Dynamical differential covariance recovers directional network structure in multiscale neural systems

Yusi Chen,[1,2*] Burke Q. Rosen,[4] Terrence J. Sejnowski[1,2,3*]

[1]Computational Neurobiology Laboratory,
Salk Institute for Biological Sciences, La Jolla, CA, 92037

[2]Section of Neurobiology, Division of Biological Sciences,
University of California San Diego, La Jolla, CA, 92093

[3]Institute for Neural Computation, University of California San Diego, La Jolla, CA, 92093

[4]Neurosciences Graduate Program, University of California San Diego, La Jolla, CA, 92093

*To whom correspondence should be addressed: cyusi@ucsd.edu(Y.C.); terry@salk.edu(T.J.S.)

Investigating causal neural interactions are essential to understanding subsequent behaviors. Many statistical methods have been used for analyzing neural activity, but efficiently and correctly estimating the direction of network interactions remains difficult (*1*). Here, we derive dynamical differential covariance (DDC), a new method based on dynamical network models that detects directional interactions with low bias and high noise tolerance without the stationary assumption. The method is first validated on networks with false positive motifs and multiscale neural simulations where the ground truth connectivity is known. Then, applying DDC to recordings of resting-state functional magnetic resonance imaging (rs-fMRI) from over 1,000 individual subjects, DDC consistently detected regional interactions with strong structural connectivity. DDC can be generalized to a wide range of dynamical models and recording techniques.

1

Understanding the neural mechanisms underlying behavior requires knowing how neurons interact with each other. Direct circuit tracing by connectomics studies (*2–4*) provides ground truth, but at high cost and the resulting static connection matrix does not by itself reveal the dynamical aspects of neural communication. This has motivated statistical methods (*1, 5*) to estimate functional connectivity (FC). (We refer here to all methods based on correlation or causation as FC (*6*)).

FC is often evaluated by estimating pairwise correlations, a symmetric measure that cannot detect directional coupling or disambiguate two unconnected nodes confounded with high correlation due to a common input (*1, 7*). Though far from causality, correlation provides an intuitive description of co-variations between time series averaged across all samples. DDC was also motivated by such measure of co-variations. More sophisticated methods including Granger causality (*8*), cross convergent mapping (CCM) (*9*) and cross-dynamical delay differential analysis (DDA) (*10*) use predictability to identify causal relationship between time series. Other generative methods such as dynamic causal modeling (DCM) (*11*) and Bayes net models (*12, 13*) search all possible causal graphs and fit the entire dataset to every hypothesis. These mathematically involved methods work under particular assumptions, require a prodigious amount of computation and does not scale well.

We previously introduced differential covariance (dCov) (*14, 15*), a directed FC estimation method, and highlighted the performance of two matrices, Δc, which calculates the correlation between the derivative signal and the signal itself, and Δp, which evaluates the partial covariance between them. In simulated test cases, they detected network connections with higher sensitivity than many of the methods reviewed in Smith et al (*1*). However, like many other FC methods, dCov is empirical and does not depend on how the network activity was generated. In this paper, we derive a direct link between dCov and dynamical network models, which leads to a new class of estimators called dynamical differential covariance (DDC).

Models of neural systems span a wide range of scales. At the microscopic level, the voltage trace, calcium dynamics and firing rate of a single neuron are highly nonlinear. These dynamics are often modeled using biophysical models based on voltage-gated ion channels. In contrast, at the macroscopic level the collective activity of a population of neurons and interactions between brain regions is often be approximated by linear dynamics because of ensemble averaging (*5, 16, 17*). For example, Abdelnour et al (*16*), used a linear forward model to predict correlations in resting-state fMRI activity based on connectivity from diffusion MRI (dMRI). Their linear model significantly outperformed nonlinear neural mass models, arguing for the linearity of large-scale brain activities, especially when measured by fMRI BOLD signals dominated by low frequencies.

We first derive DDC from a linear dynamical model in Eq. 1 for global recordings and a nonlinear dynamical model in Eq. 2 for local neural recordings:

$$\frac{d\mathbf{x}}{dt} = \mathbf{W}\mathbf{x} \qquad (1) \qquad \frac{d\mathbf{x}}{dt} = \mathbf{W}R(\mathbf{x}) \qquad (2)$$

The column vector $\mathbf{x}$ is the neural activity, such as the membrane voltage or fMRI signal, $\mathbf{W}$ is the square connectivity matrix and $R(\mathbf{x})$ is a nonlinear response function. Combining the above equations with $\mathbf{x}$ yields:

$$
\langle \frac{d\mathbf{x}}{dt}, \mathbf{x} \rangle = \mathbf{W}\langle \mathbf{x}, \mathbf{x} \rangle \qquad\qquad \langle \frac{d\mathbf{x}}{dt}, \mathbf{x} \rangle = \mathbf{W}\langle F(\mathbf{x}), \mathbf{x} \rangle
$$
$$
(3) \qquad\qquad\qquad\qquad (4)
$$
$$
\mathbf{\Delta L} := \langle \frac{d\mathbf{x}}{dt}, \mathbf{x} \rangle \langle \mathbf{x}, \mathbf{x} \rangle^{-1} \qquad\qquad \mathbf{\Delta R} := \langle \frac{d\mathbf{x}}{dt}, \mathbf{x} \rangle \langle R(\mathbf{x}), \mathbf{x} \rangle^{-1}
$$

where $\mathbf{\Delta L}$ and $\mathbf{\Delta R}$ are DDC estimators for $\mathbf{W}$ and the operator $\langle , \rangle$ takes the outer product of two vectors and performs time averaging.

The origin of the linear DDC estimator $\mathbf{\Delta L}$ from a dynamical model provides an intuition for its effectiveness in estimating $\mathbf{W}$ as the product of two matrices: The first is differential covariance, which carries information about inputs. In a neuron this is the inward current from synaptic inputs and in brain imaging it is related to changes in surrogates for local brain ac-

3

tivity; In the second term, an entry in the partial covariance matrix is zero if and only if the partial correlation between x_i and x_j is zero, which cancels the influence of common sources. By combining information from input sinks and canceling information coming from common sources it becomes possible to robustly estimate directional interactions.

A family of estimators arises from the DDC estimator ΔR for nonlinear dynamical systems depending on $R(\mathbf{x})$. Estimators can be adapted to the filtering effects from different recording techniques, such as the slow kinetics of calcium signals, by choosing the nonlinear function $R(\mathbf{x})$ appropriately. Here, we use the rectified linear unit (ReLU), often used in artificial neural networks (*18*), parameterized by a threshold (θ), and ΔReLU is the corresponding nonlinear DDC estimator. Intuitively, this threshold function should rectify low magnitude "noise" and retain large signals.

In the Supplementary Material 1.3.1, we show analytically that in systems governed by stochastic differential equations, DDC gives unbiased estimates of $\mathbf{W}$ and is robust to added Brownian motion noise. Even when the simulated dynamics do not match the actual dynamics, we find empirically that the DDC only has a small bias. This suggests that DDC can decrease the systematic inference error due to model mismatch (*19*). We also prove that DDC can be used to analyze nonstationary data whose higher-order statistics vary with time.

As a proof of principle, we first applied DDC to three-node networks with varying dynamics and network structures (Fig. 1). The chain motif (Fig. 1A,B) and confounder motif (Fig. 1C) were chosen because they both have a node pair (red dashed line) that is highly correlated but with no physical connection, which is an ideal test of whether DDC can "explain away" spurious correlations. We simulated both linear and sigmoid based nonlinear dynamics (Methods). Both ΔL and ΔReLU correctly inferred the existence and direction of the ground truth connections (Fig. 1) while the covariance matrix (Cov) failed to explain away false positive connections and partial covariance (P) was not able to determine the directionality of connections (Fig. S2).

We also quantified the variance and bias over a range of data size and observational noise (Fig. S1). DDC consistently had the least estimation error regardless of the size of the dataset (Fig. S1A). In contrast, inference bias for Cov, P and Δp diverged as data volume increased (Fig. S1A), introducing a systematic error. Regarding noise tolerance (Fig. S1B), the performance of dCov matrices (Δc and Δp) rapidly deteriorated with increasing noise, but DDC (ΔL and ΔReLU) remained robust. For models with highly nonlinear dynamics and low randomly driven input, the performance of the DDC estimators (Fig. S1C) was significantly better than Δc and Δp (Fig. S1D).

DDC was also applied to a larger network consisting of 50 nodes and structured by a combination of confounder and chain motifs (Fig. S2). As in the small network case, ΔL and ΔReLU cleanly estimated the existence and direction of connections (Fig. S2B). Estimation performance increased with larger data sets (Fig. S2C).

Can DDC track the information flow in a nonlinear Rössler system, which is deterministic but chaotic? The three equations for this system in Fig. 1D have a nonlinear bidirectional confounder motif. ΔL and ΔReLU correctly identified direct connections, and correctly ignored the strong correlations between x_2 and x_3.

Next, we tested DDC on spiking data from a network model with 200 Leaky Integrate-and-Fire (LIF) neurons (*20*). These neurons integrate exponentially filtered synaptic inputs until the membrane potential reaches a threshold, which triggers a spike and a reset to resting membrane potential. The connectivity matrix was a Erdős–Rényi random graph with uniform connection strengths, emphasizing the existence and direction of network edges (Methods). Graphs with a range of sparsity and connection strengths were simulated and DDC was applied to the sub-threshold membrane potentials (Methods). Performance was quantified by the area under the curve (AUC) of specificity versus sensitivity (Fig. 2A,B). ΔL and ΔReLU were significantly (p<0.001, rank-sum test) better than all other methods. This was partly due to DDC's additional

5

A: Chain motif + linear dynamics

B: Chain motif + nonlinear dynamics

C: Confounder motif + linear dynamics

D: Chaotic Rössler system

$$\dot{x_1} = -x_2 - x_3$$
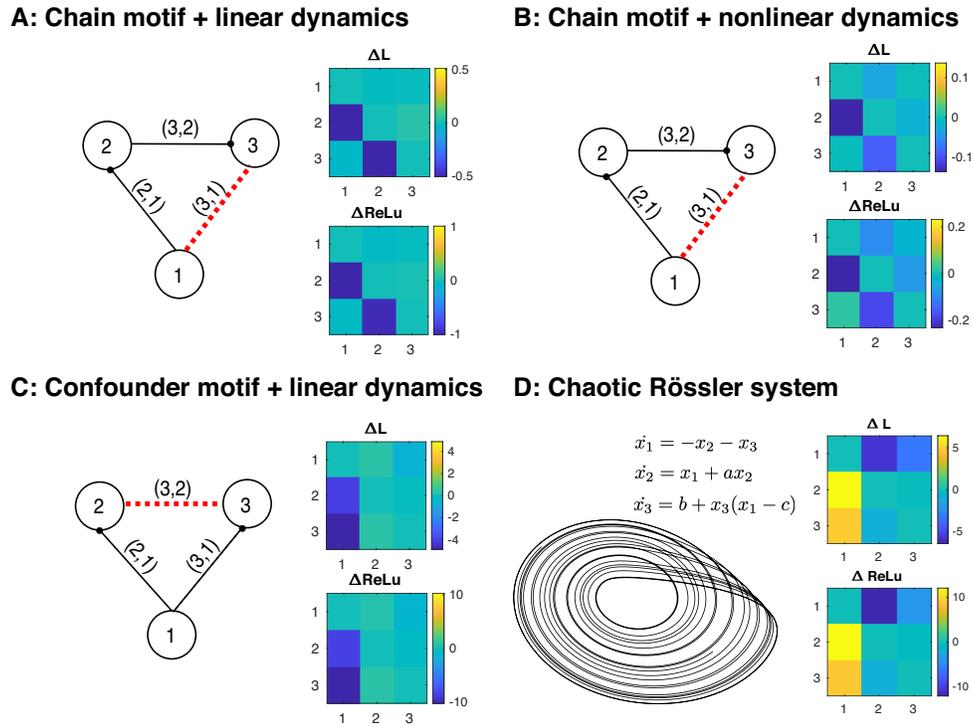$$\dot{x_2} = x_1 + ax_2$$
$$\dot{x_3} = b + x_3(x_1 - c)$$

Figure 1: **DDC recovers ground truth connectivity across multiple 3-node networks.** (A)(B)(C) Left: ground truth network structure. Black solid lines are directed physical connections and red dashed lines are false positive connections commonly inferred by covariance estimation; The edges label (i, j) stands for the matrix entry at the i-th row and j-th column. Right: estimated ΔL and ΔReLU. (D) Left: phase diagram of x_1 and x_2 of the Rosseler system governed by system equations shown above; Right: estimated ΔL and ΔReLU. For a clear illustration, we removed the diagonal values from estimated matrices.

6

capacity to estimate direction since the ground truth matrices are not symmetric. Interestingly, Δp had very high sensitivity (true positive rate) even when very few connections were thresholded as positive. This might due to the sparse estimation of Δp (*14*). DDC was robust to a broad range of network configurations (Fig. S3D).

We also tested the methods on simulated recordings of macroscopic neural activities based on the reduced Wong-Wang model of the resting state (*21*). The connectivity matrix used short-range local interactions and DSI (diffusion spectrum imaging) measurements of long-range structural connectivity (*22*) (Fig. 2C). The overall performance was quantified by c-sensitivity (Methods), which measured the separation of estimated values between true positive connections and true negative connections. C-sensitivity = 1 means the estimated matrix completely separated true positives from the others. ΔL had the highest performance followed by ΔReLU (Fig. 2D), probably because the reduced Wong-Wang model exhibited linear fluctuations around the stable point (*21*). The raw ΔL and ΔReLU matrices uncovered the strongest connections (red arrows) in the ground truth matrix (Fig. S4B). (Only the strongest long-range connections were included, because all methods failed to reach significance for graded anatomical connectivity (Fig. S5).)

A critical test of DDC is performance on real data. We applied DDC to resting-state fMRI (rs-fMRI) recordings obtained from the Human Connectome Project (HCP). The imaging voxels were parcellated through group ICA (Methods), where each independent component (IC) parcellation, shared across subjects, is composed of voxels with similar dynamics. ICs are mainly composed of spatially proximate voxels, forming anatomically recognizable brain regions (Fig. S7). In addition, we focused on the first 46 ICs that had over 40% cortical voxels (Fig. S6A) to match dMRI cortical measurements. Dual regression (Methods) assigned unique ICA-parcellated BOLD signals to each subject, which were treated as nodes for DDC analysis.

The average and standard deviation of the estimated DDC matrices across subjects are
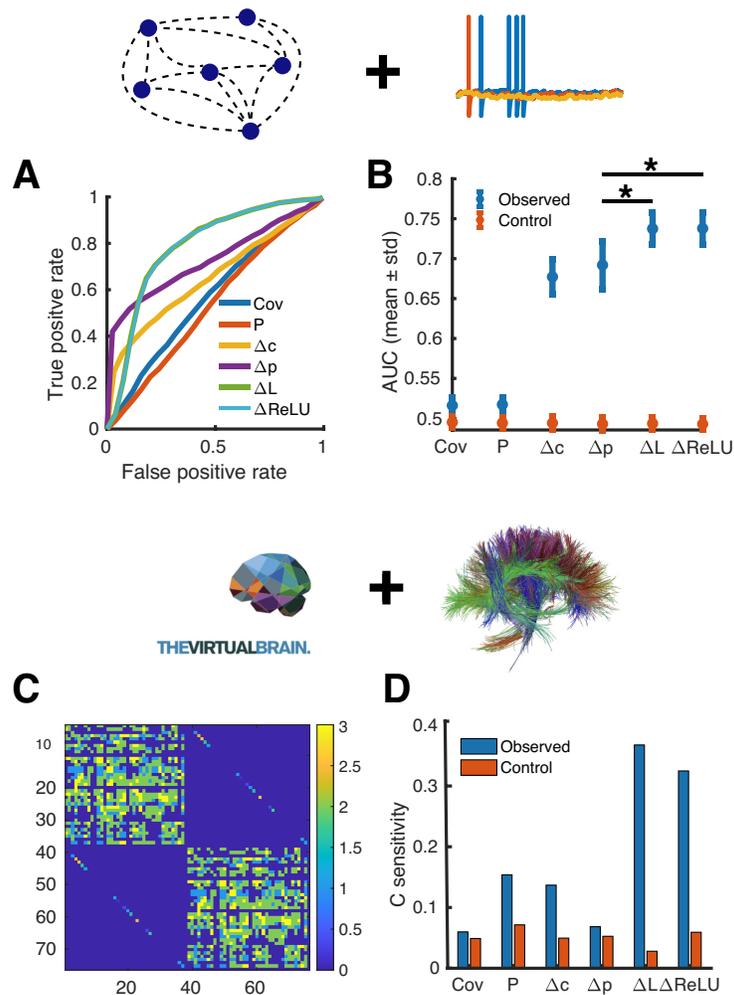
Figure 2: **Microscopic level:** LIF network (A) Receiver operating curves (ROC) quantifying classification performance of true connections. The curves for ΔL and ΔReLU were identical due to threshold selection process (Methods). (B) Area under ROC (AUC) across 50 realizations of the random graph. * p<0.001, rank-sum test. **Macroscopic level:** Resting-state brain surface model (C) Long-range structural connectivity. The block structure represents the left and the right hemisphere. (D) C-sensitivity of the estimators. Control values in (B)(D) were calculated using a different random graph realization served as GT.

shown in Fig. 3A and B where ΔL and ΔReLU were sparser than the covariance matrix. Two nodes (indicated by red arrow) in ΔL appeared to have a broader range of interactions. They were anatomically registered as "occipital pole" and "medial occipitotemporal gyrus", reflecting the large proportion of visual ICs in the network (Fig. S7 and Table S2). To increase noise tolerance, we binarized the estimated matrices based on significance levels as determined by an auto-regressive bootstraping procedure that preserved the signature power spectrum properties of BOLD signals (Methods). The significant ΔL connections (yellow entries, $p < 0.01$) across subjects were shown in Fig. 3C. These "backbone connections" shared across a majority of the subjects could be flexibly tuned for network sparsity level. We adopted a strict criterion because we were interested in the most conserved connections shared by over 90% subjects (red dashed vertical line in Fig. 3C). Their IC parcellations were registered on an MRI template and plotted in Fig. 3D. In this case, "Backbone connections" were identified between ICs from the same anatomical region (marked in red) as well as inter-regional interactions.

To quantify the extent to which estimated FCs matched the structural connectivity, we further processed dMRI measurements from the HCP dataset (*23*) to obtain individual-level IC-based dMRI matrices (Fig. S6, Methods). At the IC level, dMRI strengths were bimodal (Fig. 4A), indicating a clear separation between the strong and weak connections. ΔL identified connections with higher dMRI strength compared those chosen by the covariance matrix (Fig. 4B). Fig. 4C shows the increasing average dMRI strength for decreasing binarization threshold, linking the significance of rs-fMRI to dMRI connectivity for all methods and confirming their biological relevance. DDC uncovered connections with significantly higher dMRI strength values than covariance-based methods and also identified a larger proportion of strong connections (Fig. 4D).

Dynamical Differential Covariance is a promising new family of estimators for analyzing the structural connectivity of neural interactions underlying large-scale brain recordings. Be-
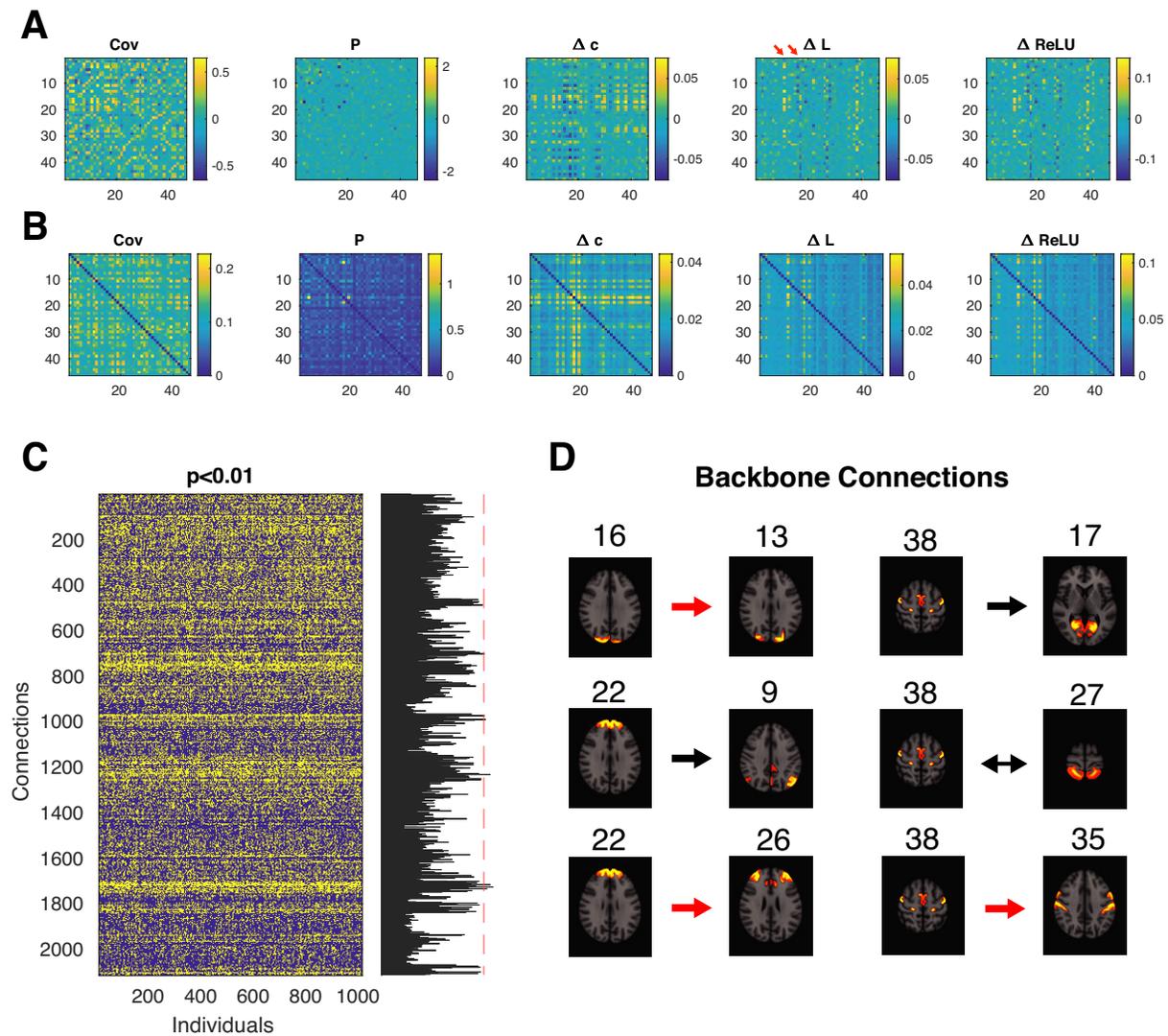
9

Figure 3: **DDC consistently recovered known connections across HCP subjects.** (A)(B) Average/Standard deviation of estimated FC matrices averaged across HCP subjects. (C) Individual level binarized significant (p-value<0.01) ΔL connections. Each column is a binarized matrix reshaped to a column vector. The summation of each row, shown on the right, represents the number of individuals that highlighted a specific connection. "Backbone connection" are those shared by most individuals (example threshold indicated by the red dashed line). (D) ΔL "backbone connection" shared by over 90% subjects and their IC parcellations registered on an MRI template. Arrows indicated the estimated connection direction and the red ones emphasized the IC pairs that are anatomically close to each other.
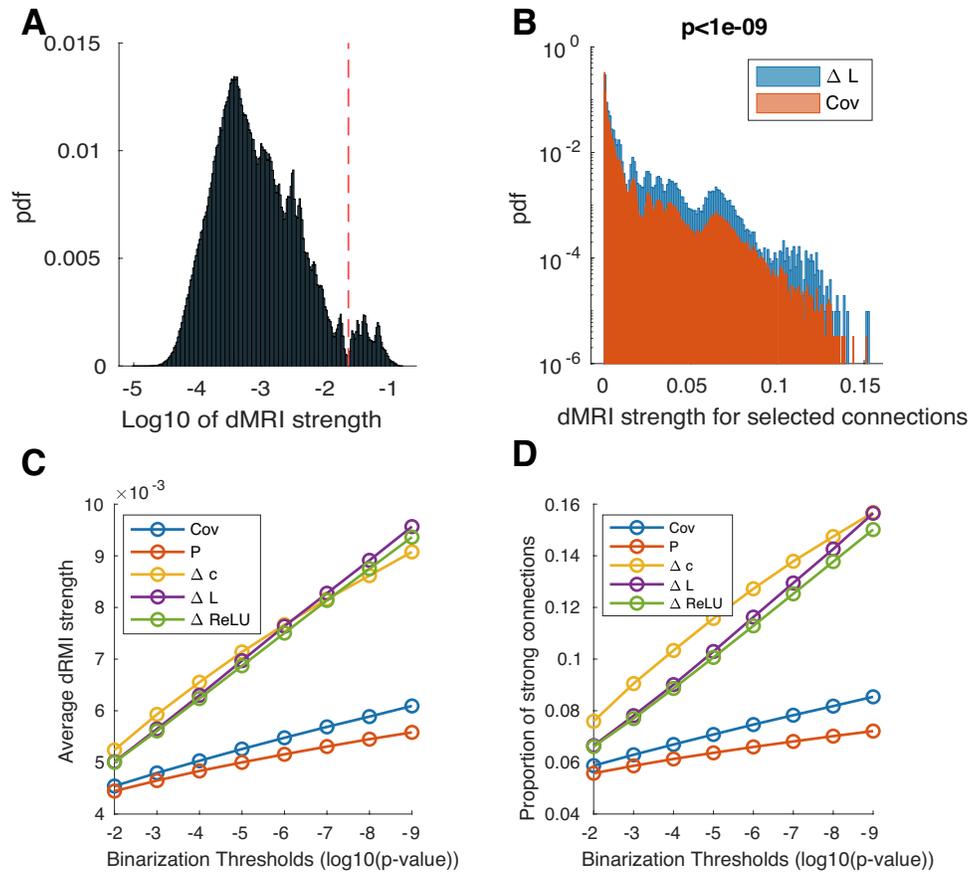
Figure 4: **DDC picked up connections with strong dMRI values.** (A) Distribution of IC-level dMRI strengths. Connections to the right of the cutoff value (red dashed line) were classified as strong connections. (B) The dMRI strength distribution of significant ΔL and Cov connections (p-value<1e-9). Note the log scale on the y axis presented due to the large abundance of weak connections. (C) Average dMRI strength value of significant connections picked by different methods with stricter binarization thresholds. (D) Proportion of strong connections. Since dMRI strength distribution is bimodal, the proportion of strong connections was also used as a supplementary statistics to compare the distributions.

cause DDC is derived directly from dynamical system equations that govern neural interactions, no optimization or model fitting is required. DDC is a practical and intuitive method that can be computed rapidly and scales well with the number of recording. Unlike methods based on covariance, which are inherently symmetrical, DDC can detect directional interactions and obtain statistical estimates of causality. DDC uncovered ground truth when applied to dynamical simulations of network models and significantly improved estimates of dMRI connectivity from rs-fMRI recordings.

In conclusion, DDC has a number of favorable mathematical properties that should ensure robust estimation of FC for a wide range of recordings. Access to the directionality of neural connections opens new avenues for interpreting the causal flow of information through networks. Identifying functional connectivity based on dynamical systems models makes direct contact with similar approaches in many other disciplines such as bioengineering, control theory and network science. DDC should have a broad impact on studies in these areas whenever there is need for estimates of directional network connectivity.

Declaration of Competing Interests

The authors declare no competing interests.

Acknowledgements

Center for Systems Neuroscience at Washington University.

Data and materials availability

Implementation of DDC, network simulation and HCP processing scripts are all available through `https://github.com/yschen13/DDC`.

References

1. S. M. Smith, *et al.*, *Neuroimage* **54**, 875 (2011).

2. S. W. Oh, *et al.*, *Nature* **508**, 207 (2014).

3. C. S. Xu, *et al.*, *BioRxiv* (2020).

4. K. Eichler, *et al.*, *Nature* **548**, 175 (2017).

5. S. Frässle, *et al.*, *Neuroimage* **155**, 406 (2017).

6. A. T. Reid, *et al.*, *Nature Neuroscience* **22**, 1751 (2019).

7. I. H. Stevenson, J. M. Rebesco, L. E. Miller, K. P. Körding, *Current opinion in neurobiology* **18**, 582 (2008).

8. C. W. J. Granger, *Econometrica* **37**, 424– (1969).

9. G. Sugihara, *et al.*, *Science* **338**, 496– (2012).

10. C. Lainscsek, C. E. Gonzalez, A. L. Sampson, S. S. Cash, T. J. Sejnowski, *Chaos* **29**, 101103 (2019).

11. K. J. Friston, L. Harrison, W. Penny, *Neuroimage* **19**, 1273 (2003).

12. R. Sanchez-Romero, *et al.*, *Network Neuroscience* **3**, 274 (2019).

13. J. D. Ramsey, S. J. Hanson, C. Glymour, *NeuroImage* **58**, 838 (2011).

14. T. W. Lin, A. Das, G. P. Krishnan, M. Bazhenov, T. J. Sejnowski, *Neural Computation* **29**, 2581 (2017).

15. T. W. Lin, *et al.*, *Neural Computation* **32**, 2389–2421 (2020).

16. F. Abdelnour, H. U. Voss, A. Raj, *Neuroimage* **90**, 335 (2014).

17. F. Abdelnour, M. Dayan, O. Devinsky, T. Thesen, A. Raj, *NeuroImage* **172**, 728 (2018).

18. I. Goodfellow, Y. Bengio, A. Courville, Y. Bengio, *Deep learning*, vol. 1 (MIT press Cambridge, 2016).

19. A. Das, I. R. Fiete, *Nature Neuroscience* **23**, 1286 (2020).

20. W. Nicola, C. Clopath, *Nature Communications* **8**, 1 (2017).

21. G. Deco, *et al.*, *Journal of Neuroscience* pp. 11239–11252 (2013).

22. P. Hagmann, *et al.*, *PLoS biology* **6**, e159 (2008).

23. B. Q. Rosen, E. Halgren, *Eneuro* **8** (2021).

Supplementary materials

Supplementary Materials and Methods

Figs. S1 to S7

Tables S1 to S2

References *(1-20)*

Supplementary Methods and Materials

Dynamical differential covariance recovers directional network structure in multiscale neural systems

Yusi Chen, Burke Rosen and Terrence J. Sejnowski

1 Functional Connectivity Estimators

All estimators and abbreviations are summarized in Table S1.

Table S1: Summary of estimators

Estimator	Notation
Cov	Covariance matrix
P	Partial covariance matrix
Δc	Differential covariance matrix
Δp	Partial differential covariance matrix
ΔL	Linear DDC
ΔR	General nonlinear DDC
$\Delta ReLU$	Nonlinear DDC with ReLU nonlinearity

1.1 Covariance based estimators

The covariance (Cov) and partial covariance (P) matricies are:

$$\text{Cov} = \langle \mathbf{x}, \mathbf{x} \rangle \tag{1}$$

$$\text{P} = \text{Cov}^{-1} \tag{2}$$

where $\mathbf{x}$ is a column vector of the system variable and the operation $\langle , \rangle$ takes the outer product of two vectors and averages across time. In this paper, all time traces were z-scored, thus, the

1

covariance matrix is equivalent to correlation. The calculation of partial covariance matrix (P) regressed out the confounding factors and is equivalent to the partial covariance matrix [2]. The covariance matrix only reveals pairwise correlations but the partial covariance matrix controls for confounding effects, one step closer to causal estimation.

1.2 Differential Covariance (dCOV) estimators

Differential covariance (Δc) was calculated as Eq.3 where $\frac{d\mathbf{x}}{dt}$ was numerically computed using a symmetric difference quotient [1]. The evaluation of partial differential covariance (Δp) was derived in parallel to partial covariance. The calculation was performed element-wise as in Eq.4 where Cov refers to the covariance matrix, K denotes the set of all nodes except i and j.

$$\Delta c = \langle \frac{d\mathbf{x}}{dt}, \mathbf{x} \rangle \tag{3}$$

$$\Delta p_{ij} = \Delta c_{ij} - \text{Cov}_{jK}\text{Cov}_{KK}^{-1}\Delta c_{iK}^T \tag{4}$$

1.3 Dynamical Differential Covariance (DDC)

The definitions of ΔL, ΔR and ΔReLU can be found in the main text. The parameter θ for ΔReLU, was varied from the 5-th percentile to the 95-th percentile of the z-scored data. The optimal value was chosen based on either the estimation errors (Fig.1, Fig.S1 and Fig.S2) or the AUC values (Fig.2). In the brain surface model θ was set to zero.

1.3.1 DDC derivation for stochastic network models

To model the randomness in the recorded neural activities, we used stochastic differential equations (SDE) and evaluated DDC in this stochastic framework:

$$\frac{d\mathbf{x}}{dt} = \mathbf{W}\mathbf{x} + \mathbf{D}\frac{d\beta}{dt} \tag{5}$$

2

where β is a multi-dimensional Brownian motion with $\mathbf{Q}$ unit variance and noise structure $\mathbf{D}$ influencing the state variable $\mathbf{x}$. See Eq. 1 in the main text for the definitions of the other terms. The time averages may be different from ensemble averages for $\mathbb{E}(\mathbf{x}\mathbf{y}^{\mathbf{T}})$ and $\langle \mathbf{x}, \mathbf{y} \rangle := \sum_{t=0}^{T} \mathbf{x}(t)\mathbf{y}(t)^{\mathbf{T}}$ under nonstationary conditions, as analyzed in the next section.

Operating on both sides of this equation with $\langle, \mathbf{x} \rangle$:

$$
\begin{aligned}
\langle \frac{d\mathbf{x}}{dt}, \mathbf{x} \rangle &= \mathbf{W}\langle \mathbf{x}, \mathbf{x} \rangle + \mathbf{D}\langle \frac{d\beta}{dt}, \mathbf{x} \rangle \\
\mathbb{E}\langle \frac{d\mathbf{x}}{dt}, \mathbf{x} \rangle &= \mathbf{W}\mathbb{E}\langle \mathbf{x}, \mathbf{x} \rangle + \mathbf{D}\mathbb{E}\langle \frac{d\beta}{dt}, \mathbf{x} \rangle
\end{aligned}
\tag{6}
$$

To evaluate $\langle \frac{d\beta}{dt}, \mathbf{x} \rangle$, we first write down the explicit solution of the linear SDE starting at $t = 0$, then time average both sides.

$$
\begin{aligned}
\mathbf{x}_t &= \exp(\mathbf{W}t)\mathbf{x}_0 + \int_0^t \exp(\mathbf{W}(t-\tau))\mathbf{D}d\beta_\tau \\
\langle \frac{d\beta}{dt}, \mathbf{x} \rangle &= \langle \frac{d\beta}{dt}, \exp(\mathbf{W}t)\mathbf{x}_0 \rangle + \langle \frac{d\beta}{dt}, \int_0^t \exp(\mathbf{W}(t-\tau))\mathbf{D}d\beta_\tau \rangle \\
&= \mathbf{W}_T + \mathbf{B}_T
\end{aligned}
\tag{7}
$$

The first term $\mathbf{W}_T$ is the summation of time-dependent linear Brownian increments, thus the mean is zero and the variance is a time-dependent scaling of the Brownian variance:

$$
\begin{aligned}
\mathbb{E}(\mathbf{W}_T) &= \mathbf{0} \\
\mathbf{Var}(\mathbf{W}_T) &= \mathbf{Q}\langle \exp(\mathbf{W}t)\mathbf{x}_0, \exp(\mathbf{W}t)\mathbf{x}_0 \rangle
\end{aligned}
\tag{8}
$$

The second term $\mathbf{B}_T$ was evaluated using the Ito integral. Because Brownian motion is nowhere differentiable on its path we numerically approximated the time derivative as we used in the simulations. If we assume $\{t_k\}_{k=1}^{\infty}$ as a partition of $[0, t]$ whose partition size is infinitesimal as $n \to \infty$, we can compute Ito integral in the limit. For simplicity, define $\mathbf{\Phi}_\tau = \exp(\mathbf{W}(t-\tau))\mathbf{D}$. Then the

two terms are composed of nonoverlapping Brownian increments:

$$(\frac{d\beta}{dt})_t = \frac{\beta_{t+dt} - \beta_t}{dt}$$

$$\int_0^t \mathbf{\Phi}_\tau d\beta_\tau = \lim_{n \to \infty} \sum_k \mathbf{\Phi}(t_k)[\beta_{t_{k+1}} - \beta_{t_k}]$$

(9)

from which it follows that $\mathbb{E}(\mathbf{B}_t) = 0$ because Brownian motion has independent and stationary increments.

$$\mathbf{B}_T \propto \sum_t \lim_{n \to \infty} \sum_k (\beta_{t+dt} - \beta_t)[\mathbf{\Phi}(t_k)(\beta_{t_{k+1}} - \beta_{t_k})]^T$$

$$\mathbb{E}(\mathbf{B}_T) = \mathbf{0}$$

(10)

Taken together, the first order statistics of our linear DDC estimator become:

$$\mathbb{E}\langle \frac{d\mathbf{x}}{dt}, \mathbf{x} \rangle = \mathbf{W} \mathbb{E}\langle \mathbf{x}, \mathbf{x} \rangle$$

$$\mathbb{E}(\Delta L) = \mathbb{E}(\langle \frac{d\mathbf{x}}{dt}, \mathbf{x} \rangle \langle \mathbf{x}, \mathbf{x} \rangle^{-1}) = \mathbf{W}$$

(11)

This derivation confirms that DDC is unbiased in the presence of noise from Brownian motion. Simulations of the linear three neuron model revealed that $\langle \frac{d\beta}{dt}, \mathbf{x} \rangle$ is at least ten times smaller in magnitude than $\langle \mathbf{x}, \mathbf{x} \rangle$ even for very $\mathbf{Q}$. This result is remarkable and means that even if the recordings have correlated noise structure ($\mathbf{D}$ in Eq.5), DDC can still recover the ground truth connectivity.

1.3.2 Nonstationary conditions

A continuous time stochastic solution of the SDE, $\{x_t\}_{t=0}^T$, is stationary when its finite-dimensional joint distribution is time-invariant, which implies that its mean and covariance remain constant across time. The covariance matrix can then be estimated by the time-averaged sample covariance.

The above SDE framework allows the mean and covariance of state variables to vary with time

4

according to the Ito formula:

$$\frac{d\mathbf{m}}{dt} = \mathbf{W}\mathbf{m}$$
$$\frac{d\mathbf{P}}{dt} = \mathbf{W}\mathbf{P} + \mathbf{P}\mathbf{W}^{\mathbf{T}} + \mathbf{D}\mathbf{Q}\mathbf{D}^{\mathbf{T}} \tag{12}$$

where $\mathbf{m} = \mathbb{E}(\mathbf{x})$ and $\mathbf{P} = \mathbf{Var}(\mathbf{x})$. The process is stationary if the right hand sides are zero.

Under nonstationary conditions, $\langle \mathbf{x}, \mathbf{x} \rangle$ is no longer a valid estimate of the covariance matrix. Because we did not assume stationarity in the derivation of DDC in the SDE framework, our DDC estimators remain valid and unbiased. These properties make DDC a robust and efficient estimator of functional conectivity.

2 Simulations of neural systems

2.1 Neural motif dynamics

We tested the performance of these methods in networks structured to have typical false positive motifs - chain (Fig.1A,B) and confounder (Fig.1C) - with different dynamics and in another Rosseler chaotic system. To stabilize the simulation, all nodes had decaying dynamics and they were linked by inhibitory connections. Specifically, the diagonal entries in the ground truth matrix was set to -1 and connection strength was set to -0.5. We tested connection strength from -0.1 to -1 and it didn't affect the estimation results qualitatively.

For linear dynamics, system variable $\mathbf{x}$ was simulated through Euler integration according to Eq.13 where $\mathbf{u}$ the is Gaussian distributed random drive and ϵ is the Gaussian distributed observational noise, both independent from $\mathbf{x}$. The integration step is 0.01 seconds and the length of simulation is 1000 seconds unless otherwise specified.

$$\frac{d\mathbf{x}}{dt} = \mathbf{W}\mathbf{x} + \mathbf{u}, \quad \mathbf{u} \sim \mathcal{N}(0, \sigma^2)$$
$$\mathbf{x}_{\text{obs}} = \mathbf{x} + \epsilon, \quad \epsilon \sim \mathcal{N}(0, \sigma_{\text{obs}}^2) \tag{13}$$

5

For nonlinear dynamics, simulation was governed by Eq.14 where we used a centered sigmoid function to simulate the nonlinearity. The sigmoid function was shifted to have mean of zero because otherwise the inhibitory signal would be too strong in the network and the signal would decay to zero in a short time interval. In the expression of $R(x)$, slope α controls the level of nonlinearity in the network. Note the mismatch between simulation nonlinearity and the estimation nonlinearity. The integration step is 0.1 ms and signals were downsampled to 100 Hz after estimation. Simulation length is 1000 seconds unless otherwise mentioned.

$$\frac{d\mathbf{x}}{dt} = \mathbf{W}R(\mathbf{x}) + \mathbf{u}, \quad \mathbf{u} \sim \mathcal{N}(0, \sigma^2)$$
$$R(x) = \frac{1}{1 + e^{-\alpha x}} - \frac{1}{2} \tag{14}$$

The equations for the Rössler system are:

$$\frac{dx_1}{dt} = -x_2 - x_3$$
$$\frac{dx_2}{dt} = x_1 + ax_2 \tag{15}$$
$$\frac{dx_3}{dt} = b + x_3(x_1 - c)$$

where $\mathbf{x} = [x_1, x_2, x_3]^T$, $a = b = 0.2$ and $c = 5.7$. This set of parameters was originally used by Rössler to study the behavior of its chaotic dynamics. The signal was integrated at the step of 0.01 seconds for 1000 seconds. The first 100 seconds of transient dynamics was discarded.

2.2 Sparse Leaky Integrate-and-Fire (LIF) network

The connectivity matrix was constructed as an Erdős–Rényi random graph: two nodes being connected has probability equal to network sparsity. All connected edges were assigned to have the same strength. The connectivity matrices were parameterized by only sparsity level and connection strength. Leaky-Integrate-Fire neurons could be described by Eq.16 with double-exponential filtered synapses [13]. Once membrane voltage V reaches a threshold V_{thres}, the neuron will emit a

spike and reset the membrane potential to V_{reset}. The spike train was described by $\sum_{t_k<t} \delta(t - t_k)$ and then filtered to generate synaptic current r_i. We used sub-threshold membrane potential as the system variable (**x**) of interest. We simulated networks with 200 neurons. The integration process was performed at the step of 0.05 ms, down-sampled to 2000 Hz and simulated for 20 seconds.

$$
\begin{aligned}
\tau \frac{dV}{dt} &= -V + \mathbf{W}r + I_{\text{BIAS}} \\
\frac{dr_i}{dt} &= -\frac{1}{\tau_d}r_i + h_i \\
\frac{dh_i}{dt} &= -\frac{1}{\tau_r}h_i + \frac{1}{\tau_d\tau_r}\sum_{t_k<t}\delta(t - t_k)
\end{aligned}
\tag{16}
$$

2.3 Anatomically supported brain surface model

The anatomic connection used here is the group average structural connectivity obtained through diffusion spectrum imaging (DSI) [10]. It involves 78 cortical regions from both hemispheres. The connectivity matrix contained both local proximity based connections and long-range DSI measured connections. It was simulated using a reduced Wong-Wang model [4] (Eq.6-8 in the reference) using the Virtual Brain Simulator [17]. All physiological parameters were followed from the reference [4]. Simulated population firing rates were down-sampled to 1000Hz and the simulation length was 100 seconds.

3 Estimator performance quantification

3.1 Variance and bias

Following Das et al [3], we decomposed the estimation error into variance and bias (Fig.S1). In most cases, the estimation is different from the ground truth matrix by a scale. So we normalized both estimated and ground truth matrices between -1 and 1. In addition, dCov based estimators are directed estimators while covariance based ones are not. For fair comparison, we only considered

the estimation of the lower triangle part where all ground truth connections are located.

After scaling and lower triangle restriction, estimation error, variance and bias were calculated as Eq.17 where $\mathbf{W}$, $\hat{\mathbf{W}}$ and $\bar{\mathbf{W}}$ are ground truth matrix, estimated matrix and the average of estimated matrices across trials and $||.||$ is the vector L2 norm. It's easy to verify that $\text{Error}^2 = \text{Bias}^2 + \text{Variance}^2$ and the vector form of bias and variance are orthogonal to each other. We measure the relative contribution of bias by the angle (θ_b, Eq.18) between the vectors associated with bias and variance. 50 repetitive trials were used across all simulations.

$$
\begin{aligned}
\text{Error} &= \frac{||\mathbf{W} - \hat{\mathbf{W}}||}{||\mathbf{W}||} \\
\text{Bias} &= \frac{||\mathbf{W} - \bar{\mathbf{W}}||}{||\mathbf{W}||} \\
\text{Variance} &= \frac{||\hat{\mathbf{W}} - \bar{\mathbf{W}}||}{||\mathbf{W}||}
\end{aligned}
\tag{17}
$$

$$
\theta_b = \tan^{-1}\left(\frac{\text{Bias}}{\text{Variance}}\right)
\tag{18}
$$

3.2 Sensitivity and specificity

To evaluate the estimation performance in LIF networks, connection recovery sensitivity and specificity were calculated since the networks have sparse connection and the connection strength are uniform. To be more specific, the estimated matrices were binarized based on their absolute values to determine the existence of connections, which were then compared with the ground truth connections. We used the absolute value because we only cared about the presence of a connection. Sensitivity and specificity were calculated as the true positive rate and one minus false positive rate. Varying the binarization threshold gave rise to the receiver operator curve (ROC). The area under ROC, calculated by trapezoidal integration, indicates the methods general performance in classifying connections.

For performance evaluation in the brain surface model, c-sensitivity [20] (Eq.17 in [11]) was adopted. It's defined as the fraction of the true positives that are estimated with higher connections

strengths than the 95th percentile of the false positive distribution. Like ROC, c-sensitivity quantitatively estimated how sensitive methods are to the estimating the presence of a connections. Thus, the absolute value of the estimated matrices were used here.

4   HCP dataset

4.1   Extracting time traces from rs-fMRI recordings

We used the extensively processed "HCP1200 Parcellation + Timeseries + Netmats (1003 Subjects)" dataset available through the website (`https://www.humanconnectome.org`). Detailed pre-processing and study design could be easily accessed through the website. In this release, 1003 healthy adult human subjects (ages 22-37 years, 534 females) were scanned on a 3-T Siemens connectome-Skyra scanner (customized to achieve 100 mT m^{-1} gradient strength). Each subject underwent 4×15 minutes recording sessions with temporal resolution of 0.73 second and spatial resolution of 2 mm isotropic.

For imaging data processing, each 15-minute run of each subject's rs-fMRI data was preprocessed according to Smith et al [18]; it was minimally-preprocessed [8], and had artefacts removed using ICA+FIX [16] [9]. Inter-subject registration of cerebral cortex was carried out using areal-feature-based alignment and the Multimodal Surface Matching algorithm ('MSMAll') [14] [7]. Each dataset was temporally demeaned and had variance normalization and then fed into the MIGP algorithm, whose output is the top 4,500 weighted spatial eigenvectors from a group-averaged PCA (a very close approximation to concatenating all subjects' timeseries and then applying PCA) [19]. The MIGP output was fed into group-ICA using FSL's MELODIC tool, applying at several different dimensionalities (D = 25, 50, 100, 200, 300). In our analysis, we used the 100-dimension decomposition.

For a given parcellation (group-ICA map), the ICA spatial maps were used to derive one representative time series per IC per subject. This process was fulfilled by the standard "dual-regression stage-1" approach, in which the full set of ICA maps was used as spatial regressors against the full

data [6]. This results in an $N \times T$ (number of components $\times$ number of time points) matrix for each subject. Thus, we consider each IC as a network node.

4.2 Significance test of the estimated connections

To assess the statistical significance of the estimated connection, we used an autoregressive (AR) bootstrap procedure [5, 12] to preserve the power spectrum density (PSD) of BOLD signals. For a specific estimated connection, denoted as element (i, j), our null hypothesis was that signal x_i and x_j are independent regardless of other nodes' influence. To generate null time series, we fit separate AR processes of model order q to node-specific time traces. The model order q was determined according to the Bayesian information criterion (BIC). A higher order model was rejected if it could not decrease BIC by more than 2. Using the estimated AR coefficients of empirical time series, we generated 1000 surrogate null time series and then computed the associated functional connectivity corresponding to the null hypothesis. For each connection, we assumed a Gaussian distribution of the null connectivity values generated from null time traces. P value was calculated as the probability of the empirical value appeared under the null Gaussian distribution. In this paper, we adopted a sequence of significance level to binarize the matrix so that we could investigate the network behavior asymptotically.

4.3 Individual level dMRI strength

In order to compare the functional connectivity metrics to the underlying corticocortical white matter connectivity, we reorganized our previously published diffusion-MRI based structural connectome [15] in which connectivity was assessed among the 360 cortical areas of the HCP-MMP1.0 atlas [7]. Of the 100 IC nodes, 46 are composed of at least 40% cortical voxels (Fig.S6A) and as the dMRI connectome was restricted to corticocortical relationships, we limited the scope of our analyses to these nodes. Because the IC nodes have a greater spatial extent than the atlas areas, each is composed of several areas, in whole or in part (mean = 28.3 areas). For each IC node

10

pair, dMRI connectivity was assessed by obtaining the average of the nodes' constituent interareal connectivity weighted by fraction of the node pair's voxels assigned to each areal pair. In cases where an atlas area was partially present in both IC nodes of a pair, that area was excluded from the mean as short-range intra-areal anatomical connectivity was not available.

Table S2: **Anatomic annotations of the 46 ICs.** First column is the index number of ICs. Second column is the manually registered anatomical region. DMN: default mode network; NA: no reasonable region identified

| Component number	Anatomical Region	Sub-network
1	Occipital Pole	Visual Network
2	Inferior Parietal Lobe	DMN
3	Lateral Occipital Cortex	Visual Network
4	Cuneal Cortex/Occipital Pole	Visual Network
5	Supramarginal Gyrus	DMN
6	Lateral Occipital Cortex	Visual Network
7	Supramarginal Gyrus	DMN
8	Lateral Occipital Cortex	Visual Network
9	Inferior Parietal Lobe	DMN
10	Medial Prefrontal Cortex/Anterior Cingulate Cortex/Lateral Temporal Cortex	DMN
11	Lingual Gyrus/medial occipitotemporal gyrus	Visual Network
12	Angular Gyrus	DMN
13	Occipital Pole	Visual Network
14	Lateral Occipital Cortex - Left	Visual Network
15	Precuneous Cortex	Other Network
16	Occipital Pole	Visual Network
17	Lingual Gyrus/medial occipitotemporal gyrus	Visual Network
18	Lateral Occipital Cortex	Visual Network
19	Occipital Pole	Visual Network
20	Inferior Parietal Lobe	DMN
21	Precentral Gyrus	sensorimotor network
22	Orbital Frontal Cortex	DMN
23	Postcentral Gyrus	sensorimotor network
24	Lateral Occipital Cortex - Right	Visual Network
25	Occipital Pole - Left	Visual Network
26	Frontal Pole	Attention Network
27	Superior Parietal Lobule	Attention Network
28	Hippocampus/Parahippocampal Cortex	DMN
29	Lateral Temporal Cortex	DMN
30	Lateral Occipital Cortex	Visual Network
31	Orbital Frontal Cortex/Lateral Temporal Cortex	Other Network
32	Occipital Pole	Visual Network
33	Inferior Parietal Lobe – Left	DMN
34	Lateral Occipital Cortex	Visual Network
35	Postcentral Gyrus	sensorimotor network
36	Medial Prefrontal Cortex/Anterior Cingulate Cortex	DMN
37	Orbital Frontal Cortex/Lateral Temporal Cortex	DMN
38	Precentral Gyrus / Juxtapositional Lobule Cortex (formerly Supplementary Motor Cortex)	sensorimotor network
39	Occipital Pole	Visual Network
40	Middle Frontal Gyrus	Attention Network
41	Postcentral Gyrus - Left	sensorimotor network
43	Postcentral Gyrus - Right	sensorimotor network
44	Orbital Frontal Cortex	DMN
45	Superior Temporal Gyrus	sensorimotor network
47	NA	NA
48	Frontal Pole	Attention Network
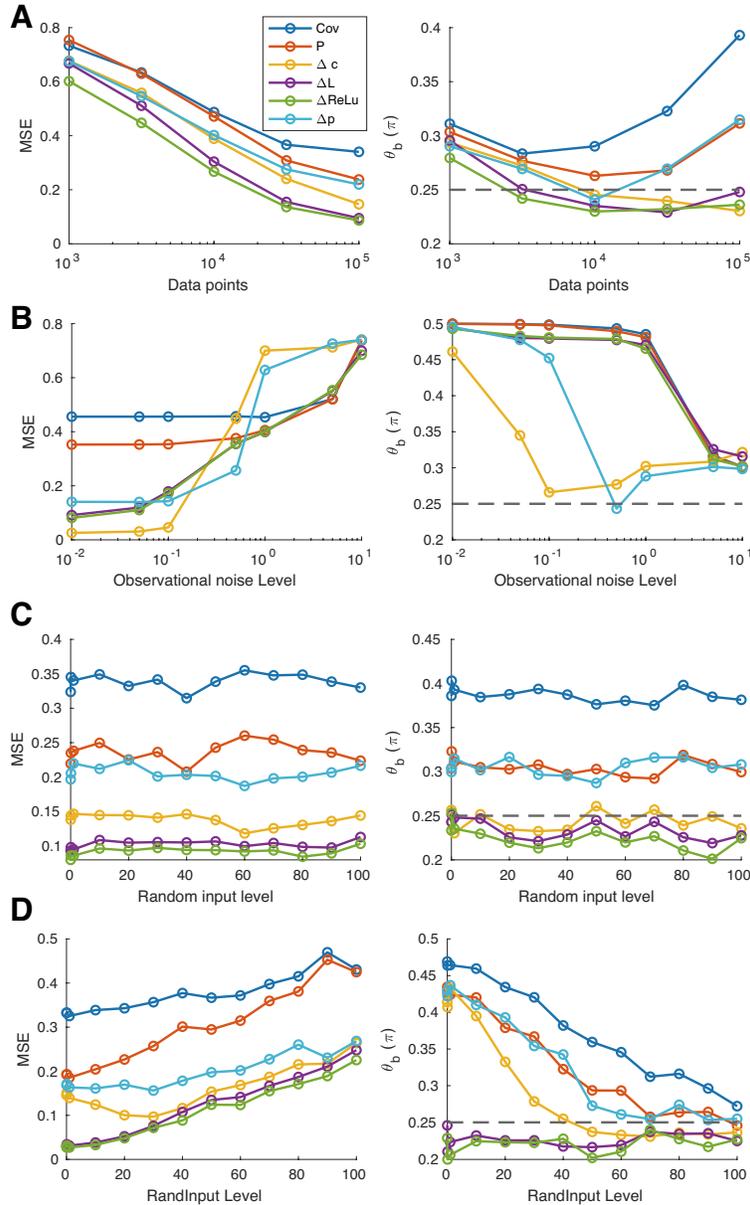
Figure S1: **Estimator variance and bias towards data limit and noise limit.** (A) Left: Influence of data volume calculated across 50 trials simulated using linear dynamics and the confounder motif; Right: Contribution of estimation bias (θ_b). $\theta_b = 0.25\pi$ refers to equal bias and variance. In general estimation errors decreased with data volume and covariance based estimators are more biased. (B) Influence of observational noise (horizontal axis is σ_{obs} in Eq.13) across 50 trials simulated using linear dynamics and the confounder motif. Δc and Δp performance deteriorated with larger amount of noise while ΔL and $\Delta ReLU$ remained robust under the influence of observational noise. (C) Influence of random input strength (horizontal axis is σ in Eq.13) across 50 trials simulated using linear dynamics and the confounder motif. Input strength had limited effect on estimator performance since DDC is an unbiased estimator. (D) Influence of random input strength across 50 trials simulated using highly nonlinear dynamics (α=50 in Eq.14) and the confounder motif. In general, the higher the input strength, the worse the estimation performance. Note the superior performance of ΔL and $\Delta ReLU$ over all other estimators when the random input strength was relatively low.
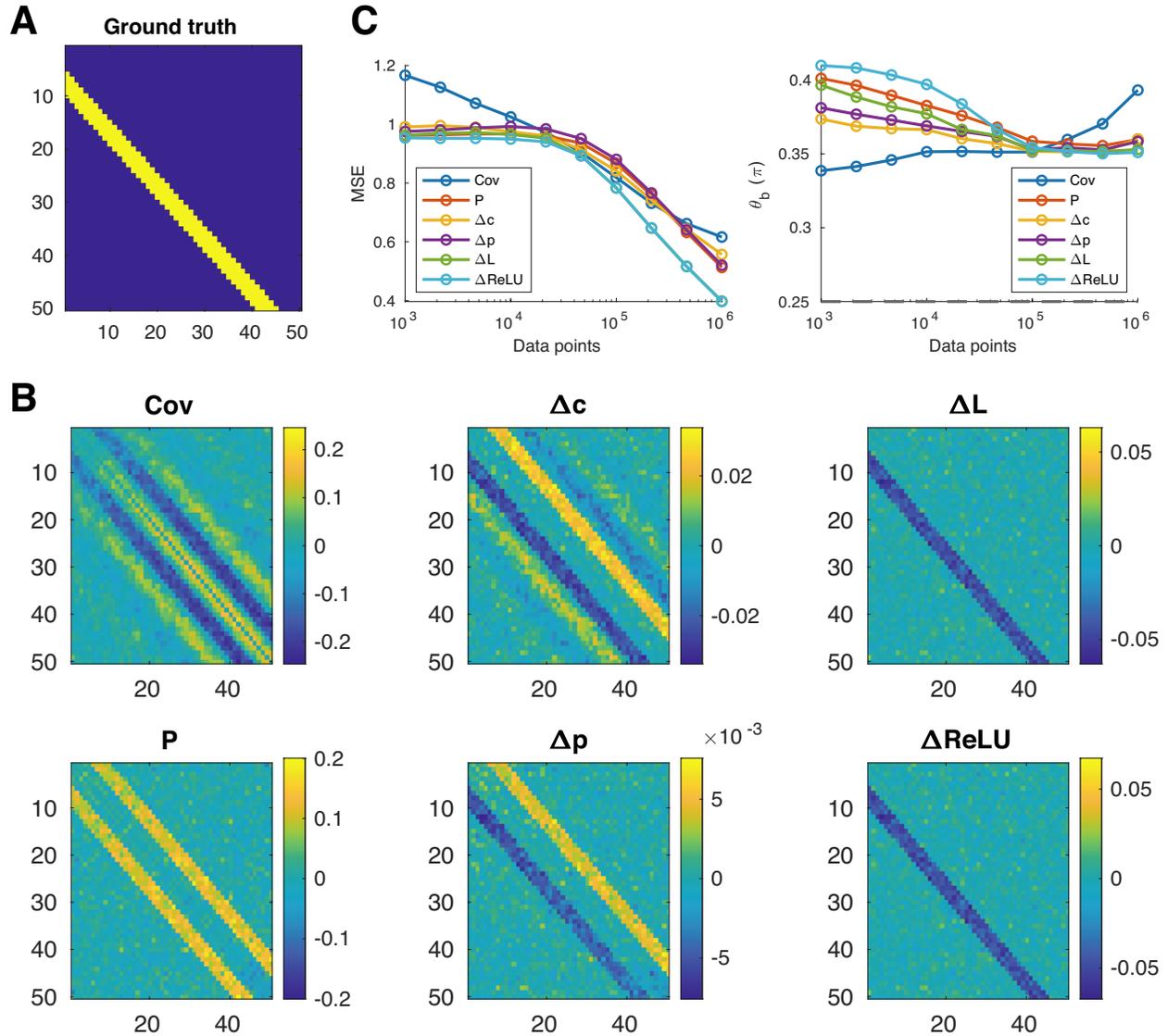
13

Figure S2: **Estimator performance scaled up to larger networks.** (A) Ground truth connectivity pattern of a 50-node network. The connected edges were shown as yellow entries and they were assigned negative strength to stabilize simulation. Time traces were simulated using nonlinear dynamics (Eq.14), which is the harder scenario according to Fig.S1. (B) Estimated matrices using a sufficiently large data volume (10^6 data points). ΔL and $\Delta ReLU$ had cleaner estimation of the ground truth matrix. (C) Influence of data volume
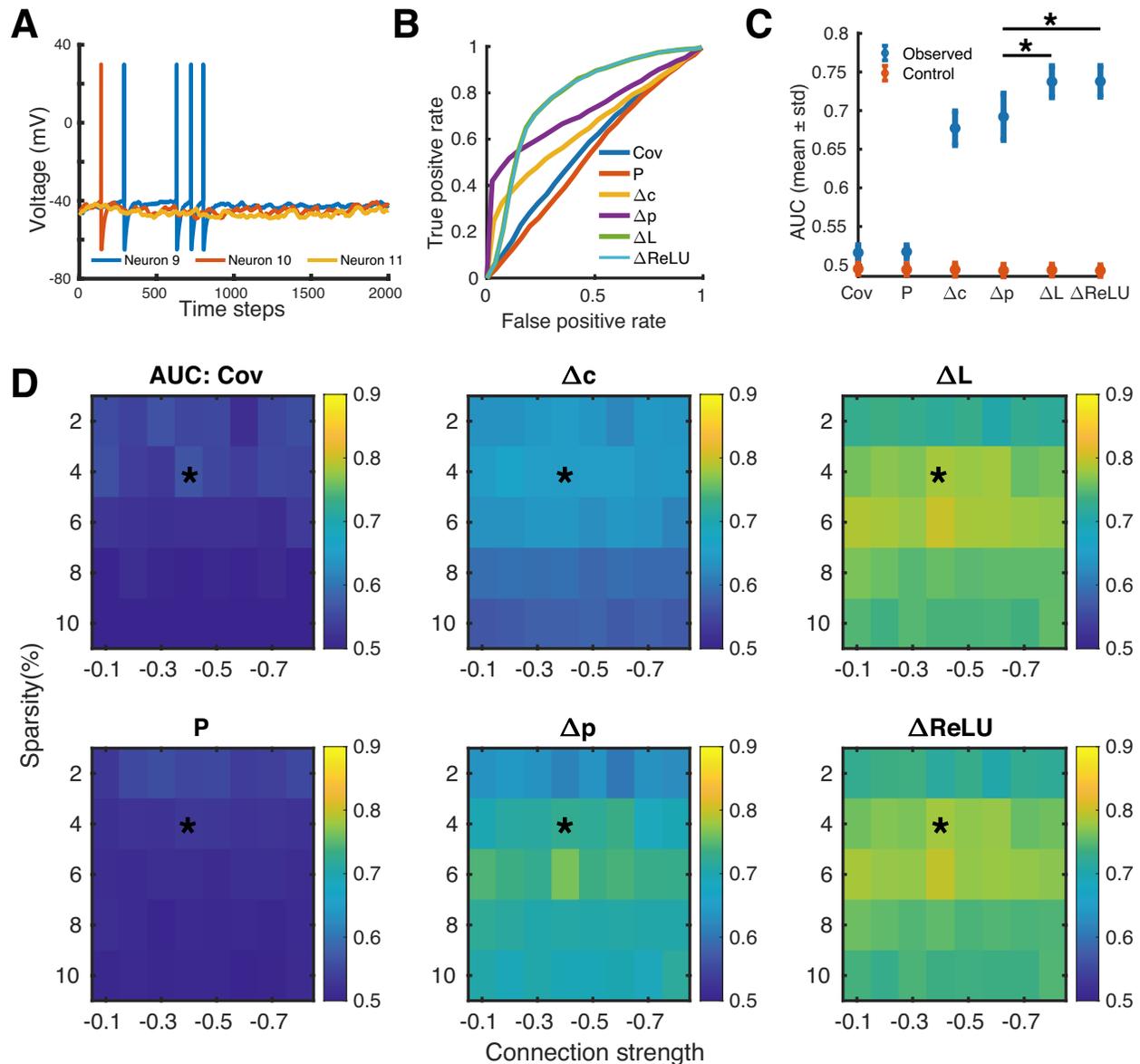
Figure S3: **DDC has high sensitivity and specificity in recovering sparse connections in spiking neural networks.** (A) Selected membrane potential traces simulated using LIF neurons with a sparse connectivity matrix. Connections were generated randomly using a Bernoulli process with probability equal to sparsity level (Erdős–Rényi random graph). Connection strength was set to a fixed value. (B) Receiver operating curves (ROC) of different methods. True positive rate and false positive rate were calculated by comparing absolute value thresholded matrices with the ground truth matrix. (C) Area under ROC (AUC) across 50 realizations of the random graph. Control values were calculated by comparing estimated matrices to a different realization of Erdős–Rényi random graph with same sparsity. Note the superior performance of ΔL and Δ ReLU. * p<0.001, rank-sum test. (D) AUC values of networks with various sparsity and connection strength. Asterisks indicate the scenario shown in (C). Note ΔL and Δ ReLU are robust to different network configurations.
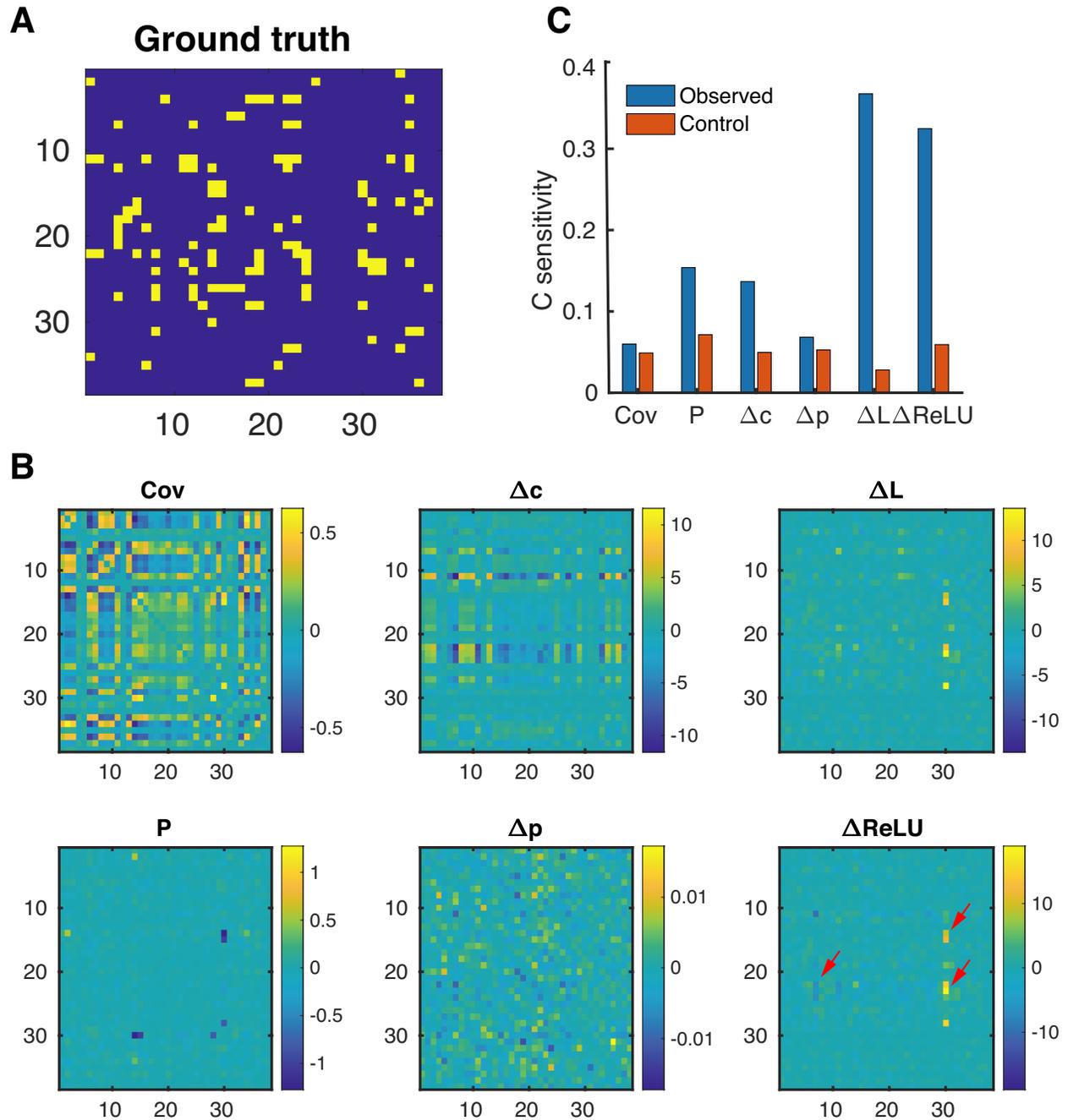
15

Figure S4: **DDC is sensitive to strong anatomic connections at macroscopic level** (A) Intra-hemisphere strong connections (yellow entries) identified by Diffusion Spectrum Imaging (Node definition followed from Hagmann et al [10]). Time traces were generated using using both inter- and intra-hemisphere connections. (B) Estimated intra-hemisphere connections. Note the esti- mated strong connections (red arrow) in ΔL and $\Delta ReLU$ correspond to true connections in (A). (C) Performance quantified by c-sensitivity (Observed values). Control values were evaluated by comparing the estimation to a shuffled ground truth matrix. Note the superior performance of ΔL and $\Delta ReLU$
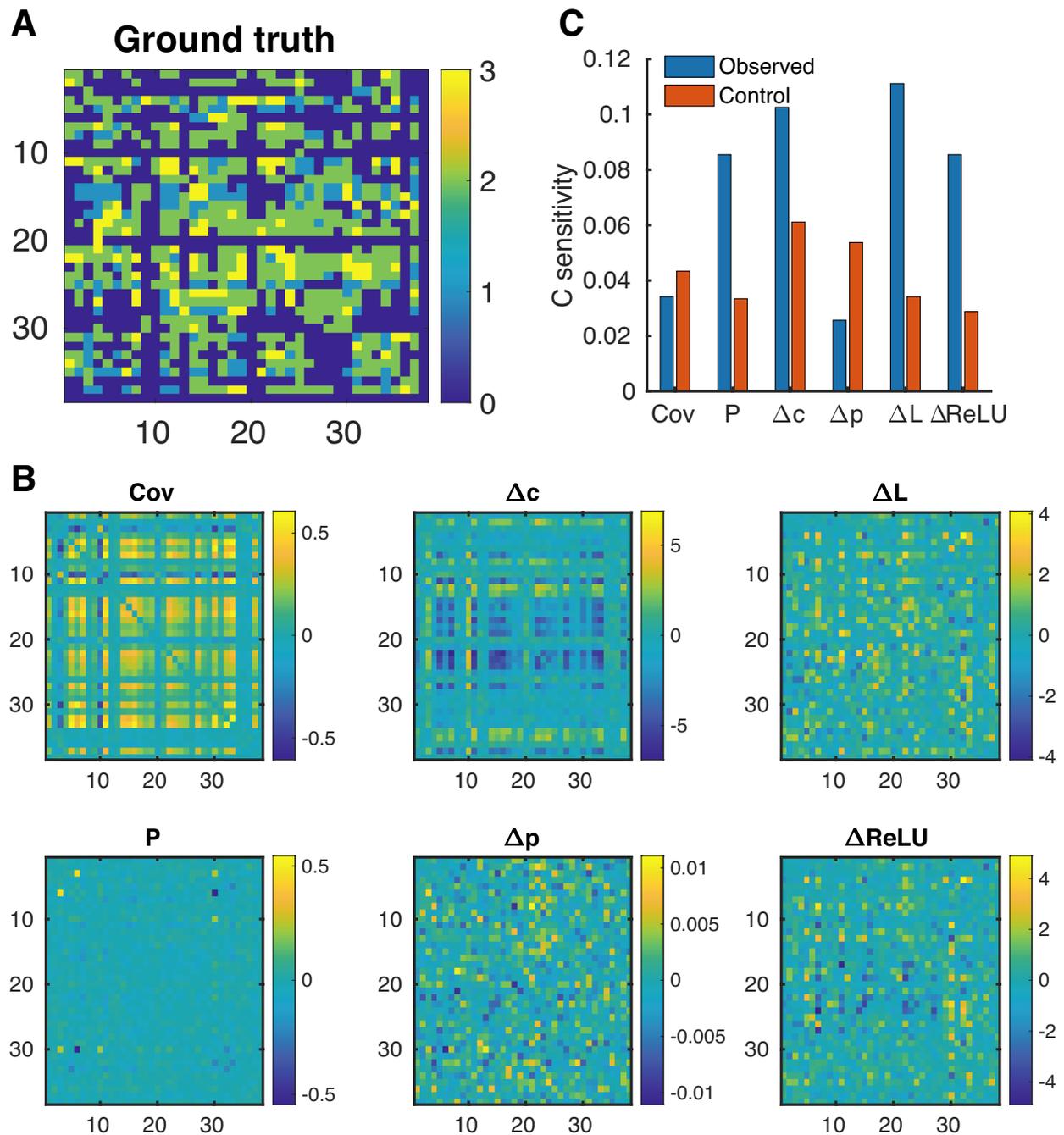
16

Figure S5: **Macroscopic model supported by graded anatomical connections** (A) Intra-hemisphere anatomical connectivity. (B) Estimated intra-hemisphere connections. (C) Performance quantified by c-sensitivity. The overall low c-sensitivity indicates the difficulty of connection estimation.
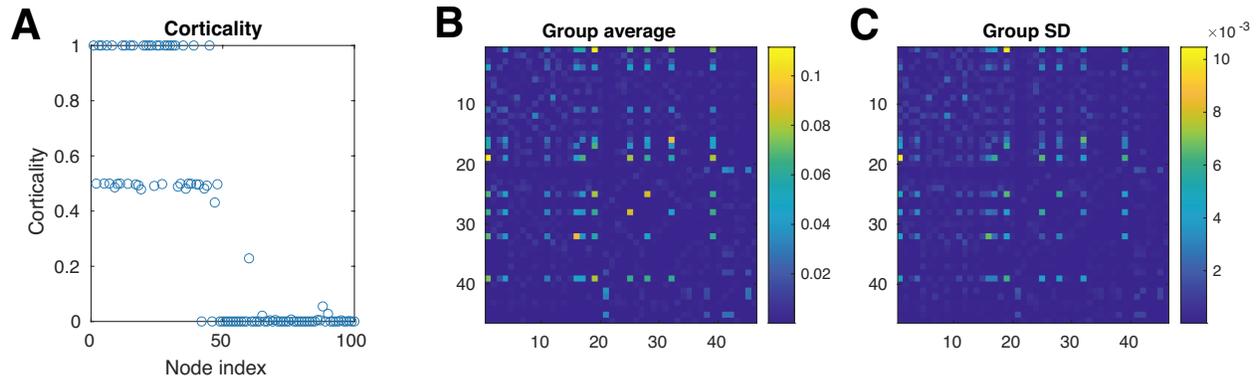
Figure S6: **Individual level dMRI statistics.** (A) Corticality was defined as the proportion of cortical voxels within each IC. Since dMRI measurements are only available for cortical surface voxels. Our analysis was restricted to the first 46 ICs with corticality greater than 40%. (B) Average of the dMRI matrices across the entire 998 subjects. (C) Standard deviation.
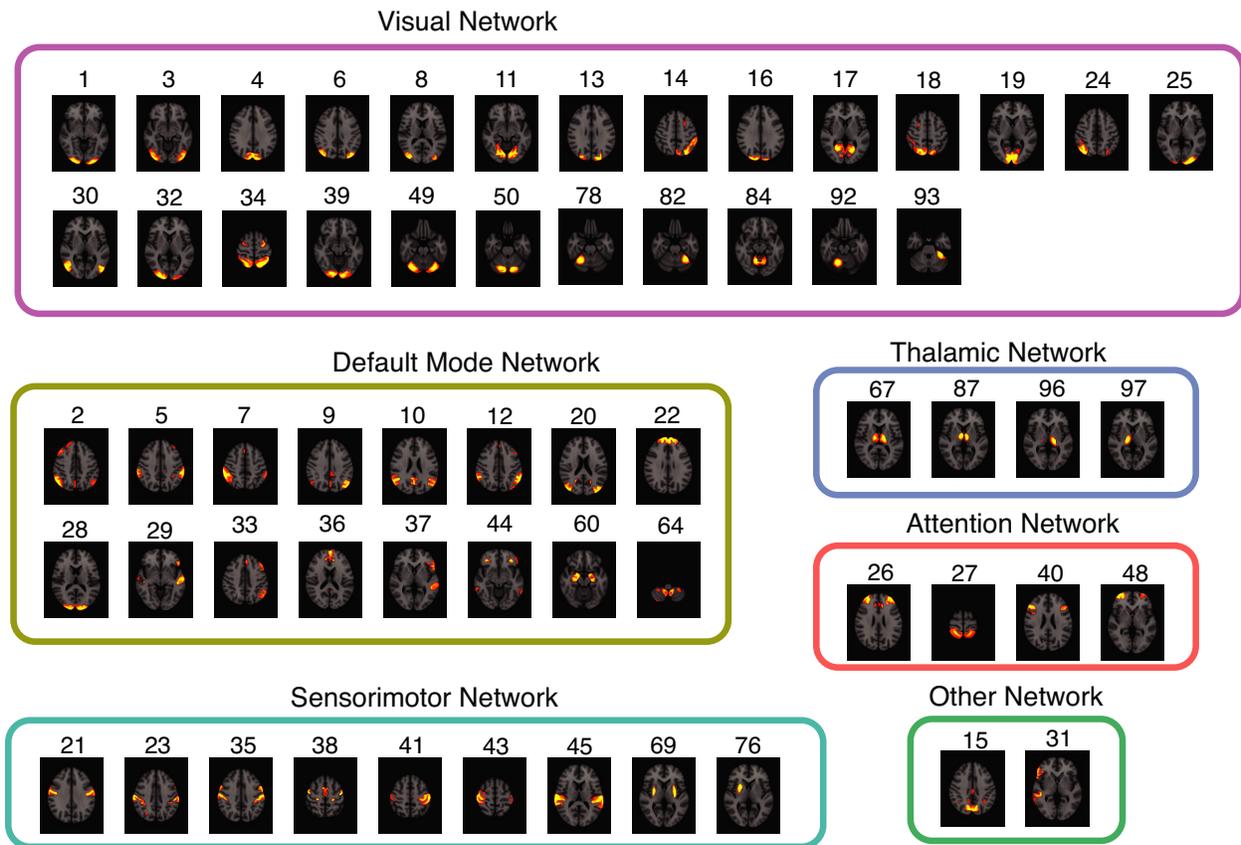

Figure S7: **Group ICA parcellation shown on an MRI template**

References

[1] Symmetric Derivative (Wikipedia).

[2] David Roxbee Cox and Nanny Wermuth. *Multivariate dependencies: Models, analysis and interpretation*. Chapman and Hall/CRC, 2014.

[3] Abhranil Das and Ila R Fiete. Systematic errors in connectivity inferred from activity in strongly recurrent networks. *Nature Neuroscience*, 23(10):1286–1296, 2020.

[4] Gustavo Deco, Adrián Ponce-Alvarez, Dante Mantini, Gian Luca Romani, Patric Hagmann, and Maurizio Corbetta. Resting-state functional connectivity emerges from structurally and dynamically shaped slow linear fluctuations. *Journal of Neuroscience*, (27):11239–11252, 2013.

[5] Bradley Efron and Robert Tibshirani. Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. *Statistical science*, pages 54–75, 1986.

[6] Nicola Filippini, Bradley J MacIntosh, Morgan G Hough, Guy M Goodwin, Giovanni B Frisoni, Stephen M Smith, Paul M Matthews, Christian F Beckmann, and Clare E Mackay. Distinct patterns of brain activity in young carriers of the apoe-ε4 allele. *Proceedings of the National Academy of Sciences*, 106(17):7209–7214, 2009.

[7] Matthew F Glasser, Timothy S Coalson, Emma C Robinson, Carl D Hacker, John Harwell, Essa Yacoub, Kamil Ugurbil, Jesper Andersson, Christian F Beckmann, Mark Jenkinson, et al. A multi-modal parcellation of human cerebral cortex. *Nature*, 536(7615):171, 2016.

[8] Matthew F Glasser, Stamatios N Sotiropoulos, J Anthony Wilson, Timothy S Coalson, Bruce Fischl, Jesper L Andersson, Junqian Xu, Saad Jbabdi, Matthew Webster, Jonathan R Polimeni, et al. The minimal preprocessing pipelines for the human connectome project. *Neuroimage*, 80:105–124, 2013.

[9] Ludovica Griffanti, Gholamreza Salimi-Khorshidi, Christian F Beckmann, Edward J Auerbach, Gwenaëlle Douaud, Claire E Sexton, Enikő Zsoldos, Klaus P Ebmeier, Nicola Filippini, Clare E Mackay, et al. Ica-based artefact removal and accelerated fmri acquisition for improved resting state network imaging. *Neuroimage*, 95:232–247, 2014.

[10] Patric Hagmann, Leila Cammoun, Xavier Gigandet, Reto Meuli, Christopher J Honey, Van J Wedeen, and Olaf Sporns. Mapping the structural core of human cerebral cortex. *PLoS biology*, 6(7):e159, 2008.

[11] Tiger W. Lin, Yusi Chen, Qasim Bukhari, Giri P. Krishnan, Maxim Bazhenov, and Terrence J. Sejnowski. Differential covariance: A new method to estimate functional connectivity in fmri. *Neural Computation*, 32(12):2389–2421, 2020.

[12] Alican Nalci, Bhaskar D Rao, and Thomas T Liu. Nuisance effects and the limitations of nuisance regression in dynamic functional connectivity fmri. *NeuroImage*, 184:1005–1031, 2019.

[13] Wilten Nicola and Claudia Clopath. Supervised learning in spiking neural networks with force training. *Nature Communications*, 8(1):1–15, 2017.

[14] Emma C Robinson, Saad Jbabdi, Matthew F Glasser, Jesper Andersson, Gregory C Burgess, Michael P Harms, Stephen M Smith, David C Van Essen, and Mark Jenkinson. Msm: a new flexible framework for multimodal surface matching. *Neuroimage*, 100:414–426, 2014.

[15] Burke Q Rosen and Eric Halgren. A whole-cortex probabilistic diffusion tractography connectome. *Eneuro*, 8(1), 2021.

[16] Gholamreza Salimi-Khorshidi, Gwenaëlle Douaud, Christian F Beckmann, Matthew F Glasser, Ludovica Griffanti, and Stephen M Smith. Automatic denoising of functional mri data: combining independent component analysis and hierarchical fusion of classifiers. *Neuroimage*, 90:449–468, 2014.

[17] Paula Sanz Leon, Stuart A Knock, M Marmaduke Woodman, Lia Domide, Jochen Mersmann, Anthony R McIntosh, and Viktor Jirsa. The virtual brain: a simulator of primate brain network dynamics. *Frontiers in neuroinformatics*, 7:10, 2013.

[18] Stephen M Smith, Christian F Beckmann, Jesper Andersson, Edward J Auerbach, Janine Bijsterbosch, Gwenaëlle Douaud, Eugene Duff, David A Feinberg, Ludovica Griffanti, Michael P Harms, et al. Resting-state fmri in the human connectome project. *Neuroimage*, 80:144–168, 2013.

[19] Stephen M Smith, Aapo Hyvärinen, Gaël Varoquaux, Karla L Miller, and Christian F Beckmann. Group-pca for very large fmri datasets. *Neuroimage*, 101:738–749, 2014.

[20] Stephen M Smith, Karla L Miller, Gholamreza Salimi-Khorshidi, Matthew Webster, Christian F Beckmann, Thomas E Nichols, Joseph D Ramsey, and Mark W Woolrich. Network modelling methods for fmri. *Neuroimage*, 54(2):875–891, 2011.