Directed neural interactions in whole-brain resting-state fMRI: a comparison between Granger Causality and Effective Connectivity

April 27, 2026

Michele Allegra[1, 2], Matthieu Gilson[3, *], Andrea Brovelli[3, *]
1 Dipartimento di Fisica e Astronomia "Galileo Galilei", University of Padova, Padova, Italy

2 Padova Neuroscience Center, University of Padova, Padova, Italy
3 Institut de Neurosciences de la Timone, Aix Marseille Université, UMR 7289 CNRS, 13005, Marseille, France

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

Abstract

A key challenge for network neuroscience is to understand the role of interactions between brain regions and how they contribute to the encoding and broadcasting of information within cognitive processes. This requires computational tools to infer directional relations between brain regions from neural time series. For fMRI, the most common approaches are based on Granger causality (GC) analysis and effective connectivity (EC) models. Despite their different conceptual framing, GC and EC models for fMRI are based on similar mathematical assumptions, grounded on continuous- and discrete-time linear stochastic models. Based on a mapping between multivariate Ornstein-Uhlenbeck (MOU) and multivariate autoregressive (MVAR) processes, we analytically obtain an approximately quadratic relation between EC and GC, after rescaling the GC to compensate for unequal noise variances of the source and target. Simulations show that these relations can be observed in finite time series only if a large amount of data is available, implying that they may emerge only at a group level in real fMRI experiments. We verified this prediction by systematically comparing EC and GC in a large-scale fMRI data set from the Human Connectome Project. Overall, our findings can provide methodological and interpretational guidance in the usage of GC and EC for brain network reconstruction by clearly elucidating what is common between the two methods, but also their respective specificities and biases.

Introduction

A central hypothesis in neuroscience posits that cognitive functions arise from the coordinated activity of neural populations distributed over large-scale brain networks [1, 2] and from the emergence of network-wide and self-organized information routing patterns [3, 4]. Therefore, the estimation of interactions between brain regions is a key challenge for modern neuroscience, and indispensable for understanding the encoding and broadcasting of information relevant for cognitive processes in

the brain. IN recent decades, functional magnetic resonance imaging (fMRI) has been increasingly used to study large-scale brain networks, due to its fine spatial resolution [5,6]. The most widely used approach to identify brain networks from ongoing activity observed by resting-state fMRI (rs-fMRI) [7] is based on functional connectivity (FC), i.e., the correlation of the rs-fMRI signals of pairs of brain areas [8]. Despite its simplicity, FC analysis leads to reproducible and reliable findings [9] about brain network activation during cognitive tasks, alteration due to neuropathologies, and specificity for individuals [10]. It has provided a considerable wealth of information about the large-scale organization of brain activity, its relation to the underlying anatomy, and the neural signatures of individual behavioral traits [10–16].

However, FC methods do not provide information about potential directional or asymmetric interactions between brain areas and are highly sensitive to third-party effects. To bypass these limitations, there exist several methods that exploit the temporal structure of ongoing fMRI signals to capture direct (non-mediated) and directional connections between brain areas [17]. Although a wide array of methods have been proposed in the literature [18], most applications rely on two classes of methods.

The first class is based on on the Wiener-Granger principle [1,19,20], which equates the presence of a directional connection with the capability of a time series to forecast another. The degree of forecasting is quantified information-theoretically, via the transfer entropy [21], or assuming an autoregressive model, via *Granger causality* (GC) [19]. GC is interpreted in terms of influence or information flow between brain regions, and is equivalent to TE for Gaussian variables [22]. Applied to rs-fMRI, most studies assume a first-order multivariate autoregressive (MVAR) model to calculate GC [23–31].

The second class relies on the concept of *effective connectivity* (EC), assuming a generative network model whose dynamics are optimized through the model connectivity to reproduce statistics of empirical data [8]. This concept was widely popularized through dynamic causal modeling (DCM) [32]. To address whole-brain rs-fMRI connectivity, several simplifications of the standard DCM are required, allowing for the analysis of large networks ($n \gtrsim 50$ areas; $n^2 \gtrsim 2500$ connections) in absence of external input (i.e., of events that allow for an easy reconstruction of the hemodynamic response function). The most popular approach, regression DCM (rDCM [33]) assumes linear generative dynamics and a linearized hemodynamic response [34] (for a similar approach, see also [35]). The estimated EC weights are still interpreted as the strength and sign of directional interactions between neuronal populations, as reflected by the corresponding time series. An alternative approach for whole-brain EC modelling is based on the Lyapunov optimization of multivariate Ornstein-Uhlenbeck processes (MOU-EC) [36, 37]. In this approach, hemodynamics is neglected and EC weights are inferred via gradient descent from non-lagged and lagged cross-covariances. In fact, rDCM and MOU-EC mainly differ in terms of inference methodology (Lyapunov optimization vs. variational Bayes), rather than in the underlying generative model. Another popular EC model [38, 39] considers each area as a weakly nonlinear oscillator coupled to other areas through EC, which is estimated by heuristically optimizing model parameters to best reproduce observed cross-correlations. Whole-brain EC is commonly used to uncover sub-networks and hierarchies of brain regions specific to cognitive tasks or pathological conditions, both at the group level or at the single-subject level [40–44].

On one hand, these two classes of methods provide conceptually different metrics. GC is a non-signed measure of *influence*, quantifying the amount of directional interaction between nodes. Instead, EC is a signed metric depicting the effect of a local activity increase on downstream regions, which can be both positive ('excitatory') or negative ('inhibitory'). On the other hand, the two classes of methods are based on similar mathematical assumptions. Although GC is usually seen as data-driven, most applications actually assume a generative model in the form of an MVAR, which is closely related to linear models used for EC estimation.

In this study, we set out to thoroughly discuss the consistency between these two classes of methods in the context of fMRI. Our first main goal was to trace the common mathematical formalism under-

lying popular approaches for GC and EC in the context of fMRI. This has led us to uncover analytic relations between MVAR-based GC and MOU-based EC. Although these relations are derived for specific versions of GC and EC, and they are expected to hold only in the limit of vanishing sampling noise, they nevertheless provide a clear theoretical framework for discussing GC-EC alignment. The second main goal was to investigate the GC-EC alignment in realistic conditions (in terms of time series length, which determines that amount of sampling noise). This step was accomplished through numerical experiments and served to properly interpret the alginment/misalignment observed in real data. Finally, the third goal was to investigate the EC-GC alignment in real fMRI data. To this end, we leveraged the standard rs-fMRI data provided by the Human Connectome Project [45,46]. In short, our analysis will reveal that a consistency between EC and GC, both in terms of connection strength and connection asymmetry, can be obtained only at a group level, and when properly accounting for unequal node variances and the existence of signed connections. Overall, our analysis is intended to provide methodological guidance in the usage of GC and EC for brain network reconstruction, by clearly elucidating what is common between the two methods, but also their respective specificities and biases.

Materials and Methods

Multivariate Ornstein–Uhlenbeck (MOU) and autogressive (MVAR) models

In this work, we set out to discuss the relation between Granger Causality (GC) and Effective Connectivity (EC) approaches to characterize directed interactions in fMRI. Our first step is therefore to show that, under specific assumptions, one can find *analytic relations* between effective connectivity and Granger causality. These relations may hold, up to some approximation, also when the assumptions are relaxed. Therefore, a discussion of these relations should be the starting point if one wishes to correctly frame a systematic comparison between GC and EC.

The core of EC methods is assuming that neural signals are generated by a continuous-time dynamical system with unknown inter-areal couplings. The observed (discrete) time series are assumed to correspond to observations of the continuous system taken with a given sampling time Δt (corresponding to the repetition time for fMRI), and the methods try to reconstruct the couplings from the available data. The resulting estimates are called *effective connections*. Different methods differ depending on the choice of dynamical system and inference method, but a large majority assume a linear dynamical system. Here we focus our attention on MOU-EC method [36], which maximally facilitates discussion of the GC-EC relation. This method assumes that the neural time series $\mathbf{x}(t)$ are produced by a 1st order, linear, Gaussian, stochastic dynamical system, i.e., a multivariate Ornstein-Uhlenbeck process (MOU) given by

$$d\mathbf{x} = J\ \mathbf{x}(t)dt + d\mathbf{w} \tag{1}$$

where J is the system's Jacobian and w a noise term. The term 'Jacobian' is used here because linearization of any dynamical system with noise $d\mathbf{x} = \mathbf{f}(\mathbf{x}(t))dt + d\mathbf{w}$ will lead to an equation like (1), with $J = \partial\mathbf{f}/\partial\mathbf{x}$ the Jacobian of the vector field $\mathbf{f}$. The MOU-EC method assumes that $J = \text{diag}(-\frac{1}{\tau_i}) + C$, i.e., the Jacobian includes a diagonal term with typically homogeneous or nearly homogeneous time constant $\tau_i \approx \tau$, and an off-diagonal connectivity term (C). The main role of the diagonal term is to stabilize dynamics, which (in the absence of external input or noise) would produce a relaxation towards a quiescent state ($\mathbf{x} = 0$) with a typical time scale τ. Even in the presence of noise, τ effectively determines the typical autocorrelation time of the dynamics. The off-diagonal terms determine the strength of inter-areal couplings in the model, and they can be both positive (effectively 'excitatory' interaction) and negative (effectively 'inhibitory' interaction). The noise is assumed to be white, i.e., $d\mathbf{w}$ is a Wiener process with covariance matrix $\langle d\mathbf{w} d\mathbf{w}^T\rangle = \Sigma\ dt$.

GC methods assume that the observed (discrete) time series, with sampling time Δt, are generated by a multivariate autoregressive (MVAR) process. In the scope of fMRI, due to the slow sampling frequency, a common choice is using a 1st order multivariate MVAR process

$$\mathbf{x}(t+\Delta t) = A\ \mathbf{x}(t) + \boldsymbol{\epsilon}(t) \tag{2}$$

where $\boldsymbol{\epsilon}(t) \sim \mathcal{N}(0, S)$ is a Gaussian noise with zero mean and covariance matrix S. To quantify the Granger-causal effect of node i on node j, one can measure the relevance of $x_i(t)$ in predicting $x_j(t+\Delta t)$. To this aim, one can define a a "reduced" MVAR process where the influence of node i is removed:

$$\mathbf{x}(t+\Delta t) = A'\mathbf{x}_{\neg i}(t) + \boldsymbol{\epsilon}'(t) \tag{3}$$

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'. The Granger causal effect of node i on node j is given by the log-ratio of the variances,

$$G_{ij} = \log\left(\frac{S'_{jj}}{S_{jj}}\right) \tag{4}$$

Note that we consider the conditional version of GC, which means that the linear regression in Eq. (3) includes all remaining nodes in the network. To compute GC, one should estimate the variances S'_{jj} and S'_{jj} from the observed time series. In addition ((4)), Geweke [47] defined the 'instantaneous' Granger causality I_{ij}, a non-directional connectivity metric that measures the correlations in the innovations ϵ_i and ϵ_j:

$$I_{ij} = -\log(1 - R^2_{ij}), \qquad R^2_{ij} = \frac{S^2_{ij}}{S_{ii}S_{jj}} \tag{5}$$

The name 'instantaneous' comes from the fact that, from the point of view of the observer, these correlations cannot be predicted autoregressively, i.e., from past values of the time-series, and thus appear as instantaneous. Although the instantaneous causality is often discarded in Granger causality analysis, it may capture a large part of the interdependence between two time series.
As linear-feedback dynamic systems, the Ornstein-Uhlenbeck process and the 1st order MVAR rely on mutually consistent assumptions about the nature of the underlying network process, only differing by the continuous vs. discrete nature of time. As shown in the Appendix, the respective connectivity terms (J and A) can be formally related, and this translates into a mathematical relationship between Granger causality estimated from the MVAR and the EC estimated from the MOU.

Regression DCM

For resting-state fMRI, alongside MOU-EC, one of the most used methods to infer EC is Regression DCM [33, 48]. This is a computationally efficient variant of DCM designed for whole-brain network analysis. Briefly, rDCM incorporates several key modifications to the standard DCM framework that enable large-scale connectivity estimation involving the whole brain. The approach distinguishes between the neural time series $\mathbf{x}(t)$ and the observed BOLD signal $\mathbf{y}(t)$, which is assumed to be related to the neural signal through convolution with a hemodynamic filter. rDCM assumes a linear dynamic model and a linearized observation model,

$$\dot{\mathbf{x}} = J\ \mathbf{x}(t) + C\mathbf{u}(t), \tag{6a}$$

$$\mathbf{y}(t) = (\mathbf{h} \otimes \mathbf{x})(t) + \mathbf{v}(t) \tag{6b}$$

where $\otimes$ denotes a convolution, $\mathbf{u}(t)$ is external input, $\mathbf{h}$ is a (linearized) hemodynamic filter, and $\mathbf{v}$ is Gaussian observation noise. Note that $\mathbf{y}$ is observed only at discrete times t, $t+\Delta t$, $t+2\Delta t$, etc. When modeling ongoing activity, the external input is absent, $\mathbf{u}(t)=0$. Passing to the frequency domain and discretizing, the system becomes

$$(\mathrm{diag}(e^{2\pi i\mathbf{k}}) - \mathbb{I})\ \tilde{\mathbf{y}} = J\ \tilde{\mathbf{y}} + \mathbf{v}' \tag{7}$$

where $\tilde{\mathbf{y}}$ denotes the discrete Fourier transform, $\mathbf{k} = (\frac{1}{N}, \frac{2}{N}, \ldots, \frac{N-1}{N})$, and $\mathbf{v}'$ is an (in principle correlated) Gaussian noise term. The simplifying assumption is made that noise affecting different regions is independent, $\mathbf{v}' \sim \mathcal{N}(0, \mathrm{diag}(\boldsymbol{\rho}^2))$ where $\boldsymbol{\rho}$ is the vector of noise variances.
Given Eq. (7), the estimation of J can proceed via Bayesian linear regression. The assumption of independent noise inputs translates into the fact that the posterior factorizes into the product of independent terms for afferent connections to each region, allowing for efficient inference. Note that Eq. (7) is completely independent of $\mathbf{h}$ (this would not hold in the presence of external inputs, $\mathbf{u}(t) \neq 0$). In other words, we could safely replace $\mathbf{h}$ with the identity filter, which would reduce the system to

$$\dot{\mathbf{x}} = J\ \mathbf{x}(t), \quad \mathbf{y}(t) = \mathbf{x}(t) + \mathbf{v}(t) \tag{8}$$

which brings it in a form very similar to (1). Regression DCM is implemented in the TAPAS toolbox for Matlab [49], but a Python implementation is also available [50].

Network simulations

The relations (9), which imply a piecewise monotonic relation, and hence a perfect alignment, between GC and IC, are theoretical: they hold between the model connectivity C and G, I computed on the true (population) covariance matrix. In finite data, these relations should reflect in analogous relations between the effective connectivity $\hat{C}$ reconstructed from the data and the estimated Granger causality measures $\hat{G}, \hat{I}$ computed from empirical (sample) covariance matrices. Goal #2 of our study is assessing whether, and to what extent, the predicted relations hold in finite-length time series. We generated time series using the MOU model (1) with process time scales $\tau \in [0.1, 10]$ and a fixed sampling period $\Delta t = 1$. 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 and their sign was 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 $\frac{1+r}{1-r}$, generating 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, aligning the noise time scale and the process time scale (stated otherwise, to align the magnitude of Σ and 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

Goal #3 of our study was to investigate relations between effective connectivity and Granger causality in real fMRI data and test predictions previously obtained through analytical derivations and simulation results. We used the subset of 100 unrelated subjects (54 females, 46 males) from the Human

Connectome Project (HCP) [45]. 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, comprising ≈ 15 minutes of scanning. The full description of the imaging parameters and minimal preprocessing pipeline is given by Ref. [51]. 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 through 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 0.01 and 0.1 Hz to retain the relevant frequency range for further analysis of the BOLD signal. Functional data can be mapped to different spatial resolutions using the Schaefer parcellation [52], which optimizes local gradient and global similarity measures of the fMRI signals. Here, we selected the parcellation consisting of 100 regions, each being linked to one of the 7 networks: visual (VN), somatosensory (SMN), dorsal attention (DAN), ventral attention (DAN), limbic (LIM), xxx (CN), default-mode network (DMN). For both fMRI datasets, regional time series were extracted using Workbench Command provided by the HCP.

Statistical analysis of directional connections

For each participant, we computed $\hat{G}_{ij}$, $\hat{I}_{ij}$ using the covariance-based approach [28,53] and $\hat{C}_{ij}$ using the Lyapunov optimization method [36]. 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 [54]. 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 $p < 0.05$, non significant if $p < 0.05$. P-values were corrected for 4950 multiple comparisons using the false discovery rate approach [54].

Results

Theoretical relations between MOU connectivity and Granger causality.

As linear-feedback dynamic systems, the Ornstein-Uhlenbeck process and the 1st order MVAR rely on mutually consistent assumptions about the nature of the underlying network process, only differing by the continuous vs. discrete nature of time. As shown in S.I. Appendix, the respective connectivity terms (J and A) can be formally related, and this translates into a mathematical relationship between Granger causality estimated from the MVAR and the EC estimated from the MOU. Given that the MOU and the MVAR are related by a (non-infinitesimal) discretization step, their relation is critically modulated by the relation between the timescale of the MOU process (τ) and that of the sampling (Δt). We can distinguish three regimes: $\tau \ll \Delta t$, *fast* dynamics; $\tau \approx \Delta t$, *matched* dynamics; $\tau \gg \Delta t$, *slow* dynamics. To understand the difference between these regimes, one should consider that in a MOU process the influence of $x_j(t)$ is felt by $x_i(t')$, with $t' > t$, only for $t' \lesssim \tau$. Therefore, one can directly observe this influence only if $\Delta t < \tau$ (slow dynamics). If $\Delta t > \tau$ (fast dynamics), one can only observe a sort of 'time-integrated' effect of the reciprocal influences between x_i and x_j, which appears as a non-directional feedback between the two time series. In the Granger-Geweke formalisms, this is captured the instantaneous causality. Our theoretical analysis (S1 Text) leads to a relationship that is

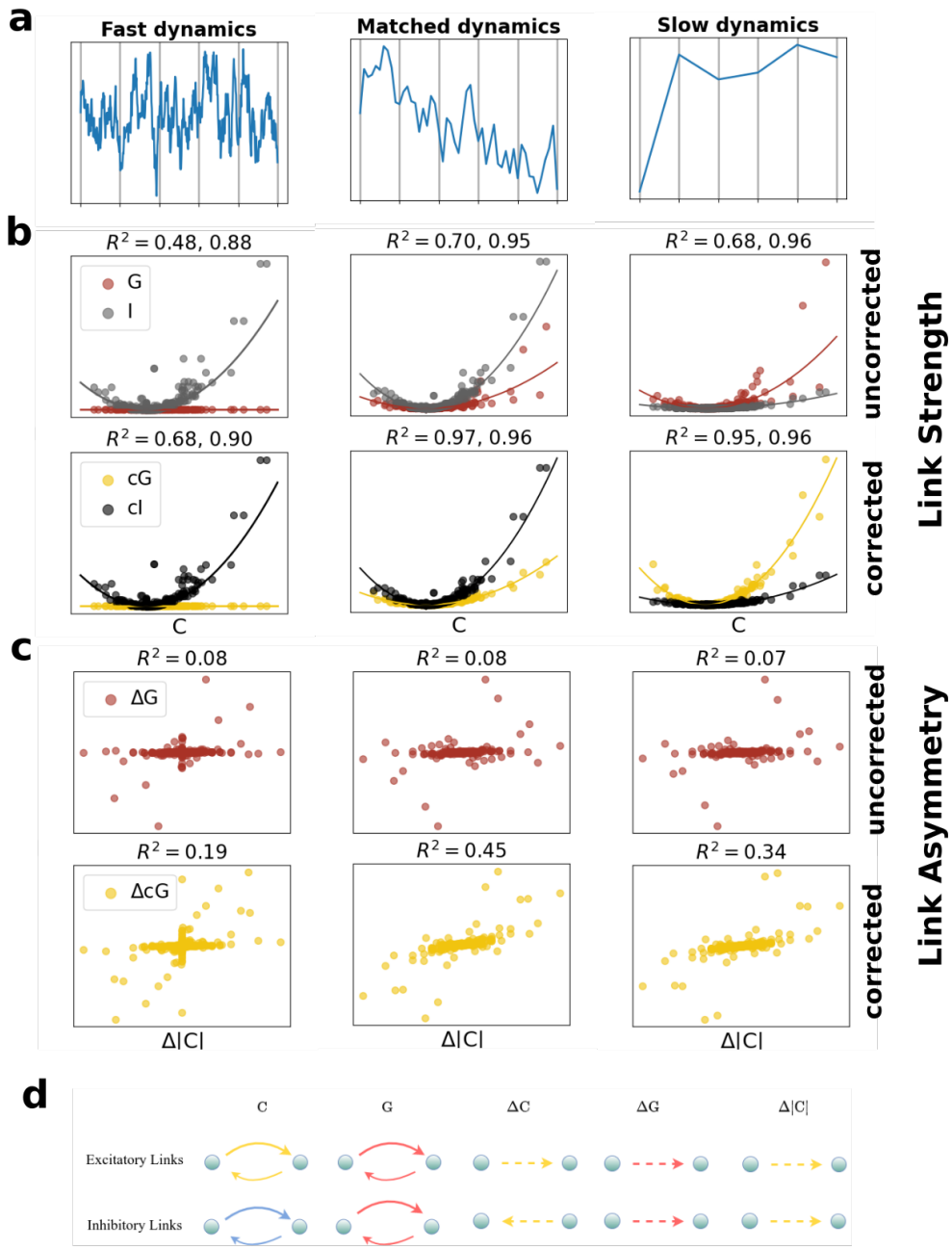

Figure 1: **Theoretical relations between model connectivity and Granger causality measures**. We considered a random network of $N = 40$ nodes evolving according to the MOU dynamics (1). The corresponding MVAR process has a sampling time $\Delta t = 1$. **(a)** Depending on the characteristic time τ of the MOU, we obtained three regimes: slow ($\tau = 0.1$), matched ($\tau = 1$) and fast ($\tau = 10$). **(b)** Relation between the model weights C and the corresponding values of G (in red) and I (in black) for each connection, without correcting for unequal input noise variances (top row) and correcting (bottom row). Each dot corresponds to a pair (i, j), and solid lines to the predicted scalings. **(c)** Relation between the asymmetry in (the absolute strength of) the model weights $\Delta|C|$ and the corresponding asymmetries in ΔG (in red), without correcting for unequal input noise variances (top row) and correcting (middle row). Each dot corresponds to a pair (i, j). **(d)** Conceptual relation between ΔG, ΔC and $\Delta|C|$ for the three cases in which C_{ij}, C_{ji} are positive ('excitation') or negative ('inhibition').

valid in the limit of weak coupling for slow dynamics and (to a certain extent) for matched dynamics, but not necessarily for fast dynamics. For those regimes, the analytical relation is given by:

$$G_{ij} \propto \frac{\sigma_i^2}{\sigma_j^2} C_{ij}^2, \tag{9a}$$

$$I_{ij} \propto \tilde{C}_{ij}^2 \tag{9b}$$

Here, $\sigma_i = \Sigma_{ii}$ is the noise variance of node i; $\tilde{C}_{ij} = \frac{\sigma_i^2 C_{ij} + \sigma_j^2 C_{ji}}{2\sigma_i \sigma_j}$ is a mean of the reciprocal Effective Connectivity weights between i and j, weighted by the noise variances C_{ij}; G_{ij} is Granger causality; and I_{ij} is the instantaneous (Granger) causality" (IC)
If the input noise is homogeneous across all nodes ($\sigma_i = \sigma$), these relations simplify as:

$$G_{ij} \propto C_{ij}^2 \qquad I_{ij} \propto \bar{C}_{ij}{}^2 \tag{10}$$

$\bar{C}_{ij} = (C_{ij} + C_{ji})/2$ being the symmetrized effective connectivity. These relations mean that standard and instantaneous Granger causality provide an estimate of the *square* of the node coupling in a MOU process. In the general case of inhomogeneous noise variance, we can still retrieve an approximately quadratic relation. Since, approximately, $Q_{ii}^0 \propto \sigma_i$, where $Q^0 = \langle \mathbf{x}\mathbf{x}^T \rangle$ is the covariance matrix of $\mathbf{x}$, we can introduce "corrected" versions of G and I,

$$cG_{ij} \equiv \frac{Q_{jj}^0}{Q_{ii}^0} G_{ij} \propto C_{ij}^2, \qquad cI_{ij} \equiv I_{ij} \cdot \frac{Q_{ii}^0 Q_{jj}^0}{Q_{ii}^0 + Q_{jj}^0} \propto C_{ij}^2 \tag{11}$$

Finally, the ratio between Granger causality and instantaneous causality is approximated as:

$$\rho = \frac{cG_{ij}}{cI_{ij}} \approx \frac{1}{2} \frac{\tau^2}{\Delta t^2} \tag{12}$$

This means that for slow dynamics $cG \gg cI$, whereas $cG \ll cI$ for fast dynamics.
Thus, assuming that a continuous MOU model is a valid generative model for the observed time series, the information about the model connectivity C is differentially captured by G and I depending on the sampling regime. When the dynamics is faster than sampling, the values of the 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 [55]. The presence of strong effective connections still creates a strong statistical dependency between the two time series, but most of this interdependence appears as 'instantaneous' correlation that cannot be predicted on the basis of previous time points.

Validation of the theoretical relation in simulated data

To test the analytical relations derived in the previous sections, we generated a MOU process with known ("ground truth") connectivity and tested the above quadratic relations between the model connectivity (C in equations) and Granger causality measures (G and I, respectively). We considered a network of $N = 40$ nodes, with probability $p = 0.3$ or a directional link 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 and negative weights (see Methods for details). We also considered the case of variable noise across nodes (spanning more than one order of magnitude). We fixed the sampling time $\Delta t = 1$ and considered different values of τ to explore all the three dynamical regimes (fast, matched and slow). The results

are shown in Fig. 1b. We compared the consistency of C with the uncorrected (G, I) and corrected (cG and cI) versions of Granger causality. The quadratic relations 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 degraded. Furthermore, as expected, we observed that I prevailed over G in the fast regime, while the opposite occurs in the slow one. Although the theoretical relations were derived for the conditional versions of G and I, very similar results were obtained by considering the respective unconditional versions (Fig. S1).
The theoretical relation suggests that G could also provide information about the link asymmetries in the strength of reciprocal connections C_{ij} and C_{ji}. In other words, when considering asymmetries

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

one would expect a relation between ΔG_{ij} and ΔC_{ij}. This intuition is correct, but there are important caveats. Firstly, the quadratic relation (11) implies an relation for asymmetries only if the noise variances coincide, $\sigma_i = \sigma_j$. Therefore, correcting for unequal noise variances becomes truly essential to get concordance. Secondly, the relation (11) is quadratic (not monotonic), and implies that G is insensitive to whether the underlying connections are 'excitatory' ($C_{ij} > 0$) or 'inhibitory' ($C_{ij} < 0$). If all connections were excitatory, ΔG_{ij} and ΔC_{ij} would always be concordant. However, when effective connections are signed, this is not necessarily the case. This point is exemplified in Fig. 1d. A positive value of $\Delta C_{ij} > 0$ associated to a net asymmetry $i \to j$ of the effective connection can correspond to different situations, depending on the sign of C_{ij}, C_{ji}:

- $C_{ij} > 0, C_{ji} > 0, \Delta C_{ij} > 0$: both connections are excitatory, and the influence of i on 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 on i that prevails. ΔC_{ij} and ΔG_{ij} are discordant.
- $C_{ij} > 0, C_{ji} < 0, \Delta C_{ij} > 0$: one connection is excitatory and the other inhibitory. In this case, ΔC_{ij} and ΔG_{ij} are not necessarily concordant; ΔG_{ij} could even be zero. This case is quite rare in real data (see the following section).

To retrieve an alignment between C and G one should compare the net influence asymmetry measured by ΔG with a net difference in the *absolute strength* (positive or negative) of effective connections, measured by

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

In Fig. 1c) we show the consistency of $\Delta|C|$ and ΔG, ΔcG. We observe an alignment between $\Delta|C|$ and ΔcG only in the matched and fast regimes, where interaction unfolds over a time scale longer than the sampling time, and interaction asymmetries are more easily measurable.

Robustness of the theoretical relation with respect to sampling noise

In the previous section, we have shown that when data are generated by a continuous, linear system with Gaussian noise (i.e., a MOU process), Granger causality is related to the model connectivity by a quadratic relation. In real data, however, this relation may be hidden under sampling noise, as (i) Granger causality must be estimated from finite samples; we will denote by $\hat{G}, \hat{I}$ the resulting estimates, which will differ from the theoretical values G, I obtained in the infinite sample limit (ii) the model connectivity C is generally unknown; effective connectivity $\hat{C}$ provides an estimate of it. To test the effects of sampling noise, we considered simulations of the same model as in the previous section. We varied the characteristic time of the MOU model, which determines the dynamical regime,

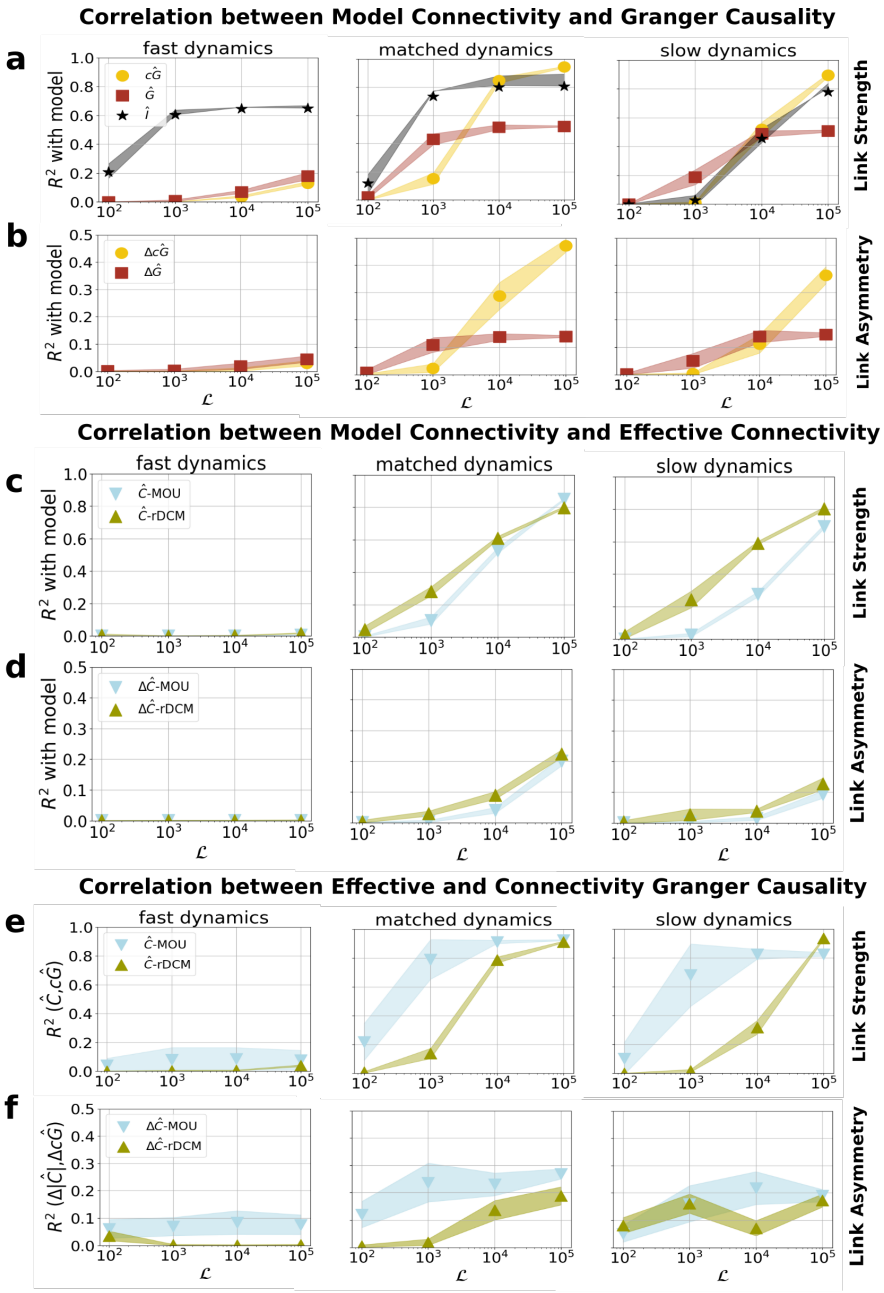

Figure 2: **Relations between model connectivity, effective connectivity, and Granger causality measures for finite sampling time.** A random network was simulated for different durations $\mathcal{L}$ and different dynamical regimes: fast ($\tau = 0.1$), matched ($\tau = 1$) and slow ($\tau = 10$). **(a)** Variance of the model connectivity C explained by uncorrected ($\hat{G}$) and corrected ($c\hat{G}$) Granger causality, and instantaneous Granger causality ($\hat{I}$). **(b)** Same as (a), but considering asymmetry in the absolute link strength $\Delta|C|$ vs $\Delta\hat{G}, \Delta c\hat{G}$. **(c)** Variance of the model connectivity C explained effective connectivity estimated from the dat via MOU-EC ($\hat{C}$-MOU) and rDCM ($\hat{C}$-rDCM), **(d)** Same as (c), but considering asymmetry in the absolute link strength $\Delta|C|$ vs $\Delta\hat{C}$-MOU,$\Delta\hat{C}$-rDCM. **(e)** Variance of the effective connectivity (estimated with MOU and rDCM) explained by corrected Granger causality **(f)** Same as (e), but considering link asymmetries $\Delta|\hat{C}|$-MOU, $\Delta\hat{C}$-rDCM vs $\Delta c\hat{G}$.

and the simulation length $\mathcal{L}$, which determines the amount of sampling noise. We estimated GC using the covariance-based approach, which does not try to fit an MVAR but computes GC directly from the non-lagged and lagged covariance matrices $\hat{Q}^0, \hat{Q}^1$. EC is estimated using two approaches, rDCM and MOU-EC. In this way, we can guarantee that conclusions are not contingent on the usage of a specific method (MOU-EC).

We first tested if relations between the model connectivity and Granger causality would also hold when considering sampling noise affecting Granger causality.

To this end we performed two fits: (i) a quadratic fit of the model connectivity C on $\hat{cG}$,$\hat{G}$, $\hat{I}$, to test the predicted quadratic relations for link strengths (ii) a linear fit of $\Delta|C|$ on $\Delta\hat{cG}$, $\Delta\hat{G}$, to test the predicted linear relations for link asymmetries. The goodness of fit was quantified through the R^2 coefficient ($R^2 \to 1$ corresponding to a perfect fit). The results are shown in Fig. 2a,b.

For *fast* dynamics, $\hat{G}$ and $\hat{cG}$ are poorly related to C, consistent with the fact that one cannot appreciate the directionality of interactions in this regime. Correspondingly, there is also no correlation between link asymmetries in $\Delta\hat{G}$,$\Delta\hat{cG}$ and $\Delta|C|$. Instead, the instantaneous causality I reaches high values of $R^2 \approx 0.6$ as long as the simulation is sufficiently long ($\mathcal{L} \gtrsim 10^3$). For *matched* dynamics, we observed a strong alignment between $\hat{cG}$ and C^2 ($R^2 > 0.8$) but only for very long sampling length ($\mathcal{L} \gtrsim 10^4$). Alignment between $\hat{G}$ and C^2 was better ($R^2 \approx 0.4$) for $\mathcal{L} = 10^3$ (a quite realistic length for fMRI), but did not improve for long sampling lengths. In terms of asymmetry we obtained a similar picture, as $\Delta\hat{cG}$ aligned with $\Delta|C|$ but only for very long sampling lengths, reaching $R^2 \approx 0.5$ for $\mathcal{L} = 10^5$. Instantaneous causality aligned well with C^2 for $\mathcal{L} \gtrsim 10^3$. Finally, for *slow* dynamics, results were similar to those obtained for matched dynamics, but the values of R^2 decreased, especially for the instantaneous causality that was completely uncorrelated to C for a realistic sampling length ($\mathcal{L} = 10^3$. Overall, the predicted relations between model connectivity and Granger causality, in terms of link strengths and asymmetries, were best observable when the characteristic time of the MOU process corresponded to the sampling time (matched dynamics); the relation between C and I was robust to sampling noise, while the relation between C and cG (and between ΔC and ΔcG) was not, emerging only for long sampling lengths.

Secondly, we asked whether relations would still hold when replacing the model connectivity C with effective connectivity estimates $\hat{C}$. To this aim, we first performed a general assessment of the quality of estimates obtained with MOU-EC and rDCM, by performing a linear fit of C on $\hat{C}$. Results are shown in Fig. 2c-d. For fast dynamics, estimates are completely unreliable, remaining around $R^2 \approx 0$ for all sampling lengths, consistent with the fact that directional interactions cannot be properly observed in this regime. For matched and slow dynamics, good estimates generally required large sampling length ($\mathcal{L} \gtrsim 10^4$). For link strenghts, the maximum alignment was $R^2 \approx 0.8$, for link asymmetry $R^2 \approx 0.2$. RDCM yielded more accurate estimates than MOU for realistic sampling lengths ($\mathcal{L} = 10^3$) and slow dynamics. From these results, we might expect a clear alignment between effective connectivity and Granger causality only if (i) the sampling lengths is large, or (ii) estimates are affected by common biases (e.g., note that covariance based GC and MOU-EC both employ the zero-lag and lagged empirical covariance matrices). We tested this by performing (i) a quadratic fit of effective connectivity $\hat{C}$ on $\hat{cG}$,$\hat{G}$, $\hat{I}$ (ii) a linear fit of $\Delta|\hat{C}|$ on $\Delta\hat{cG}$. The results are shown in Fig. 2e,f. For fast dynamics, Granger causality and DCM-$\hat{C}$ are very weakly correlated. Unsurprisingly, for fast dynamics we obtained very weak correlations for both link strength and asymmetry. For matched and slow dynamics, we generally obtained a strong alignment between $\hat{cG}$ and $\hat{C}$-MOU for all $\mathcal{L} \gtrsim 10^3$, reaching $R^2 \approx 0.9$, even larger than the correlation between model connectivity and effective connectivity. The same occurred for link asymmetry, which reached $R^2 \approx 0.3$ (compared to $R^2 \approx 0.2$ for model and effective connectivity). Instead, $\hat{C}$-rDCM aligned well with $\hat{cG}$ only for a large sampling length ($\mathcal{L} \gtrsim 10^4$).

EC and GC alignment in rs-fMRI data

We considered human resting-state fMRI data of 100 unrelated participants from the Human Connectome Project [45]. For each participant, two recording sessions of 1200 time points (TR=0.7) were available. We mainly considered the first session, using the second session for validation. 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. [56]. Effective connectivity $\hat{C}$, that was estimated individually for each participant using MOU-EC and rDCM. Granger causality was estimated individually using the covariance-based approach.

Effective connectivity analysis. At a group level, MOU-$\hat{C}$ featured 3804 significant connections (two-tailed T test over subjects, $p < 0.01$, FDR corrected for 9900 multiple comparisons), among which we found both significantly positive (64%) and significantly negative connections (36%) (Fig. 3a). Reciprocal connections ($i \to j$ and $j \to i$), tended to be either unidirectional (i.e., when one connection was significant the reciprocal one was non-significant), concordant (both connections were significant and had the same sign; Fig. 3c), or both non-significant. We found only $\approx 1\%$ of pairs of reciprocal connections with discordant sign. Positive-positive reciprocal connections were frequently found within different areas of the same resting state networks (RSNs), both within and across hemispheres. Connections between nodes of different RSNs were predominantly unidirectional (64%). Negative connections were nearly exclusively found between different RSNs (S3 a).
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 (S3 b). 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 (S3 b). Only a negligible fraction (2%) of significant asymmetries involved discordant connections.
The estimated rDCM-$\hat{C}$ showed a much larger number (7676) of significant connections at the group level, and a much lower proportion of negative links (83% positive, 17% negative connections). Reciprocal connections tended to be either unidirectional or concordant; most of them were concordantly positive (Fig. 3c). Roughly 3% of the pairs of significant connections had a discordant sign. Within-RSN connections were almost exclusively positive. Negative connections were almost exclusively found between different RSNs (S3 c). We found a significant asymmetry for the majority (58%) of link pairs (S3C). Asymmetries between positive connections were found both within the same and across different RSNs. Asymmetries between negative connections were only found across different RSNs (S3 d).
We averaged $\hat{C}$ 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. 3c we show significant positive and negative network-wise links (T test over subjects, $p < 0.01$ FDR corrected for 182 multiple comparisons). We observed strong positive connections within each RSN, both within the same hemisphere and across the two hemispheres. On the other hand, negative connections were found between different RSNs, and were strongest for outgoing connections from the control network (CON), limbic network (LIM), and default mode network (DMN). Positive connections were mostly observed within RSNs. For rDCM-$\hat{C}$, we mostly observed positive connections. The strongest connections were observed within each RSN. We also see that the RSNs split in two blocks, one formed by 'task-positive' networks (visual, somato-motor, dorsal and ventral attention) and the other by 'task-negative' network (control, limbic and default mode network). Within-block connections were positive, between-block ones null or slightly negative.

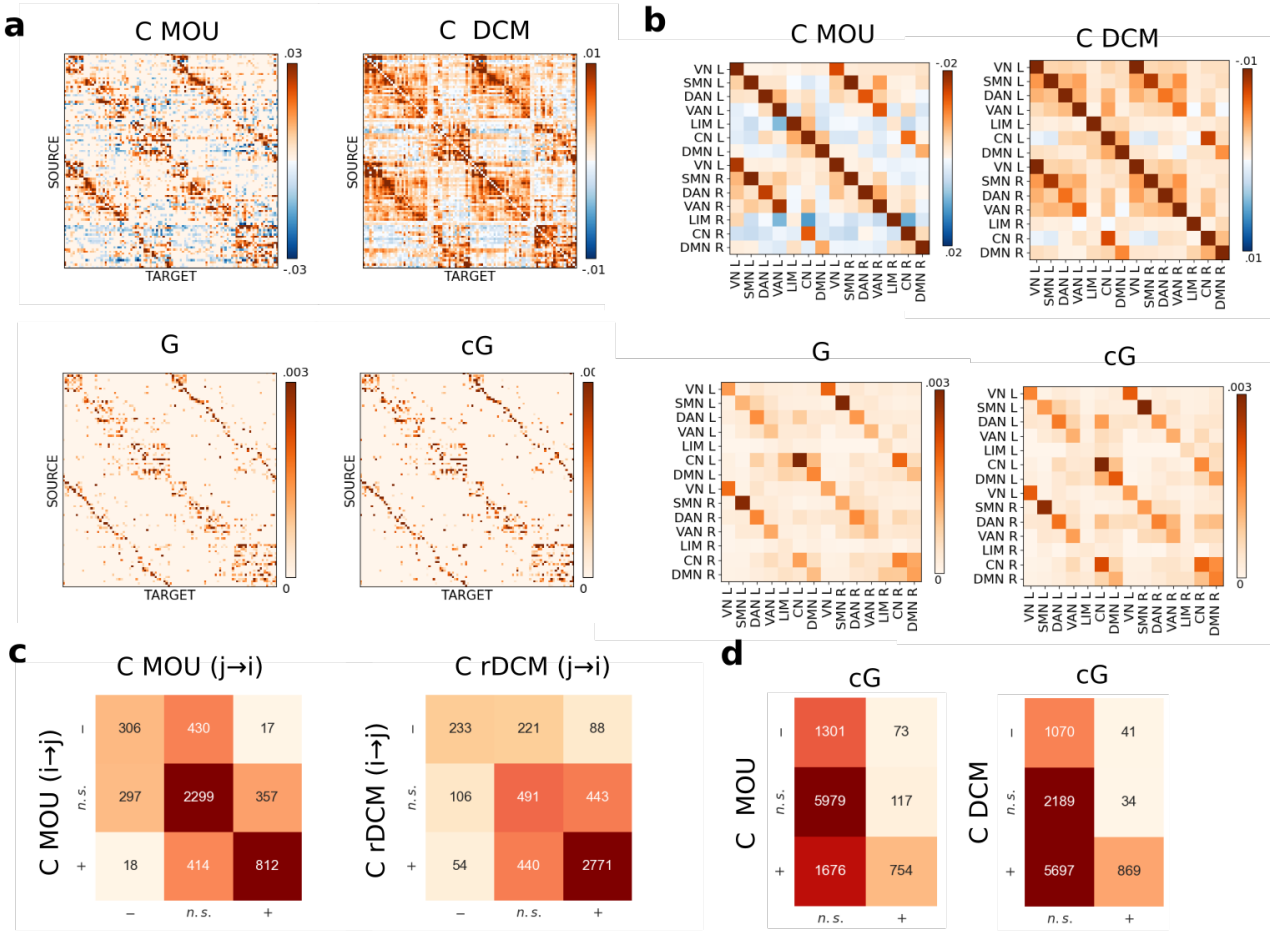

Figure 3: **Group-level results of Granger Causality and Effective connectivity in HCP. (a)** Group-level (average) matrices of $\hat{G}$, $\hat{cG}$, $\hat{C}$-MOU, $\hat{C}$-rDCM. Only links that are significantly different from zero are shown (T test across subjects, $p < 0.05$ FDR-corrected), in red for positive weights and blue for negative weights. **(b)** Network-wise group-level matrices averaged in 7 functional networks: visual (VN), somatosensory (SMN), dorsal attention (DAN), ventral attention (DAN), limbic (LIM), control (CN), default-mode network (DMN). **(c)** For all pairs of reciprocal connections ($i \to j$ and $j \to i$), we counted the number of pairs for which connections are significantly positive (+), significantly negative (−) or non-significant (n.s.) for $\hat{C}$-MOU and $\hat{C}$-DCM **(d)** We counted the number of connections that are significantly positive/negative or non-significant (n.s.) for $\hat{C}$ (MOU and rDCM) and and $\hat{cG}$.

Granger causality analysis. Next, we estimated Granger causality. Since the noise variances affecting different nodes were found to be widely different among nodes, spanning roughly an order of magnitude (S4), we expect potential differences between the corrected and non-corrected versions of $\hat{G}$. At the group level, the estimated $c\hat{G}$ had only 944 significant connections (Fig. 3B) (T test over subjects, $p < 0.01$ FDR corrected for 9900 multiple comparisons). Of these, the overwhelming majority were identified as significantly positive connections by both $\hat{C}$-MOU and $\hat{C}$-DCM (respectively, 754 and 869). Less than 10% were associated with negative connections in $\hat{C}$-MOU and $\hat{C}$-DCM (respectively, 73 and 41) (Fig. 3d). The strongest connections were found between homologous regions of the left and right hemishperes. We averaged the $c\hat{G}$ 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. 3d) we show significant network-wise links (T test over subjects, $p < 0.01$ FDR corrected for 182 multiple comparisons). The strongest links were again observed within the LH and RH parts of each RSN. We tested pairs of reciprocal links for connection asymmetry: only a small number of link pairs (222, $\sim 5\%$) presented a significant asymmetry. Of these, most also presented a significant asymmetry in terms of $\hat{C}$-MOU and (respectively, 171 and 160). Overall, $c\hat{G}$ appears much more conservative than $\hat{C}$ in detecting significant connections and significant asymmetries at a group level. Results for the non-corrected $\hat{G}$ are nearly equivalent.

Comparison between effective connectivity and Granger causality. Having analyzed EC and GC separately, we now proceed to a systematic comparison between the two methods. MOU-EC estimates the process time as $\tau = 2.3 \pm 0.6$ (mean $\pm$ st.dev. over subjects). This suggests that fMRI data are approximately in the 'matched dynamics' regime. Based on the simulation results, we thus expect to observe an alignment between C-MOU and cG, and, to a weaker extent, between C-rDCM and cG. The alignment should become tighter as the sampling noise is reduced. The length of fMRI recordings is fixed ($\sim 10^3$ time points per participant), but an effective way to reduce sampling noise is to combine data of different subjects to produce group averages. Thus, we expect to see a tighter alignment when averaging $\hat{C}$, $c\hat{G}$ estimates over an increasing number of participants. We considered subgroups of participants of increasing size n. For each n, we considered $M = 100$ randomly selected subgroups of size n. For each subgroup, we averaged $\hat{C}$ and $c\hat{G}$ and computed the R^2 between $\langle \hat{C} \rangle_n^2$ and $\langle c\hat{G} \rangle_n$ (here, $\langle \cdot \rangle_n$ indicates a subgroup average over n participants). We finally averaged the resulting R^2 over the M random subgroup choices.
The results are shown in Fig. 4a for $\hat{C}$-MOU: R^2 increased rapidly with n, reaching values $R^2 \gtrsim 0.6$ for $n > 20$. Considering the whole group ($n = 100$), we obtained $R^2 = 0.72$. Fig. 4a also shows the scatter plot of $\hat{C}$ and $c\hat{G}$, averaged over the whole group of $n = 100$ subjects.

When considering the non-corrected version of Granger causality ($\hat{G}$), we obtained values of values $R^2 \gtrsim 0.5$ for $n > 20$, up to $R^2 = 0.6$ for $n = 100$. In Fig. 4b, we show results for $\hat{C}$-rDCM: R^2 again increases with n, reaching values $R^2 \gtrsim 0.3$ for $n > 20$. Considering the whole group ($n = 100$), we obtained $R^2 = 0.35$. The alignment with the non-corrected $\hat{G}$ was slightly better, reaching $R^2 = 0.4$ for $n = 100$.
We also expected some level of alignment between the asymmetry of connections. We thus repeated the same analysis, considering the R^2 between (subgroup-averaged) $\Delta|\hat{C}|$ and $\Delta c\hat{G}$. When considering $\hat{C}$-MOU, The alignment increases with group size n, reaching values $R^2 \approx 0.3$ for $n > 20$ (Fig. 4c) and $R^2 = 0.39$ for $n = 100$. The alignment between $\Delta|\hat{C}|$ and the non-corrected $\Delta\hat{G}$ is notably inferior ($R^2 = 0.13$ for $n = 100$). When considering $\hat{C}$-rDCM, we observed a very weak alignment (Fig. 4d), remaining at the level of $R^2 \approx 0.03$ independently of subgroup size. In summary, the theoretically expected alignment between effective connectivity and Granger causality at the level of link strength is visible in fMRI data at a group level, and it is much stronger when EC is obtained via MOU-EC. Alignment at the level of link asymmetry was also visible at group level, but only for the corrected

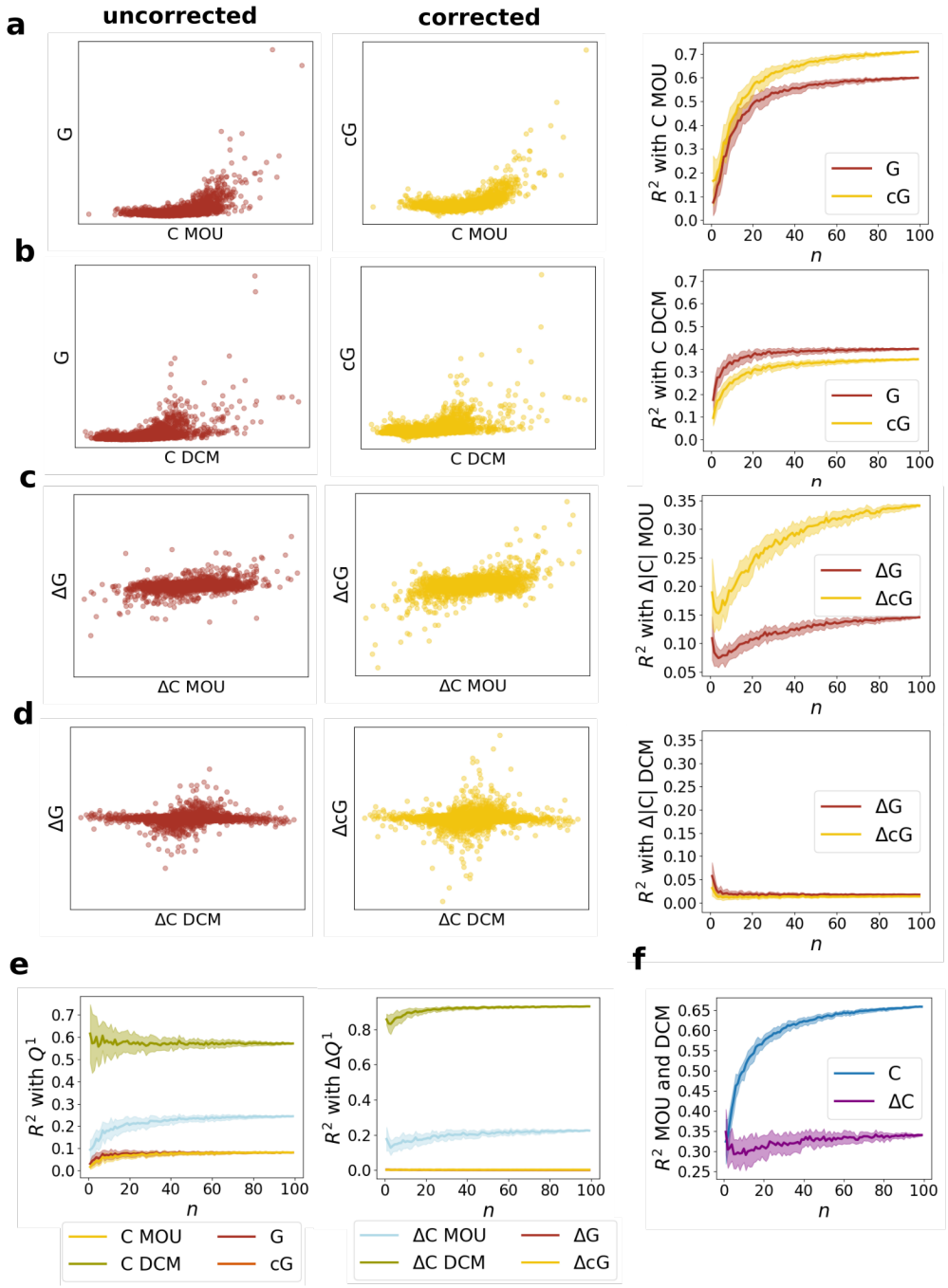

Figure 4: **(a)** effective connectivity $\hat{C}$ (group average over 100 subjects) vs corrected Granger causality $\hat{cG}$ (group average). **(b)** squared Pearson correlation R^2 between group $\hat{C}$ and group $\hat{cG}$ for groups of increasing size n. **(c)** effective connectivity $\hat{C}$ (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.

version of Granger causality and for MOU-EC.
Some comments about the discrepancies between $\hat{C}$-MOU and $\hat{C}$-rDCM are in order. At a group level, we observe good alignment between $\hat{C}$-MOU and $\hat{C}$-rDCM. As long as we consider subgroups of $\gtrsim 20$ participants, we observe $R^2 > 0.55$ at the level of link strength and $R^2 > 0.3$ at the level of link asymmetry (Fig. 4e). For the whole group of 100 participants, we obtain $R^2 = 0.64$ at the level of link strength and $R^2 > 0.31$ at the level of link asymmetry. Despite this partial convergence, the two EC methods seems to be affected by different biases. $\hat{C}$-MOU estimates are biased towards $c\hat{G}$, both at the level of link strength and link asymmetry. As for $\hat{C}$-rDCM, we observed a strong alignment with the lagged cross-covariance $Q^1 = \langle \mathbf{x}(t)\mathbf{x}(t+1)\rangle$ (Fig. 4f). In particular, for the whole group of $n = 100$ participants we observed $R^2 = 0.58$ between $\hat{C}$-rDCM and $\hat{Q}^1$; the alignment was even stronger at the level of link asymmetry ($R^2 = 0.85$).

Sources and Sinks for EC and GC. Directed connectivity measures such as effective connectivity and Granger causality are often used to infer hierarchies between brain regions, based on a reconstruction of 'sources' (nodes with predominantly outgoing connections) and 'sinks' (nodes with predominantly incoming connections). We tested to what extent different measures would lead to a consistent identification of source and sink regions. To this aim, we considered the 'net outgoing strength' as the difference between the out-strength and the in-strength of each nodes, i.e., the total difference between outgoing and incoming link strengths. For instance, for region i, the net outgoing strength measured by $\hat{G}$ is

$$\Delta\hat{G}_i \equiv \sum_j (\hat{G}_{ij} - \hat{G}_{ji}) = \sum_j \Delta\hat{G}_{ij}$$

Analogous metrics can be defined for $c\hat{G}$, $\hat{C}$-MOU and $\hat{C} - rDCM$. For $\hat{C}$, we consider two versions of this metric, one accounting for the presence of excitatory and inhibitory connections (thus using the signed link values $\hat{C}_{ij}$), and one considering the absolute link strength (hence using the absolute values $|C_{ij}|$). These resulting maps are shown in Fig. 5a. All maps exhibit some commonalities: for example, sensorimotor regions are consistently identified as sources by all measures, while the precuneus and posterior cingulate are consistently identified as sources. However, correspondence between all maps is limited. In Fig. 5b we show the correlation (Pearson R) between all pairs of maps. Results are clearly in line with previous considerations. The maps obtained with rDCM taking the signed vs. absolute value are nearly equivalent (given the minor relevance of negative connections). The maps obtained with MOU are instead quite different depending on taking or not the absolute value: the maps obtained with the signed values aligns best with the rDCM map, while the map obtained with absolute values aligns with the cG map. Discrepancies between the maps are visible from Fig. 5c, where we show the net outgoing strength for nodes of different RSNs. While all methods find a positive net outgoing strength for SMN and DAN, and a negative net outgoing strength for CN, results for other networks strongly depend on the chosen method.

Reproducibility. We replicated the same analysis using the second fMRI session available for all subjects. In Fig. 6 we show a quantitative comparison between the results of the two sessions. In Fig. 6a we show the consistency (R^2) of the estimates of $\hat{C}$-MOU,$\hat{C}$-rDCM,$\hat{G}$, $c\hat{G}$,$\hat{I}$ 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 by $\hat{C}$-MOU and $c\hat{G}$. For $\hat{G}$ we obtained a weak $R^2 \approx 0.2$. Instead, for $\hat{C}$-rDCM we obtained a good consistency $R^2 = 0.5$, and an even better one for $\hat{I}$ ($R^2 \approx 0.6$). The result for $\hat{I}$ is in line with those obtained for the non-lagged and lagged covariance matrices $\hat{Q}^0$ and $\hat{Q}^1$. When considering link asymmetries, results of $\Delta\hat{C}$-MOU, $\hat{G}$ and $c\hat{G}$ were also poorly consistent $R^2 < 0.05$, while some consistency was obtained by $\Delta\hat{C}$-rDCM ($R^2 = 0.15$). We hypothesize that

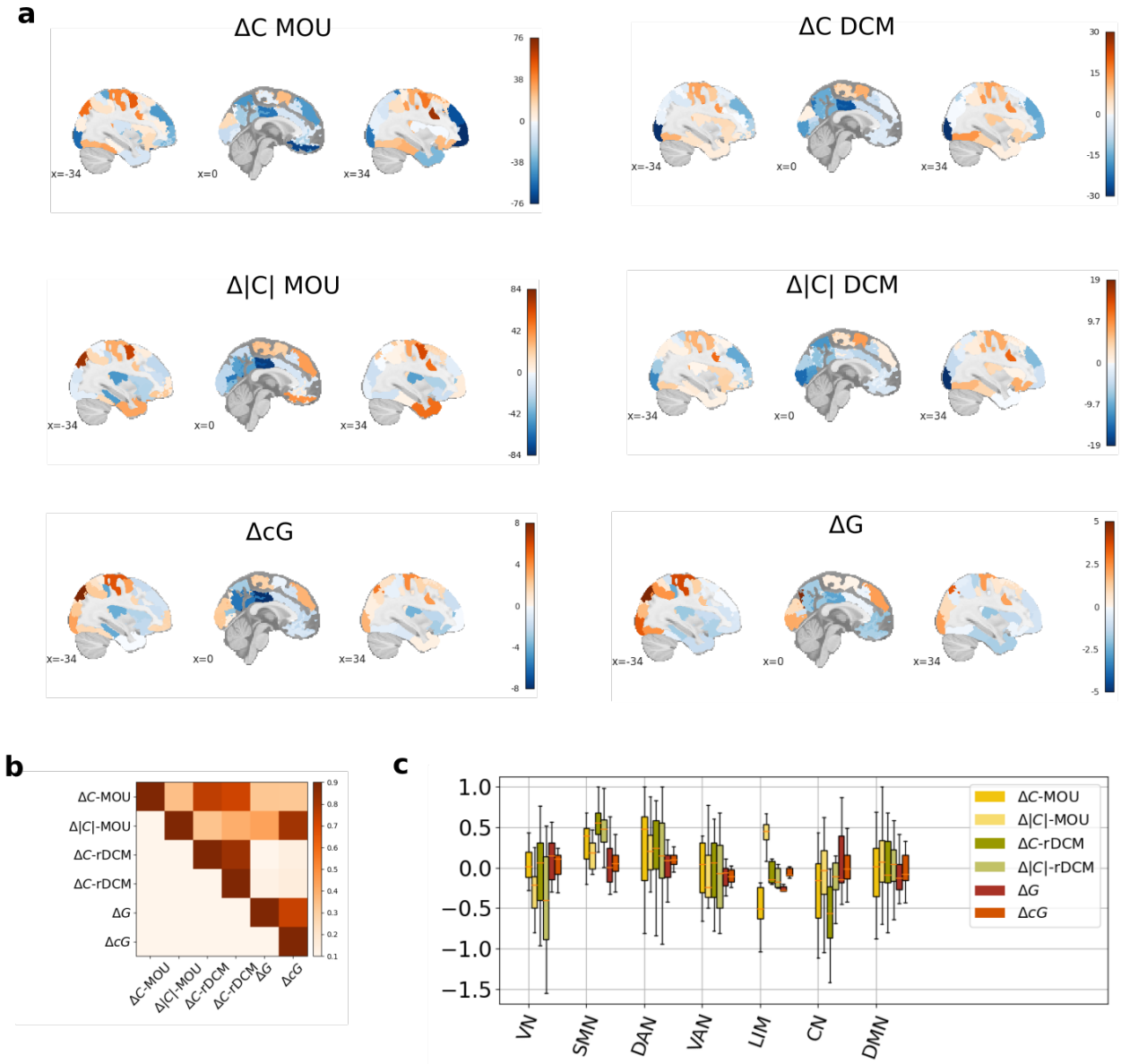

Figure 5: **Sinks and sources for EC and GC (a)** Group-level maps of the difference between the out-strength and the in-strength for different metrics **(b)** Correlation between the maps **(c)** Difference between out-strength and in-strength in different resting state networks.

the much larger consistency of $\hat{C}$-rDCM, compared to $\hat{C}$-MOU and $\hat{G}$,$\hat{cG}$ is due to the previously noticed strong convergence with Q^1. In Fig. 6b we show group-level results obtained when considering subgroups of different size n. Averaging over at least 20 subjects yielded reliable estimates ($R^2 > 0.6$) for all measures. Overall, the degree of consistency showed a hierarchy $\hat{C}$-MOU ,$\hat{cG} < \hat{G} < \hat{C}$-rDCM ,$\hat{I}$. Consistency at the level of link asymmetry also obeyed the same trend, with reliable estimates ($R^2 > 0.4$) starting from 20 subjects (Fig. 6C), and the same hierarchy between the different measures. When using the whole group of $n = 100$ subjects, estimates of connection strength and asymmetry were consistent at a group level ($R^2 > 0.8$) for all measures.

We also tested for consistency in the identification of significant links. In Fig. 6c we compare the links identified as significant in the two sessions by several measures. For $\hat{C}$-MOU, $\approx 80\%$ of the links identified as significantly positive in session one was also labeled as significantly positive in session two, with the remaining 20% labeled as non-significant (and vice versa); $\approx 70\%$ of the links identified as significantly negative in session one was also labeled as significantly negative in session two, with the remaining 30% labeled as non-significant (and vice versa); no link was labeled as significantly positive in one session and significantly negative in the other. An even more consistent picture was obtained for rDCM: $\approx 95\%$ of the links identified as significantly positive in session one was also labeled as significantly positive in session two, with the remaining 5% labeled as non-significant (and vice versa); $\approx 90\%$ of the links identified as significantly negative in session one was also labeled as significantly negative in session two, with the remaining 10% labeled as non-significant (and vice versa); no link was labeled as significantly positive in one session and significantly negative in the other. For $\hat{G}$ and $\hat{cG}$, $\approx 80\%$ of the links identified as significantly positive in session one was also labeled as significantly positive in session two.

Finally, we tested the consistency in the identification of links with significant asymmetry (Fig. 6d). For $\hat{C}$-MOU, $\approx 65\%$ of the links displaying a significant link asymmetry in session 1 also presented a significant asymmetry in the same direction in session 2, with the remaining 35% displaying a non-significant asymmetry (and vice versa). For $\hat{C}$-rDCM, $\approx 85\%$ of the links displaying a significant link asymmetry in session 1 also presented a significant asymmetry in the same direction in session 2, with the remaining 15% displaying a non-significant asymmetry (and vice versa). For $\hat{G}$ and $\hat{cG}$, approximately 65% of the links displaying a significant link asymmetry in session 1 also presented a significant asymmetry in the same direction in session 2, with the remaining 35% displaying a non-significant asymmetry (and vice versa). No measure identified asymmetries in opposite directions in session 1 and 2.

In summary, results are highly replicable, but only at a group level. Overall, rDCM produced more consistent results than MOU-EC and Granger causality.

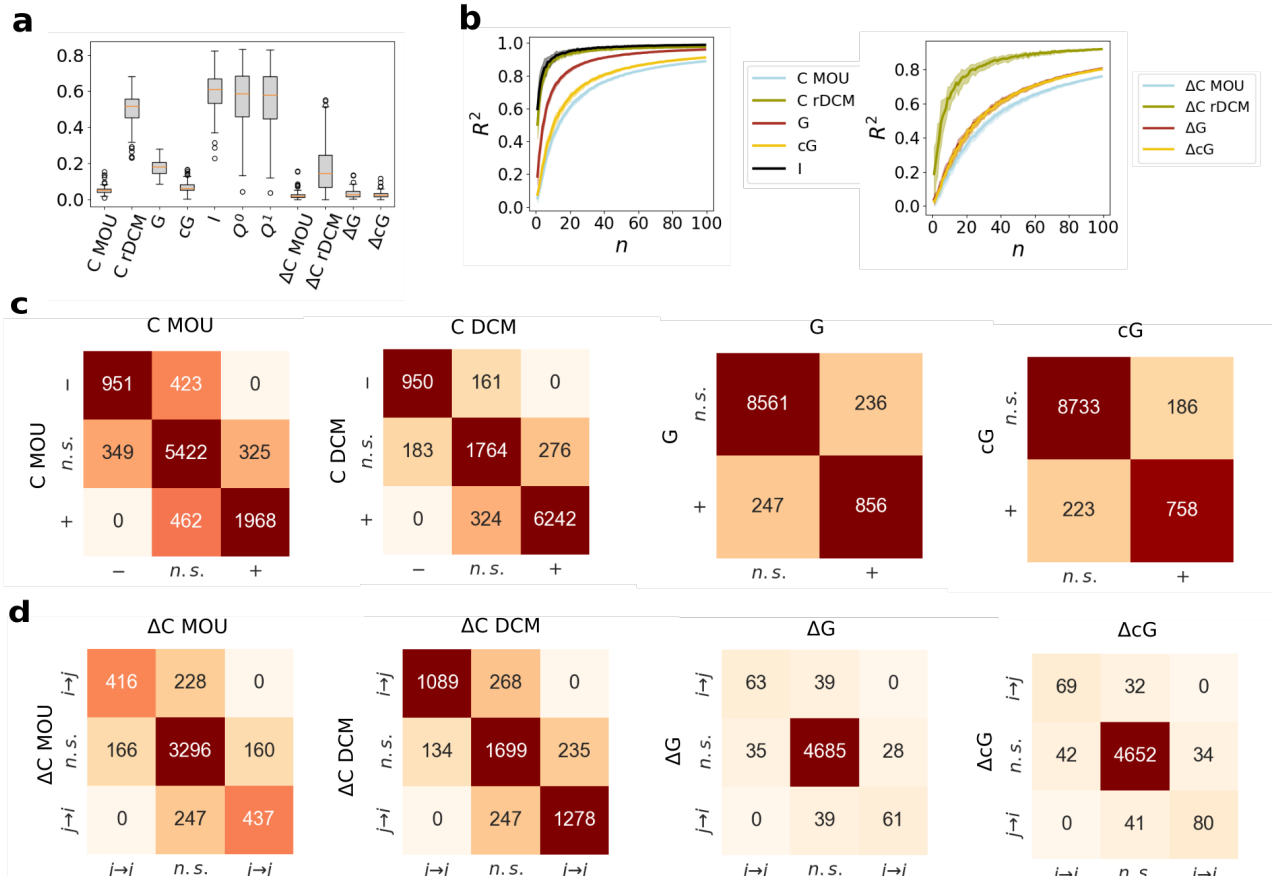

Figure 6: Consistency of results in two independent independent recording sessions **(A)** consistency of *single-subject* estimates across the two sessions (squared Pearson correlation). **(B)** Squared Pearson correlation R^2 between group estimates of $\hat{C}$, $\hat{cG}$, $\hat{G}$, $\hat{I}$ for groups of increasing size n. **(C)** Squared Pearson correlation R^2 between group estimates of $\Delta\hat{C}, \Delta\hat{cG}, \Delta\hat{G}, \hat{I}$ 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 ($n = 100$).

Discussion

A fundamental objective of system and network neuroscience is to uncover the graph of causal interconnections between areas of the human brain. Substantial information about directional connections in the brain can be learned from time-lagged relationships between brain areas observed in fMRI [17]. Traditional, non-directional functional connectivity cannot characterize asymmetries between reciprocal connections, which underlie a key principle of brain organization, hierarchical processing [57, 58]. Directional connectivity methods overcome these limitations, elucidating the task-related modulation of functional interactions [59], hierarchical relations underlying cognitive functions (such as working memory [60, 61], cognitive control [62] and language [43]), and pathological alteration of functional hierarchies [28, 63] - leading to improved biomarkers for clinical classification [42].
Despite the availability of several methods to analyze directional connectivity [64], the majority of studies employ effective connectivity (EC) or Granger causality (GC). The state-of-the-art wisdom about these two methods is somewhat contradictory. On the one hand, the methodological literature has often stressed their differences [32], insisting that they rest on different premises: EC aims to retrieve a latent dynamical model, while GC is based on statistical dependencies. On the other hand, in applications the two techniques are used quite interchangeably, EC or GC links being generally interpreted as directional couplings between brain areas, with an implicit intuition that large effective connections should be reflected into large values of Granger causality and vice versa.

Theoretical relations explaining similarities and differences between GC and EC. The first contribution of this work is to provide a clear theoretical framework to understand similarities and differences between EC and GC in the context of brain network reconstruction in fMRI. Common EC approaches for whole-brain fMRI assume a linear, stochastic dynamical model driven by Gaussian noise, i.e., a multivariate Ornstein-Uhlenbeck process. Similarly, GC approaches assume a linear multivariate autoregressive (MVAR) process with Gaussian noise. Assuming the MVAR to be first-order (as commonly done in fMRI studies [20]), we leveraged an explicit mapping between the continuous-time MOU and the discrete time MVAR to derive analytical relations between EC and GC (Eqs.(32),(33) and Fig. 1), valid between *model* values of the two quantities (which of course differ from *estimates* obtained from finite samples). These relations allowed us to formulate clear predictions about the EC-GC correspondence.
Firstly, the implication large EC $\Rightarrow$ large GC critically depends on the relation between two time scales: the sampling time (Δt; interval between two successive observations, corresponding to TR in fMRI) and the process time (τ; the time scale of the stochastic ODE system underlying EC analysis, related to the autocorrelation time of the dynamical system). When the process is fast ($\tau \ll \Delta t$), 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 (as underscored by previous work [55]). The presence of a strong effective connection reflects in an "instantaneous" correlation that cannot be predicted on the basis of previous time points, which is 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 [47]. When the process is slow ($\tau \gg \Delta t$), the values of the 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 reflects in large values of GC (while IC is very small). In fact, the analytical relations (Eqs.(32),(33)) show that both GC and IC are (roughly) proportional to the square of EC, but for $\tau \ll \Delta t$ the proportionality constant is much stronger for IC than GC (and vice versa for $\tau \gg \Delta t$). Thus, our study shows that large effective connections are reflected in large values of Granger causality *or* instantaneous causality.
Secondly, a direct relation between EC and GC link strengths holds only if all areas have equal signal

variance, corresponding to a homogeneous level of random input. Otherwise, the relation is modulated by the ratio of the input noise affecting the two endpoint nodes. As an example, for symmetric effective connections between two areas, the GC from the area with higher variance will be larger than in the reverse direction. This is not necessarily a pitfall of GC, that was conceived as a measure of "influence": in presence of equal connections, the influence of an area with larger variance (signal power) is stronger. However, it implies that GC cannot be considered a measure of coupling, like EC. In presence of non-homogeneous signal power, it is still possible to recover a quadratic relation between EC and GC/IC (Eqs.(34),(35)) by "correcting" GC estimates by simple factors accounting for the heterogeneity in nodal input.
Thirdly, the EC-GC relation is non-monotonic (approximately quadratic, Eqs.(32),(33)). Effective connections can be positive or negative, and larger values of GC correspond to stronger values of EC in *modulus* (i.e, strongly positive or strongly negative). This is again consistent with the interpretation of GC as a measure of influence: strongly "excitatory" or strongly "inhibitory" effective connections both determine a large causal effect from the source node to the target node. Correspondingly, EC-GC alignment at the level of link asymmetry can be observed only if the definition of the EC asymmetry involves the modulus of the connection strength (Fig. 1), identifying the direction of the strongest connection (whether positive or negative).

Effects of sampling. The second contribution of this work is to give predictions about the level of the EC-GC alignment that one may expect in finitely sampled time series. The analytical relations obtained hold at the theoretical level, i.e., between model couplings of a continuous-time MOU process and model GC values of the corresponding discrete-time MVAR process. In practice, EC and GC *estimates* obtained from finite samples differ from the model values.
With the aid of numerical simulations of artificial networks, we discussed to what extent the theoretical EC/GC relations could be observed from finite-length data, depending on the sampling rate and the sampling length (Fig. 2). For estimation, we used the two most common methods for whole-brain EC estimation (MOU-EC and regression DCM) and a popular approach for GC estimation (the covariance-based approach). The expected relations were clearly observed only for a sufficiently fast sampling rate that matches the process time scale ($\Delta t \leq \tau$). Moreover, a sampling length of $\mathcal{L} = 10^4$ was required, which amounts to 10 times the length of a typical recording session. For $\mathcal{L} = 10^3$ (the typical sampling time of an fMRI session), we observed a strong EC/GC correlation only when using MOU-EC, but not rDCM. This is probably due to common estimation biases between covariance-based GC and MOU-EC, which both depend on empirical covariance matrices. This analysis shows that GC-EC correlation can emerge only in unusually long recordings (such as [65]) or, more plausibly, at the group level (neglecting inter-individual variability). Thus, we predicted that group-level analyses using at least 10 subjects (order of magnitude) should be considered to observe the EC-GC alignment in fMRI data. To observe alignment at the level of connection asymmetry, this requirement becomes even stricter: at least $N \sim 100$ subjects should be used to reach the theoretically predicted level of correlation ($R^2 \approx 0.4$).

EC/GC comparison in fMRI data. The third contribution of this work is an assessment of the consistency of EC and GC in fMRI data. Although some methodological studies restricted attention to simulated network activity to compare directional connectivity methods [55, 64], using real data is crucial, as certain assumptions behind synthetic data (e.g. Gaussianity of generated signals) might only be loosely respected, if not entirely violated.
In agreement with predictions from simulated data, we observed EC-GC correlation *at the group level* (Fig. 4) when pooling at least ~ 15 subjects together. Consistently with the simulation results, the GC-EC relation was much stronger for MOU-EC than rDCM ($R^2 > 0.6$ vs $R^2 > 0.25$). Overall, EC

analysis detected a much larger number of significant connections at a group level (38% for mou-EC, 77% for rDCM, compared to 9% for GC; fig. 3). The overwhelming majority of significant GC links were identified as significant by EC (both my MOU-EC and rDCM), while the converse did not hold. This demonstrates that GC is more conservative than EC in identifying significant connections at a group-level. At the level of link asymmetry, consistency between the two methods was achieved only for MOU-EC, and only if a group of 20 subjects or more was employed (fig. 4). Congruently, when aggregating asymmetry at the nodal level (by computing the difference of the out-strength and in-strength of each node), we obtained a good alignment between GC and mou-EC ($R^2 \sim 0.6$), but not between EC and rDCM-EC ($R^2 \approx 0$); Fig. 5). This implies that GC and EC could give widely different pictures of functional hierarchies, depending on the method chosen for EC estimation.
Our test-retest analysis (Fig. 6 also allowed us to assess the reliability of single-subject and group estimates of EC and GC. GC estimates were not reliable at the single-subject level, but we obtained a good test-retest consistency $R^2 = 0.8$ at the group level for groups of at least ~ 20 subjects. The reliability of EC estimates depended strongly on the chosen estimation method. RDCM yielded relatively consistent estimates ($R^2 = 0.5$) even at the individual level and well consistent estimates ($R^2 > 0.8$) for groups of at least 5 subjects, whereas MOU-EC was poorly consistent individually ($R^2 < 0.1$) and achieved good group consistency ($R^2 < 0.8$) only for groups of at least 50 subjects. This striking difference should be attributed to the different estimation procedures of the two methods. In the absence of a ground truth, we hesitate to conclude that rDCM yields "better" estimates. In fact, we observed that rDCM estimates correlated very strongly with the lagged covariance matrix ($R^2 \sim 0.6$). This strong bias may improve the test-retest consistency, but also impoverish the meaningfulness of rDCM EC, which appears to be barely different from a time-lagged FC. We stress that the lack of reliability for MOU-EC concerns individual connection estimates; this does not necessarily contradict previous work that leverages information from the whole network of connections to obtain robust prediction of clinical condition [40–42].

Future work and potential limitations. In this study, we did not investigate or model hemodynamics. The effect of hemodynamics could be multiple: it could introduce deviations from the model at the level of observed BOLD time series (such that a MOU or first order MVAR no longer provide an accurate description of the data), or bias timing relationships due to regional differences in the hemodynamic response, hence biasing EC/GC estimates (in particular estimates of connection asymmetry). In principle, one should estimate EC and GC from "neural" time series obtained after a deconvolution of the hemodynamic response. Additional work is needed to appreciate the effect of the hemodynamic response on the GC/EC relation, e.g. applying blind approaches to deconvolve the hemodynamic response from resting-state fMRI data (such as [66, 67]). Another possible limitation of this study is the assumption of a diagonal noise covariance matrix in the MVAR model: in the 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 (large) IC values observed in fMRI data, despite the fact that $\tau \sim \Delta t$. 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 the case of correlated noise.

Conclusion. Our study addressed the consistency of GC and EC in the reconstruction and analysis of directional 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. Moreover, the relation can emerge only in the limit of large sampling - which, due to practical limitations in fMRI acquisition, means that in practice

it can be observed only at the group level for at least $10 - 20$ subjects. Furthermore, the GC-EC alignment is strongly affected by the method used for EC estimation (MOU-EC giving much stronger alignment than rDCM). Taking these factors into account in the analysis of real fMRI recordings, we showed that the two methods yield a relatively consistent description of directional brain networks at a group level.

Data and Code Availability

All simulated data and code used in the present manuscript is available at https://github.com/micheleallegra/GC-and-MOU. The HCP data are part of the 1200 Subjects Release (S1200) and are publicly available on the ConnectomeDB database https://db.humanconnectome.org.

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

2. Varela F, Lachaux JP, Rodriguez E, Martinerie J. The brainweb: phase synchronization and large-scale integration. Nature reviews neuroscience. 2001;2(4):229–239.

3. Deco G, Tononi G, Boly M, Kringelbach ML. Rethinking segregation and integration: contributions of whole-brain modelling. Nature Reviews Neuroscience. 2015;16(7):430–439.

4. Battaglia D, Brovelli A. Functional connectivity and neuronal dynamics: insights from computational methods. The Cognitive Neurosciences, Sixth Edition. 2020;.

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

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

7. 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.

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

9. 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.

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. 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.

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

13. 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.

14. 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.

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

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

17. 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.

18. Bielczyk NZ, Uithol S, van Mourik T, Anderson P, Glennon JC, Buitelaar JK. Disentangling causal webs in the brain using functional magnetic resonance imaging: A review of current approaches. Network Neuroscience. 2019;3(2):237–273.

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

20. 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.

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

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

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

24. 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.

25. 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.

26. 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.

27. 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.

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

29. Tang W, Bressler SL, Sylvester CM, Shulman GL, Corbetta M. Measuring Granger causality between cortical regions from voxelwise fMRI BOLD signals with LASSO. PLoS computational biology. 2012;8(5):e1002513.

30. Bressler SL, Tang W, Sylvester CM, Shulman GL, Corbetta M. Top-down control of human visual cortex by frontal and parietal cortex in anticipatory visual spatial attention. Journal of Neuroscience. 2008;28(40):10056–10061.

31. Asemi A, Ramaseshan K, Burgess A, Diwadkar VA, Bressler SL. Dorsal anterior cingulate cortex modulates supplementary motor area in coordinated unimanual motor behavior. Frontiers in Human Neuroscience. 2015;9:309.

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

33. 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.

34. Friston KJ, Kahan J, Biswal B, Razi A. A DCM for resting state fMRI. NeuroImage. 2014;94:396–407. doi:10.1016/j.neuroimage.2013.12.009.

35. 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.

36. 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.

37. 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.

38. Deco G, Kringelbach ML, Jirsa VK, Ritter P. The dynamics of resting fluctuations in the brain: metastability and its dynamical cortical core. Scientific reports. 2017;7(1):3095.

39. Jobst BM, Hindriks R, Laufs H, Tagliazucchi E, Hahn G, Ponce-Alvarez A, et al. Increased stability and breakdown of brain effective connectivity during slow-wave sleep: mechanistic insights from whole-brain computational modelling. Scientific reports. 2017;7(1):4634.

40. 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.

41. 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.

42. 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.

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

44. Idesis S, Allegra M, Vohryzek J, Perl YS, Metcalf NV, Griffis JC, et al. Generative whole-brain dynamics models from healthy subjects predict functional alterations in stroke at the level of individual patients. Brain Communications. 2024;6(4):fcae237.

45. 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.

46. 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.

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

48. Frässle S, Harrison SJ, Heinzle J, Clementz BA, Tamminga CA, Sweeney JA, et al. Regression dynamic causal modeling for resting-state fMRI. Human brain mapping. 2021;42(7):2159–2180.

49. Frässle S, Aponte EA, Bollmann S, Brodersen KH, Do CT, Harrison OK, et al. TAPAS: an open-source software package for translational neuromodeling and computational psychiatry. Frontiers in psychiatry. 2021;12:680811.

50. Jianxiao W. rDCM toolbox in Python; 2023. Available from: `https://github.com/jadecci/rDCM_py`.

51. 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.

52. 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.

53. 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.

54. 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.

55. 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.

56. 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;.

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

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

59. 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).

60. 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.

61. 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.

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

63. 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;.

64. 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.

65. 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.

66. 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.

67. 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.

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

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

72. 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.

Supporting information - Appendix

Multivariate autoregressive (MVAR) models and Granger Causality

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

$$\mathbf{x}(t+\Delta t) = A\mathbf{x}(t) + \boldsymbol{\epsilon}(t) \tag{13}$$

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. (2) corresponds to the residual of the linear regression

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

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, $x_j(t+\Delta t)$, namely

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

.

To quantify the Granger causal effect of node i on node j, one can measure the relevance of $x_i(t)$ in predicting $x_j(t+\Delta t)$. To this aim, one defines a a "reduced" MVAR process where the influence of node i is removed:

$$\mathbf{x}(t+\Delta t) = A'\mathbf{x}_{\neg i}(t) + \boldsymbol{\epsilon}'(t) \tag{14}$$

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

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

The effect of node i on node j defined by *Granger causality* (GC) [23, 47] is given by the log-ratio of the variances

$$G_{ij} = \log\left(\frac{S'_{jj}}{S_{jj}}\right) \tag{15}$$

Note that we consider the conditional version of GC, meaning that the linear regression in Eq. (3) includes all remaining nodes in the network. Also note that Eq. (4) 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 MVAR coefficient a_{ji}. Assuming that $A' \simeq A^{\neg i}$ (where $A^{\neg i}$ is simply A without the i-th column), one obtains

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

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

$$G_{ij} \simeq \log\left(\frac{A_{ji}^2 Q_{ii}^0 + S_{jj}}{S_{jj}}\right) = \log\left(1 + \frac{A_{ji}^2 Q_{ii}^0}{S_{jj}}\right) \tag{16}$$

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 (4), Geweke [47] 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 MVAR (2) and the reduced MVAR (3). An alternative and computationally simple approach exploits a relation between the variance of the MVAR residuals and covariance terms, assuming Gaussian innovations [53, 70, 71]. Indeed, it can be shown that

$$Var[\epsilon_j] \;= Var[x_j(t+\Delta t)|\mathbf{x}(t)] = \frac{\det(Cov[x_j(t+\Delta t), \mathbf{x}(t)])}{\det(Cov[\mathbf{x}(t)])} \tag{17}$$

$$Var[\epsilon'_j] \;= Var[x_j(t+\Delta t)|\mathbf{x}_{\neg i}(t)] = \frac{\det(Cov[x_j(t+\Delta t), \mathbf{x}_{\neg i}(t)])}{\det(Cov[\mathbf{x}_{\neg i}(t)])} \tag{18}$$

Furthermore, we have:

$$\begin{aligned} \det S^{[ij]} &= \det(Cov[x_i(t+\Delta t)x_j(t+\Delta t)|\mathbf{x}(t)]) = \frac{\det(Cov[x_i(t+\Delta t), x_j(t+\Delta t), \mathbf{x}(t)])}{\det(Cov[\mathbf{x}(t)])} \\ \det S_{ii} &= \det(Cov[x_i(t+\Delta t)|\mathbf{x}(t)]) = \frac{\det(Cov[x_i(t+\Delta t), \mathbf{x}(t)])}{\det(Cov[\mathbf{x}(t)])} \\ \det S_{jj} &= \det(Cov[x_j(t+\Delta t)|\mathbf{x}(t)]) = \frac{\det(Cov[x_j(t+\Delta t), \mathbf{x}(t)])}{\det(Cov[\mathbf{x}(t)])} \end{aligned}$$

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

$$G_{ij} = \log \frac{\det(Cov[\mathbf{x}(t)])\det(Cov[x_j(t+\Delta t), \mathbf{x}_{\neg i}(t)])}{\det(Cov[x_j(t+\Delta t), \mathbf{x}(t)])\det(Cov[\mathbf{x}_{\neg i}(t)])} \tag{19}$$

$$I_{ij} = \log \frac{\det(Cov[x_i(t+\Delta t), \mathbf{x}(t)])\det(Cov[x_j(t+\Delta t), \mathbf{x}(t)])}{\det(Cov[\mathbf{x}(t)])\det(Cov[x_i(t+\Delta t)x_j(t+\Delta t)|\mathbf{x}(t)])} \tag{20}$$

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

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

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 [22].Granger causality measures can therefore be formulated in completely information-theoretical terms, based on entropy estimates [28, 53].
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 [72].

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

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

$$d\mathbf{x} = J\mathbf{x}dt + dw \tag{21}$$

where

$$J = -\frac{1}{\tau}\mathbb{I} + C^T \tag{22}$$

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. (1) satisfies the (continuous) Lyapunov equation

$$JQ^0 + Q^0J^T + \Sigma = 0 \tag{23}$$

Furthermore, the 0-lagged and Δt-lagged covariance matrix Q^0 and $Q^{\Delta t}$ obey the relation

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

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

$$J = \frac{1}{\Delta t}\log((Q^0)^{-1}Q^{\Delta t})^T \tag{24}$$

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 $Q^{\Delta}t$ in Eq. (24) 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 [36]. In essence, it uses a partial differentiation of Eq. (24) 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 (1), we can build an equivalent 1st order MVAR process (2). Indeed, by integrating Eq. (1) for a given lag Δt, one obtains

$$\mathbf{x}(t+\Delta t) = e^{J\Delta t}\mathbf{x}(t) + \mathbf{w} \tag{25}$$

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

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

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

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

$$S = \tilde{\Sigma}, \qquad A = e^{J\Delta t} = e^{-\frac{\Delta t}{\tau}} e^{C^T \Delta t} \tag{26}$$

which implies the following constraints on S and A:

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

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

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 (1) with real coefficients is associated with an equivalent MVAR process (2) 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 MVAR* when the process can be associated to a real-valued J.

Theoretical relation between GC and EC

Assuming compatibility, 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:

$$e^{C^T \Delta t} \simeq \mathbb{I} + C^T \Delta t \tag{27}$$

Eq. (27) holds under the following condition

$$||C|| \ll \frac{1}{\Delta t} \tag{28}$$

where $||\cdot||$ denotes matrix 2-norm. Note that this condition can be formulated considering the dominant eigenvalue of C: $\mu_1 \ll \frac{1}{\Delta t}$ 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,

$$\begin{aligned} \tau \ll \Delta t \qquad & \text{process faster than sampling = fast dynamics} \\ \tau \approx \Delta t \qquad & \text{process as fast as sampling = matched dynamics} \\ \tau \gg \Delta t \qquad & \text{process slower than sampling = slow dynamics} \end{aligned}$$

In the slow case, Eq. (28) is always satisfied since $|\mu_i| \leq \frac{1}{\tau} \ll \frac{1}{\Delta t}$. In the matched case, Eq. (28) is satisfied whenever

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

becoming more constraining in the fast case

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

In practice, condition (28) can often be considered to be satisfied in both the slow and the matched cases, essentially because large values of C_{ij} are incompatible with stability, so $||C|| \ll 1/\tau$. However, fast dynamics might break the assumption.

We can show that, provided Eq. (27) holds, there exist quadratic relations between IC, GC and EC. As detailed in Eqs. (29) and (31) in S1Text.

If Eq.(27) holds, we have quadratic relations between IC, GC and EC. Indeed, $A_{ij} \simeq \Delta t e^{-\frac{\Delta t}{\tau}} C_{ji}$ and

$$G_{ij} = \log(1 + \Delta t^2 \frac{e^{-\frac{2\Delta t}{\tau}} C_{ji}^2 Q_{ii}^0}{S_{jj}}) \simeq \Delta t^2 \frac{e^{-\frac{2\Delta t}{\tau}} C_{ji}^2 Q_{ii}^0}{S_{jj}} \tag{29}$$

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

$$\begin{aligned} S = \int_0^{\Delta t} dt\ e^{Jt} \Sigma e^{J^T t} \simeq \int_0^{\Delta t} dt\ e^{-\frac{2t}{\tau}} (\mathbb{I} + Ct) \Sigma (\mathbb{I} + C^T t) \simeq \int_0^{\Delta t} dt\ e^{-\frac{2t}{\tau}} (\Sigma + C\Sigma t + \Sigma C^T t) = \\ = \frac{\tau}{2} \Big((1 - e^{-\frac{2\Delta t}{\tau}}) \Sigma + (-\Delta t e^{-\frac{2\Delta t}{\tau}} + \frac{\tau}{2}(1 - e^{-\frac{2\Delta t}{\tau}}))(C\Sigma + \Sigma C^T) \Big) = \\ = \frac{\tau}{2}(1 - e^{-\frac{2\Delta t}{\tau}}) \Big(\Sigma + \big(\frac{\tau}{2} - \Delta t \frac{e^{-\frac{2\Delta t}{\tau}}}{1 - e^{-\frac{2\Delta t}{\tau}}} \big)(\Sigma C + \Sigma C^T) \Big) \end{aligned}$$

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

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

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 ($A \approx e^{-\frac{-\Delta t}{\tau}}$)

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

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

$$G_{ij} \simeq \Delta t^2 \frac{C_{ji}^2}{e^{\frac{2\Delta t}{\tau}} - 1} \frac{\sigma_i^2}{\sigma_j^2} \tag{30}$$

For the instantaneous causality, one obtains

$$I_{ij} \simeq \frac{S_{ij}^2}{S_{ii} S_{jj}} \simeq \tau^2 \big(1 - \frac{2\Delta t}{\tau} \frac{e^{-\frac{2\Delta t}{\tau}}}{1 - e^{-\frac{2\Delta t}{\tau}}}\big)^2 \tilde{C}_{ij}{}^2 \tag{31}$$

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

, we have

$$G_{ij} \simeq \Delta t^2 \frac{1}{e^{\frac{2\Delta t}{\tau}} - 1} \frac{\sigma_i^2}{\sigma_j^2} C_{ij}^2 \tag{32}$$

$$I_{ij} \simeq \tau^2 \big(1 - \frac{2\Delta t}{\tau} \frac{e^{-\frac{2\Delta t}{\tau}}}{1 - e^{-\frac{2\Delta t}{\tau}}}\big)^2 \frac{\sigma_i^2 C_{ij} + \sigma_j^2 C_{ji}}{2\sigma_i \sigma_j} \tag{33}$$

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:

$$G_{ij} \simeq \Delta t^2 \frac{C_{ij}^2}{e^{\frac{2\Delta t}{\tau}} - 1}$$

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

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:

$$cG_{ij} = \frac{Q_{jj}^0}{Q_{ii}^0} G_{ij} \simeq \Delta t^2 \frac{C_{ij}^2}{e^{\frac{2\Delta t}{\tau}} - 1} \tag{34}$$

$$cI_{ij} = I_{ij} \cdot \frac{4 Q_{ii}^0 Q_{jj}^0}{Q_{ii}^0 + Q_{jj}^0} \simeq \tau^2 \big(1 - \frac{2\Delta t}{\tau} \frac{e^{-\frac{2\Delta t}{\tau}}}{1 - e^{-\frac{2\Delta t}{\tau}}}\big)^2 C_{ij}{}^2 \tag{35}$$

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 (23) 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:

$$\rho = \frac{cG_{ij}}{cI_{ij}} \approx \frac{(\Delta t^2/\tau^2)}{\big(e^{\frac{2\Delta t}{\tau}} - 1 - \frac{2\Delta t}{\tau}\big)^2} \simeq \frac{(\Delta t^2/\tau^2)}{\big(2\frac{\Delta t^2}{\tau^2}\big)^2} = \frac{1}{2}\frac{\tau^2}{\Delta t^2} \tag{36}$$

This means that for slow dynamics we have $cG \gg cI$, whereas $cG \ll cI$ for fast dynamics (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 dynamical regime.

Supplementary figures

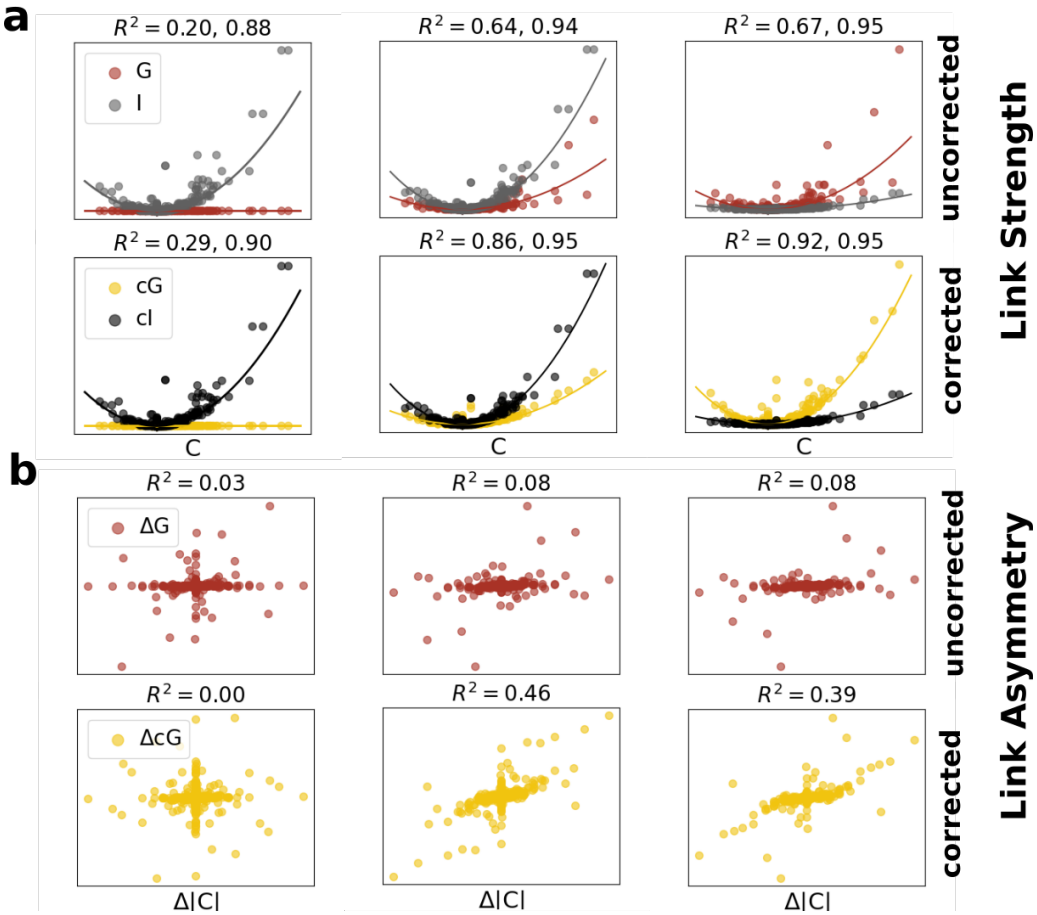

Figure S1: **Theoretical relations between model EC and non-conditional GC/IC.** We considered a random network of $N = 40$ nodes evolving according to the MOU of $\tau = 0.1, 1, 10$. In **(A-C)** we show the relation between the EC weights and the corresponding values of GC and IC. The GC and IC were 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.

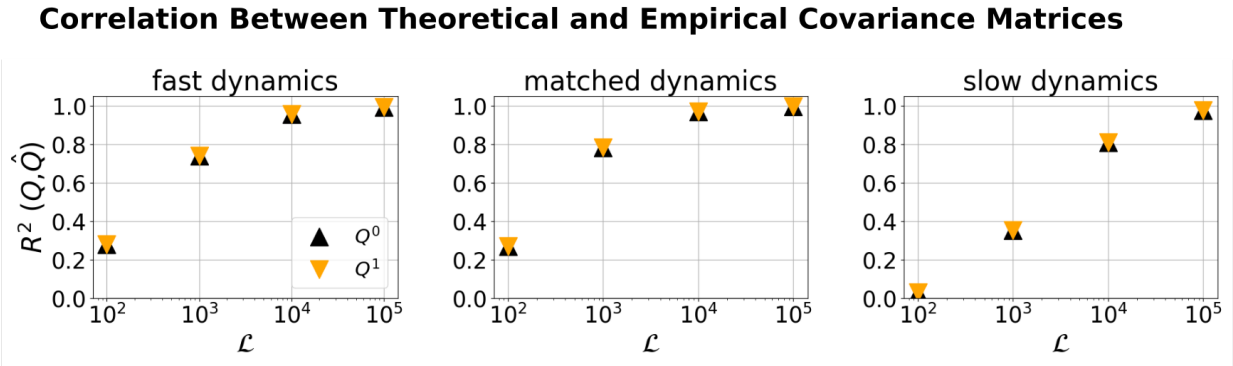

Figure S2: **Error in covariance matrices for finite sampling time.** We considered the same random network as in Fig. 2. For finite sampling length $\mathcal{L} < \infty$, the covariance matrices Q^0, Q^1 are affected by an estimation error, whose magnitude decreases with $\mathcal{L}$. We show here the relation between the ideal Q^0, Q^1 and the empirical $\hat{Q}^0, \hat{Q}^1$, for different values of τ and $\mathcal{L}$, measured in terms of squared Pearson correlation R^2.

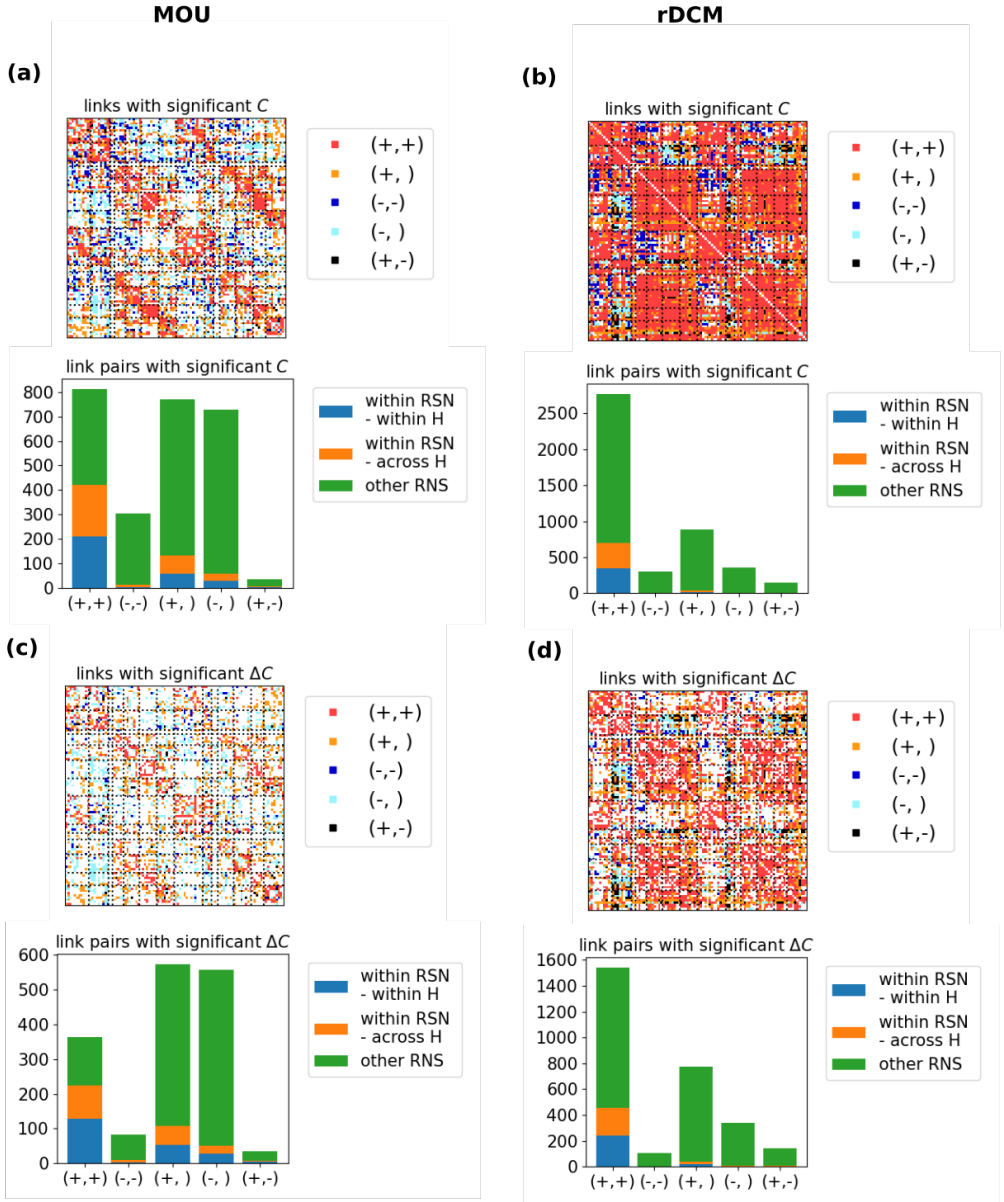

Figure S3: **Effective Connection sign and asymmetries**. We considered all pairs of reciprocal connections, and divided them into *sign categories* according to their sign: $(+,+)$ both connections are significantly positive $(+,)$ one connection is significantly positive and the reciprocal is non-significant $(-,)$ 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.

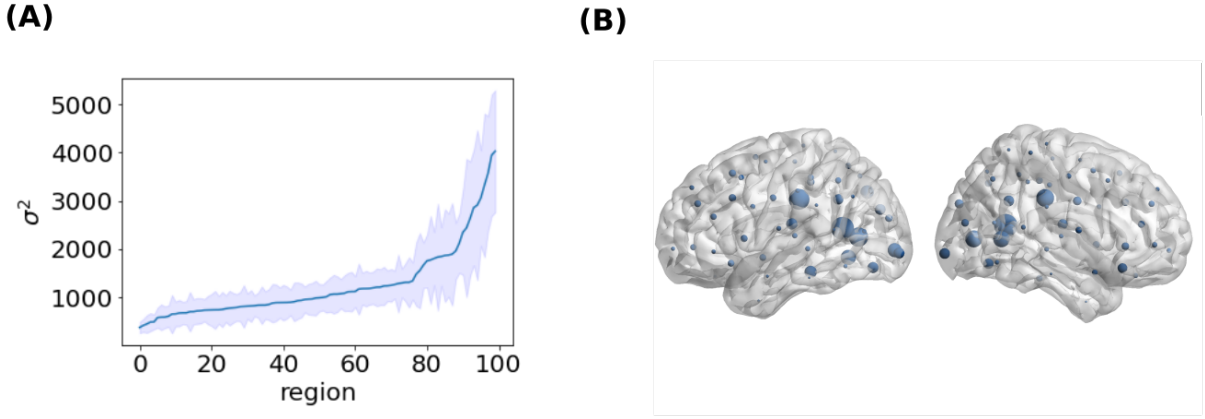

Figure S4: **Noise affecting different nodes**. **(A)** For each region, we computed the average (over subjects) noise σ^2 and sorted regions by this value. The blue line represents the average (over subjects) noise σ^2, the shaded area represents standard deviation (over subjects). **(B)** In this rendering, the node size is proportional to the average (over subjects) noise σ^2.

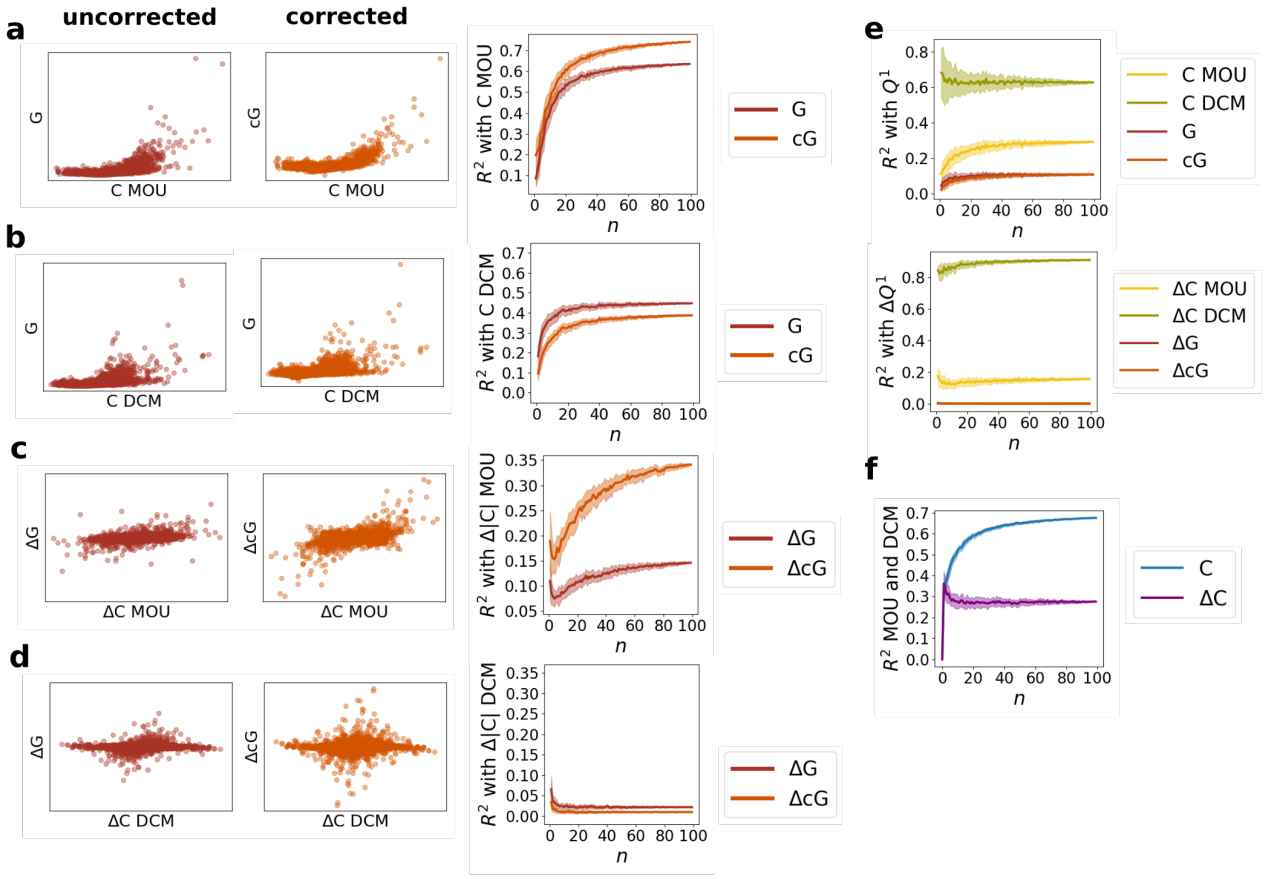

Figure S5: Replication of fig. 4 with independent recording sessions **(A)** effective connectivity $\hat{C}$ (group average over 100 subjects) vs corrected Granger causality $\hat{cG}$ (group average). **(B)** squared Pearson correlation R^2 between group $\hat{C}$ and group $\hat{cG}$ for groups of 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.