, we used a diagonal noise matrix $\Sigma = \mathrm{diag}\{\sigma_i^2\}$. We made the values of σ_i dependent on τ, as to align the noise time scale to the process time scale (stated otherwise, to align the magnitude of Σ to that of C). We generated random numbers $0.2 \leq s_i \leq 5$, and considered $\sigma_i = \sqrt{5 s_i / \tau}$, i.e., different nodes were affected by noise of different magnitude, with noise magnitudes spanning roughly an order of magnitude.

Human resting-state fMRI data

We used the 100 unrelated subjects' subset (54 females, 46 males) from the Human Connectome Project (HCP) [35]. For the main analysis, we used the left-right (LR) phase- encoding runs from the first session resting state fMRI data. We later replicated the analysis for the right-left (RL) phase-encoding runs. Time series had 1200 time points with a TR of 0.72 sec, meaning ≈ 15 mins duration. The full description of the imaging parameters and minimal preprocessing pipeline is given in Ref. [42]. In short, after correction for motion, gradient, and susceptibility distortions the fMRI data was aligned to an anatomical image. The aligned functional image was corrected for intensity bias, demeaned, and projected to a common surface space, which resulted in a cifti-file. Artifacts were removed though Independent Component Analysis (ICA) using the FSL's MELODIC tool paired with the FMRIB'S ICA-based X-noisefilter. No additional global signal regression was applied. All fMRI data were filtered between 0.1 and 0.01 Hz to retain the relevant frequency range for further analyses of the BOLD signal. Functional data can be mapped to different spatial resolutions using the Schaefer parcellation [43], which optimizes local gradient and global similarity measures of the fMRI signals. Here, we selected the parcellation consisting of 100 regions. For both fMRI datasets, regional time series were extracted using Workbench Command provided by the HCP.

For each participant, we computed $\hat{G}_{ij}, \hat{I}_{ij}$ using the covariance-based approach and $\hat{C}_{ij}$ using the Lyapunov optimization method. For each individual connection (in $\hat{I}$,$\hat{G}$,$\hat{C}$), we estimated a group-level significance using a t test. Connections were considered significantly positive if $T > 0$ and $p < 0.05$, significantly negative if $T < 0$ and $p < 0.05$, non-significant if $p > 0.05$. P-values were corrected for 9900 multiple comparisons using the false discovery rate approach [44]. For each pair of reciprocal connections ($i \to j$, $j \to i$) we computed connection asymmetries as $\Delta\hat{C} = C_{ij} - C_{ji}$, $\Delta|\hat{C}| = |C|_{ij} - |C|_{ji}$, $\Delta\hat{G} = G_{ij} - G_{ji}$. We estimated the group-level significance of connection asymmetry using a t test. Asymmetry was considered significant in the direction $i \to j$ if $T > 0$ and $p < 0.05$, significant in the direction $j \to i$ if $T < 0$ and $p < 0.05$, nonsignificant if $p < 0.05$. P-values were corrected for 4950 multiple comparisons using the false discovery rate approach [44].

Results

Relation between EC and GC on numerical simulations

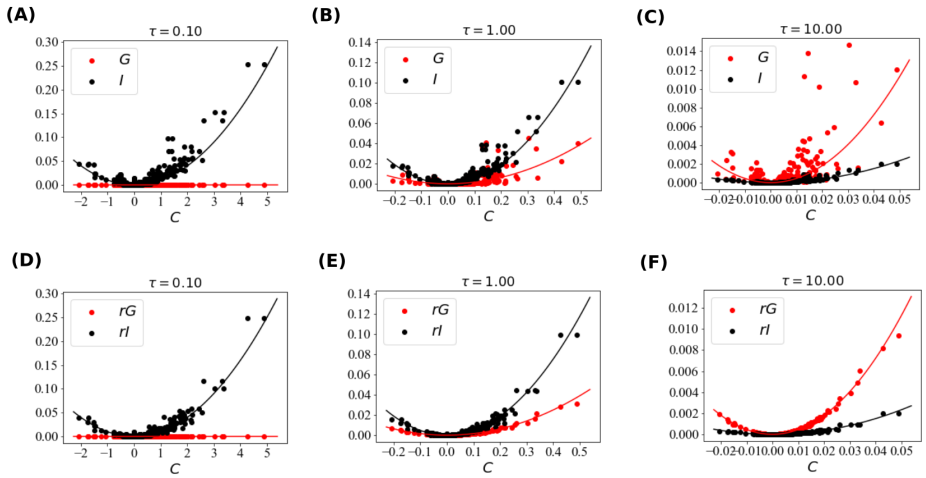

Fig 1. Theoretical relations between model EC and conditional GC/IC. We considered a random network of $N = 40$ nodes evolving according to the MOU dynamics (9) for different values of τ: 0.1 for SSR, 1 for MSR and 10 for FSR. Panels **(A-C)** show the relation between the C weights and the corresponding values of G (in red) and I (in black) for each connection, *without correcting for unequal input noise variances*. Panels **(D-F)** show the relation between the C weights and the corresponding values of rG (in red) and rI (in black) for each connection, i.e., *correcting for unequal input noise variances*. G and I (as well as rG, rI) were obtained from the theoretical covariance matrices Q^0, Q^1, corresponding to their empirical counterparts in the limit of infinite observation length, $\mathcal{L} \to \infty$. Each dot corresponds to a pair (i, j), and solid lines to the approximate quadratic scalings; note that correction on GC/IC is not necessary here as the generative model has homogeneous input noise variance. For all valued of τ, the approximate quadratic relation are well satisfied. The relative importance of I vs. G depends on the value of τ, with I prevailing at low τ and G at large τ.

Validation of the analytical relations using simulated MOU network dynamics In order to test the analytical relations derived in the previous sections, we generated a MOU process with known ("ground truth") connectivity and we computed the Granger causality measures (GC and IC) from simulated time series. We tested the quadratic relations ((22) and (24) between the model connectivity (C in equations) and Granger causality measures (G and I, respectively), which were derived in the "Models and analytical derivation sections and reported here for convenience:

$$G_{ij} \simeq \Delta^2 \frac{1}{e^{\frac{2\Delta}{\tau}} - 1} \frac{\sigma_i^2}{\sigma_j^2} C_{ij}^2$$
$$I_{ij} \simeq \tau^2 \big(1 - \frac{2\Delta}{\tau} \frac{e^{-\frac{2\Delta}{\tau}}}{1 - e^{-\frac{2\Delta}{\tau}}}\big)^2 \frac{\sigma_i^2 C_{ij} + \sigma_j^2 C_{ji}}{2\sigma_i \sigma_j}$$

where τ is the timescale of the MOU process, Δ the sampling time, σ_i, σ_j the standard deviations of the noise affecting i,j. These relations are *theoretical*, as they emerge in the limit of infinite sampling time where G and I are reconstructed with no error from the data. The relation for G is relatively easy to interpret: the influence from i to j is proportional to the square of the connection weight C_{ij} and the noise ratio $\frac{\sigma_i^2}{\sigma_j^2}$. This implies two sources of possible imbalance between i and j: coming from $C_{ij} \neq C_{ji}$ and/or $\sigma_i \neq \sigma_j$. For I, we have $\tilde{C}_{ij} = \frac{\sigma_i^2 C_{ij} + \sigma_j^2 C_{ji}}{2\sigma_i \sigma_j}$ that corresponds to the mean of the reciprocal connection strengths weighted by the input/noise variances. As discussed in Methods, one can obtain simpler, (approximately) quadratic relations by defining "corrected" quantities cG (19) and cI (20) that obliterate the effect of the noise imbalance between the two nodes:

$$cG_{ij} \simeq \Delta^2 \frac{1}{e^{\frac{2\Delta}{\tau}} - 1} C_{ij}^2 \qquad I_{ij} \simeq \tau^2 \big(1 - \frac{2\Delta}{\tau} \frac{e^{-\frac{2\Delta}{\tau}}}{1 - e^{-\frac{2\Delta}{\tau}}}\big)^2 C_{ij}^2$$

We tested these relations by considering a MOU process with random connectivity. We considered networks of $N = 10$, $N = 40$ and $N = 100$ nodes, with directed links randomly with probability $p = 0.3$ between each pair of nodes. We randomly selected link weights from a power-law distribution to reproduce the heavy-tailed weight distribution observed in typical brain networks, and we allowed for connectivity asymmetries (see Methods for details). We considered the case of unequal noise on all nodes (spanning more than one order of magnitude). We fixed the sampling time $\Delta = 1$ and considered three different cases where $\Delta \ll \tau, \Delta \approx \tau$ and $\Delta \gg \tau$.
Results are shown in Fig. 1. We compared the consistency of C with the uncorrected (G, I) and corrected (cG and cI) versions of Granger causality. The quadratic relations (22) and (24) were satisfied to a very good accuracy for the corrected case. When the correction was not applied, the quadratic relation between C and G was significantly degraded. Furthermore, as expected, we observed that I prevails over G at low τ, while the opposite occurs at high τ. Although the theoretical relations were derived for the conditional version of G and I, very similar results were obtained by considering the unconditional version of G and I (Supp. Fig. S1Fig).

Influence of sampling rate and time series duration. Following previous work [45, 46], we then examined how the GC estimates are affected by the finite length $\mathcal{L}$ of time series, as well as the sampling rate of the signal. For increasing values of $\mathcal{L}$, we computed a quadratic fit of $\hat{cG}, \hat{cI}$ over the model connectivity C, estimating the goodness of fit through the R^2 coefficient. Values of R^2 equal to 1 signify that the quadratic scaling is perfectly satisfied. Results are shown in Fig. 2, Panels (A) to (F). As expected, the quadratic relations are well satisfied for large sampling time $\mathcal{L}$. For $\mathcal{L} \gtrsim 10^4$, the

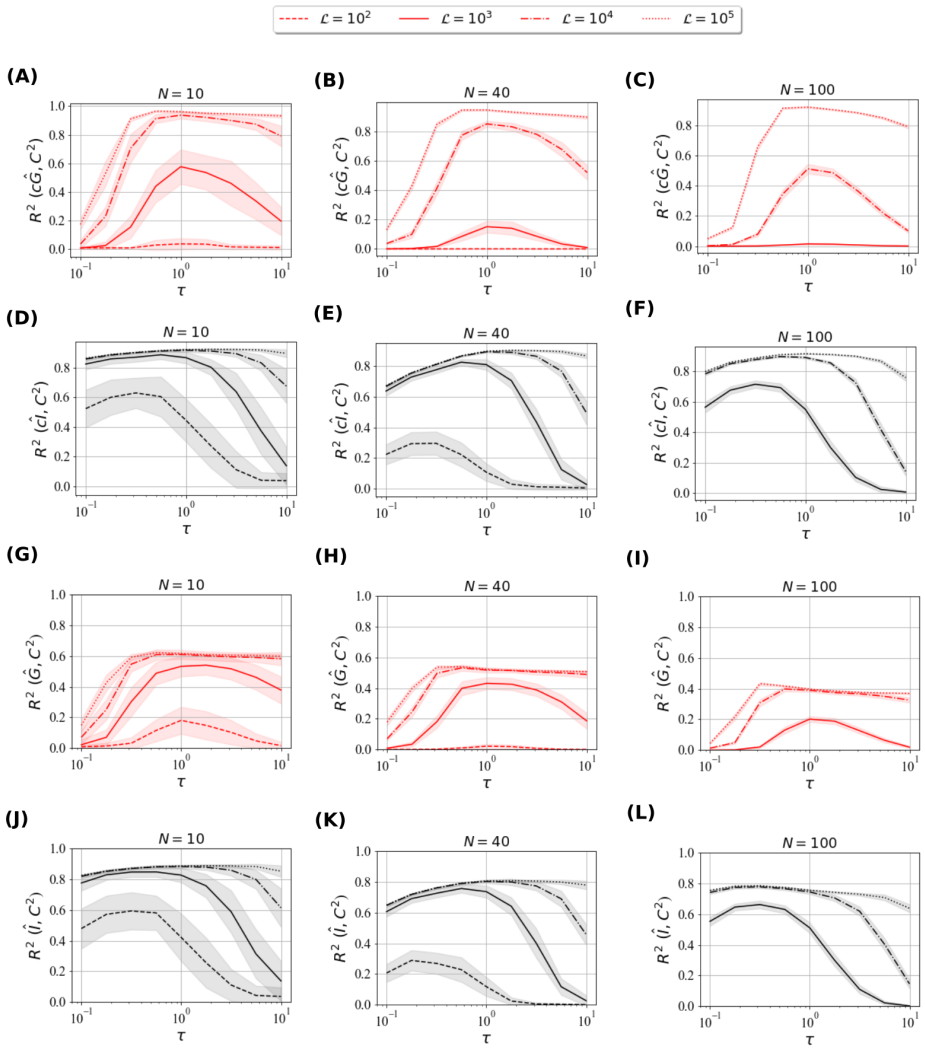

Fig 2. Relations between model connectivity (MC) and empirical conditional Granger measures (GC/IC) for finite sampling time. We considered the same random network as in Fig. 1, which we simulated for different values of τ (x-axis) for $\mathcal{L} \in \{10^2, 10^3, 10^4, 10^5\}$ (dashed, solid, dashed-dotted, and dotted, resp.). The columns correspond to several network sizes, from $N = 10$ to 100. **(A-C)** Variance explained by the fit of $c\hat{G}$ as a quadratic function of C (the model connectivity). The match between MD and GC is bad for small τ (SSR), but becomes better for larger values (MSR and FSR). Somewhat surprisingly, the match for short duration $\mathcal{L}$ becomes worse for FSR. **(D-F)** Same as above panels for $c\hat{I}$. A contrary trend is seen with a better match of IC with MC for SSR and MSR, becoming worse for FSR.In both cases, larger duration $\mathcal{L}$ implies a better match. **(G-L)** Same as (A-F), but using non-corrected quantities G,I. In absence of the correction, the relation between $\hat{G}$ and C is significantly degraded.

relations are well satisfied with R^2 close to 1. However, the relations are not observed at very low sampling time $\mathcal{L} = 10$ ($R^2 < 0.1$ for $\hat{cG}$; $R^2 < 0.2$ for $\hat{cI}$) and only partially observed at intermediate sampling time $\mathcal{L} = 10^3$. More generally, the quadratic relation between $\hat{cG}$ and C is not observed in the SSR condition ($\tau \ll \Delta$), congruently with the fact that $\hat{cG}$ is small and more difficult to estimate. Analogously, the quadratic relation between $\hat{cI}$ and C is not observed in FSR ($\tau \gg \Delta$), congruently with the fact that $\hat{cI}$ is small and difficult to estimate in this regime. Somewhat unexpectedly, the quadratic relation between $\hat{cG}$ and C is also degraded in the FSR condition ($\tau \gg \Delta$) for low values of $\mathcal{L}$. This depends on the fact that in this regime, as the process time scale is long, the empirical $\hat{Q}^1$ is less accurate since the system has not evolved enough to appreciate lagged dependencies. In S2Fig (D-F) we show the Pearson's R^2 between the theoretical Q^1 and the estimated $\hat{Q}^1$. The most accurate estimate $\hat{Q}^1$ occurs when $\tau \approx \Delta$. We observed that I was less affected by the problem of inaccuracies of $\hat{Q}^1$, as it mostly depends on $\hat{Q}^0$, whose estimation was much more robust (S2FigA-C). In Panels (G) to (L) we considered the quadratic fit of the uncorrected quantities $\hat{G}, \hat{I}$ over the model connectivity C. The behavior of the fit with respect to the sampling length $\mathcal{L}$ was qualitatively similar to what observed for $\hat{cG}, \hat{cI}$. Hover, without the correction the relation between C and cG was degraded, as reflected by much lower values of R^2.
Finally, in S4Fig we showed the ratio between G and I. While the ratio monotonically increased with τ as expected, we noted that only in the limit of large sampling time we observed a strong variation with τ. Generally, in the case of finite sampling it may not be easy to accurately identify the dynamical regime from this ratio.

Asymmetric interactions . The quadratic (non-monotonic) relation (19) implies that G is insensitive to whether the underlying connections are "excitatory" or "inhibitory", $C_{ij} > 0$ or $C_{ij} < 0$, respectively. This complicates the interpretation of G in terms of *asymmetry* of reciprocal connections, aiming to quantify a net directional effect between pairs of nodes. For the typical situation of excitatory weights, asymmetry can be measured by taking the difference between the weights of reciprocal connections $i \to j$ and $j \to i$. For C and G, this leads to e.g.,

$$\Delta C_{ij} = C_{ij} - C_{ji} \qquad \Delta G_{ij} = G_{ij} - G_{ji}$$

However, when effective connections are either positive or negative, these quantities are not necessarily concordant. This is exemplified in Fig. 3A. A positive value of $\Delta C_{ij} > 0$ associated to a net asymmetry $i \to j$ of the effective connection can correspond to three different situations, depending on the sign of C_{ij} and C_{ji}:

- $C_{ij} > 0, C_{ji} > 0, \Delta C_{ij} > 0$: both connections are excitatory, and the influence of i over j is larger than the reverse. ΔC_{ij} and ΔG_{ij} are concordant.
- $C_{ij} < 0, C_{ji} < 0, \Delta C_{ij} > 0$: both connections are inhibitory, but in this case it is the influence of j over i that prevails. ΔC_{ij} and ΔG_{ij} are discordant.
- $C_{ij} > 0, C_{ji} < 0, \Delta C_{ij} > 0$: effective connections are discordant - one is excitatory and the other inhibitory. In this case, ΔC_{ij} and ΔG_{ij} are not necessarily concordant; ΔG_{ij} could even be zero.

In all cases, the net influence asymmetry measured by ΔG aligns with the net difference in the *absolute strength* (positive or negative) of the effective connection, measured by

$$\Delta |C|_{ij} = |C_{ij}| - |C_{ji}|$$

This is further illustrated in Fig. 3B-D. We considered the network in Fig.1 for $N = 40$ and $\tau = 10$. We divided pairs of effective connections depending on the sign of connections.

If both signs are positive, ΔC and ΔG are concordant (3B; If both signs are negative, ΔC and ΔG are discordant (3C; If signs are discordant, ΔC and ΔG can be either concordant or discordant (3D. If we combine all types of connections, no clear relation emerges (Fig. 3E). In order to observe a meaningful relation, one must consider separately $\Delta|C|_{ij}$ and ΔG_{ij}. As shown in Fig. 3F, these two quantities obey a monotonic relation.

In Fig. 4 we analyzed the concordance between ΔG_{ij} and $\Delta|C_{ij}|$ for finite sampling length $\mathcal{L}$, using the same setting as in Fig. 2. Note that, due to the unequal noise variances affecting different nodes, we used the corrected quantity cG. In panels A to C we show the Pearson R^2 between ΔcG_{ij} and $\Delta|C_{ij}|$ as a function of $\mathcal{L}$ and τ. Qualitatively, we have the same behavior observed for cG: R^2 is monotonically increasing as a function of $\mathcal{L}$, and peaks for $\tau \approx 1$. However, a good detection of the asymmetry of connections requires large sampling time $\mathcal{L}$ ($\gtrsim 10^4$). This implies that the asymmetry of connections (in terms of strength of influence) estimated with GC is not accurate for low and moderate sampling time. This implies severe limitations in reliably inferring the asymmetry of connections using individual-level data in fMRI (with, typically, $\mathcal{L} \sim 10^3$). We note that the correction for unequal noise variances is critical to have a concordance in the estimation of connection asymmetry. In panels D to F we show the concordance between $\Delta|C|$ and the uncorrected ΔG. Without the correction, the estimates of the asymmetry of the connection are much less consistent between the two methods.

Comparing GC/IC estimates with EC estimates In general, the ground truth C is not known, but it can be inferred from the data by minimizing the discrepancy between the model Q^0, Q^1 and the observed $\hat{Q}^0, \hat{Q}^1$. Here, we rely on a gradient-descent approach developed in [13], which is more efficient than directly solving Eq. (12). We thus investigated the relation between the estimated $\hat{G}, \hat{I}$ and the estimated $\hat{C}$. Results are shown in S3Fig. We observed that quadratic relations hold better for larger sampling time $\mathcal{L}$. The values of R^2 between $\hat{C}$ and $\hat{G}$ are generally larger than those between C and $\hat{G}$. This suggests that both $\hat{G}, \hat{I}$ and $\hat{C}$ are affected by similar biases via their estimation procedures, which are based on the same data. We also observed that the quadratic relation between I and $\hat{C}$ is not observed in SSR ($\tau \ll \Delta$), congruently with the fact that C is difficult to estimate in this regime.

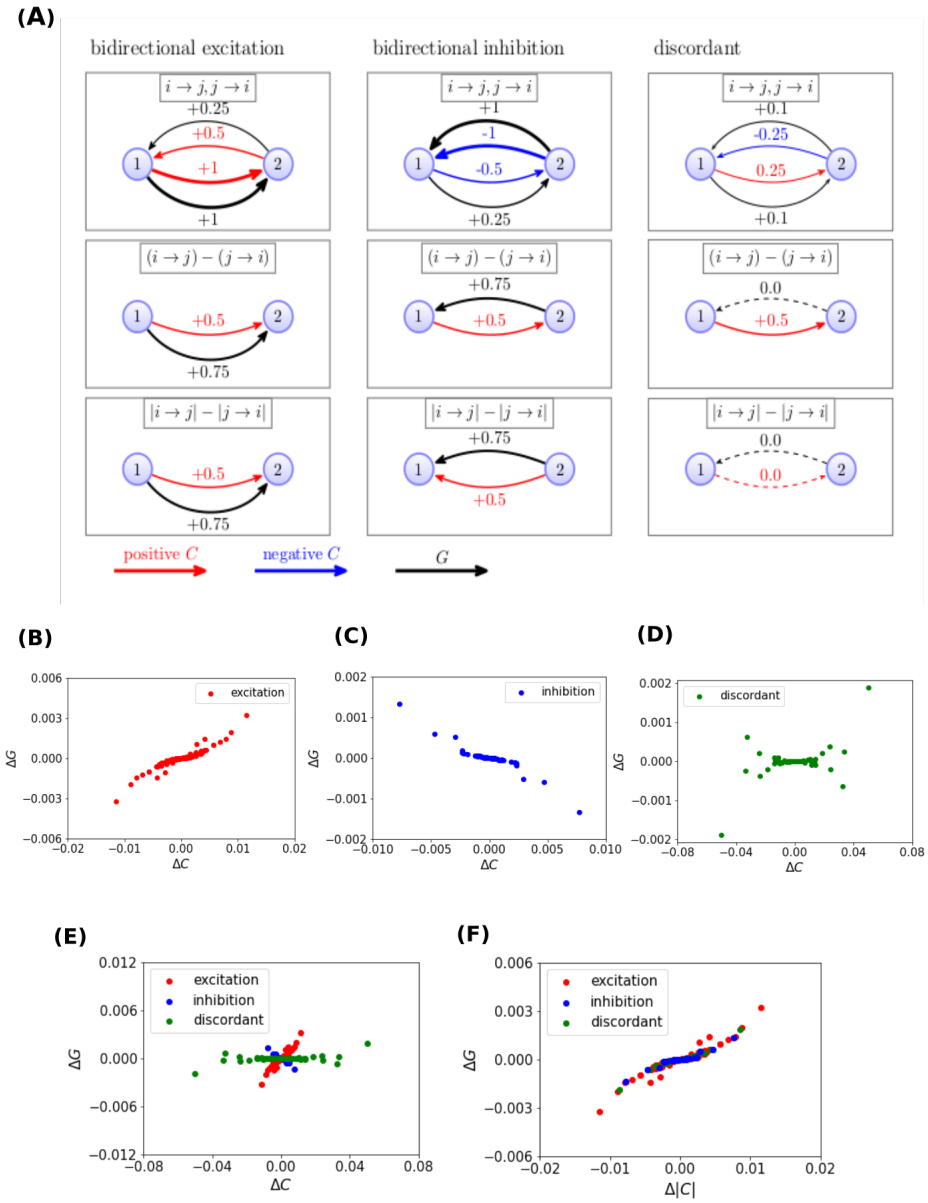

Fig 3. Theoretical relations between the asymmetry of connections for EC and GC Panel **(A)** displays the conceptual relation between ΔG, ΔC and $\Delta|C|$ for the three cases in which C_{ij}, C_{ji} are both positive ("bidirectional excitation"), both negative ("bidirectional inhibition") or discordant. Panels **(B-D)** display the relations between ΔC and ΔG for the same random network as in Fig. 1 with $N = 40$ and $\tau = 10$, separately by categorizing the pairs of reciprocal connections according to the three cases depicted in (A). Combining those, panel **(E)** pools all types of pairs together for ΔG as a function of ΔC, to be compared with ΔG as a function of $\Delta|C|$ in panel **(F)**. It illustrates that a monotonic relation appears between ΔC and ΔG, but not $\Delta|C|$ and ΔG.

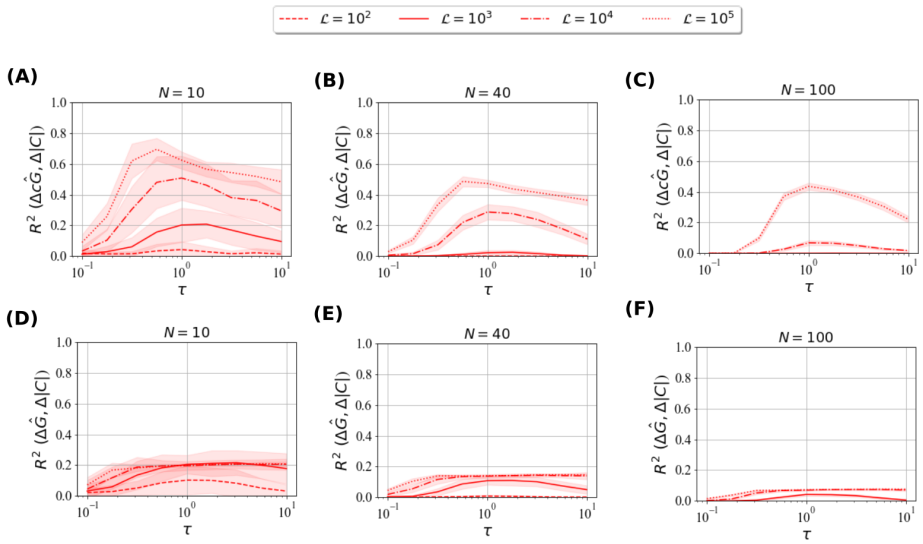

Fig 4. Relations between EC asymmetry and GC asymmetry for finite sampling time. We considered the same random networks as in Fig. 1, which we simulated for different values of τ and $\mathcal{L}$. **(A-C)** We show the R^2 between the model asymmetry in EC, $|\Delta C|$, and the asymmetry in the corrected Granger causality, ΔcG. **(D-F)** We show the R^2 between the model asymmetry in EC, $|\Delta C|$, and the asymmetry in the uncorrected Granger causality, ΔcG.

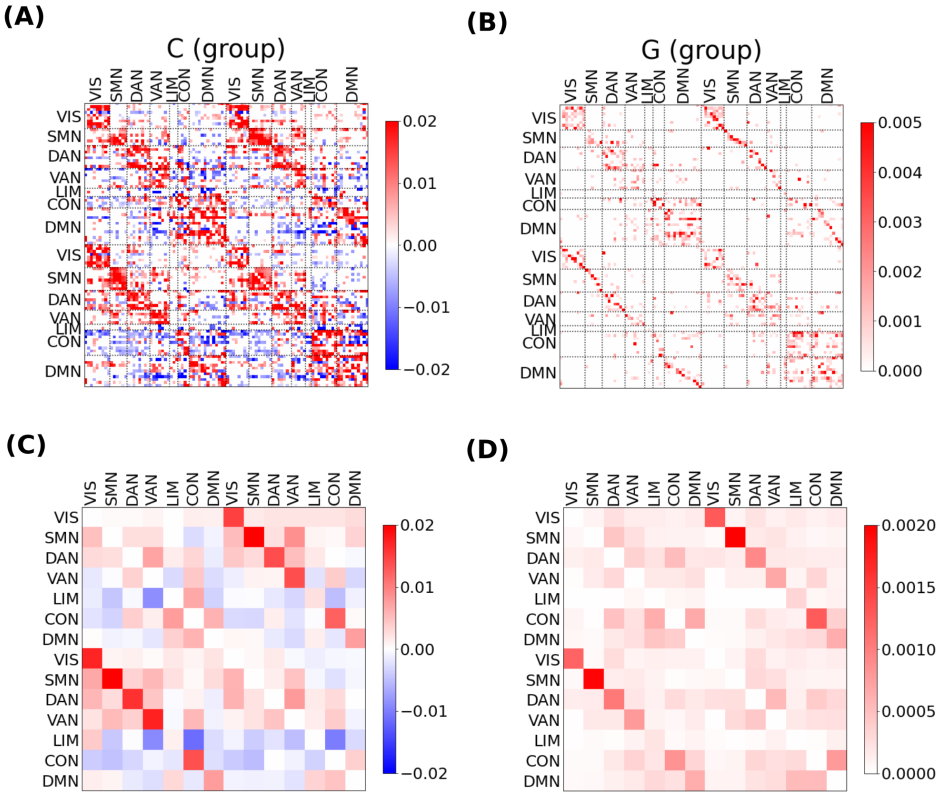

Fig 5. **(A)** $\hat{C}$ matrix (group average over 100 subjects). Only significant links (t test over subjects, $p < 0.05$ FDR-corrected) are shown. **(B)** $\hat{C}$ matrix (group average over 100 subjects). Only links that are significantly different from zero (t test across subjects, $p < 0.05$ FDR-corrected) are shown, in red for positive weights and blue for negative weights. **(C)** Network-wise effective connectivity $\hat{C}^{(net)}$ (group average over 100 subjects). Only significant links (t test over subjects, $p < 0.05$ FDR-corrected) are shown. **(D)** network-wise Granger causality $\hat{G}^{(net)}$ (group average over 100 subjects). Only significant links (t test over subjects, $p < 0.05$ FDR-corrected) are shown.

EC and GC relations in rs-fMRI data

We considered human resting-state fMRI data of 100 unrelated participants from the Human Connectome Project [35]. Upon surface projection and minimal preprocessing, time series were projected onto the Schaefer-100 atlas (100 regions). Regions were divided into seven resting state networks, (RSNs) according to the well-established division proposed by Yeo et al. [47].

Effective connectivity analysis. We first considered the effective connectivity $\hat{C}$, that was estimated individually for each participant. We observed that the estimated $\hat{C}$ displayed both positive and negative effective connections, on average across subjects. In particular, we found 3804 significant connections at a group level (t test over subjects, $p < 0.05$, FDR corrected for 9900 multiple comparisons). Among all significant connections, we found both significantly positive (64%) and significantly negative connections (36%). There is a wide debate about the spurious or genuine nature of negative correlations of BOLD signals [48], with several authors arguing that negative correlations are due

$C_{ij} > 0, C_{ji} > 0$	$C_{ij} > 0, C_{ji} = 0$	$C_{ij} > 0, C_{ij} < 0$	$C_{ij} < 0, C_{ji} < 0$	$C_{ij} < 0, C_{ji} = 0$
812	771	35	306	727

Table 1. Sign concordance of reciprocal connections

to the use of global signal regression [49, 50]. Here, we find negative connections with statistical significance and without global signal regression (Fig. 5A). Looking at pairs of reciprocal connections ($A \to B$ and $B \to A$), significant connections tended to be either unidirectional (i.e., the reciprocal connection is non-significant), or concordant (the reciprocal connection is also significant and has the same sign; Table 1). We found only $\approx 1\%$ of pairs of significant connections with discordant sign (S5FigA-B). Positive-positive reciprocal connections were frequently found within different areas of the same resting state networks (RSNs), while connections between nodes of different RSNs were predominantly unidirectional. Negative connections were nearly exclusively found between different RSNs (S5Figb). For all pairs of reciprocal connections, we tested for significant connection asymmetry, $\Delta C_{ij} = C_{ij} - C_{ji}$ (t test over subjects, FDR corrected for 4950 multiple comparisons). We found a significant asymmetry for 27% of link pairs (S5FigC). Asymmetries between positive connections were found within areas of the same resting state network (RSNs), while asymmetries between inter-RSN links were mostly due to one connection being positive/negative and the reciprocal one being non-significant (S5FigD). Only a negligible fraction (2%) of significant asymmetries involved discordant connections.

We averaged the $\hat{C}$ values over the nodes belonging to each of seven resting state networks (RSNs) for both hemispheres, therefore obtaining 14×14 network-wise effective connectivity matrices. In Fig. 5C we show significant positive and negative network-wise links (t test over subjects, $p < 0.05$ FDR corrected for 182 multiple comparisons). We observed strong positive connections within each RSN, between its LH and RH parts. On the other hand, negative connections were found between different RSNs, more strongly within the RH. Negative connections were found between nodes of the control network (CON), especially the right CON and other RSNs; and between the default mode network (DMN) and other cognitive networks such as the ventral attention network (VAN), the dorsal attention network (DAN) and the CON. Positive connections were mostly observed within RSNs. Along with $\hat{C}$, we also estimated the noise covariance $\hat{\Sigma}$, assuming diagonal covariance. The noise variances affecting different nodes wee found to be widely different among nodes, spanning roughly an order of magnitude (S6Fig), with ventral and posterior nodes displaying the larges values of σ^2.

Granger causality analysis. Next, we estimated Granger causality. At a group level, the estimated $\hat{cG}$ displayed 944 significant connections (Fig. 5B) (t test over subjects, $p < 0.05$ FDR corrected for 9900 multiple comparisons). Of these, 827 were also significant $\hat{C}$ connections. Most (744, 90%) of these significant $\hat{cG}$ connections are associated with positive $\hat{C}$ connections (Table 2). The strongest connections were found between homologous regions of the LH/RH. We averaged the $\hat{cG}$ values over the LH/RH nodes belonging to each of seven resting state networks (RSNs), obtaining 14×14 network-wise Granger causality matrices. In Fig. 5d) we show significant network-wise links (t test over subjects, $p < 0.05$ FDR corrected for 182 multiple comparisons). The strongest links were again observed within the LH and RH part of each RSN. We tested pairs of reciprocal links for connection asymmetry: only a small number of link pairs (217, $\sim 5\%$) presented a significant asymmetry. Of these, most (146) also presented a significant asymmetry in terms of $\hat{C}$. Thus, $\hat{cG}$ appears generally less effective than $\hat{C}$ in detecting both significant connections and significant asymmetries.

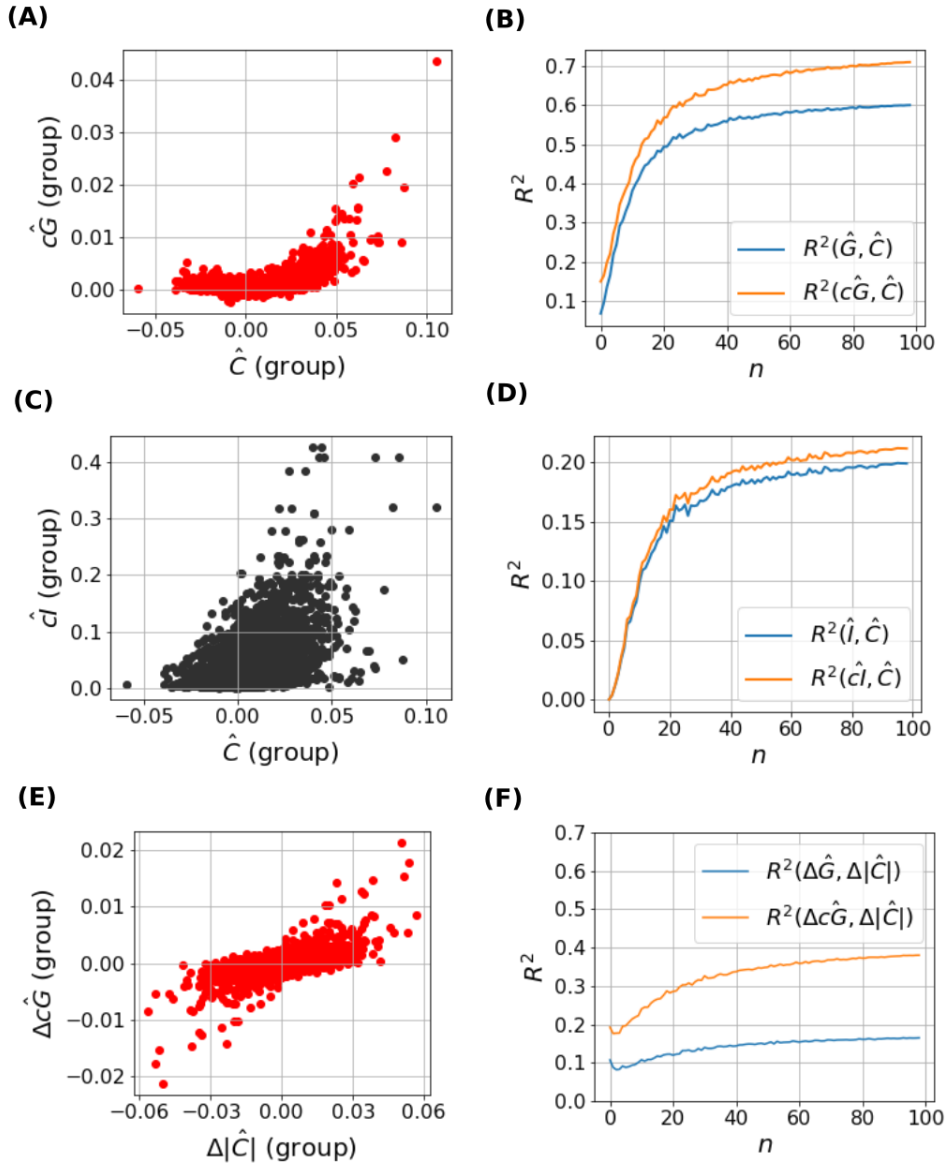

Fig 6. **(A)** effective connectivity $\hat{C}$ (group average over 100 subjects) vs corrected Granger causality $\hat{cG}$ (group average). **(B)** squared Pearson correlation R^2 between group $\hat{C}$ and group $\hat{cG}$ for groups of increasing size n. **(C)** effective connectivity $\hat{C}$ (group average over 100 subjects) vs corrected instantaneous Granger causality $\hat{cI}$ (group average). **(D)** squared Pearson correlation R^2 between group $\hat{C}$ and group $\hat{cI}$ for groups of increasing size n. **(E)** asymmetry in effective connectivity $\Delta\hat{C}$ (group average over 100 subjects) vs asymmetry in corrected Granger causality $\Delta\hat{cG}$ (group average). **(F)** squared Pearson correlation R^2 between group $\Delta\hat{C}$ and group $\Delta\hat{cI}$ for groups of increasing size n.

	$\hat{C} > 0$	$\hat{C} < 0$	$\hat{cG}$
$\hat{C} > 0$	2430	0	754
$\hat{C} < 0$		1374	73
$\hat{cG}$			944

Table 2. Significant connections shared among effective connectivity $\hat{C}$ and Granger causality $\hat{G}$

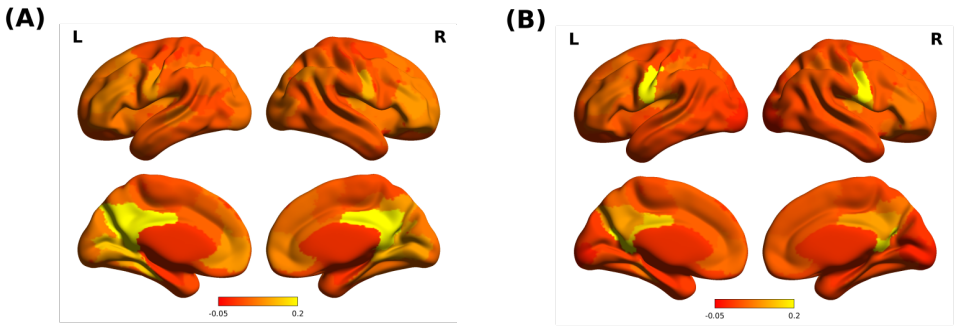

Fig 7. Principal eigenvector of the average C^2 matrix **(A)** and the average G matrix **(B)**, projected onto the cortical surface.

Comparison between effective connectivity and Granger causality . We systematically assessed the degree of group-level consistency between $\hat{C}$ and $\hat{cG}$ by computing averaged group-wise matrices $\hat{C}$ and $\hat{cG}$. Fig. 6A shows the scatter plot of $\hat{C}$ and $\hat{cG}$, averaged over the whole cohort of $n = 100$ subjects. The theoretically expected quadratic relation is approximately matched. We computed the squared Pearson correlation R^2 between $\hat{C}^2$ and $\hat{cG}$, obtaining a large degree of consistency, $R^2 = 0.72$. We tested group-level consistency also for smaller groups of subjects. To this aim, we considered subgroups of increasing size n. For each n, we averaged the R^2 over $N = 100$ random subgroups of size n. Results are shown in Fig. 6B: R^2 rapidly increases with n, reaching values $R^2 \gtrsim 0.6$ for $n > 20$. For comparison, we also show the R^2 between $\hat{C}$ and the non-corrected version of Granger causality, $\hat{G}$. The degree of consistency is notably inferior, with values $R^2 \gtrsim 0.5$ for $n > 20$, up to $R^2 = 0.6$ for $n = 100$. We assessed the degree of consistency between the two measures at the level of *asymmetry* of connections. Fig. 6E shows the scatter plot between $\Delta|\hat{C}|$ and $\Delta\hat{cG}$, for links with $\hat{C} > 0$. We computed the Pearson correlation R^2 between $\Delta|\hat{C}|$ and $\Delta\hat{cG}$, obtaining $R^2 = 0.39$. Also in this case, the degree of consistency increases with group size n, reaching values $R^2 \approx 0.3$ for $n > 20$ (Fig. 6F). For comparison, we also show the corresponding results for the non-corrected version of Granger causality, $\hat{G}$. In this case, the degree of consistency is much lower ($R^2 \approx 0.18$).

Model inconsistencies. Fig. 6C shows the scatter plot of $\hat{C}$ and $\hat{cI}$, averaged over subjects. In this case, the squared Pearson correlation R^2 between the two measures is much lower than for $\hat{G}$, $R^2 = 0.21$. Fig. 6D shows that R^2 rapidly increases with n, reaching values $R^2 \gtrsim 0.15$ for $n > 20$. This analysis shows several discrepancies with respect to theoretical expectations. First of all, the values of $\hat{cI}$ are much larger than the values of $\hat{G}$. In fact, on average, $\hat{G}/\hat{I} \approx 10^{-2}$. This ratio would be expected in case of an extremely rapid process time scale; however, the average (estimated) process time is $\tau = 2.3 \pm 0.6$ (mean $\pm$ st.dev. over subjects). Furthermore, the correlation between $\hat{C}$ and $\hat{G}$ is larger than the correlation between $\hat{C}$ and $\hat{I}$, consistently with this process

time estimate.

Hub structure. We compared C and G in terms of the consistency of the hub structure related to the C and G network respectively. To this aim, we computed the principal eigenvectors ot the average G matrix and of the average C^2 matrix. We find a correlation $R = 0.41$ between the two principal eigenvectors, pointing at a good consistency of the hub structure. The two eigenvectors are projected on the cortical surface and shown in Fig. 7. The important regions indicated by this centrality measure slightly differ between C^2 and G, presumably because of the empirical noise.

Replication. We replicated the same analysis using the second fMRI session available for all subject. In Fig. 8 we perform a quantitative comparison between the results of the two sessions. In Fig. 8A we show the consistency (R^2) of the estimates of $\hat{C}$,$\hat{G}$, $\hat{cG}$,$\hat{I}$ and the corresponding asymmetries at the individual level. At this level, we observed a very poor consistency ($R^2 < 0.1$) between the estimates of the two sessions, except for $\hat{G}$ ($R^2 \approx 0.2$) and, especially, $\hat{I}$ ($R^2 \approx 0.6$). However, averaging over at least 20 subjects (Fig. 8B) yielded reliable estimates ($R^2 > 0.6$) for all of $\hat{C}$,$\hat{G}$, $\hat{cG}$,$\hat{I}$. Moreover, also the asymmetries $\Delta\hat{C}$,$\Delta\hat{G}$, $\Delta\hat{cG}$ obeyed the same trend, with reliable estimates ($R^2 > 0.4$) starting from 20 subjects (Fig. 8C). Last, the confusion matrices in Fig. 8D-E showed that when using the whole sample ($n = 100$), the detection of links with significant weights in $\hat{C}$ exhibits more than 80% replicability, with no single connection being identified as significantly positive in one session and significantly negative in the other. Moreover, the detection of pairs with significant asymmetry achieved 65% replicability, with no detection of opposite asymmetries in the two sessions. Thus, estimates of connection strength and asymmetry were reliable at a group level.
Consistently with these results, the relations between C and G observed in Fig. 6 at a group level were well replicated in the second session. Fig. S7Fig shows results analogous to those of Fig. 6 for the second session. It is apparent that results are fully consistent with those obtained for the first session.

Discussion

Substantial (if partial) information about causal interconnections can be learned from time-lagged relationships between brain areas observed in neural activity recordings [12]. Research in this direction has mostly relied on fMRI recordings, due to their fine spatial resolution. In this context, techniques going beyond traditional, undirected functional connectivity can remove spurious connections arising from indirect effects and reveal asymmetries between reciprocal connections (that underlie a key principle of brain organization, hierarchical processing [51, 52]. Thus, these methods have contributed to elucidating hierarchical relations underlying specific cognitive functions such as working memory [53, 54], cognitive control [55], language [56], and revealing task-related modulation of functional interactions [57]. From a clinical standpoint, these methods are essential to characterize pathological alterations of functional hierarchies, [34, 58] and obtain improved biomarkers for clinical classification [59].

While several directed connectivity analysis methodologies have been proposed [60], the majority of studies focus on two classes of methods: effective connectivity (EC) analysis, and Granger causality (GC) analysis. GC reflects the amount of directional interactions (or "information flow" if interpreted according to the notion of transfer entropy), whereas EC quantifies both the amount and sign of the directed interaction. The state-of-the-art wisdom about using these two methods is somewhat contradictory. On one hand, methodological literature has often stressed the difference between GC and EC [27], insisting that i) they rest on different assumptions: EC aims to "retrieve

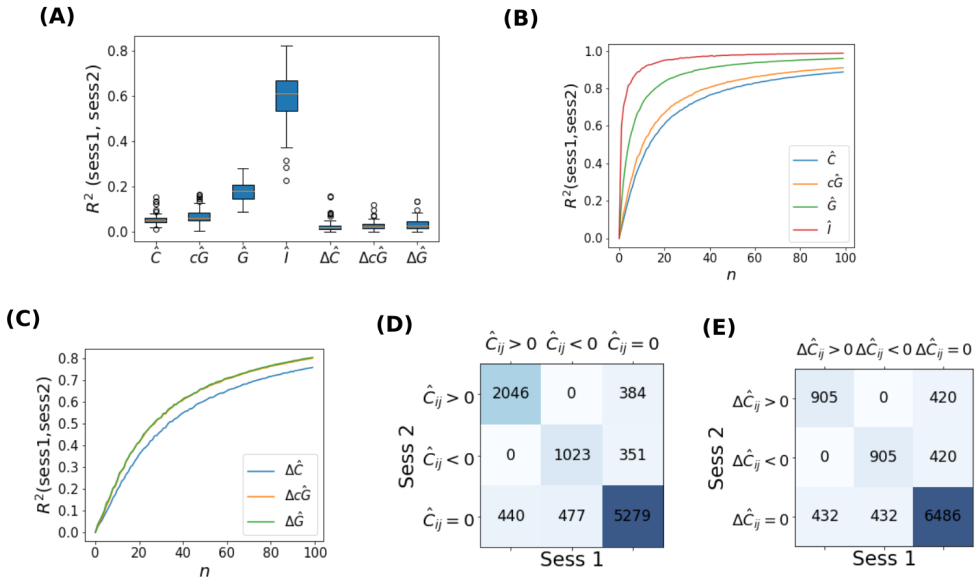

Fig 8. Consistency of results in two independent independent recording sessions **(A)** consistency of *single-subject* estimates across the two sessions (squared Pearson correlation). **(B)** Squared Pearson correlation R^2 between group estimates of $\hat{C}, c\hat{G}, \hat{G}, \hat{I}$ for groups of increasing size n. **(C)** Squared Pearson correlation R^2 between group estimates of $\Delta\hat{C}, \Delta c\hat{G}, \Delta\hat{G}, \hat{I}$ for groups of increasing size n. **(D)** Confusion matrix between the connection category identified in fMRI session 1 (significantly positive, significantly negative, non significant/null) and the category identified in fMRI session 2, at a group level ($n = 100$). **(E)** Confusion matrix between the connection asymmetry category identified in fMRI session 1 (significant asymmetry from $i \to j$, significant asymmetry from $j \to i$, non significant asymmetry) and the category identified in fMRI session 2, at a group level ($n = 100$).

the minimal circuit explaining the observed timing relationships", while GC is based on statistical dependencies ii) they differ in important technical details: EC often (but not always [13]) takes into account hemodynamics in the modeling, while GC often (but not always [61]) doesn't. On the other hand, in applications the two techniques are used quite interchangeably, in that the results of an EC or GC analysis on fMRI are often *interpreted* in the same way: EC or GC links are thought of as directed couplings between brain areas. In particular, many researchers in the field share an implicit intuition that large effective connections between two areas should be reflected into large values of Granger causality. In this work, we have emphasized the methodological similarity of the two approaches, which are based on common assumptions: essentially, that observed time series are generated by a linear process. In fact, at the theoretical level, if Gaussian noise is assumed (as it often happens for most GC and EC variants), we explicitly derived that GC and EC are mathematically linked by a monotonic (approximately quadratic) relation (Eqs.(17),(18) and Fig. 1).

With the aid of numerical simulations of artificial networks, we discussed to what extent the theoretical EC/GC relations can be observed from finite-length data (Fig. 2,4). The main limitation arises from the sampling rate and the length of empirical time series, in line with previous work [17,62]. For a network of 100 areas (a typical number for usual fMRI parcellation schemes), the theoretical relations can be clearly observed only for a sampling time of $\mathcal{L} = 10^4$, which amounts to 10 times the length of a typical recording session. In case of the typical sampling time of an fMRI session ($\mathcal{L} = 10^3$), the quadratic relation is satisfied quite roughly (correlation between C^2 and G yields $R^2 \approx 0.4$). Therefore, the relation can be neatly observed only for unusually long recordings (such as [63]), or, more plausibly, by concatenating the data of several subjects (neglecting inter-individual variability). At least 10 subjects (order of magnitude) should be concatenated. While the relations predict a concordance of EC/GC in the estimation of connection asymmetry, observing it in finite-length data is considerably difficult.For a network of 100 nodes, only a very large sampling time ($T = 10^5$) can allow observing a certain degree of consistency ($R^2 \approx 0.4$).

Importantly, we did not limit ourselves to testing the EC/GC relation on simulated network activity (as in [45, 60]. We also tested the similarity between EC and GC with real fMRI data, where some of the assumptions (e.g. Gaussianity of generated signals) are either only loosely or simply not satisfied. Our results show that GC and EC yield similar connectivity *at a group level*(Fig. 6), meaning that they provide consistent and reliable estimates for directional connectivity across brain regions when pooling at least 15 subjects together. To observed consistency between the two methods in the analysis of pairwise asymmetry, a group of 20 subjects or more is needed. These numbers are necessary to average the empirical session-to-session noise inherent to fMRI. In fact, we observed that individual EC and GC estimates are not reliable, as seen from a test-retest analysis (Fig. 8. We stress that lack of reliability concerns individual connection estimates, while reliable information may emerge at a network level. Indeed, it has been previously shown that EC can be used for robust prediction at the level of individual subjects [14, 59, 64]. In that case, the prediction power comes from using the whole network (all EC values jointly) for classification. In other words, there is an unavoidable trade-off: accuracy cannot be achieved at both individual session and individual connection level, unless sufficient data is collected over several subjects or connections.

The analytical relation between EC and GC derived for the simple case of covariance-based, first-order Granger causality highlights subtle difference between EC and GC, suggesting several important caveats to consider when interpreting both quantities. We now review the main possible incongruences between the two methods when analysing real fMRI data, where first-order GC is often used [17, 34] in parallel to EC.

The first difference stems from the non-monotonicity of the EC-GC relation, which is (approximately) quadratic (Eqs.(17),(18): both strongly positive and strongly negative effective weights between two areas are reflected in large values of Granger causality. Negative weights generally result from anticorrelations, whose presence in fMRI data has been hotly debated: negative FC can be artificially induced when using global signal regression [48, 65], but recent studies have shown that they it predicts brain states and cognitive functions [66, 67]. Here, we find negative EC values in rs-fMRI in approximately a third of connections consistently across subjects (Fig. 5) without global signal regression. This suggest that the analysis of brain networks using graph theoretical approaches, such as community detection or identification of hubs, would benefit from taking into account the presence of links displaying negative (i.e., "repulsive") effects that are not immediately indentified in GC analysis. The combination of GC and EC estimates in network analysis may be a promising direction for future studies. The quadratic relation between EC and GC also determines the agreement (or lack thereof) of the two methods in assessing the *asymmetry* of connections. The two methods agree only if the asymmetry for EC is defined by taking into account the modulus of the connection strength (Fig. 3): i.e., if a stronger negative $A \to B$ than $B \to A$ implies that the asymmetry is $A \to B$.

The second difference is that quadratic EC/GC relations hold only if all areas have equal signal variance, corresponding to the same level of random input noise affecting or received by areas or nodes). However, when the signal power (variance) differs across areas, the quadratic relation fails. As an example, for symmetric connections between two areas, the GC from the area with higher power will be larger than in the reverse direction. This is not necessary a pitfall of GC, that was conceived as a measure of "influence": in presence of equal connections, the influence of an area with larger power is stronger. However, this implies that GC cannot be considered a measure of coupling like EC. In "theoretical relation between EC and GC", we showed that in presence of non-homogeneous signal power, it is still possible to recover a quadratic relation between EC and GC/IC (Eqs.(19),(20)) by "correcting" GC estimates by a simple factors accounting for the heterogeneity in nodal input.

The third difference is that the validity of the implication large EC $\Rightarrow$ large GC depends on the relation between two time scales: the sampling time and the process time. The process time (τ) is the time scale of the stochastic ODE system underlying the observed time series, related to the autocorrelation time of the dynamical system. The sampling time (Δ) is the interval between two successive observations, which, in fMRI, corresponds to the repetition time. When the process is fast ($\tau \ll \Delta$), the values of the (discrete) observed time series at the previous time point are poorly predictive of the next time point. Thus, regardless of the strength of effective connections, GC remains low, in line with previous work [45]. The presence of strong effective connection still creates a strong statistical dependency between the two time series, but most of this interdependence appears as correlation that cannot be predicted on the basis of previous time points. This is instead captured by the "instantaneous (Granger) causality" (IC), a non-directional connectivity measure that is often neglected but is part of the original Granger-Geweke formalism [37]. In "theoretical relation between EC and GC", we also derived an analytical relation between EC and IC (20). While both measures are then proportional to the square of EC, for $\tau \ll \Delta$ the proportionality constant is much stronger for the IC than GC. Instead, when the process is slow, ($\tau \gg \Delta$), the values of the (discrete) observed time series at the previous time point are strongly predictive of the next time point. In this case, the presence of a large effective connection creates a strong statistical dependency between the two time series that can be predicted on the basis of previous time points, reflecting into large values of GC (while IC is very small). More generally, our study shows that large (in modulus) effective connections

reflect large values of Granger causality *or* instantaneous causality. Detecting the regime ($\tau \ll \Delta$ or $\tau \gg \Delta$) from data is not trivial. In principle, one can locate the regime based on the average GC/IC ratio, but this requires long sampling (Fig. S4Fig). More generally, our study stresses the importance of considering the regime at which the neural data are sampled. In this respect, a possible limitation of this study is the assumption of a diagonal noise covariance matrix in the MAR model: in presence of a large common input between regions, the noise should be modeled as having non-diagonal covariance Σ, which could reflect into larger values of the instantaneous Granger causality I. This could explain the anomalous IC values observed in fMRI data. Reliably inferring a non-diagonal noise covariance poses additional challenges [68,69]. Future work will better investigate the relation between C, Σ and G, I in case of correlated noise.

We finally mention a possible limitation of the application to fMRI in this study: we did not investigate or model hemodynamics. Assuming that the model (MOU/MAR) holds at the level of neural time series, the effect of hemodynamics can be multiple. Firstly, it can introduce deviations from the model at the level of observed BOLD time series (such that a MOU or first order MAR no longer yield an accurate description of the data). Secondly, regional differences in the hemodynamic response may bias timing relations observed at the BOLD level, hence biasing EC/GC estimates, in particular estimates of connection asymmetry. In principle, one should then estimate EC/GC from "neural" time series obtained after a deconvolution of the hemodynamic response. Further work is need to appreciate the effect of the hemodynamic response on the GC/EC relation, and the application of blind approaches (e.g., [61, 70]) to deconvolve the hemodynamic response from resting-state fMRI data.

To conclude, our study addressed the consistency of GC and EC in the reconstruction and analysis of directed brain networks from neuroimaging time series. In the context of the first-order autoregressive models typically used in fMRI, GC and EC share common assumptions and are mathematically related. The relation is non-trivial due to the presence of negative effective connections and unequal noise variance on different network nodes. When these factors are properly taken into account, the two methods yield a largely consistent description of directed brain networks at a group level.

Acknowledgments

M.A. and A.B. were supported by FLAG-ERA JTC 2017, grant ANR-17-HBPR-0001, "Brainsynch-Hit". MA was supported by the Italian Ministry for University, through the 2022 PRIN program, project 2022HSKLK9, "Unveiling the role of low dimensional activity manifolds in biological and artificial neural networks". MG received support from the French government under the France 2030 investment plan, under the Agence Nationale de la Recherche grant (ANR-22-CPJ2-0020-01) and as part of the Initiative d'Excellence d'Aix-Marseille Université – A*MIDEX (AMX-22-CPJ-01). MG was also partly supported by the European Union Horizon 2020 Grant No. 945539 (Human Brain Project SGA3). AB was supported by the Agence National de la Recherche, Grant ANR-18-CE28-0016 and by the European Union's Horizon 2020 Framework Programme for Research and Innovation under the Specific Grant Agreement No. 945539 (Human Brain Project SGA3).

References

1. Bassett DS, Bullmore ET. Human brain networks in health and disease. Current opinion in neurology. 2009;22(4):340.

2. Bressler SL, Menon V. Large-scale brain networks in cognition: emerging methods and principles. Trends in cognitive sciences. 2010;14(6):277–290.

3. Van Den Heuvel MP, Pol HEH. Exploring the brain network: a review on resting-state fMRI functional connectivity. European neuropsychopharmacology. 2010;20(8):519–534.

4. Power JD, Cohen AL, Nelson SM, Wig GS, Barnes KA, Church JA, et al. Functional network organization of the human brain. Neuron. 2011;72(4):665–678.

5. Uddin LQ, Yeo B, Spreng RN. Towards a universal taxonomy of macro-scale functional human brain networks. Brain topography. 2019;32(6):926–942.

6. Damoiseaux JS, Rombouts S, Barkhof F, Scheltens P, Stam CJ, Smith SM, et al. Consistent resting-state networks across healthy subjects. Proceedings of the national academy of sciences. 2006;103(37):13848–13853.

7. Fox MD, Snyder AZ, Vincent JL, Corbetta M, Van Essen DC, Raichle ME. The human brain is intrinsically organized into dynamic, anticorrelated functional networks. Proceedings of the National Academy of Sciences. 2005;102(27):9673–9678.

8. Fox MD, Greicius M. Clinical applications of resting state functional connectivity. Frontiers in systems neuroscience. 2010; p. 19.

9. Greicius M. Resting-state functional connectivity in neuropsychiatric disorders. Current opinion in neurology. 2008;21(4):424–430.

10. Finn ES, Shen X, Scheinost D, Rosenberg MD, Huang J, Chun MM, et al. Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity. Nature neuroscience. 2015;18(11):1664–1671.

11. Gratton C, Laumann TO, Nielsen AN, Greene DJ, Gordon EM, Gilmore AW, et al. Functional brain networks are dominated by stable group and individual factors, not cognitive or daily variation. Neuron. 2018;98(2):439–452.

12. Reid AT, Headley DB, Mill RD, Sanchez-Romero R, Uddin LQ, Marinazzo D, et al. Advancing functional connectivity research from association to causation. Nature neuroscience. 2019;22(11):1751–1760.

13. Gilson M, Moreno-Bote R, Ponce-Alvarez A, Ritter P, Deco G. Estimation of directed effective connectivity from fMRI functional connectivity hints at asymmetries of cortical connectome. PLoS computational biology. 2016;12(3):e1004762.

14. Gilson M, Zamora-López G, Pallarés V, Adhikari MH, Senden M, Campo AT, et al. Model-based whole-brain effective connectivity to study distributed cognition in health and disease. Network Neuroscience. 2020;4(2):338–373.

15. Brovelli A, Ding M, Ledberg A, Chen Y, Nakamura R, Bressler SL. Beta oscillations in a large-scale sensorimotor cortical network: Directional influences revealed by Granger causality. Proceedings of the National Academy of Sciences. 2004;101(26):9849–9854. doi:10.1073/pnas.0308538101.

16. Bressler SL, Seth AK. Wiener-Granger causality: a well established methodology. Neuroimage. 2011;58(2):323–329.

17. Seth AK, Barrett AB, Barnett L. Granger Causality Analysis in Neuroscience and Neuroimaging. The Journal of Neuroscience. 2015;35(8):3293–3297. doi:10.1523/jneurosci.4399-14.2015.

18. Granger C. Investigating Causal Relations by Econometric Models and Cross-Spectral Methods. Econometrica. 1969;37(3):424–38.

19. Schreiber T. Measuring Information Transfer. Physical Review Letters. 2000;85(2):461–464. doi:10.1103/physrevlett.85.461.

20. Massey JL. CAUSALITY, FEEDBACK AND DIRECTED INFORMATION; 1990.Available from: `https://api.semanticscholar.org/CorpusID:9433943`.

21. Roebroeck A, Formisano E, Goebel R. Mapping directed influence over the brain using Granger causality and fMRI. Neuroimage. 2005;25(1):230–242.

22. Goebel R, Roebroeck A, Kim DS, Formisano E. Investigating directed cortical interactions in time-resolved fMRI data using vector autoregressive modeling and Granger causality mapping. Magnetic resonance imaging. 2003;21(10):1251–1261.

23. Deshpande G, LaConte S, James GA, Peltier S, Hu X. Multivariate Granger causality analysis of fMRI data. Human brain mapping. 2009;30(4):1361–1373.

24. Barnett L, Seth AK. The MVGC multivariate Granger causality toolbox: a new approach to Granger-causal inference. Journal of neuroscience methods. 2014;223:50–68.

25. Liao W, Mantini D, Zhang Z, Pan Z, Ding J, Gong Q, et al. Evaluating the effective connectivity of resting state networks using conditional Granger causality. Biological cybernetics. 2010;102(1):57–69.

26. Friston KJ. Functional and effective connectivity: a review. Brain connectivity. 2011;1(1):13–36.

27. Friston KJ, Harrison L, Penny W. Dynamic causal modelling. Neuroimage. 2003;19(4):1273–1302.

28. Razi A, Kahan J, Rees G, Friston KJ. Construct validation of a DCM for resting state fMRI. Neuroimage. 2015;106:1–14.

29. Prando G, Zorzi M, Bertoldo A, Corbetta M, Zorzi M, Chiuso A. Sparse DCM for whole-brain effective connectivity from resting-state fMRI data. NeuroImage. 2020;208:116367. doi:10.1016/j.neuroimage.2019.116367.

30. Frässle S, Lomakina EI, Kasper L, Manjaly ZM, Leff A, Pruessmann KP, et al. A generative model of whole-brain effective connectivity. Neuroimage. 2018;179:505–529.

31. Gilson M, Deco G, Friston KJ, Hagmann P, Mantini D, Betti V, et al. Effective connectivity inferred from fMRI transition dynamics during movie viewing points to a balanced reconfiguration of cortical interactions. Neuroimage. 2018;180:534–546.

32. Brovelli A, Chicharro D, Badier JM, Wang H, Jirsa V. Characterization of cortical networks and corticocortical functional connectivity mediating arbitrary visuomotor mapping. Journal of Neuroscience. 2015;35(37):12643–12658.

33. Brovelli A, Badier JM, Bonini F, Bartolomei F, Coulon O, Auzias G. Dynamic reconfiguration of visuomotor-related functional connectivity networks. Journal of Neuroscience. 2017;37(4):839–853.

34. Allegra M, Favaretto C, Metcalf N, Corbetta M, Brovelli A. Stroke-related alterations in inter-areal communication. NeuroImage: Clinical. 2021;32:102812.

35. Van Essen DC, Smith SM, Barch DM, Behrens TE, Yacoub E, Ugurbil K, et al. The WU-Minn human connectome project: an overview. Neuroimage. 2013;80:62–79.

36. Smith SM, Beckmann CF, Andersson J, Auerbach EJ, Bijsterbosch J, Douaud G, et al. Resting-state fMRI in the human connectome project. Neuroimage. 2013;80:144–168.

37. Geweke J. Measurement of linear dependence and feedback between multiple time series. Journal of the American statistical association. 1982;77(378):304–313.

38. Lütkepohl H. New introduction to multiple time series analysis. Springer Science & Business Media; 2005.

39. On the spectral formulation of Granger causality. Biological cybernetics. 2011;105:331–347.

40. Barnett L, Barrett AB, Seth AK. Granger causality and transfer entropy are equivalent for Gaussian variables. Physical review letters. 2009;103(23):238701.

41. Ince RA, Giordano BL, Kayser C, Rousselet GA, Gross J, Schyns PG. A statistical framework for neuroimaging data analysis based on mutual information estimated via a gaussian copula. Human brain mapping. 2017;38(3):1541–1573.

42. Glasser MF, Sotiropoulos SN, Wilson JA, Coalson TS, Fischl B, Andersson JL, et al. The minimal preprocessing pipelines for the Human Connectome Project. Neuroimage. 2013;80:105–124.

43. Schaefer A, Kong R, Gordon EM, Laumann TO, Zuo XN, Holmes AJ, et al. Local-global parcellation of the human cerebral cortex from intrinsic functional connectivity MRI. Cerebral cortex. 2018;28(9):3095–3114.

44. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal statistical society: series B (Methodological). 1995;57(1):289–300.

45. Seth AK, Chorley P, Barnett LC. Granger causality analysis of fMRI BOLD signals is invariant to hemodynamic convolution but not downsampling. Neuroimage. 2013;65:540–555.

46. Lin FH, Ahveninen J, Raij T, Witzel T, Chu YH, Jääskeläinen IP, et al. Increasing fMRI sampling rate improves Granger causality estimates. PloS one. 2014;9(6):e100319.

47. Yeo BT, Krienen FM, Sepulcre J, Sabuncu MR, Lashkari D, Hollinshead M, et al. The organization of the human cerebral cortex estimated by intrinsic functional connectivity. Journal of neurophysiology. 2011;.

48. Murphy K, Birn RM, Handwerker DA, Jones TB, Bandettini PA. The impact of global signal regression on resting state correlations: are anti-correlated networks introduced? Neuroimage. 2009;44(3):893–905.

49. Liu TT, Nalci A, Falahpour M. The global signal in fMRI: Nuisance or Information? Neuroimage. 2017;150:213–229.

50. Murphy K, Fox MD. Towards a consensus regarding global signal regression for resting state functional connectivity MRI. Neuroimage. 2017;154:169–173.

51. Zeki S, Shipp S. The functional logic of cortical connections. Nature. 1988;335(6188):311–317.

52. Gazzaniga MS. Cerebral specialization and interhemispheric communication: does the corpus callosum enable the human condition? Brain. 2000;123(7):1293–1326.

53. Jung K, Friston KJ, Pae C, Choi HH, Tak S, Choi YK, et al. Effective connectivity during working memory and resting states: A DCM study. NeuroImage. 2018;169:485–495.

54. Cai W, Ryali S, Pasumarthy R, Talasila V, Menon V. Dynamic causal brain circuits during working memory and their functional controllability. Nature Communications. 2021;12(1):3314.

55. Nee DE, D'Esposito M. The hierarchical organization of the lateral prefrontal cortex. Elife. 2016;5:e12112.

56. Rolls ET, Deco G, Huang CC, Feng J. The human language effective connectome. NeuroImage. 2022;258:119352.

57. Ajmera S, Jain H, Sundaresan M, Sridharan D. Decoding task-specific cognitive states with slow, directed functional networks in the human brain. Eneuro. 2020;7(4).

58. Panda R, López-González A, Gilson M, Gosseries O, Thibaut A, Frasso G, et al. Whole-brain analyses indicate the impairment of posterior integration and thalamo-frontotemporal broadcasting in disorders of consciousness. Human Brain Mapping. 2023;.

59. Adhikari MH, Griffis J, Siegel JS, Thiebaut de Schotten M, Deco G, Instabato A, et al. Effective connectivity extracts clinically relevant prognostic information from resting state activity in stroke. Brain communications. 2021;3(4):fcab233.

60. Smith SM, Miller KL, Salimi-Khorshidi G, Webster M, Beckmann CF, Nichols TE, et al. Network modelling methods for FMRI. Neuroimage. 2011;54(2):875–891.

61. Wu GR, Liao W, Stramaglia S, Ding JR, Chen H, Marinazzo D. A blind deconvolution approach to recover effective connectivity brain networks from resting state fMRI data. Medical image analysis. 2013;17(3):365–374.

62. Barnett L, Seth AK. Detectability of Granger causality for subsampled continuous-time neurophysiological processes. Journal of neuroscience methods. 2017;275:93–121.

63. Gordon EM, Laumann TO, Gilmore AW, Newbold DJ, Greene DJ, Berg JJ, et al. Precision functional mapping of individual human brains. Neuron. 2017;95(4):791–807.

64. Pallarés V, Insabato A, Sanjuán A, Kühn S, Mantini D, Deco G, et al. Extracting orthogonal subject-and condition-specific signatures from fMRI data using whole-brain effective connectivity. NeuroImage. 2018;178:238–254.

65. Weissenbacher A, Kasess C, Gerstl F, Lanzenberger R, Moser E, Windischberger C. Correlations and anticorrelations in resting-state functional connectivity MRI: a quantitative comparison of preprocessing strategies. Neuroimage. 2009;47(4):1408–1416.

66. Keller JB, Hedden T, Thompson TW, Anteraper SA, Gabrieli JD, Whitfield-Gabrieli S. Resting-state anticorrelations between medial and lateral prefrontal cortex: association with working memory, aging, and individual differences. Cortex. 2015;64:271–280.

67. Nir T, Jacob Y, Huang KH, Schwartz AE, Brallier JW, Ahn H, et al. Resting-state functional connectivity in early postanaesthesia recovery is characterised by globally reduced anticorrelations. British Journal of Anaesthesia. 2020;125(4):529–538.

68. Tam H, Ching ES, Lai PY. Reconstructing networks from dynamics with correlated noise. Physica A: Statistical Mechanics and its Applications. 2018;502:106–122.

69. Cheng CH, Lai PY, et al. Efficient reconstruction of directed networks from noisy dynamics using stochastic force inference. Physical Review E. 2022;106(3):034302.

70. Wu GR, Deshpande G, Laureys S, Marinazzo D. Retrieving the hemodynamic response function in resting state fMRI: methodology and application. In: 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE; 2015. p. 6050–6053.

Supporting information

S1Text: Approximate formulas for GC, IC

If Eq.(15) holds, we have quadratic relations between IC, GC and EC. Indeed, $A_{ij} \simeq \Delta e^{-\frac{\Delta}{\tau}} C_{ji}$ and

$$G_{ij} = \log(1 + \Delta^2 \frac{e^{-\frac{2\Delta}{\tau}} C_{ji}^2 Q_{ii}^0}{S_{jj}}) \simeq \Delta^2 \frac{e^{-\frac{2\Delta}{\tau}} C_{ji}^2 Q_{ii}^0}{S_{jj}} \tag{22}$$

Under the same approximation we also have $e^{Ct} \simeq \mathbb{I} + Ct$ for $t \leq \Delta$ so

$$S = \int_0^\Delta dt\, e^{Jt} \Sigma e^{J^T t} \simeq \int_0^\Delta dt\, e^{-\frac{2t}{\tau}} (\mathbb{I} + Ct) \Sigma (\mathbb{I} + C^T t) \simeq \int_0^\Delta dt\, e^{-\frac{2t}{\tau}} (\Sigma + C\Sigma t + \Sigma C^T t) =$$

$$= \frac{\tau}{2} \Big((1 - e^{-\frac{2\Delta}{\tau}}) \Sigma + (-\Delta e^{-\frac{2\Delta}{\tau}} + \frac{\tau}{2}(1 - e^{-\frac{2\Delta}{\tau}}))(C\Sigma + \Sigma C^T) \Big) =$$

$$= \frac{\tau}{2}(1 - e^{-\frac{2\Delta}{\tau}}) \Big(\Sigma + (\frac{\tau}{2} - \Delta \frac{e^{-\frac{2\Delta}{\tau}}}{1 - e^{-\frac{2\Delta}{\tau}}})(\Sigma C + \Sigma C^T) \Big)$$

Note that $[\Sigma C]_{ij} = \sigma_i^2 C_{ij}$, $[C^T \Sigma]_{ij} = \sigma_j^2 C_{ji}$. Hence,

$$S_{ii} = \sigma_i^2 \frac{\tau}{2}(1 - e^{-\frac{2\Delta}{\tau}}), \qquad S_{ij} = \frac{\tau}{2}(1 - e^{-\frac{2\Delta}{\tau}}) \Big((\frac{\tau}{2} - \Delta \frac{e^{-\frac{2\Delta}{\tau}}}{1 - e^{-\frac{2\Delta}{\tau}}})(\sigma_i^2 C_{ij} + \sigma_j^2 C_{ji})$$

Furthermore, from Lyapunov equation one gets

$$\sum_{kl} A_{ik} Q_{kl}^0 A_{li} - Q_{ii}^0 + S_{ii} = 0$$

and to 1st order ($A \approx e^{-\frac{-\Delta}{\tau}}$)

$$Q_{ii}^0 (1 - e^{-\frac{-2\Delta}{\tau}}) \simeq S_{ii} \qquad \Rightarrow \qquad Q_{ii}^0 \simeq \frac{S_{ii}}{1 - e^{-\frac{2\Delta}{\tau}}} = \sigma_i^2 \frac{\tau}{2}$$

which yields, using the previous expression of S_{ii},

$$G_{ij} \simeq \Delta^2 \frac{C_{ji}^2}{e^{\frac{2\Delta}{\tau}} - 1} \frac{\sigma_i^2}{\sigma_j^2} \tag{23}$$

For the instantaneous causality, one obtains

$$I_{ij} \simeq \frac{S_{ij}^2}{S_{ii}S_{jj}} \simeq \tau^2\big(1 - \frac{2\Delta}{\tau}\frac{e^{-\frac{2\Delta}{\tau}}}{1-e^{-\frac{2\Delta}{\tau}}}\big)^2 \tilde{C}_{ij}{}^2 \tag{24}$$

with $\tilde{C}_{ij} = \frac{\sigma_i^2 C_{ij} + \sigma_j^2 C_{ji}}{2\sigma_i\sigma_j}$.

Supplementary figures

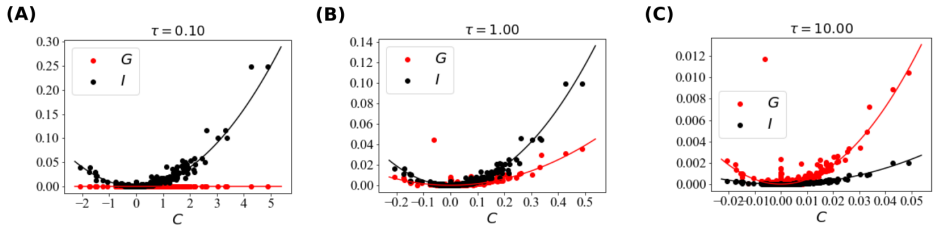

Fig S1Fig. Theoretical relations between model EC and non-conditional GC/IC. We considered a random network of $N = 40$ nodes evolving according to the MOU of $\tau = 0.1, 1, 10$. In **(A-C)** we show the relation between the EC weights and the corresponding values of GC and IC. The GC and IC were obtained from the ideal covariance matrices Q^0_∞, Q^1_∞ obtained in the limit of infinite observation length, $T \to \infty$. Each dot corresponds to a pair (i, j), and straight lines to the approximate quadratic scalings, $GC \sim \frac{C^2}{e^{2/\tau}-1}$, $IC \sim \tau^2 C^2 \big(1 - \frac{2/\tau}{e^{2/\tau}-1}\big)^2$. For all valued of τ, the approximate quadratic relation are well satisfied. The relative importance of IC vs. GC depends on the value of τ, with IC prevailing at low τ and GC at large τ.

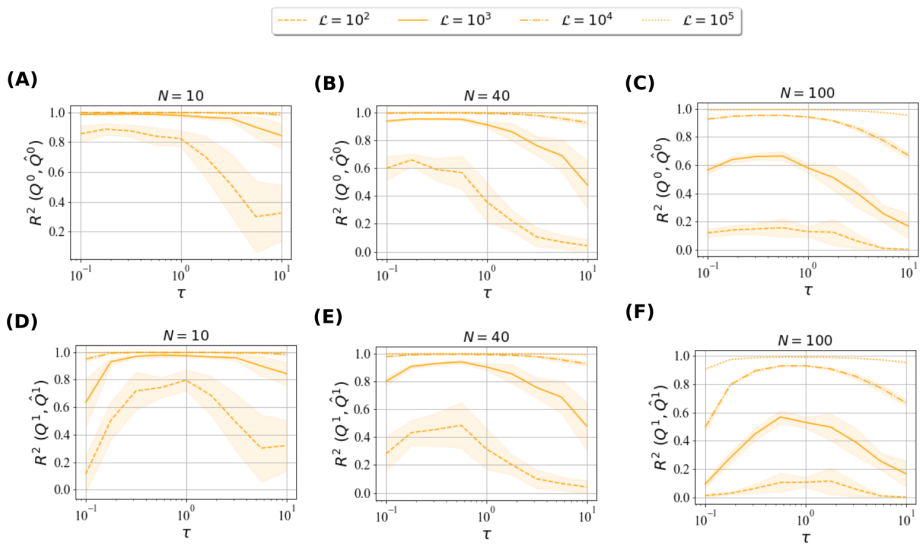

Fig S2Fig. Error in covariance matrices for finite sampling time. We considered the same random network as in Fig. 1. In real cases $T < \infty$, the covariance matrices Q^0, Q^1 are affected by an estimation error (whose magnitude decreases with T). We show here the relation between the ideal Q^0, Q^1 and the empirical $Q^{0,emp}, Q^{1,emp}$, for different values of τ and T, measured in terms of squared Pearson correlation R^2. R^2 is an increasing function of T. Note that R^2 is not a monotonic function of τ, but reaches a maximum for $\tau \approx 1$, i.e., when sampling time and process time are of the same order.

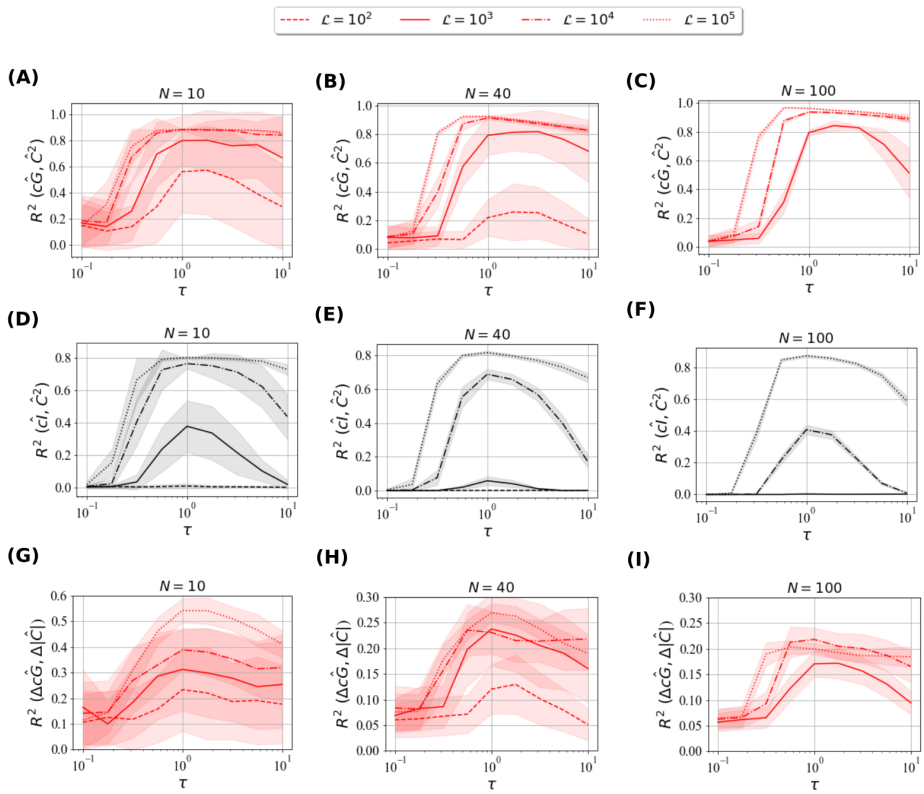

Fig S3Fig. Relations between empirical EC and empirical conditional GC/IC for finite sampling time. We considered the same random network as in Fig. 1, which we simulated for different values of τ and T. In real cases $T < \infty$, the covariance matrices $Q^{0,emp}, Q^{1,emp}$ are affected by an estimation error (whose magnitude decreases with T). In the presence of estimation noise, the relation between EC and GC, IC become much less tight. For each T,τ, we fitted GC and IC as a quadratic function of EC_0 (the ideal EC). The R^2 of the fit is an increasing function of T, with relatively low values of R^2 obtained for $T \leq 1000$. Estimation noise also affects the IC/GC relation as a function of τ. In panel (g-i) we show the average ratio GC/IC (averaged over all pairs i, j) as a function of τ for different values of T. The monotonic increase of GC/IC as a function of τ becomes less sharp for low values of T.

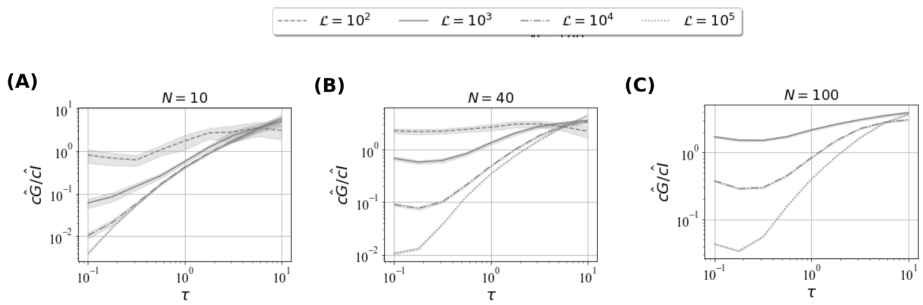

Fig S4Fig. Average ratio between conditional Granger measures (GC/IC) for finite sampling time. We considered the same random network as in Fig. 1, which we simulated for different values of τ (x-axis) for $\mathcal{L} \in \{10^2, 10^3, 10^4, 10^5\}$ (dashed, solid, dashed-dotted, and dotted, resp.). The columns correspond to several network sizes, from $N = 10$ to 100. **(a-c)** Average ratio cG/cI (over all pairs i, j) as a function of τ. The monotonic increase of cG/cI as a function of τ becomes sharper for larger values of $\mathcal{L}$.

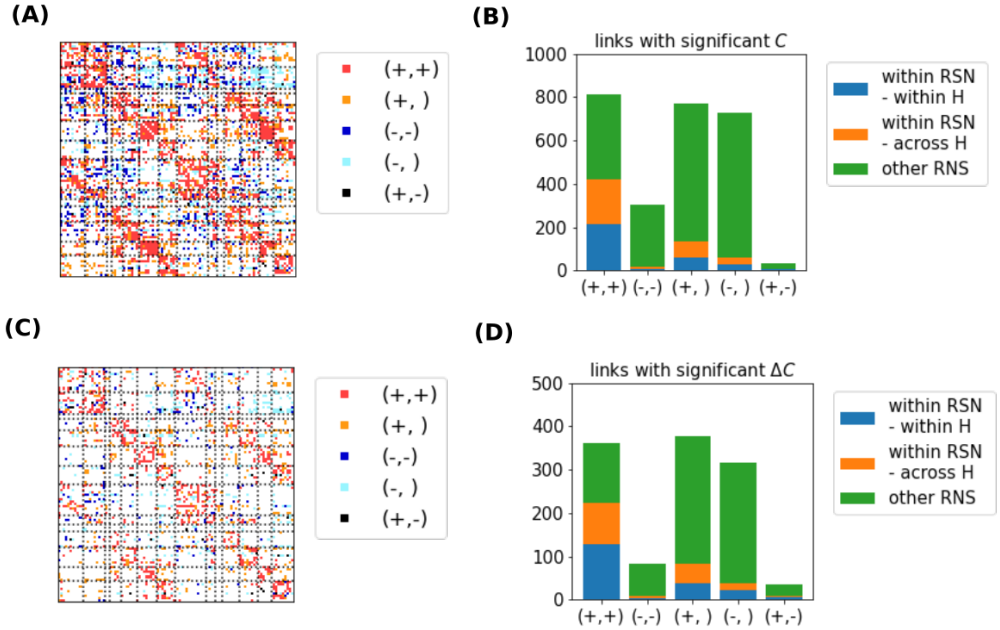

Fig S5Fig. Effective Connection sign and asymmetries. We considered all pairs of reciprocal connections, and divided them into *sign categories* according to their sign: $(+,+)$ both connections are significantly positive $(+,)$ one connection is significantly positive and the reciprocal is non-significant $(-,)$ both connections are significantly negative $(-,)$ one connection is significantly negative and the reciprocal is non-significant $(+,-)$ one connection is significantly positive and the reciprocal is significantly negative. Furthermore, connections pairs were divided in *network categories* depending on the connected areas: i) areas belonging to the same RSN and the same hemisphere ii) areas belonging to the same RSN but different hemispheres iii) areas belonging to different RSNs. **(A)** We show all significant link pairs, with color depending on their sign category **(B)** For each sign category, we show the fraction of significant connections belonging to each network category. **(C)** We show all link pairs with significant asymmetry, with color depending on their sign category **(D)** For each sign category, we show the fraction of connections with significant asymmetry belonging to each network category.

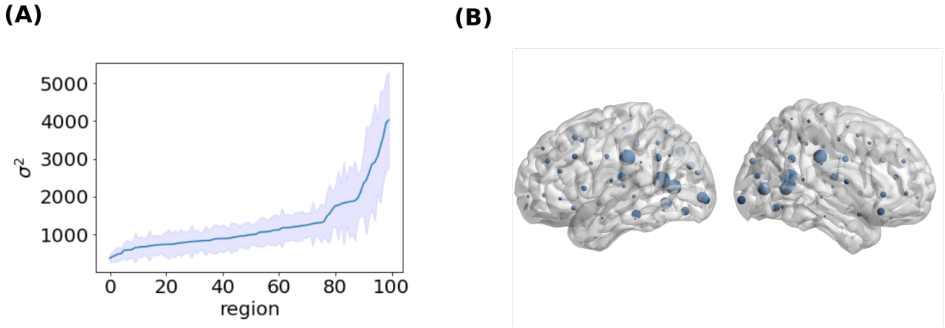

Fig S6Fig. Noise affecting different nodes. **(A)** For each region, we computed the average (over subjects) noise σ^2 and sorted regions by this value. The blue line represents the average (over subjects) noise σ^2, the shaded area represents standard deviation (over subjects). **(B)** In this rendering, the node size is proportional to the average (over subjects) noise σ^2.

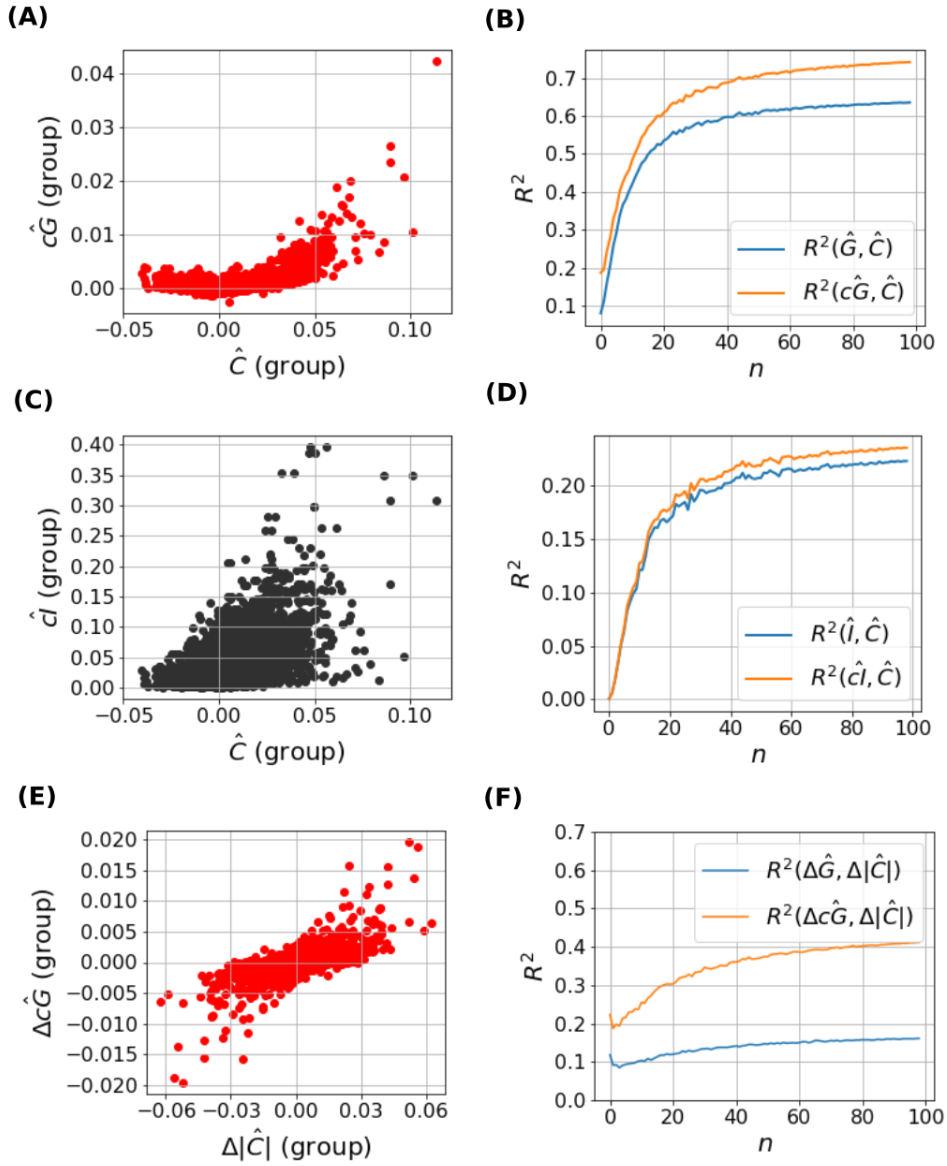

Fig S7Fig. Replication of fig. 6 with independent recording sessions **(A)** effective connectivity $\hat{C}$ (group average over 100 subjects) vs corrected Granger causality $\hat{cG}$ (group average). **(B)** squared Pearson correlation R^2 between group $\hat{C}$ and group $\hat{cG}$ for groups of increasing size n. **(C)** effective connectivity $\hat{C}$ (group average over 100 subjects) vs corrected instantaneous Granger causality $\hat{cI}$ (group average). **(D)** squared Pearson correlation R^2 between group $\hat{C}$ and group $\hat{cI}$ for groups of increasing size n. **(E)** asymmetry in effective connectivity $\Delta\hat{C}$ (group average over 100 subjects) vs asymmetry in corrected Granger causality $\Delta\hat{cG}$ (group average). **(F)** squared Pearson correlation R^2 between group $\Delta\hat{C}$ and group $\Delta\hat{cI}$ for groups of increasing size n.