Regression dynamic causal modeling for resting-state fMRI

Stefan Frässle[a,*], Samuel J. Harrison[a], Jakob Heinzle[a], Brett A. Clementz[b], Carol A. Tamminga[c], John A. Sweeney[d], Elliot S. Gershon[e], Matcheri S. Keshavan[f], Godfrey D. Pearlson[g,h], Albert Powers[h], Klaas E. Stephan[a,i]

[a] *Translational Neuromodeling Unit (TNU), Institute for Biomedical Engineering, University of Zurich & ETH Zurich, 8032 Zurich, Switzerland*

[b] *Departments of Psychology and Neuroscience, Bio-Imaging Research Center, University of Georgia, Athens, USA*

[c] *Department of Psychiatry, University of Texas Southwestern Medical Center, Dallas, TX 75390, USA*

[d] *Department of Psychiatry, University of Cincinnati, Cincinnati, USA*

[e] *Department of Psychiatry, Department of Human Genetics, University of Chicago, USA*

[f] *Department of Psychiatry, Beth Israel Deaconess Medical Center and Harvard Medical School, Boston, MA, USA*

[g] *Olin Neuropsychiatry Research Center, Hartford Hospital, Institute of Living, Hartford, CT, USA*

[h] *Departments of Psychiatry & Neuroscience, Yale University School of Medicine, New Haven, CT, USA*

[i] *Max Planck Institute for Metabolism Research, Cologne, Germany*

* to whom correspondence should be addressed

Correspondence:

Stefan Frässle

University of Zurich & ETH Zurich

Translational Neuromodeling Unit (TNU)

Institute for Biomedical Engineering

Wilfriedstrasse 6

8032 Zurich, Switzerland

Phone: +41 44 634 91 14

E-mail: stefanf@biomed.ee.ethz.ch

1

1 **ABSTRACT**

2 "Resting-state" functional magnetic resonance imaging (rs-fMRI) is widely used to study brain

3 connectivity. So far, researchers have been restricted to measures of functional connectivity that are

4 computationally efficient but undirected, or to effective connectivity estimates that are directed but

5 limited to small networks.

6 Here, we show that a method recently developed for task-fMRI – regression dynamic causal modeling

7 (rDCM) – extends to rs-fMRI and offers both directional estimates and scalability to whole-brain

8 networks. First, simulations demonstrate that rDCM faithfully recovers parameter values over a wide

9 range of signal-to-noise ratios and repetition times. Second, we test construct validity of rDCM in

10 relation to an established model of effective connectivity, spectral DCM. Using rs-fMRI data from

11 nearly 200 healthy participants, rDCM produces biologically plausible results consistent with estimates

12 by spectral DCM. Importantly, rDCM is computationally highly efficient, reconstructing whole-brain

13 networks (>200 areas) within minutes on standard hardware. This opens promising new avenues for

14 connectomics.

15

16 **KEYWORDS**

17 Regression dynamic causal modeling, rDCM, generative model, effective connectivity, resting state,

18 hierarchy, connectomics

1 INTRODUCTION

"Resting-state" functional magnetic resonance imaging (rs-fMRI) has long been used to examine the functional organization of the brain during unconstrained cognition when no specific experimental manipulations are present (Biswal et al., 1995; Biswal et al., 1997). Specifically, resting-state fMRI has revealed spontaneous (endogenous) fluctuations in the blood oxygen level dependent (BOLD) signal that are highly structured across the brain (Fox et al., 2005; Greicius et al., 2009; Raichle et al., 2001). Over the last two decades, rs-fMRI has become one of the most vibrant fields in neuroimaging (for reviews, see Smith et al., 2013; van den Heuvel and Hulshoff Pol, 2010) and now plays a central role in disciplines such as connectomics (Craddock et al., 2013) and network neuroscience (Bassett and Sporns, 2017). This is partly due to its practical simplicity, which renders rs-fMRI amenable to subject populations that may otherwise struggle to adhere to complex cognitive tasks, such as neuropsychiatric patients, elderly people, or infants.

So far, rs-fMRI data has been analyzed predominantly in terms of functional connectivity, which represents statistical interdependencies amongst BOLD signal time series from spatially distinct regions of the brain (Friston, 2011). The simplest way to assess functional connectivity is by computing Pearson's correlation coefficient between the respective BOLD signal time series. Other measures of functional connectivity include partial correlation, coherence, mutual information and independent component analysis (for a comprehensive review, see Karahanoglu and Van De Ville, 2017).

Regardless of the exact approach, these statistical techniques have revealed a number of large-scale networks of correlated temporal patterns in the resting brain. These *resting-state networks (RSN)* are characterized by fairly coherent time courses across its set of inherent components, whereas time courses are sufficiently distinct between networks (Beckmann et al., 2005; Fox and Raichle, 2007; Smith et al., 2009). The most prominent RSN is arguably the default mode network (DMN) which is also frequently referred to as the task-negative network (Raichle et al., 2001; Shulman et al., 1997). The DMN was initially discovered as a consistent set of brain regions that showed deactivations across various attention-demanding, goal-oriented, non-self-referential tasks (for a comprehensive review, see Raichle, 2015). By now, it is clear that the DMN can also be active in specific goal-oriented paradigms such as autobiographical tasks (Spreng, 2012), thinking about others and oneself, or remembering the past and planning the future (Buckner et al., 2008). Since the discovery of the DMN, similar patterns of resting-state coherence have been identified in most cortical systems, giving rise to other RSNs such as the dorsal attention network (DAN; Corbetta and Shulman, 2002; Fox et al., 2006), salience network (SAN; Dosenbach et al., 2007; Menon and Uddin, 2010), and central executive network (CEN; Spreng et al., 2013; 2010). All RSNs persist even during sleep and under various forms of anaesthesia (Picchioni et al., 2013; but see Tagliazucchi and Laufs, 2014), and show high heritability (Fornito et al., 2011; Glahn et al., 2010), suggesting that they might represent fundamental organizing principles of the human brain.

1 Importantly, mapping the macroscopic functional connectome from rs-fMRI data has not only shed light

2 on the organizational principles in the healthy human brain, but also in disease. Aberrant resting-state

3 functional connectivity has been observed in most psychiatric and neurological disorders (Baker et al.,

4 2019; Buckholtz and Meyer-Lindenberg, 2012; Bullmore and Sporns, 2009; Fornito et al., 2015; Stam,

5 2014). For example, psychiatric diseases including schizophrenia (Baker et al., 2014; Friston et al.,

6 2016; Lui et al., 2015), depression (Greicius et al., 2007; Wang et al., 2012a) and autism (Courchesne

7 et al., 2007; Hahamy et al., 2015) have all been associated with pathological alterations in resting-state

8 functional connectivity.

9 Despite these profound contributions to our understanding of the organizing principles of the human

10 brain, measures of functional connectivity essentially describe statistical properties of the data and do

11 not reveal how the data were caused or generated. As such, functional connectivity does not directly

12 capture mechanisms at a latent level of neuronal interactions. Furthermore, measures of functional

13 connectivity are undirected and thus do not capture asymmetries in reciprocal connections. This is

14 problematic because asymmetries in the coupling amongst brain regions have been repeatedly found

15 both in terms of anatomy (Felleman and Van Essen, 1991; Markov et al., 2014a) and function (Frässle

16 et al., 2016; Gazzaniga, 2000; Stephan et al., 2007; Zeki and Shipp, 1988).

17 In contrast, effective connectivity refers to directed interactions amongst brain regions and can be

18 assessed by exploiting a generative model of the latent (hidden) neuronal states and how these give rise

19 to the observed measurements (Friston, 2011). One of the most widely used generative modeling

20 frameworks for inferring effective connectivity from fMRI data is dynamic causal modeling (DCM;

21 Friston et al., 2003), and variants of DCM have been established to model the resting state, including

22 stochastic DCM (Daunizeau et al., 2009; Li et al., 2011) and spectral DCM (Friston et al., 2014a). While

23 capable of more mechanistic accounts of functional integration during the resting state, these models

24 are restricted to relatively small networks due to computational limitations.

25 We recently introduced a novel variant of DCM for fMRI – termed regression dynamic causal modeling

26 (rDCM; Frässle et al., 2018a; 2017) – that differs from previous DCMs in several aspects. Most

27 important, rDCM is computationally highly efficient and scales gracefully to very large networks

28 including hundreds of nodes, paving the way for whole-brain effective connectivity analyses (Frässle et

29 al., 2020a). Furthermore, the model can exploit structural connectivity information to constrain

30 inference on directed functional interactions or, where no such information is available, infer optimally

31 sparse representations of whole-brain connectivity patterns. While rDCM was initially designed to work

32 with experimentally controlled perturbations (i.e., task data), the current implementation can, in

33 principle, also by applied to rs-fMRI data. However, this has not been tested thoroughly so far. Here,

34 we assess the face validity and construct validity of rDCM for inferring effective connectivity during

35 the resting state. To this end, we conducted comprehensive simulation analyses and compared rDCM

36 against spectral DCM as an alternative and well-established generative model of rs-fMRI data for small

1   networks. This comparison made use of a large empirical resting-state fMRI dataset acquired by the

2   Bipolar-Schizophrenia Network on Intermediate Phenotypes (B-SNIP) consortium (Tamminga et al.,

3   2013). Finally, we showed that rDCM gracefully scales to large networks with more than 200 areas and

4   enables fast inference on whole-brain effective connectivity patterns from rs-fMRI data at the time-scale

5   of minutes.

5

2 METHODS AND MATERIALS

2.1 Regression dynamic causal modeling

2.1.1 *General overview*

Regression DCM (rDCM) is a novel variant of DCM for fMRI that enables effective connectivity analyses in large (whole-brain) networks (Frässle et al., 2018a; 2017). This can be achieved by applying several modifications and simplifications to the original DCM framework. In brief, these include (i) translating state and observation equations from time to frequency domain using the Fourier transformation (under stationarity assumptions), (ii) replacing the nonlinear biophysical model of hemodynamics with a linear hemodynamic response function (HRF), (iii) applying a mean field approximation across regions (i.e., connectivity parameters targeting different regions are assumed to be independent), and (iv) specifying conjugate priors on neuronal (i.e., connectivity and driving input) parameters and noise precision. These modifications essentially reformulate the linear DCM in the time domain as a Bayesian linear regression in the frequency domain, resulting in the following likelihood function:

$$
\begin{aligned}
p(Y|\theta, \tau, X) &= \prod_{r=1}^{R} \mathcal{N}\left(Y_r; X\theta_r, \tau_r^{-1} I_{N \times N}\right) \\
Y_r &= \left(e^{2\pi i \frac{m}{N}} - 1\right)\frac{\hat{y}_r}{T} \\
X &= \left[\hat{y}_1, \hat{y}_2, \dots, \hat{y}_R, \hat{h}\hat{u}_1, \hat{h}\hat{u}_2, \dots, \hat{h}\hat{u}_K\right] \\
\theta_r &= \left[a_{r,1}, a_{r,2}, \dots, a_{r,R}, c_{r,1}, c_{r,2}, \dots, c_{r,K}\right]
\end{aligned}
\tag{1}
$$

Here, Y_r is the dependent variable in region r that is explained as a linear mixture of afferent connections from other regions and direct (driving) inputs. Specifically, Y_r is the Fourier transformation of the temporal derivative of the measured signal in region r. Furthermore, y_r represents the measured BOLD signal in region r, X is the design matrix (comprising a set of regressors and explanatory variables), u_k is the kth experimental input, and the hat symbol denotes the discrete Fourier transform (DFT). Additionally, θ_r represents the parameter vector comprising all connections $a_{r,1}, \dots, a_{r,R}$ and all driving input parameters $c_{r,1}, \dots, c_{r,K}$ targeting region r. Finally, τ_r denotes the noise precision parameter for region r and $I_{N \times N}$ is the identity matrix (where N denotes the number of data points). Choosing appropriate priors on the parameters and hyperparameters in Eq. (1) (see Frässle et al., 2017) results in a generative model that can be used for inference on the directed connection strengths and inputs.

Under this formulation, inference can be done very efficiently by (iteratively) executing a set of analytical Variational Bayes (VB) update equations concerning the sufficient statistics of the posterior density. In addition, one can derive an expression for the negative (variational) free energy (Friston et al., 2007). The negative free energy represents a lower-bound approximation to the log model evidence that accounts for both model accuracy and complexity. Hence, the negative free energy offers a sensible metric for scoring model goodness and thus serves as a criterion for comparing competing hypotheses

6

1   (Bishop, 2006). We have recently further augmented rDCM by introducing sparsity constraints into the

2   likelihood of the model in order to allow automatic pruning of fully connected network structures

3   (Frässle et al., 2018a). Notably, the present paper focuses on the original implementation of rDCM

4   without sparsity constraints. A comprehensive description of the generative model of rDCM, including

5   the mathematical details of the neuronal state equation underlying the likelihood function in Eq. (1), can

6   be found elsewhere (Frässle et al., 2018a; 2017).

7

8   *2.1.2   Application to resting state data*

9   While rDCM was originally designed to work with experimentally controlled perturbations (i.e., task

10   data), it is also possible to fit the model to rs-fMRI data. This can be achieved by "switching off" driving

11   inputs (i.e., setting all input parameters $c_{i,j}$ to zero in the absence of experimental inputs u_k) and, in

12   order to explain activity in any given region, relying on measured data (in the Fourier domain) in regions

13   from which afferent connections are received; compare the likelihood function in Eq. (1). This means

14   that, in contrast to stochastic variants of DCM, the generative model does not explicitly represent

15   endogenous fluctuations in neuronal activity and has no concept of stochastic "innovations" or similar

16   ways noise can drive neuronal activity intrinsically. Instead, endogenous fluctuations of BOLD signal

17   in any given region are predicted as a linear mixture of intrinsic fluctuations from other regions.

18   In this study, we examine the face validity and construct validity of rDCM in its specific application to

19   rs-fMRI data. Concerning face validity, we investigate whether the particular formulation of rDCM

20   allows for veridical inference on effective connectivity, using simulated BOLD data fluctuations that

21   were generated by means of a deterministic generative model with stochastic driving inputs (see below).

22   With regard to construct validity, we apply rDCM to empirical fMRI data and compare the results to

23   those obtained by spectral DCM (Friston et al., 2014a), an alternative and established model of effective

24   connectivity for rs-fMRI data.

25

26   **2.2   Synthetic data**

27   First, we evaluated the face validity of rDCM in simulation studies by generating synthetic rs-fMRI data

28   for which the ground truth (i.e., the data-generating parameter values) were known. In particular, we

29   assessed model parameter recovery in four different linear DCMs and evaluated performance as a

30   function of the repetition time (TR) and signal-to-noise ratio (SNR) of the synthetic data. Each of the

31   models comprised four regions, yet differed with regard to the degree of sparsity of the network

32   architecture (**Fig. 1A**). More precisely, model 1 represented a full network where all regions were

33   connected via reciprocal connections, model 2 comprised 75% of all possible connections, model 3

34   comprised 50% of all possible connections, and model 4 was the sparsest network with only 25% of all

35   connections present. Please note that this simulation work was restricted to a small network with only

1   four regions in order to allow comparability of rDCM with spectral DCM (Friston et al., 2014a), which

2   represents an alternative variant of DCM for rs-fMRI data (see below).

3   For each model, synthetic fMRI data were then generated to resemble the characteristic amplitude and

4   low-frequency fluctuations seen in resting-state BOLD signal time series. To this end, we followed the

5   procedure described in previous work (Friston et al., 2014a; Razi et al., 2015) and implemented in the

6   MATLAB function *DEM_demo_induced_fMRI.m* as part of SPM12 (version R7487, Wellcome Centre

7   for Human Neuroimaging, London, UK, http://www.fil.ion.ucl.ac.uk). In brief, endogenous neuronal

8   fluctuations were generated independently for each region by using an AR(1) process with an

9   autoregression coefficient of one half. The endogenous neuronal fluctuations were scaled to a standard

10   deviation of 1/4. These values were identified by Friston et al. (2014a) to produce realistic BOLD signal

11   changes of about 2%. Region-wise endogenous fluctuations were then utilized as driving inputs to a

12   classical deterministic DCM (Friston et al., 2003). Similarly, measurement noise was generated

13   independently for each region from another AR(1) process and added on top of the noise-free BOLD

14   signal time courses. Synthetic resting-state fMRI data were simulated to mimic an experiment of 10

15   minutes duration. Importantly, BOLD signal time series were generated under a variety of different

16   signal-to-noise ratios (SNR = [0.5 1 3]) and repetition times (TR = [2s, 1s, 0.5s]) in order to evaluate

17   the performance of rDCM as a function of data quality and sampling rate, respectively. Here, SNR was

18   defined as the ratio between the standard deviation of the signal and the standard deviation of the AR(1)-

19   type measurement noise (i.e., SNR = $\sigma_{signal}/\sigma_{noise}$).

20   For each of the four models and each combination of SNR and TR, 20 different sets of observations

21   ("synthetic subjects") were generated. To this end, the generating ("true") parameter values of each

22   simulation were sampled from the prior distributions of the endogenous parameters. Parameter values

23   too close to zero (i.e., $|a_{i,j}| < 0.05$) were discarded and re-sampled in order to ensure sufficient

24   information transfer between regions.

25   Examples of endogenous fluctuations and ensuing synthetic BOLD signal time series are shown in **Fig.**

26   **1B**. Note that, as a result of the low-pass filtering of the hemodynamic response function, the BOLD

27   signal is much smoother than the underlying endogenous fluctuations. The resulting BOLD signal time

28   series are biologically realistic and show the characteristic slow (low-frequency) fluctuations typically

29   observed during the resting state.

30   Based on the synthetic data, model inversion was performed by making use of the function

31   *tapas_rdcm_estimate.m* from the rDCM toolbox as implemented in the **T**ranslational **A**lgorithms for

32   **P**sychiatry **A**dvancing **S**cience (TAPAS) toolbox (www.translationalneuromodeling.org/software).

33   Model parameter recovery was then assessed by computing (i) the root-mean-squared-error (RMSE)

34   and (ii) the Pearson's correlation coefficient (r) between the inferred and the generating ("true")

35   parameter values in order to quantify how much the estimated posterior parameter values differed from

36   the ground truth.

A   model space for simulation study

model 1   model 2   model 3   model 4
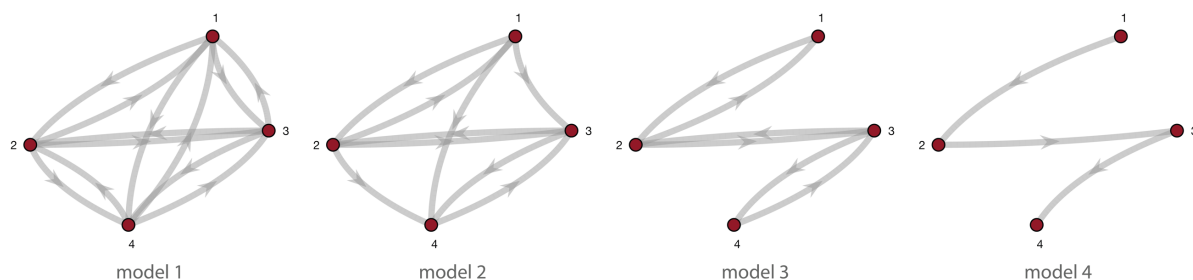
B   generating synthetic resting-state BOLD signal time series

endogenous fluctuations   sampling "true" parameter values   resting-state BOLD signal
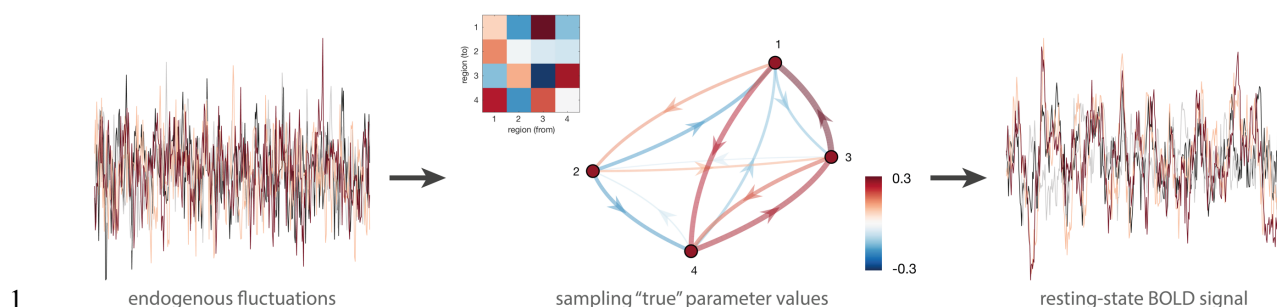
Figure 1. Pipeline for simulation studies on regression DCM (rDCM). (A) Model space for the simulation studies dedicated to assessing the face validity of rDCM for resting-state data. Four different network structures were utilized that varied in their degree of sparsity. Model 1 represents a fully connected model where all regions are coupled via reciprocal connections (i.e., 12 connectivity parameters), model 2 represents a network with 75% of all connections present (i.e., 9 connectivity parameters), model 3 represents a network with 50% of all connections present (i.e., 6 connectivity parameters), and model 4 represents the sparsest network with only 25% of all connectivity present (i.e., 3 connectivity parameters). **(B)** Regional endogenous fluctuations were generated from an AR(1) process and then utilized as driving inputs in the respective model. Specifically, we here used a classical deterministic DCM to generate synthetic data. Model parameters were sampled from their prior distributions and then utilized to generated synthetic BOLD signal time series that resembled the characteristic amplitudes and slow fluctuations observed during the resting state. Measurement noise, generated from another AR(1) process, was added on top of the predicted BOLD signal time series.

2.3   Empirical data

Second, we examined the construct validity of rDCM in relation to spectral DCM. For this, we applied both models to an empirical rs-fMRI dataset and compared the results. Specifically, we made use of a large dataset collected by the Bipolar-Schizophrenia Network on Intermediate Phenotypes (B-SNIP-1) consortium, with one measure being rs-fMRI (Tamminga et al., 2013; 2014). Overall, participants with psychosis (including schizophrenia, schizoaffective disorder and psychotic bipolar I disorder), their first-degree relatives, and demographically matched healthy participants were collected for the B-SNIP-1 dataset.

1   *2.3.1   Participants*

2   The data utilized in the present study were acquired by the B-SNIP-1 consortium as part of a large cross-

3   sectional study of intermediate phenotypes in psychosis (Tamminga et al., 2013). The B-SNIP-1

4   consortium included five sites in the United States: (i) Baltimore, (ii) Chicago, (iii) Dallas, (iv) Detroit

5   and Boston, and (v) Hartford. The sites used identical study protocols, and considerable effort was taken

6   to harmonize recoding and testing conditions across sites by utilizing identical stimulus presentation and

7   recording equipment. Furthermore, experimenters across sites were cross-trained and frequently

8   monitored to ensure comparable data collection procedures. A detailed study description is provided

9   elsewhere (Tamminga et al., 2013; 2014).

10   Given the methodological focus of this work, here we focus only on the healthy control participants (a

11   comprehensive analysis of the full clinical sample will be presented in future work). Healthy participants

12   were recruited from the local communities. They were without lifetime psychotic disorders and had no

13   first-degree relatives with a history of psychotic or bipolar disorder according to the Family History

14   Research Diagnostic Criteria (Andreasen et al., 1977). Overall, 459 healthy participants were assessed

15   by the B-SNIP-1 consortium. For the current study, only those participants were included that had: (i) a

16   complete data set comprising all relevant demographical information and neuroimaging data, and (ii)

17   sufficiently high data quality of the fMRI measurements (exclusion criteria related to data quality are

18   listed below). This yielded a final sample of 196 healthy participants that were included in our study (80

19   females, 116 males; age: 38.3±12.5 years; age range: 15-64 years). The study protocol was approved by

20   the institutional review board at each local site. All participants provided written informed consent prior

21   to inclusion and after they obtained a complete description of the study.

22

23   *2.3.2   Experimental procedure*

24   While in the MR scanner, participants were asked to fixate on a small cross presented on the screen, and

25   to remain alert with eyes open and head still. These instructions were designed to help prevent

26   participants from falling asleep, provide an experimental control over visual inputs and eye movements,

27   and reduce head movements. Head movements were further restricted by having participants' heads

28   placed in a custom-build head-coil cushion.

29

30   *2.3.3   Data acquisition*

31   Structural and functional MRI data were acquired at the five sites of the B-SNIP-1 consortium. These

32   included the University of Maryland School of Medicine (Baltimore), Commonwealth Research Center

33   at Harvard Medical School (Boston), University of Illinois Medical Center (Chicago), University of

34   Texas Southwestern Medical Center (Dallas), and Olin Neuropsychiatry Research Center at the Institute

35   of Living (Hartford). Participants were scanned on 3 Tesla MR scanners of different manufacturers,

1   including GE Signa, Siemens Trio, Philips Achieva, and Siemens Allegra. For comprehensive details

2   on data acquisition parameters across the different sites, see Supplementary Material S1.

3

4   *2.3.4   Data preprocessing*

5   Preprocessing of the MR data was performed using fMRIPrep (version 1.4.1; Esteban et al., 2019) which

6   is based on Nipype (version 1.2.0; Gorgolewski et al., 2011) and was specifically designed for automated

7   high-quality preprocessing in the context of large-scale datasets. We here utilized the standard

8   preprocessing pipeline from fMRIPrep which comprised the following steps.

9   The T1-weighted (T1w) anatomical image was corrected for intensity non-uniformity (INU) with

10   *N4BiasFieldCorrection* (Tustison et al., 2010), distributed with ANTs 2.2.0 (Avants et al., 2008), and

11   used as T1w reference throughout the workflow. The T1w reference was then skull-stripped with a

12   Nipype implementation of the *antsBrainExtraction* workflow. Brain tissue segmentation of

13   cerebrospinal fluid (CSF), white matter (WM) and gray matter (GM) was performed on the brain-

14   extracted T1w image using *fast* (Zhang et al., 2001), distributed with FSL 5.0.9 (Jenkinson et al., 2012).

15   Volume-based spatial normalization to the MNI standard space was performed through nonlinear

16   registration with *antsRegistration* (ANTs 2.2.0), using brain-extracted versions of both T1w reference

17   and T1w template (MNI152NLin2009cAsym; Fonov et al., 2009).

18   Functional images were preprocessed by first generating a reference volume (BOLD reference) and its

19   skull-stripped version by making use of a custom methodology of fMRIPrep. Head-motion parameters

20   with respect to the BOLD reference (transformation matrices, and six corresponding rotation and

21   translation parameters) were estimated before any spatiotemporal filtering using *mcflirt* (Jenkinson et

22   al., 2002). Slice-timing correction was performed to account for differences in acquisition timing across

23   slices. A deformation field to correct for susceptibility distortions was estimated based on the *fieldmap-

24   less* approach in fMRIPrep (Wang et al., 2017). This procedure utilizes the T1w reference from the same

25   subject as the undistorted target in a nonlinear registration scheme. To maximize the similarity between

26   T2* contrast of the EPI scan and the T1w reference, the intensities of the latter are inverted. To regularize

27   the optimization of the deformation field, displacements are constrained to be non-zero only along the

28   phase-encoding direction and the magnitude of displacements is modulated with an average fieldmap

29   template (Treiber et al., 2016). Based on the estimated susceptibility distortion, unwarped BOLD images

30   were calculated for a more accurate co-registration with the anatomical reference by resampling

31   functional images onto their original, native space through a single, composite transform to correct for

32   head-motion and susceptibility distortions. This was followed by co-registration to the T1w reference

33   using *flirt* (Jenkinson and Smith, 2001) with the boundary-based registration cost function (Greve and

34   Fischl, 2009), implemented in FSL 5.0.9 (Jenkinson et al., 2012). Co-registration was configured with

35   nine degrees of freedom to account for remaining distortions in the BOLD images. The corrected and

36   coregistered   functional   images   were   then   normalized   into   MNI   standard   space

11

1   (MNI152NLin2009cAsym; Fonov et al., 2009) by making use of the warps obtained from the spatial

2   normalization of the T1w reference (see above).

3

2.3.5   *Estimation of confound regressors*

5   From the functional images, fMRIPrep computes several time series of potentially confounding

6   influences that can be used subsequently (during the extraction of BOLD signal time series) as confound

7   regressors to correct for variance of no interest. Here, we will focus only on those confound regressors

8   that were utilized in the present analysis; for a comprehensive list of all confound regressors estimated

9   by fMRIPrep, see elsewhere (Esteban et al., 2019).

10   In brief, framewise displacement (FD) and DVARS (Power et al., 2012) were calculated by making use

11   of their implementations in Nipype, following the definitions by Power et al. (2014). Furthermore, three

12   global signals were extracted within the CSF, WM, and a whole-brain mask. Additionally, the head

13   motion estimates (translation and rotation) obtained during preprocessing were utilized as confound

14   regressors. The confound regressors derived from head motion estimates and global signals were

15   expanded with the inclusion of temporal derivatives and quadratic terms for each (Friston et al., 1996;

16   Satterthwaite et al., 2013). Furthermore, frames that exceeded a threshold of 0.5mm for FD or 1.5 for

17   standardized DVARS were annotated as motion outliers and marked using a stick regressor. Finally,

18   regressors encoding a discrete cosine set with a 128s cut-off were utilized for high-pass filtering.

19

2.3.6   *Exclusion criteria*

21   In addition to fMRIPrep, we utilized MRIQC (version 0.15.1; Esteban et al., 2017) for quality control

22   of the neuroimaging data. Specifically, we ran MRIQC with the default parameters, with the exception

23   of the framewise displacement threshold being adjusted to 0.5mm. Participants were excluded from

24   further effective connectivity analysis if any of the following image quality metrics derived by MRIQC

25   were above/below the specified threshold. For the structural images, exclusion criteria were defined as:

26   (i) "QI_1" > 0.005, (ii) "overlap_tpm_csf" < 0.1, (iii) "overlap_tpm_gm" < 0.3, and (iv)

27   "overlap_tpm_wm" < 0.45. In brief, "QI_1" represents the proportion of artifact-corrupted voxels as

28   assessed using the algorithm described by Mortamet et al. (2009). Furthermore, "overlap_tpm_*"

29   represents the overlap between the tissue probability maps (TPMs) estimated from the anatomical image

30   and the corresponding maps from the MNI152NLin2009cAsym template. For the functional images,

31   exclusion criteria were specified as: (i) "aqi" > 0.025, and (ii) "fd_perc" > 20.0. In brief, "aqi" is AFNI's

32   mean quality index computed by the *3dTqual* routine, and "fd_perc" represents the percentage of scans

33   with a FD above the FD threshold with respect to the entire time series. A full description of the metrics

34   is available from https://mriqc.readthedocs.io/en/latest/measures.html.

1   *2.3.7   Time series extraction: small networks*

2   To evaluate the construct validity of rDCM in application to resting-state fMRI data, we examined

3   functional integration during the resting state from two perspectives: modes and nodes in rs-fMRI

4   (Friston et al., 2014b). First, we examined effective connectivity among key intrinsic networks of the

5   human brain (i.e., modes). Specifically, we investigated the established anticorrelations between the

6   task-negative default mode network, engaged during rest and internally directed tasks, and task-positive

7   networks, engaged during externally oriented tasks. These anticorrelations have been proposed to serve

8   as a fundamental characteristic of endogenous or spontaneous fluctuations during the resting state (Fox

9   et al., 2005; Raichle, 2015; Smith et al., 2009; but see Discussion and Fox et al., 2009).

10   To this end, we selected in a first step four modes of interest, comprising the default mode network

11   (DMN; Buckner et al., 2008; Raichle, 2015), dorsal attention network (DAN; Corbetta and Shulman,

12   2002; Fox et al., 2006), salience network (SAN; Dosenbach et al., 2007; Menon and Uddin, 2010), and

13   central executive network (CEN; Spreng et al., 2013; 2010). Masks for each of the four RSNs were

14   created by making use of previously published templates (Shirer et al., 2012) and are displayed in **Fig.**

15   **2A**.

16   In a second step, we investigated effective connectivity during the resting state not only at the level of

17   entire networks, but at the level of the regional components (nodes) of those networks. To this end, we

18   restricted ourselves to the DMN, DAN, and SAN, in line with previous work on the hierarchical

19   organization of RSNs by Zhou et al. (2018). Consistent with Zhou et al. (2018), we subdivided the three

20   key intrinsic networks as follows: For the (core) DMN, we identified 4 regions: posterior cingulate

21   cortex (PCC), anterior medial prefrontal cortex (aMPFC), and left and right angular gyrus (lAG and

22   rAG). For the DAN, we identified 6 regions: left and right frontal eye fields (lFEF and rFEF), left and

23   right inferior frontal gyrus (lIFG and rIFG), and left and right superior parietal lobe (lSPL and rSPL).

24   For the SAN, we identified 5 regions: dorsal anterior cingulate cortex (dACC), left and right anterior

25   insula (lAI and rAI), and left and right anterior prefrontal cortex (lPFC and rPFC). This yielded a total

26   of 15 regions of interest (ROIs) from which BOLD signal time series were extracted. Again, we utilized

27   the templates provided by Shirer et al. (2012) as masks for the respective ROIs (**Fig. 2B**).

28   Delineating the components of the intrinsic networks allowed for a more fine-grained analysis of

29   effective connectivity during the resting state and for disentangling functional integration within *and*

30   between networks. It also served as a more challenging test scenario for rDCM given the increase in

31   network size and, thus, number of free connectivity parameters. Having said this, the resulting network

32   sizes are still an order of magnitude smaller than what is possible using rDCM. However, for the

33   assessment of construct validity, we deliberately restricted ourselves to relatively small networks in

34   order to allow for a comparison between rDCM and spectral DCM (but see below for an application of

35   rDCM to rs-fMRI data from a whole-brain network with more than 200 regions).

13

1 For either of the two analysis pipelines (i.e., "modes" and "nodes"), BOLD signal time series were

2 extracted as the first eigenvariate of all voxels within the respective mask. Time series were mean

3 centered and confound-related variance was removed (by regression using all nuisance regressors

4 included in the first-level GLM).

5
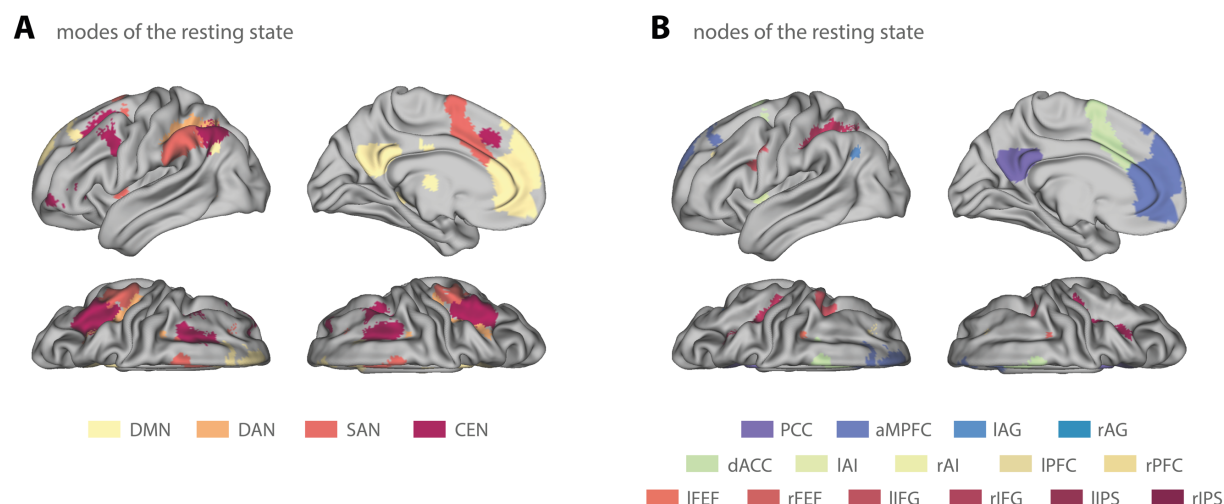
Figure 2. Masks for the empirical analysis of the B-SNIP-1 resting-state data. (A) Masks comprising four key
intrinsic networks (i.e., modes) of the resting state. Specifically, we included the default mode network (DMN;
yellow), dorsal attention network (DAN; *orange*), salience network (SAN; *red*), and the central executive network
(CEN; *purple*). **(B)** Masks comprising 15 subcomponents of three key intrinsic networks (DMN, DAN, and SAN)
of the resting state. For the DMN (*blueish colors*): posterior cingulate cortex (PCC), anterior medial prefrontal
cortex (aMPFC), and left and right angular gyrus (lAG and rAG). For the SAN (*yellowish colors*): dorsal anterior
cingulate cortex (dACC), left and right anterior insula (lAI and rAI), and left and right anterior prefrontal cortex
(lPFC and rPFC). For the DAN (*reddish colors*): left and right frontal eye fields (lFEF and rFEF), left and right
inferior frontal gyrus (lIFG and rIFG), and left and right superior parietal lobe (lSPL and rSPL). All masks were
taken from the templates published in Shirer et al. (2012).

17

18 *2.3.8 Time series extraction: whole-brain network*

19 Finally, to demonstrate that rDCM can infer effective connectivity patterns during the resting state at

20 the whole-brain level, we utilized the Human Brainnetome atlas as a whole-brain parcellation scheme

21 (Fan et al., 2016). The Brainnetome atlas represents a connectivity-based parcellation derived from

22 diffusion weighted imaging (DWI) data. In brief, the atlas comprises 246 distinct parcels (123 per

23 hemisphere), including 210 cortical and 36 subcortical regions. Due to signal dropouts in the raw

24 functional images (especially in inferior temporal regions near the skull base), BOLD signal time series

25 could not be extracted for all regions defined by the Brainnetome atlas. Overall, 221 regions could be

26 extracted in all participants. However, in order to ensure interhemispheric consistency of the network,

27 we further reduced this set by excluded regions that were present in only one hemisphere.

1   This yielded a total of 212 brain regions from which sensible BOLD signal time series could be obtained

2   in every participant. As for the small-network analyses, BOLD signal time series were then extracted as

3   the first eigenvariate of all voxels within the respective mask. Time series were mean centered and

4   confound-related variance was removed (by regression using all nuisance regressors included in the

5   first-level GLM).

6

7   *2.3.9   rDCM analysis*

8   The extracted BOLD signal time series from the previous step were then utilized for subsequent effective

9   connectivity analyses using rDCM. For this, we constructed fully connected networks where all

10 modes/regions were coupled to each other via reciprocal connections. This yielded a total of 16 free

11 parameters for the 4-mode network and 225 free parameters for the 15-node network (all possible inter-

12 regional connections and inhibitory self-connections), as well as 18,260 free parameters for the whole-

13 brain network (all inter-regional connections present in the structural connectome of the Brainnetome

14 atlas and inhibitory self-connections) to be estimated. Again, model inversion was performed by

15 utilizing the routine *tapas_rdcm_estimate.m* from the rDCM toolbox as implemented in TAPAS.

3   RESULTS

## 3.1   Simulations

3.1.1   *Face validity*

First, we performed comprehensive simulation studies to assess the face validity of rDCM for modeling resting-state fMRI data. To this end, we tested whether rDCM can recover generating ("true") parameter values in four different 4-region networks (**Fig. 1**) under various settings of SNR and TR of the synthetic data.

Overall, model parameter values could be recovered faithfully over a wide range of SNR and TR settings, for all four network architectures (**Fig. 3A** and **Supplementary Fig. 1**). Specifically, root-mean-squared-errors (RMSE) were below 0.15 in all tested cases. On average, we found the mean RMSE to range from 0.09 (95% confidence interval: [0.08 0.10]) for model 1 to 0.04 [0.037 0.045] for model 4 for the most challenging case of high measurement noise (SNR = 0.5) and low sampling rate (TR = 2s). Note that the absolute RMSE value is difficult to interpret as it will depend on the scaling of the data; hence, values should only be interpreted in a relative manner as they allow comparison across different data settings (i.e., SNR, TR) as well as across the different DCM variants. The average Pearson's correlation coefficient ($\bar{r}$) ranged from 0.70 [0.63 0.76] for model 1 to 0.95 [0.92 0.97] for model 4 for the low-SNR-slow-TR scenario, and thus showed a pattern highly consistent with the RMSE. Notably, averaging was performed by (i) transforming individual correlation coefficients to z-space using Fisher r-to-z transformation, (ii) computing mean as well as lower and upper bound of the 95% confidence interval in z-space, and finally (iii) back-transforming estimates to r-space. Overall, we observed an (expected) dependence of model parameter recovery on the sparsity of the network, with the most accurate results being observed for the simplest model 4 with the lowest number of free parameters (**Fig. 3A**). Furthermore, we observed a strong dependence of model parameter recovery performance on the data quality, suggesting that rDCM can infer upon connection strengths more faithfully for higher SNR settings of the (synthetic) BOLD signal time series. A similar effect was observed for the sampling rate (i.e., TR). Specifically, model parameter recovery performance of rDCM improved for faster TR settings, although this was somewhat less pronounced as the effect of the SNR (see **Supplementary Fig. 1**). The observed dependencies of model parameter recovery performance on SNR and TR settings are consistent with previous simulation analyses on rDCM in the context of task-based data (Frässle et al., 2018a; 2017).

16

A model parameter recovery (rDCM)
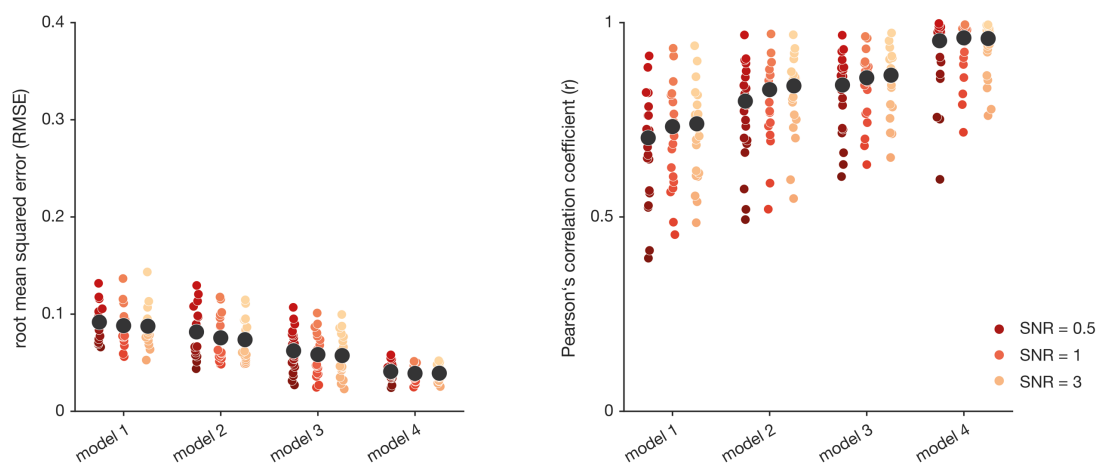
B model parameter recovery (spectral DCM)
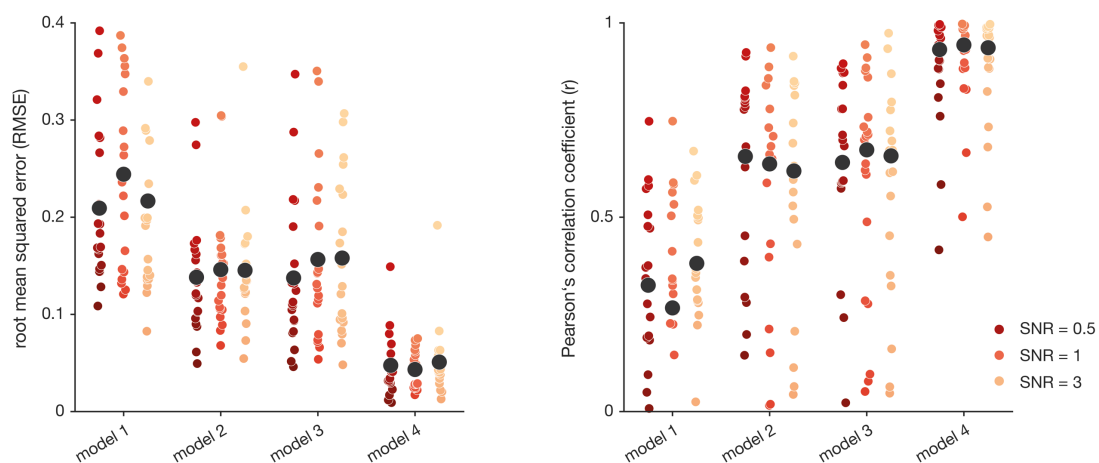
Figure 3. Model parameter recovery for regression DCM and spectral DCM. (A) Model parameter recovery for synthetic resting-state fMRI data in terms of the root mean squared error (*left*) and the Pearson's correlation coefficient (*right*) between the inferred and the "true" data-generating parameter values for regression DCM (rDCM), and **(B)** spectral DCM. Each colored dot represents the respective RMSE or Pearson's correlation coefficient for an individual synthetic subject. Black dots represent the mean RMSE or mean Pearson's correlation coefficient, averaged across all 20 synthetic subjects. Note that averaging of the individual Pearson's correlation coefficients was performed in z-space (see main text). Results are shown for four different network architectures, where model 1 is the most complex (full) model and model 4 is the sparsest network. Furthermore, results are shown for an exemplary repetition time (TR) of 2s and three different signal-to-noise ratios (SNRs): 0.5 (red), 1 (orange), and 3 (yellow).

3.1.2   *Comparison to spectral DCM*

In a second step, we compared the model parameter recovery performance of rDCM with results obtained with spectral DCM (Friston et al., 2014a), which represents an alternative variant of DCM for fMRI that is suited to model resting state data. Importantly, spectral DCM has already been assessed in

17

1   terms of face validity (Friston et al., 2014a), construct validity (Razi et al., 2015), and test-retest

2   reliability (Almgren et al., 2018). Consequently, spectral DCM serves as a useful benchmark against

3   which rDCM can be challenged.

4   Spectral DCMs were fitted to the exact same synthetic BOLD signal time series that have previously

5   been utilized for rDCM. As before, model parameter recovery of spectral DCM improved as a function

6   of network sparsity (**Fig. 3B**). However, the dependence of model parameter recovery on the

7   characteristics of the synthetic fMRI data was less clear for spectral DCM: while no obvious dependence

8   on data quality (i.e., SNR) could be observed, model parameter recovery showed a tendency to improve

9   as a function of sampling rate (see **Supplementary Fig. 2**), which is consistent with previous work

10   (Friston et al., 2014a). When comparing parameter recovery performance across the two DCM variants,

11   overall, rDCM performed better than spectral DCM. More specifically, rDCM and spectral DCM were

12   on par for the simplest model 4, whereas for all other network architectures, rDCM outperformed

13   spectral DCM in terms of RMSE and Pearson's correlation coefficient. This pattern was observed across

14   all SNR and TR settings.

15   In summary, our simulation analyses provide evidence for the face validity of rDCM in application to

16   synthetic resting-state fMRI data in a relatively broad range of SNR and TR settings (**Supplementary**

17   **Fig. 1-2**).

18

19   **3.2   Empirical analyses: 4-mode network**

20   *3.2.1   Effective connectivity during resting state*

21   Next, we investigated the construct validity of rDCM in comparison to spectral DCM, using an empirical

22   rs-fMRI dataset. We here used the 196 healthy participants from the B-SNIP-1 consortium (Tamminga

23   et al., 2013) described earlier (see Methods and Materials).

24   First, we investigated effective connectivity among four key modes of the resting state (**Fig. 2A**), namely

25   the default mode network (DMN), dorsal attention network (DAN), salience network (SAN), and central

26   executive network (CEN). Individual connectivity parameters were estimated in a fully connected

27   network using rDCM. Model inversion resulted in biologically plausible connectivity patterns (**Fig. 4A**).

28   Specifically, we observed pronounced inhibitory afferent (inward) and efferent (outward) connections

29   between the DMN, representing the task-negative network, and the DAN and SAN, representing task-

30   positive networks. Additionally, we observed excitatory reciprocal connections between DAN and SAN.

31   Interestingly, the CEN exerted a relatively strong excitatory influence on the DMN and simultaneously

32   received a (weaker) positive influence from the DMN. Furthermore, the CEN was coupled positively to

33   the DAN and negatively to the SAN.
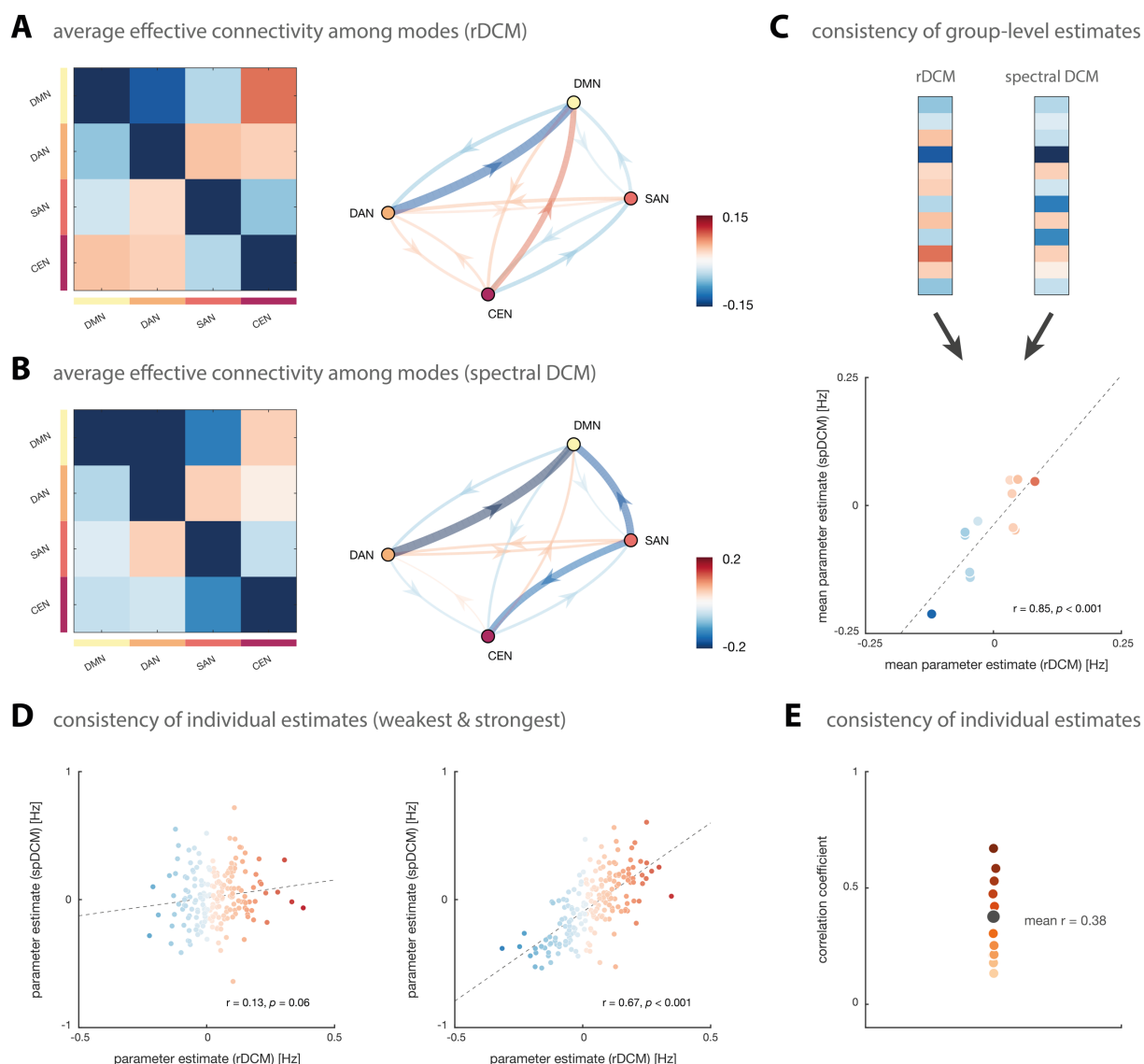
Figure 4. Effective connectivity amongst key modes of the resting state. (A) Group-averaged posterior parameter estimates during the resting state as inferred with regression DCM (rDCM), and **(B)** spectral DCM (spDCM). Note that in order to make the correlation between group-averaged connectivity estimates of the two DCM variants more easily visible, color was scaled in relation to the maximal values of each method. Hence, we utilize slightly different color scales for the two DCM variants for illustration purposes. **(C)** Consistency of group-level parameter averages between rDCM and spDCM. Shown are the group-averaged inter-regional connectivity estimates (*top*) as well as the Pearson's correlation coefficient (r) amongst them (*bottom*). Dots are colored using the rDCM color scale. **(D)** Consistency of individual parameter estimates for each connection separately; here, shown for the connection with the weakest (*left*) and strongest association (*right*) between the two DCM variants. Dots are colored using the rDCM color scale. **(E)** Range of correlation coefficients across all 12 connections of the 4-mode network. Each colored dot represents the Pearson's correlation coefficient for an individual inter-regional connection and the black dot denotes the mean Pearson's correlation coefficient, averaged across all connections. Note that averaging of individual Pearson's correlation coefficients was performed in z-space (see main text).

Based on the posterior parameter estimates, one can compute a metric of hierarchy strength as the difference between averaged absolute efferent and afferent connections for each of the four intrinsic networks of the resting state. This approach has been used by a previous DCM study on the hierarchical

19

1   organization of resting-state networks (Zhou et al., 2018), and is similar to approaches used for

2   analyzing hierarchical projections in the monkey brain (Goulas et al., 2014) and the hierarchical

3   organization of the prefrontal cortex in humans (Nee and D'Esposito, 2016). Importantly, this definition

4   of "hierarchy" should not be confused with other accounts that relate to layer-specificity of forward and

5   backward connections (Felleman and Van Essen, 1991; Markov et al., 2014b). Instead, the present

6   definition simply refers to the degree of influence that one mode exerts over others (for a comprehensive

7   review of alternative definitions of hierarchy, see Hilgetag and Goulas, 2020). The total hierarchy

8   strength was indicative of a hierarchical organization of the four networks with the DMN at the bottom

9   (-0.12), predominantly serving as a sink, and the other task-positive networks ranking higher (DAN:

10 0.05, SAN: 0.02, CEN: 0.05), predominantly serving as neuronal drivers. This finding is consistent with

11 previous observations on the hierarchical organization of resting-state networks (Sridharan et al., 2008;

12 Zhou et al., 2018).

13

14 *3.2.2   Comparison to spectral DCM*

15 As for the simulations above, we again compared our findings from rDCM to those obtained by spectral

16 DCM, using the identical BOLD signal time series. We observed that, at the group level, effective

17 connectivity patterns were highly consistent across the two DCM variants (**Fig. 4B**). In brief, spectral

18 DCM also revealed pronounced inhibitory afferent and efferent connections between the DMN and the

19 DAN and SAN, as well as excitatory reciprocal connections amongst DAN and SAN. The only

20 (qualitative) difference between spectral DCM and rDCM results was observed in how the CEN

21 integrated into this network. While, consistent with the rDCM results, efferent connections from the

22 CEN to the DMN and DAN were excitatory, afferent connections to the CEN from DMN and DAN

23 were inhibitory. With regard to the coupling between CEN and SAN, spectral DCM and rDCM again

24 yielded consistent results. This similarity between the two DCM variants was also represented by a very

25 high Pearson's correlation coefficient between the group-level inter-regional connectivity estimates (r

26 = 0.85, *p* < 0.001; **Fig. 4C**).

27 In a second step, we investigated the consistency of parameter estimates at the individual level (rather

28 than at the group level) for each connection separately. At the individual level, connections showed

29 some variability in the degree of consistency between the two DCM variants. While for some

30 connections, the Pearson's correlation between rDCM and spectral DCM parameter estimates was

31 relatively weak, as seen for the connection from CEN to DAN (r = 0.13, *p* = 0.06; **Fig. 4D**, *left*), others

32 showed much higher consistency, like the connection from DAN to CEN (r = 0.67, *p* < 0.001; **Fig. 4D**,

33 *right*). Overall, the mean Pearson's correlation coefficient, averaged across connections, was $\bar{r}$ = 0.38

34 [0.27 0.48] (**Fig. 4E**). Again, averaging was performed by (i) transforming individual correlation

35 coefficients to z-space using Fisher r-to-z transformation, (ii) computing mean as well as lower and

36 upper bound of the 95% confidence interval in z-space, and finally (iii) back-transforming estimates to

1   r-space. Despite the observed variability, our results suggest that, across connections, there is significant

2   evidence for consistency across parameter estimates from the two DCM variants at the individual level,

3   as assessed using a one-sample *t*-test after Fisher r-to-z transformation of the correlation coefficients

4   ($t_{(1,11)} = 6.40$, $p < 0.001$).

5

6   **3.3   Empirical analysis: 15-node network**
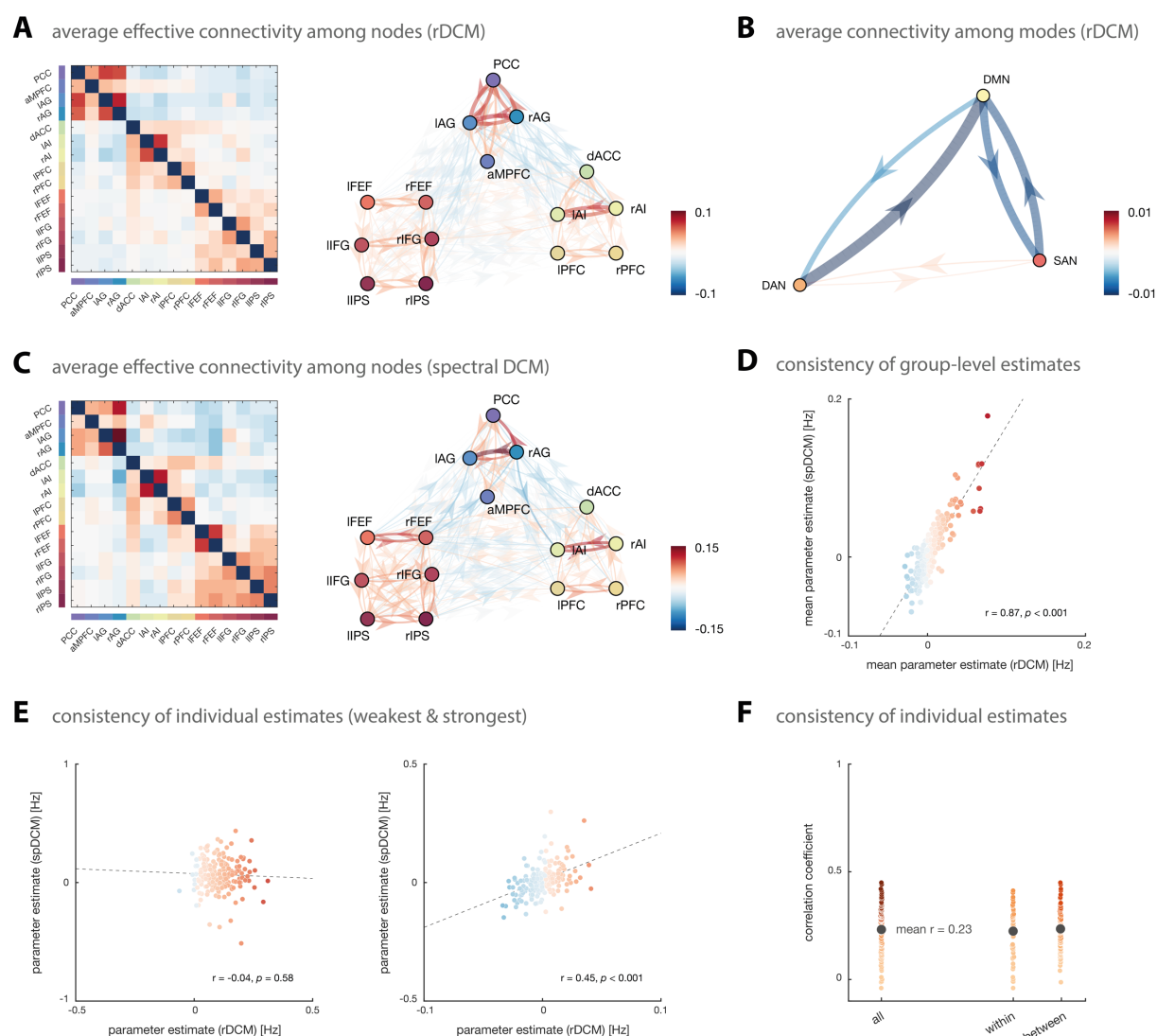
7   *3.3.1   Effective connectivity during resting state*

8   Next, we aimed to study effective connectivity during the resting state not only at the level of entire

9   networks, but at the level of the regional components of those networks. To this end, we restricted our

10   analysis to the DMN, DAN, and SAN and, consistent with Zhou et al. (2018), divided these three

11   networks into 15 subcomponents (**Fig. 2B**). This yielded four regions for the DMN (PCC, aMPFC, lAG,

12   and rAG), five regions for the SAN (dACC, lAI, rAI, lPFC, and rPFC), and six regions for the DAN

13   (lFEF, rFEF, lIFG, rIFG, lSPL, and rSPL).

14   As for the 4-mode network, individual connectivity parameters were estimated in a fully connected

15   network using rDCM, which resulted in biologically plausible connectivity patterns (**Fig. 5A**). The first

16   observation that becomes apparent from our result is the clear modular structure of the average

17   endogenous connectivity matrix where regions that belong to the same RSN group together and show

18   almost exclusively positive (excitatory) within-network interactions. Conversely, between-network

19   connections (i.e., connections originating from a region in one RSN and terminating at a region in

20   another RSN) were overall weaker and more variable, showing both inhibitory and excitatory influences.

21   More precisely, afferent and efferent connections between regions of the DMN and regions of the SAN

22   and DAN were still predominantly negative, whereas connections between the two task-positive

23   networks were more diverse.

24   In order to allow for direct comparability between this analysis of nodes and the previous analysis of

25   modes, we computed the average connection strength between networks (**Fig. 5B**). For example, the

26   strength of all connections originating from a node of the DAN and terminating at a node of the DMN

27   were averaged in order to yield the between-network connection strength from the DAN to the DMN.

28   This analysis suggested that, at the between-network level, results from the 15-node network were

29   highly consistent with the ones obtained in the previous analysis of the 4-mode network. Specifically,

30   we again observed that the DMN was reciprocally connected to the other two RSNs via strong inhibitory

31   connections, whereas the two task-positive networks (i.e., DAN and SAN) showed exhibitory coupling

32   (although very weakly).

33   Finally, based on these between-network connections (**Fig. 5B**), we again computed the hierarchy

34   strength as the difference between averaged absolute efferent and afferent connections between the three

35   RSNs. As for the previous analysis, the total hierarchy strength was indicative of a hierarchical

1   organization of the three networks with the DMN at the bottom (-0.007), predominantly serving as a

2   sink, and the two task-positive networks ranking higher (DAN: 0.006, SAN: 0.0007), predominantly

3   serving as neuronal drivers. In fact, the three RSNs displayed even the exact same ordering as in the

4   previous analysis of modes – a remarkable observation given the differences in the underlying networks

5   – and are thus again consistent with previous work (Sridharan et al., 2008; Zhou et al., 2018).

6

Figure 5. Effective connectivity amongst key brain regions of the resting state. (A) Group-averaged posterior parameter estimates during the resting state as inferred with regression DCM (rDCM). **(B)** Average connection strength between resting-state network (i.e., DMN, DAN, and SAN), computed as the average of the strength of all connections originating in one resting-state network and terminating in another one. **(C)** Group-averaged posterior parameter estimates during the resting state as inferred with spectral DCM (spDCM). Note that in order to make the correlation between group-averaged connectivity estimates of the two DCM variants more easily visible, color was scaled in relation to the maximal values of each method. Hence, we utilize slightly different color scales for the two DCM variants for illustration purposes. **(D)** Consistency of group-level parameter averages between rDCM and spDCM. Shown is the Pearson's correlation coefficient (r) amongst the group-averaged inter-regional connectivity estimates. Dots are colored using the rDCM color scale. **(E)** Consistency of individual parameter estimates for each connection separately; here, exemplarily shown for the connection with the weakest

1   (*left*) and strongest association (*right*) between the two DCM variants. Dots are colored using the rDCM color

2   scale. **(F)** Range of correlation coefficients across all 210 connections of the 15-node network, as well as separately

3   for all connections *within* a resting-state network and all connections *between* resting state network. Each colored

4   dot represents the Pearson's correlation coefficient for an individual inter-regional connection and the black dot

5   denotes the mean Pearson's correlation coefficient, averaged across all connections. Note that averaging of

6   individual Pearson's correlation coefficients was performed in z-space (see main text).

7

8   *3.3.2   Comparison to spectral DCM*

9   As before, we compared our empirical findings from rDCM against the results obtained when utilizing

10 spectral DCM to fit the identical BOLD signal time series. Group-level effective connectivity patterns

11 were again highly consistent across the two DCM variants (**Fig. 5C**). Specifically, spectral DCM

12 revealed the same modular structure of the (average) endogenous connectivity matrix with pronounced

13 excitatory connections between regions of the same RSN and a more diverse pattern for between-

14 network connections. The similarity between the two DCM variants was again represented by a very

15 high Pearson's correlation coefficient between the group-level connectivity estimates (r = 0.87, *p* <

16 0.001; **Fig. 5D**).

17 At the individual level, connections again varied in the degree of consistency between the two DCM

18 variants. Correlations between rDCM and spectral DCM parameter estimates ranged from virtually

19 absent (r = -0.04, *p* = 0.58) for the connection from left to right IFG (**Fig. 5E**, *left*) to robust (r = 0.45, *p*

20 < 0.001) for the connection from left IPS to dACC (**Fig. 5E**, *right*), representing a significant correlation.

21 Overall, the mean correlation coefficient, averaged across all connections, was r̄ = 0.23 [0.22 0.24] (**Fig.**

22 **5F**), and thus slightly decreased as compared to the 4-mode networks. Nevertheless, consistency

23 between rDCM and spectral DCM was significantly larger than zero, as assessed using a one-sample *t*-

24 test ($t_{(1,209)}$ = 34.67, *p* < 0.001). Notably, consistency between the two DCM variants was not

25 significantly different for within-module as compared to between-module connections ($t_{(1,208)}$ = -0.75, *p*

26 = 0.46), suggesting that there was no systematic difference between those two types. As before,

27 averaging of correlation coefficients as well as statistical testing were performed in z-space – that is,

28 after Fisher r-to-z transformation of the individual correlation coefficients.

29 In summary, these results were very similar to the ones obtained for the 4-mode networks, although

30 correlations at the individual level were overall somewhat lower for the 15-node networks.

31

32 **3.4   Empirical analysis: whole-brain network**

33 Having established the face validity and construct validity of rDCM for rs-fMRI data in small networks,

34 we assessed the practical utility of rDCM for inferring resting-state effective connectivity at the whole-

35 brain level. To this end, we utilized the Brainnetome atlas as a whole-brain parcellation scheme, using

36 a 212-node network covering the entire cortex (see Methods for details). Structural connectivity

1   information provided by the Brainnetome atlas was used to constrain the network, resulting in 18,260

2   directed connections to be estimated (**Fig. 6A**).

3   Applying rDCM to rs-fMRI data from these 212 regions resulted in biologically plausible connectivity

4   patterns (**Fig. 6B**). The average endogenous connectivity matrix displayed a clear modular structure

5   where regions within the same lobe (e.g., frontal, occipital) were strongly connected amongst each other

6   (**Fig. 6B**, *left*). Specifically, connections amongst regions within the same lobe were stronger and mostly

7   positive (excitatory), whereas connections amongst regions in different lobes were more variable (i.e.,

8   showing both inhibitory and excitatory influences). This modular structure at the whole-brain level is

9   consistent with our analysis of the connectivity amongst resting-state nodes. Furthermore, we found that

10   while the majority of strong connections are within hemispheres, we also observed strong

11   interhemispheric interactions, primarily amongst homotopic regions (**Fig. 6B**, *right*).

12
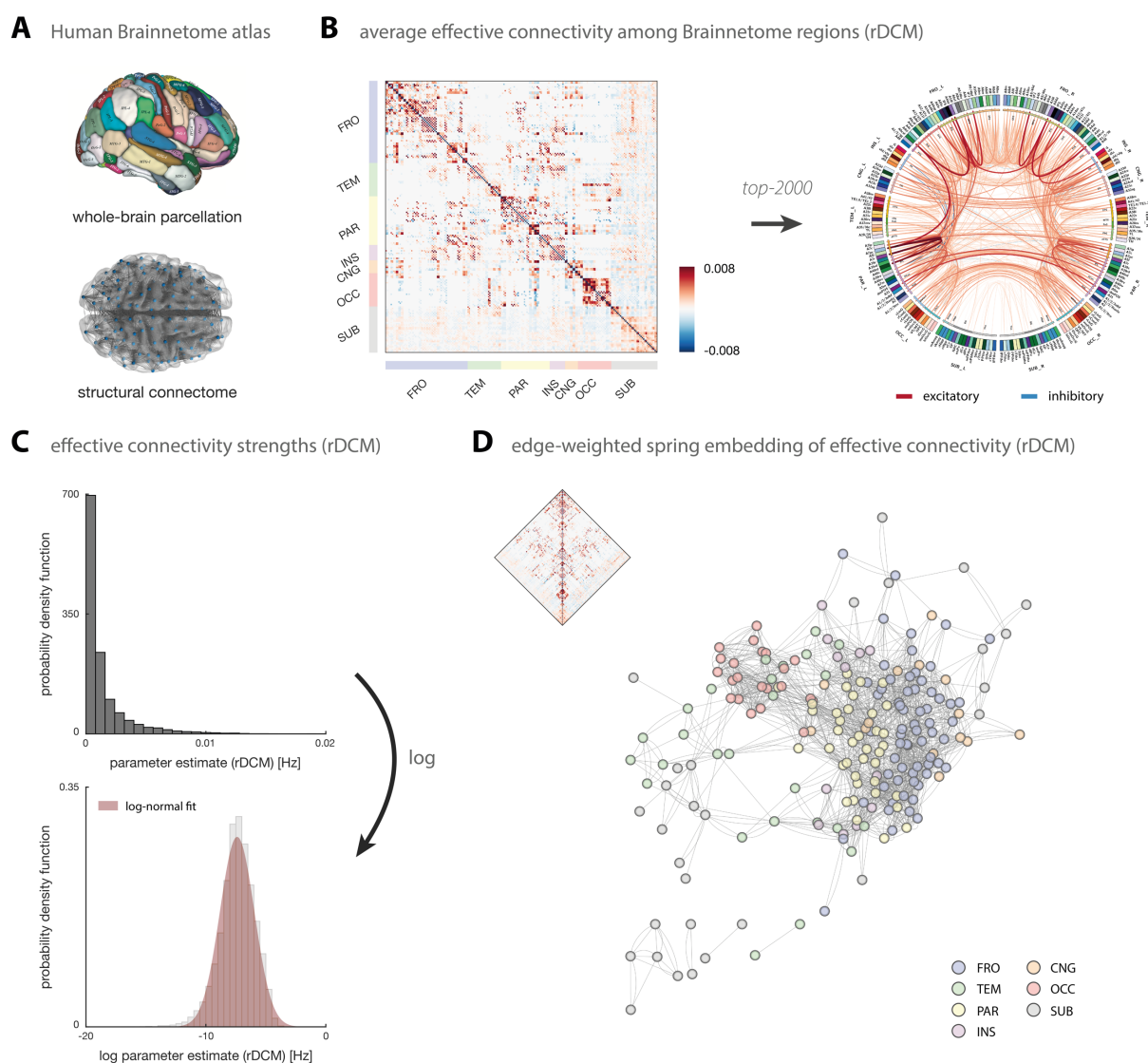
14   **Figure 6. Whole-brain effective connectivity during the resting state. (A)** The whole-brain rDCM analysis
15   used the Human Brainnetome atlas that provides both a whole-brain parcellation of cortical and subcortical regions

24

1   as well as information on the structural connectivity amongst these parcels. The image of the Brainnetome

2   parcellation was reprinted with permission from Fan et al. (2016). **(B)** Group-averaged posterior estimates of

3   directed connection strengths during the resting state as inferred with rDCM. Results are displayed as the full

4   adjacency matrix (*left*) and as a connectogram of the top 2,000 connections with highest (absolute) weights (*right*).

5   The connectogram was produced using Circos (publicly available at http://www.circos.ca/software/). Regions are

6   divided into distinct sets, including frontal (FRO), temporal (TEM), parietal (PAR), insular (INS), cingulate

7   (CNG)[1], occipital (OCC), and subcortical (SUB). **(C)** Distribution of (absolute) endogenous connectivity

8   parameters both in native (*top*) and logarithmic space (*bottom*). **(D)** Edge-weighted spring embedding of the

9   endogenous connectivity matrix based on the top 2000 connections (see above). The edge-weighted spring

10 embedding projection was computed using Cytoscape (publicly available at https://cytoscape.org/download.html).

11 Inspection of the distribution of (absolute) endogenous connectivity parameters revealed a log-normal

12 weight distribution, with connection strengths varying over several orders of magnitude (**Fig. 6C**). This

13 suggests a predominance of weak connections (corresponding to low-frequency influences) in the

14 average endogenous connectivity matrix. At the same time, the log-normal distribution highlights the

15 presence of an extended tail, indicating many strong connections (i.e., higher-frequency influences) in

16 the effective connectome. This finding is consistent with previous reports on log-normal distributions

17 of anatomical connection weights in the mouse cortex (Oh et al., 2014; Wang et al., 2012b) and in the

18 macaque cortical connectome (Markov et al., 2013; 2014a). The fall-off from lower to higher frequency

19 influences in the distribution of rDCM parameter estimates is also broadly consistent with the general

20 spectral characteristics of rs-fMRI data where low frequencies dominate (Chen and Glover, 2015;

21 Kalcher et al., 2014; Niazy et al., 2011).

22 Finally, the modular structure of the resting-state effective connectivity becomes even more apparent

23 when displaying the endogenous connectivity matrix as an edge-weighted spring-embedded projection.

24 In brief, edge-weighted spring embedding positions nodes that interact strongly next to each other, by

25 treating each node as a steel ring linked by springs (edges), thereby forming a mechanical system

26 (Kamada and Kawai, 1989). The optimal layout is then defined as one that minimizes the total energy

27 of the system. Applying edge-weighted spring embedding to the top 2000 connections of the average

28 endogenous connectivity matrix revealed a clear modular structure with a tendency for nodes from the

29 same set to cluster together (**Fig. 6D**). This was particularly clear for regions in the occipital (*red*),

30 parietal (*yellow*), and frontal lobe (*blue*). Furthermore, the spring-embedded projection of the

31 endogenous connectivity matrix was indicative of a hierarchical organization from occipital over

32 parietal to frontal regions. This is consistent with the widespread view on brain organization as a

33 hierarchy ranging from unimodal (e.g., visual) to transmodal (e.g., frontal) regions (Mesulam, 1998).

34 Regions from some lobes (e.g., temporal) did not integrate clearly into this hierarchy, probably due to a

35 dominance of weaker connection strengths below the chosen threshold. Notably, the spring-embedded

36 projection shown here should only be treated as a qualitative illustration, as the exact layout will depend

---

[1] In the Brainnetome nomenclature, this set of regions is called "LIM" (limbic). However, as the term "limbic" is not well-defined (Kötter & Stephan, 1997) and since "LIM" exclusively consists of cingulate areas, we prefer to call this set of regions "CNG" (cingulate).

on, for instance, the density of the network (and thus thresholding) and the exact initialization of the edge-weighted spring embedding algorithm. Having said this, we verified that the overall pattern observed here holds for a range of different settings.

3.5 Computational burden

To provide an overview of computational efficiency, we report the run-time for the different analyses. These values should be treated as a rough indication only, as run-times will depend on the specific hardware used. Here, for all analyses, we used a single processor core on the Euler cluster at ETH Zurich (https://scicomp.ethz.ch/wiki/Euler).

First, rDCM took roughly 1 second for the small networks, regardless of the network architecture. More specifically, for the 4-mode networks (16 free parameters: 12 connection and 4 inhibitory self-connection parameters), model inversion took on average 0.48±0.07s (range: 0.35-0.77s), whereas for the 15-node networks (225 free parameters: 210 connection and 15 inhibitory self-connection parameters), model inversion took on average 0.62±0.10s (range: 0.47-1.32s). For the whole-brain networks (18,260 free parameters: 18,048 connection and 212 inhibitory self-connection parameters), model inversion took on average 132.30±20.26s (range: 90.63-175.30s). This highlights the computational efficiency of rDCM in application to rs-fMRI data and illustrates that the model scales gracefully with the number of regions, consistent with previous work on task-fMRI (Frässle et al., 2018a; 2017).

In comparison, while spectral DCM was also computationally efficient for the small networks, run-times were considerably longer. Specifically, for the 4-mode networks (30 free parameters: 12 connection, 4 inhibitory self-connection, 6 hemodynamic and 8 endogenous noise parameters), model inversion took on average 14.07±5.99s (range: 8.88-48.26s), whereas for the 15-node networks (261 free parameters: 210 connections, 15 inhibitory self-connections, 17 hemodynamic and 19 endogenous noise parameters), model inversion already took on average 3031±1551s (range: 819-12157s).

In summary, although the number of free parameters of the small networks was comparable for the two DCM variants, run-times were up to three orders of magnitude larger for spectral DCM. Specifically, our run-time analyses suggest that rDCM scales much more gracefully to large networks as compared to spectral DCM. Here, we only applied rDCM to rs-fMRI data from a parcellation with more than 200 areas; this was because runtimes of 21-42 hours per subject were reported for spectral DCM in application to a much smaller network of 36 regions (Razi et al., 2017).

1   **4   DISCUSSION**

2   In this paper, we assessed the face validity and construct validity of regression DCM (rDCM; Frässle et

3   al., 2018a; 2017) for inferring effective connectivity during the resting state (i.e., unconstrained

4   cognition) from non-invasive fMRI data. We demonstrated in simulation studies that rDCM can

5   faithfully recover known model parameter values from synthetic rs-fMRI data. Furthermore, we

6   assessed the utility of rDCM in application to an empirical rs-fMRI dataset acquired by the B-SNIP-1

7   consortium (Tamminga et al., 2013). These analyses suggest that rDCM yields biologically plausible

8   results with regard to the functional integration during the resting state. Importantly, effective

9   connectivity results obtained using rDCM in small networks (up to 15 nodes) were highly consistent

10   with the estimates from spectral DCM when fitted to the identical BOLD signal time series.

11   Furthermore, we then demonstrated that rDCM scales gracefully to large-scale network and provides

12   sensible whole-brain effective connectivity estimates from rs-fMRI data within minutes.

13   First, we conducted comprehensive simulation studies to establish the face validity of rDCM when

14   applied to synthetic resting-state fMRI data. Specifically, we found that rDCM faithfully inferred upon

15   the data-generating ("true") parameter values over a wide range of different settings of data quality (i.e.,

16   SNR) and sampling rate (i.e., TR). While model parameter recovery of rDCM improved for higher SNR

17   and faster TR, consistent with previous observations in the context of task-based fMRI (Frässle et al.,

18   2018a; Frässle et al., 2017), performance was reasonable even for the most challenging scenario (i.e.,

19   SNR = 0.5, TR = 2s). Besides this dependence on SNR and TR, model parameter recovery performance

20   also depended on the complexity of the network with more accurate results being observed as the number

21   of free parameters decreased.

22   Second, we applied rDCM to an empirical rs-fMRI dataset to study the effective connectivity amongst

23   entire resting state networks (4 modes), as well as between their subcomponents (15 nodes). For both

24   network architectures, results were biologically plausible and revealed inhibitory afferent and efferent

25   connections for (regions of) the DMN, whereas (regions of) the task-positive networks elicited

26   predominantly excitatory coupling amongst each other. This was also represented by the fact that the

27   different intrinsic RSNs followed a hierarchical ordering (under the given metric of hierarchy strength

28   utilized in the present study), where the DMN was situated at the bottom (primarily serving as a sink)

29   and the task-positive networks (DAN, SAN, and CEN) were situated higher in the hierarchy (primarily

30   serving as drivers). These findings were consistent with previous work on the hierarchical organization

31   of RSNs (Sridharan et al., 2008; Zhou et al., 2018).

32   It is worth highlighting that several other studies have already applied DCM to resting-state fMRI data

33   and have found results consistent with observations reported here. Specifically, these studies have also

34   demonstrated a tight functional coupling amongst regions within a given RSN (Razi et al., 2015; Zhou

35   et al., 2018) as well as comparable patterns of integration between the RSNs (Zhou et al., 2018). This

36   consistency is noteworthy given the substantial differences in analysis strategies among these studies.

27

1   First, the generative model utilized ranged from regression DCM (present work) to spectral DCM (Razi
2   et al., 2015; Sharaev et al., 2016; Zhou et al., 2018), stochastic DCM (Li et al., 2012), and even
3   deterministic DCM with a discrete cosine set (DCT) input (Di and Biswal, 2014). Second, studies
4   differed in the exact choices made regarding preprocessing, the choice of regions/networks, and group
5   level analyses. Despite these methodological and practical differences, all studies (including the present
6   work) revealed tight coupling in the DMN between left angular gyrus and PCC, bilateral angular gyrus
7   and the aMPFC, as well as reciprocal connections between left and right angular gyrus. In particular,
8   the consistent involvement of the bilateral angular gyrus has been suggested to resemble its functional
9   role in domain-general automatic processing, where the angular gyrus has been claimed to function as
10   an automatic bottom-up buffer of incoming information (Humphreys and Lambon Ralph, 2017).
11   Concerning functional integration amongst the subcomponents of the task-positive networks, DCM
12   studies have so far mainly focused on task-based fMRI data. Nevertheless, these studies are still
13   consistent with our findings in that pronounced integration was observed for regions of the SAN (Ham
14   et al., 2013; Lamichhane and Dhamala, 2015) and DAN (Vossel et al., 2012).

15   Importantly, rDCM estimates were not only biologically plausible and consistent with previous findings
16   in the literature, but also quantitatively similar to the results obtained using spectral DCM (Friston et
17   al., 2014a). This is reassuring because spectral DCM represents a well-established variant of DCM for
18   resting-state fMRI which has previously been tested in terms face validity (Friston et al., 2014a),
19   construct validity (Razi et al., 2015), and test-retest reliability (Almgren et al., 2018). Hence, spectral
20   DCM here served as a useful benchmark for assessing the construct validity of rDCM.

21   While group-level estimates were highly consistent across the two variants of DCM, results were more
22   variable at the individual level. Specifically, while individual parameter estimates were highly consistent
23   for some connections, they were more variable between the two methods for others (**Fig. 4E** and **Fig.
24   5F**). An interesting question for future work will be to disentangle whether the observed variability is
25   merely due to chance or whether differences between the two methods for some of the connections
26   actually represents meaningful differences in the variance that is picked up by rDCM and spectral DCM,
27   respectively.

28   We would like to emphasize that, although rDCM and spectral DCM yielded comparable results for
29   small networks (especially at the group level), rDCM was computationally much more efficient. This
30   was indicated by run-times that were up to 3 orders of magnitude higher for spectral DCM for the 15-
31   node network. While this increase in computational efficiency comes at the cost of physiological realism
32   (e.g., fixed HRF, mean-field approximation across regions), this allows rDCM to scale gracefully to
33   very large networks with hundreds of brain regions. In contrast, the current implementation of spectral
34   DCM is unlikely to scale to whole-brain networks of this size, considering the run-times reported here
35   for the small networks as well as those reported previously, suggesting that model inversion of a single
36   DCM with 36 regions takes between 21-42h (Razi et al., 2017).

28

1   Along these lines, we here conducted an additional exploratory analysis to demonstrate the practical

2   feasibility of whole-brain effective connectivity analyses during the resting state with rDCM. These

3   analyses suggested that rDCM can indeed provide biologically plausible effective connectivity patterns

4   at the whole-brain level within minutes. Specifically, rDCM revealed an expected modular structure of

5   the effective connectivity patterns, where areas of the same lobe grouped together and showed more

6   pronounced within-lobe interactions as compared to between-lobe interactions. This is consistent with

7   our findings on the coupling between components of the RSNs ("nodes") as well as previous reports on

8   functional integration during the resting state (Fornito et al., 2016). Furthermore, (absolute) endogenous

9   connection strengths inferred using rDCM followed a log-normal distribution, consistent with previous

10   reports on anatomical connection strengths in mice (Oh et al., 2014; Wang et al., 2012b) and macaques

11   (Markov et al., 2013; 2014a).

12   The combination of computational efficiency and applicability to simple conditions like the "resting

13   state" renders rDCM particularly promising for application in the fields of Computational Psychiatry

14   and Computational Neurology. Here, computational readouts of directed connectivity in whole-brain

15   networks are of major relevance because global dysconnectivity has been postulated as a hallmark of

16   most psychiatric and neurological disorders (Deco and Kringelbach, 2014; Fornito et al., 2015; Frässle

17   et al., 2018b; Menon, 2011; Stephan et al., 2015). The ability to infer whole-brain effective connectivity

18   patterns from rs-fMRI measurements is appealing from a clinical perspective as it puts minimal burden

19   on the patient in the MR scanner. Besides these promises for clinical applications, computational

20   efficiency and application to rs-fMRI renders rDCM also a promising tool for analyses of large-scale

21   datasets (Sudlow et al., 2015; Van Essen et al., 2013). These datasets typically focus on rs-fMRI for its

22   advantage in terms of easy harmonization of multi-site acquisition efforts. The vast amount of data

23   available from these consortia calls for computational tools that are highly efficient. We anticipate that

24   the application of rDCM to these open-source datasets may represent an exciting new opportunity to

25   move human connectomics and network neuroscience towards directed measures of functional

26   inegration.

27   It is worth mentioning that, while we replicated previous findings on putative anticorrelations between

28   the task-negative network (DMN) and the task-positive networks (DAN, SAN) during the resting state

29   (Fox et al., 2005), one should be cautious with regard to the interpretation of this finding. Specifically,

30   it has been argued that the observed anticorrelations may – at least in part – be a consequence of specific

31   preprocessing choices. In particular, global signal regression has been argued to artificially induce

32   anticorrelations by removing any positive correlation between DMN and DAN/SAN that may be hidden

33   in the signal (Buckner et al., 2008; Fox et al., 2009; Murphy et al., 2009). In the present work, we did

34   perform global signal regression to correct for potential confounds like psychological, hardware, and

35   motion artifacts. Hence, results obtained with rDCM in the present study are consistent with previous

36   reports that followed comparable preprocessing strategies (e.g., global signal regression). However, our

1   methodological study cannot – and does not intend – to resolve the ongoing controversy on the validity

2   of anticorrelations between RSNs.

3   Finally, we would like to highlight that the development of rDCM is still in an early stage and the present

4   implementation is subject to several limitations. First, in its current form, rDCM does not account for

5   variability in the hemodynamic response across regions and subjects (Aguirre et al., 1998; Handwerker

6   et al., 2004). Hence, a major future development is to replace the fixed HRF with a more flexible

7   hemodynamic model that nevertheless preserves the computational efficiency of rDCM. Second, it is

8   possible that explicitly accounting for endogenous neuronal fluctuations in the generative forward model

9   could further increase the explanatory power of rDCM for rs-fMRI data. Third, the linear formulation

10   of rDCM and the stationarity assumption of the Fourier transform implies that only stationary measures

11   of functional integration can be obtained. However, functional integration changes dynamically on short

12   timescales, under the influence of synaptic plasticity and neuromodulation (Stephan et al., 2015). Hence,

13   another important future development is to introduce dynamics back into the effective connectivity

14   estimates obtained with rDCM.

15   Despite these limitations, our results demonstrate the face and construct validity of rDCM for modeling

16   rs-fMRI data: they demonstrate good recovery of known connection strengths from simulated data,

17   illustrate that rDCM can infer biologically plausible effective connectivity patterns from empirical rs-

18   fMRI data and show that rDCM estimates are consistent with results from spectral DCM, a well-

19   established DCM variant for rs-fMRI. Importantly, due to the computational efficiency of our method,

20   rDCM scales gracefully with network size and provides sensible whole-brain resting-state effective

21   connectivity patterns. This opens up promising new avenues for whole-brain connectomics based on

22   directed measures of functional integration during the resting state. Such a readout of whole-brain

23   effective connectivity will hopefully not only provide novel insights into the organizing principles of

24   the human brain in health, but might also prove valuable for unraveling pathophysiological mechanisms

25   and for enabling single-subject predictions about clinical outcomes (Frässle et al., 2020b).

5   CODE AND DATA AVAILABILITY

A Matlab implementation of the regression dynamic causal modeling (rDCM) approach is available as open source code in the **T**ranslational **A**lgorithms for **P**sychiatry-**A**dvancing **S**cience (TAPAS) toolbox (www.translationalneuromodeling.org/software). Extensions to the rDCM toolbox that facilitate its application to rs-fMRI data will be made available as part of the next release. Furthermore, following acceptance of this paper, we will publish the code for the analysis online as part of an online repository that conforms to the FAIR (Findable, Accessible, Interoperable, and Re-usable) data principles. The fMRI data were acquired by the B-SNIP-1 consortium and access to the data is possible upon reasonable request.

6   ACKNOWLEDGEMENTS

7   CONFLICT OF INTEREST

The authors declare no competing interests.

8   REFERENCES

Aguirre, G.K., Zarahn, E., D'esposito, M., 1998. The variability of human, BOLD hemodynamic responses. Neuroimage 8, 360-369.

Almgren, H., Van de Steen, F., Kuhn, S., Razi, A., Friston, K., Marinazzo, D., 2018. Variability and reliability of effective connectivity within the core default mode network: A multi-site longitudinal spectral DCM study. Neuroimage 183, 757-768.

Andreasen, N.C., Endicott, J., Spitzer, R.L., Winokur, G., 1977. The Family History Method Using Diagnostic Criteria: Reliability and Validity. Archives of General Psychiatry 34, 1229-1235.

Avants, B.B., Epstein, C.L., Grossman, M., Gee, J.C., 2008. Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. Med Image Anal 12, 26-41.

Baker, J.T., Dillon, D.G., Patrick, L.M., Roffman, J.L., Brady, R.O., Jr., Pizzagalli, D.A., Ongur, D., Holmes, A.J., 2019. Functional connectomics of affective and psychotic pathology. Proc Natl Acad Sci U S A 116, 9050-9059.

Baker, J.T., Holmes, A.J., Masters, G.A., Yeo, B.T., Krienen, F., Buckner, R.L., Ongur, D., 2014. Disruption of cortical association networks in schizophrenia and psychotic bipolar disorder. JAMA Psychiatry 71, 109-118.

Bassett, D.S., Sporns, O., 2017. Network neuroscience. Nature Neuroscience 20, 353-364.

Beckmann, C.F., DeLuca, M., Devlin, J.T., Smith, S.M., 2005. Investigations into resting-state connectivity using independent component analysis. Philosophical Transactions of the Royal Society B-Biological Sciences 360, 1001-1013.

Bishop, C.M., 2006. Pattern recognition and machine learning. Springer, New York. 12, 13, 47, 105.

Biswal, B., Yetkin, F.Z., Haughton, V.M., Hyde, J.S., 1995. Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. Magn Reson Med 34, 537-541.

Biswal, B.B., Van Kylen, J., Hyde, J.S., 1997. Simultaneous assessment of flow and BOLD signals in resting-state functional connectivity maps. NMR Biomed 10, 165-170.

Buckholtz, J.W., Meyer-Lindenberg, A., 2012. Psychopathology and the human connectome: toward a transdiagnostic model of risk for mental illness. Neuron 74, 990-1004.

Buckner, R.L., Andrews-Hanna, J.R., Schacter, D.L., 2008. The brain's default network: anatomy, function, and relevance to disease. Ann N Y Acad Sci 1124, 1-38.

Bullmore, E., Sporns, O., 2009. Complex brain networks: graph theoretical analysis of structural and functional systems. Nat Rev Neurosci 10, 186-198.

1   Chen, J.E., Glover, G.H., 2015. BOLD fractional contribution to resting-state functional connectivity
2   above 0.1 Hz. Neuroimage 107, 207-218.

3   Corbetta, M., Shulman, G.L., 2002. Control of goal-directed and stimulus-driven attention in the brain.
4   Nat Rev Neurosci 3, 201-215.

5   Courchesne, E., Pierce, K., Schumann, C.M., Redcay, E., Buckwalter, J.A., Kennedy, D.P., Morgan, J.,
6   2007. Mapping early brain development in autism. Neuron 56, 399-413.

7   Craddock, R.C., Jbabdi, S., Yan, C.G., Vogelstein, J.T., Castellanos, F.X., Di Martino, A., Kelly, C.,
8   Heberlein, K., Colcombe, S., Milham, M.P., 2013. Imaging human connectomes at the
9   macroscale. Nature Methods 10, 524-539.

10 Daunizeau, J., Friston, K., Kiebel, S., 2009. Variational Bayesian identification and prediction of
11   stochastic nonlinear dynamic causal models. Physica D-Nonlinear Phenomena 238, 2089-2118.

12 Deco, G., Kringelbach, M.L., 2014. Great expectations: using whole-brain computational connectomics
13   for understanding neuropsychiatric disorders. Neuron 84, 892-905.

14 Di, X., Biswal, B.B., 2014. Modulatory interactions between the default mode network and task positive
15   networks in resting-state. PeerJ 2, e367.

16 Dosenbach, N.U., Fair, D.A., Miezin, F.M., Cohen, A.L., Wenger, K.K., Dosenbach, R.A., Fox, M.D.,
17   Snyder, A.Z., Vincent, J.L., Raichle, M.E., Schlaggar, B.L., Petersen, S.E., 2007. Distinct brain
18   networks for adaptive and stable task control in humans. Proc Natl Acad Sci U S A 104, 11073-
19   11078.

20 Esteban, O., Birman, D., Schaer, M., Koyejo, O.O., Poldrack, R.A., Gorgolewski, K.J., 2017. MRIQC:
21   Advancing the automatic prediction of image quality in MRI from unseen sites. Plos One 12,
22   e0184661.

23 Esteban, O., Markiewicz, C.J., Blair, R.W., Moodie, C.A., Isik, A.I., Erramuzpe, A., Kent, J.D.,
24   Goncalves, M., DuPre, E., Snyder, M., Oya, H., Ghosh, S.S., Wright, J., Durnez, J., Poldrack,
25   R.A., Gorgolewski, K.J., 2019. fMRIPrep: a robust preprocessing pipeline for functional MRI.
26   Nat Methods 16, 111-116.

27 Fan, L., Li, H., Zhuo, J., Zhang, Y., Wang, J., Chen, L., Yang, Z., Chu, C., Xie, S., Laird, A.R., Fox,
28   P.T., Eickhoff, S.B., Yu, C., Jiang, T., 2016. The Human Brainnetome Atlas: A New Brain Atlas
29   Based on Connectional Architecture. Cereb Cortex 26, 3508-3526.

30 Felleman, D.J., Van Essen, D.C., 1991. Distributed hierarchical processing in the primate cerebral
31   cortex. Cereb Cortex 1, 1-47.

32 Fonov, V., Evans, A., McKinstry, R., Almli, C.R., Collins, D., 2009. Unbiased nonlinear average age-
33   appropriate brain templates from birth to adulthood. Neuroimage 47.

33

1   Fornito, A., Zalesky, A., Bassett, D.S., Meunier, D., Ellison-Wright, I., Yucel, M., Wood, S.J., Shaw,
2   K., O'Connor, J., Nertney, D., Mowry, B.J., Pantelis, C., Bullmore, E.T., 2011. Genetic
3   influences on cost-efficient organization of human cortical functional networks. J Neurosci 31,
4   3261-3270.

5   Fornito, A., Zalesky, A., Breakspear, M., 2015. The connectomics of brain disorders. Nat Rev Neurosci
6   16, 159-172.

7   Fornito, A., Zalesky, A., Bullmore, E., 2016. Fundamentals of Brain Network Analysis. Elsevier Inc.

8   Fox, M.D., Corbetta, M., Snyder, A.Z., Vincent, J.L., Raichle, M.E., 2006. Spontaneous neuronal
9   activity distinguishes human dorsal and ventral attention systems. Proc Natl Acad Sci U S A
10   103, 10046-10051.

11 Fox, M.D., Raichle, M.E., 2007. Spontaneous fluctuations in brain activity observed with functional
12   magnetic resonance imaging. Nat Rev Neurosci 8, 700-711.

13 Fox, M.D., Snyder, A.Z., Vincent, J.L., Corbetta, M., Van Essen, D.C., Raichle, M.E., 2005. The human
14   brain is intrinsically organized into dynamic, anticorrelated functional networks. Proceedings
15   of the National Academy of Sciences of the United States of America 102, 9673-9678.

16 Fox, M.D., Zhang, D., Snyder, A.Z., Raichle, M.E., 2009. The global signal and observed anticorrelated
17   resting state brain networks. J Neurophysiol 101, 3270-3283.

18 Frässle, S., Lomakina, E.I., Kasper, L., Manjaly, Z.M., Leff, A., Pruessmann, K.P., Buhmann, J.M.,
19   Stephan, K.E., 2018a. A generative model of whole-brain effective connectivity. Neuroimage
20   179, 505-529.

21 Frässle, S., Lomakina, E.I., Razi, A., Friston, K.J., Buhmann, J.M., Stephan, K.E., 2017. Regression
22   DCM for fMRI. Neuroimage 155, 406-421.

23 Frässle, S., Manjaly, Z.M., Do, C.T., Kasper, L., Pruessmann, K.P., Stephan, K.E., 2020a. Whole-brain
24   estimates of directed connectivity for human connectomics. doi:
25   https://doi.org/10.1101/2020.02.20.958124.

26 Frässle, S., Marquand, A.F., Schmaal, L., Dinga, R., Veltman, D.J., van der Wee, N.J.A., van Tol, M.J.,
27   Schobi, D., Penninx, B., Stephan, K.E., 2020b. Predicting individual clinical trajectories of
28   depression with generative embedding. Neuroimage Clin 26, 102213.

29 Frässle, S., Paulus, F.M., Krach, S., Schweinberger, S.R., Stephan, K.E., Jansen, A., 2016. Mechanisms
30   of hemispheric lateralization: Asymmetric interhemispheric recruitment in the face perception
31   network. Neuroimage 124, 977-988.

32 Frässle, S., Yao, Y., Schobi, D., Aponte, E.A., Heinzle, J., Stephan, K.E., 2018b. Generative models for
33   clinical applications in computational psychiatry. Wiley Interdiscip Rev Cogn Sci 9, e1460.

Friston, K., Brown, H.R., Siemerkus, J., Stephan, K.E., 2016. The dysconnection hypothesis (2016). Schizophr Res 176, 83-94.

Friston, K., Harrison, L., Penny, W., 2003. Dynamic causal modelling. Neuroimage 19, 1273-1302.

Friston, K., Mattout, J., Trujillo-Barreto, N., Ashburner, J., Penny, W., 2007. Variational free energy and the Laplace approximation. Neuroimage 34, 220-234.

Friston, K.J., 2011. Functional and effective connectivity: a review. Brain Connect 1, 13-36.

Friston, K.J., Kahan, J., Biswal, B., Razi, A., 2014a. A DCM for resting state fMRI. Neuroimage 94, 396-407.

Friston, K.J., Kahan, J., Razi, A., Stephan, K.E., Sporns, O., 2014b. On nodes and modes in resting state fMRI. Neuroimage 99, 533-547.

Friston, K.J., Williams, S., Howard, R., Frackowiak, R.S., Turner, R., 1996. Movement-related effects in fMRI time-series. Magn Reson Med 35, 346-355.

Gazzaniga, M.S., 2000. Cerebral specialization and interhemispheric communication: does the corpus callosum enable the human condition? Brain 123 (Pt 7), 1293-1326.

Glahn, D.C., Winkler, A.M., Kochunov, P., Almasy, L., Duggirala, R., Carless, M.A., Curran, J.C., Olvera, R.L., Laird, A.R., Smith, S.M., Beckmann, C.F., Fox, P.T., Blangero, J., 2010. Genetic control over the resting brain. Proc Natl Acad Sci U S A 107, 1223-1228.

Gorgolewski, K., Burns, C.D., Madison, C., Clark, D., Halchenko, Y.O., Waskom, M.L., Ghosh, S.S., 2011. Nipype: a flexible, lightweight and extensible neuroimaging data processing framework in python. Front Neuroinform 5, 13.

Goulas, A., Uylings, H.B., Stiers, P., 2014. Mapping the hierarchical layout of the structural network of the macaque prefrontal cortex. Cereb Cortex 24, 1178-1194.

Greicius, M.D., Flores, B.H., Menon, V., Glover, G.H., Solvason, H.B., Kenna, H., Reiss, A.L., Schatzberg, A.F., 2007. Resting-state functional connectivity in major depression: abnormally increased contributions from subgenual cingulate cortex and thalamus. Biol Psychiatry 62, 429-437.

Greicius, M.D., Supekar, K., Menon, V., Dougherty, R.F., 2009. Resting-state functional connectivity reflects structural connectivity in the default mode network. Cereb Cortex 19, 72-78.

Greve, D.N., Fischl, B., 2009. Accurate and robust brain image alignment using boundary-based registration. Neuroimage 48, 63-72.

Hahamy, A., Behrmann, M., Malach, R., 2015. The idiosyncratic brain: distortion of spontaneous connectivity patterns in autism spectrum disorder. Nat Neurosci 18, 302-309.

35

1 Ham, T., Leff, A., de Boissezon, X., Joffe, A., Sharp, D.J., 2013. Cognitive control and the salience
2   network: an investigation of error processing and effective connectivity. J Neurosci 33, 7091-
3   7098.

4 Handwerker, D.A., Ollinger, J.M., D'Esposito, M., 2004. Variation of BOLD hemodynamic responses
5   across subjects and brain regions and their effects on statistical analyses. Neuroimage 21, 1639-
6   1651.

7 Hilgetag, C.C., Goulas, A., 2020. 'Hierarchy' in the organization of brain networks. Philos Trans R Soc
8   Lond B Biol Sci 375, 20190319.

9 Humphreys, G.F., Lambon Ralph, M.A., 2017. Mapping Domain-Selective and Counterpointed
10   Domain-General Higher Cognitive Functions in the Lateral Parietal Cortex: Evidence from
11   fMRI Comparisons of Difficulty-Varying Semantic Versus Visuo-Spatial Tasks, and Functional
12   Connectivity Analyses. Cereb Cortex 27, 4199-4212.

13 Jenkinson, M., Bannister, P., Brady, M., Smith, S., 2002. Improved optimization for the robust and
14   accurate linear registration and motion correction of brain images. Neuroimage 17, 825-841.

15 Jenkinson, M., Beckmann, C.F., Behrens, T.E., Woolrich, M.W., Smith, S.M., 2012. Fsl. Neuroimage
16   62, 782-790.

17 Jenkinson, M., Smith, S., 2001. A global optimisation method for robust affine registration of brain
18   images. Med Image Anal 5, 143-156.

19 Kalcher, K., Boubela, R.N., Huf, W., Bartova, L., Kronnerwetter, C., Derntl, B., Pezawas, L., Filzmoser,
20   P., Nasel, C., Moser, E., 2014. The spectral diversity of resting-state fluctuations in the human
21   brain. Plos One 9, e93375.

22 Kamada, T., Kawai, S., 1989. An Algorithm for Drawing General Undirected Graphs. Information
23   Processing Letters 31, 7-15.

24 Karahanoglu, F.I., Van De Ville, D., 2017. Dynamics of large-scale fMRI networks: Deconstruct brain
25   activity to build better models of brain function. Current Opinion in Biomedical Engineering 3,
26   28-36.

27 Kotter, R., Stephan, K.E., 1997. Useless or helpful? The "limbic system" concept. Rev Neurosci 8, 139-
28   145.

29 Lamichhane, B., Dhamala, M., 2015. The Salience Network and Its Functional Architecture in a
30   Perceptual Decision: An Effective Connectivity Study. Brain Connect 5, 362-370.

31 Li, B., Daunizeau, J., Stephan, K.E., Penny, W., Hu, D., Friston, K., 2011. Generalised filtering and
32   stochastic DCM for fMRI. Neuroimage 58, 442-457.

1   Li, B.J., Wang, X., Yao, S.Q., Hu, D.W., Friston, K.J., 2012. Task-dependent modulation of effective
2   connectivity within the default mode network. Frontiers in Psychology 3.

3   Lui, S., Yao, L., Xiao, Y., Keedy, S.K., Reilly, J.L., Keefe, R.S., Tamminga, C.A., Keshavan, M.S.,
4   Pearlson, G.D., Gong, Q., Sweeney, J.A., 2015. Resting-state brain function in schizophrenia
5   and psychotic bipolar probands and their first-degree relatives. Psychol Med 45, 97-108.

6   Markov, N.T., Ercsey-Ravasz, M., Van Essen, D.C., Knoblauch, K., Toroczkai, Z., Kennedy, H., 2013.
7   Cortical high-density counterstream architectures. Science 342, 1238406.

8   Markov, N.T., Ercsey-Ravasz, M.M., Ribeiro Gomes, A.R., Lamy, C., Magrou, L., Vezoli, J., Misery,
9   P., Falchier, A., Quilodran, R., Gariel, M.A., Sallet, J., Gamanut, R., Huissoud, C., Clavagnier,
10   S., Giroud, P., Sappey-Marinier, D., Barone, P., Dehay, C., Toroczkai, Z., Knoblauch, K., Van
11   Essen, D.C., Kennedy, H., 2014a. A weighted and directed interareal connectivity matrix for
12   macaque cerebral cortex. Cereb Cortex 24, 17-36.

13 Markov, N.T., Vezoli, J., Chameau, P., Falchier, A., Quilodran, R., Huissoud, C., Lamy, C., Misery, P.,
14   Giroud, P., Ullman, S., Barone, P., Dehay, C., Knoblauch, K., Kennedy, H., 2014b. Anatomy
15   of hierarchy: feedforward and feedback pathways in macaque visual cortex. J Comp Neurol 522,
16   225-259.

17 Menon, V., 2011. Large-scale brain networks and psychopathology: a unifying triple network model.
18   Trends Cogn Sci 15, 483-506.

19 Menon, V., Uddin, L.Q., 2010. Saliency, switching, attention and control: a network model of insula
20   function. Brain Struct Funct 214, 655-667.

21 Mesulam, M.M., 1998. From sensation to cognition. Brain 121 (Pt 6), 1013-1052.

22 Mortamet, B., Bernstein, M.A., Jack, C.R., Jr., Gunter, J.L., Ward, C., Britson, P.J., Meuli, R., Thiran,
23   J.P., Krueger, G., Alzheimer's Disease Neuroimaging, I., 2009. Automatic quality assessment
24   in structural brain magnetic resonance imaging. Magn Reson Med 62, 365-372.

25 Murphy, K., Birn, R.M., Handwerker, D.A., Jones, T.B., Bandettini, P.A., 2009. The impact of global
26   signal regression on resting state correlations: are anti-correlated networks introduced?
27   Neuroimage 44, 893-905.

28 Nee, D.E., D'Esposito, M., 2016. The hierarchical organization of the lateral prefrontal cortex. Elife 5.

29 Niazy, R.K., Xie, J., Miller, K., Beckmann, C.F., Smith, S.M., 2011. Spectral characteristics of resting
30   state networks. Prog Brain Res 193, 259-276.

31 Oh, S.W., Harris, J.A., Ng, L., Winslow, B., Cain, N., Mihalas, S., Wang, Q., Lau, C., Kuan, L., Henry,
32   A.M., Mortrud, M.T., Ouellette, B., Nguyen, T.N., Sorensen, S.A., Slaughterbeck, C.R.,
33   Wakeman, W., Li, Y., Feng, D., Ho, A., Nicholas, E., Hirokawa, K.E., Bohn, P., Joines, K.M.,
34   Peng, H., Hawrylycz, M.J., Phillips, J.W., Hohmann, J.G., Wohnoutka, P., Gerfen, C.R., Koch,

C., Bernard, A., Dang, C., Jones, A.R., Zeng, H., 2014. A mesoscale connectome of the mouse brain. Nature 508, 207-214.

Picchioni, D., Duyn, J.H., Horovitz, S.G., 2013. Sleep and the functional connectome. Neuroimage 80, 387-396.

Power, J.D., Barnes, K.A., Snyder, A.Z., Schlaggar, B.L., Petersen, S.E., 2012. Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. Neuroimage 59, 2142-2154.

Power, J.D., Mitra, A., Laumann, T.O., Snyder, A.Z., Schlaggar, B.L., Petersen, S.E., 2014. Methods to detect, characterize, and remove motion artifact in resting state fMRI. Neuroimage 84, 320-341.

Raichle, M.E., 2015. The brain's default mode network. Annu Rev Neurosci 38, 433-447.

Raichle, M.E., MacLeod, A.M., Snyder, A.Z., Powers, W.J., Gusnard, D.A., Shulman, G.L., 2001. A default mode of brain function. Proc Natl Acad Sci U S A 98, 676-682.

Razi, A., Kahan, J., Rees, G., Friston, K.J., 2015. Construct validation of a DCM for resting state fMRI. Neuroimage 106, 1-14.

Razi, A., Seghier, M.L., Zhou, Y., McColgan, P., Zeidman, P., Park, H.J., Sporns, O., Rees, G., Friston, K.J., 2017. Large-scale DCMs for resting-state fMRI. Network Neuroscience 1, 222-241.

Satterthwaite, T.D., Elliott, M.A., Gerraty, R.T., Ruparel, K., Loughead, J., Calkins, M.E., Eickhoff, S.B., Hakonarson, H., Gur, R.C., Gur, R.E., Wolf, D.H., 2013. An improved framework for confound regression and filtering for control of motion artifact in the preprocessing of resting-state functional connectivity data. Neuroimage 64, 240-256.

Sharaev, M.G., Zavyalova, V.V., Ushakov, V.L., Kartashov, S.I., Velichkovsky, B.M., 2016. Effective Connectivity within the Default Mode Network: Dynamic Causal Modeling of Resting-State fMRI Data. Frontiers in Human Neuroscience 10.

Shirer, W.R., Ryali, S., Rykhlevskaia, E., Menon, V., Greicius, M.D., 2012. Decoding subject-driven cognitive states with whole-brain connectivity patterns. Cereb Cortex 22, 158-165.

Shulman, G.L., Fiez, J.A., Corbetta, M., Buckner, R.L., Miezin, F.M., Raichle, M.E., Petersen, S.E., 1997. Common Blood Flow Changes across Visual Tasks: II. Decreases in Cerebral Cortex. J Cogn Neurosci 9, 648-663.

Smith, S.M., Fox, P.T., Miller, K.L., Glahn, D.C., Fox, P.M., Mackay, C.E., Filippini, N., Watkins, K.E., Toro, R., Laird, A.R., Beckmann, C.F., 2009. Correspondence of the brain's functional architecture during activation and rest. Proc Natl Acad Sci U S A 106, 13040-13045.

1   Smith, S.M., Vidaurre, D., Beckmann, C.F., Glasser, M.F., Jenkinson, M., Miller, K.L., Nichols, T.E.,
2   Robinson, E.C., Salimi-Khorshidi, G., Woolrich, M.W., Barch, D.M., Ugurbil, K., Van Essen,
3   D.C., 2013. Functional connectomics from resting-state fMRI. Trends Cogn Sci 17, 666-682.

4   Spreng, R.N., 2012. The fallacy of a "task-negative" network. Front Psychol 3, 145.

5   Spreng, R.N., Sepulcre, J., Turner, G.R., Stevens, W.D., Schacter, D.L., 2013. Intrinsic architecture
6   underlying the relations among the default, dorsal attention, and frontoparietal control networks
7   of the human brain. J Cogn Neurosci 25, 74-86.

8   Spreng, R.N., Stevens, W.D., Chamberlain, J.P., Gilmore, A.W., Schacter, D.L., 2010. Default network
9   activity, coupled with the frontoparietal control network, supports goal-directed cognition.
10   Neuroimage 53, 303-317.

11   Sridharan, D., Levitin, D.J., Menon, V., 2008. A critical role for the right fronto-insular cortex in
12   switching between central-executive and default-mode networks. Proc Natl Acad Sci U S A
13   105, 12569-12574.

14   Stam, C.J., 2014. Modern network science of neurological disorders. Nat Rev Neurosci 15, 683-695.

15   Stephan, K., Marshall, J., Penny, W., Friston, K., Fink, G., 2007. Interhemispheric integration of visual
16   processing during task-driven lateralization. J. Neurosci. 27, 3512-3522.

17   Stephan, K.E., Iglesias, S., Heinzle, J., Diaconescu, A.O., 2015. Translational Perspectives for
18   Computational Neuroimaging. Neuron 87, 716-732.

19   Sudlow, C., Gallacher, J., Allen, N., Beral, V., Burton, P., Danesh, J., Downey, P., Elliott, P., Green, J.,
20   Landray, M., Liu, B., Matthews, P., Ong, G., Pell, J., Silman, A., Young, A., Sprosen, T.,
21   Peakman, T., Collins, R., 2015. UK biobank: an open access resource for identifying the causes
22   of a wide range of complex diseases of middle and old age. PLoS Med 12, e1001779.

23   Tagliazucchi, E., Laufs, H., 2014. Decoding wakefulness levels from typical fMRI resting-state data
24   reveals reliable drifts between wakefulness and sleep. Neuron 82, 695-708.

25   Tamminga, C.A., Ivleva, E.I., Keshavan, M.S., Pearlson, G.D., Clementz, B.A., Witte, B., Morris, D.W.,
26   Bishop, J., Thaker, G.K., Sweeney, J.A., 2013. Clinical phenotypes of psychosis in the Bipolar-
27   Schizophrenia Network on Intermediate Phenotypes (B-SNIP). Am J Psychiatry 170, 1263-
28   1274.

29   Tamminga, C.A., Pearlson, G., Keshavan, M., Sweeney, J., Clementz, B., Thaker, G., 2014. Bipolar and
30   schizophrenia network for intermediate phenotypes: outcomes across the psychosis continuum.
31   Schizophr Bull 40 Suppl 2, S131-137.

32   Treiber, J.M., White, N.S., Steed, T.C., Bartsch, H., Holland, D., Farid, N., McDonald, C.R., Carter,
33   B.S., Dale, A.M., Chen, C.C., 2016. Characterization and Correction of Geometric Distortions
34   in 814 Diffusion Weighted Images. Plos One 11, e0152472.

Tustison, N.J., Avants, B.B., Cook, P.A., Zheng, Y., Egan, A., Yushkevich, P.A., Gee, J.C., 2010. N4ITK: improved N3 bias correction. IEEE Trans Med Imaging 29, 1310-1320.

van den Heuvel, M.P., Hulshoff Pol, H.E., 2010. Exploring the brain network: a review on resting-state fMRI functional connectivity. Eur Neuropsychopharmacol 20, 519-534.

Van Essen, D.C., Smith, S.M., Barch, D.M., Behrens, T.E.J., Yacoub, E., Ugurbil, K., Consortium, W.-M.H., 2013. The WU-Minn Human Connectome Project: An overview. Neuroimage 80, 62-79.

Vossel, S., Weidner, R., Driver, J., Friston, K.J., Fink, G.R., 2012. Deconstructing the architecture of dorsal and ventral attention systems with dynamic causal modeling. J Neurosci 32, 10637-10648.

Wang, L., Hermens, D.F., Hickie, I.B., Lagopoulos, J., 2012a. A systematic review of resting-state functional-MRI studies in major depression. J Affect Disord 142, 6-12.

Wang, Q., Sporns, O., Burkhalter, A., 2012b. Network analysis of corticocortical connections reveals ventral and dorsal processing streams in mouse visual cortex. J Neurosci 32, 4386-4399.

Wang, S., Peterson, D.J., Gatenby, J.C., Li, W., Grabowski, T.J., Madhyastha, T.M., 2017. Evaluation of Field Map and Nonlinear Registration Methods for Correction of Susceptibility Artifacts in Diffusion MRI. Front Neuroinform 11, 17.

Zeki, S., Shipp, S., 1988. The functional logic of cortical connections. Nature 335, 311-317.

Zhang, Y., Brady, M., Smith, S., 2001. Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. IEEE Trans Med Imaging 20, 45-57.

Zhou, Y., Friston, K.J., Zeidman, P., Chen, J., Li, S., Razi, A., 2018. The Hierarchical Organization of the Default, Dorsal Attention and Salience Networks in Adolescents and Young Adults. Cerebral Cortex 28, 726-737.

1 # SUPPLEMENTARY MATERIAL

2 # Regression dynamic causal modeling for resting-state fMRI

3 Stefan Frässle, Samuel J. Harrison, Jakob Heinzle, Brett A. Clementz, Carol A. Tamminga,

4 John A. Sweeney, Elliot S. Gershon, Matcheri S. Keshavan, Godfrey D. Pearlson, Albert Powers,

5 Klaas E. Stephan

6

7

8 ## SUPPLEMENTARY MATERIAL S1: DETAILS ON DATA ACQUISITION

9 Structural and functional MRI data were acquired at the five sites of the Bipolar-Schizophrenia Network

10 on Intermediate Phenotypes (B-SNIP-1) consortium (Tamminga et al., 2013; 2014). These included the

11 University of Maryland School of Medicine (Baltimore), Commonwealth Research Center at Harvard

12 Medical School (Boston), University of Chicago Medical Center (Chicago), University of Texas

13 Southwestern Medical Center (Dallas), and Olin Neuropsychiatry Research Center at the Institute of

14 Living (Hartford). Participants were scanned on 3-Tesla MR scanners of different manufacturers,

15 including GE Signa, Siemens Trio, Philips Achieva, and Siemens Allegra. For the present

16 methodological study, neuroimaging data from the sites in Baltimore, Chicago, Dallas and Hartford was

17 utilized since not all information relevant for our preprocessing pipeline was available for the Boston

18 data.

19 For *Baltimore*, resting-state fMRI data was acquired on a 3T Siemens Tim Trio scanner. Volumes were

20 acquired using a T_2^*-weighted echo-planar imaging (EPI) sequence (36 slices, TR=2210ms, TE=30ms,

21 field of view (FoV) 224×224×144mm^3, voxel size 3.5×3.5×4mm^3, flip angle 70°). In addition to the

22 functional images, a high-resolution anatomical image was acquired for each participant using a T1-

23 weighted Magnetization Prepared Rapid Gradient Echo (MPRAGE) sequence, following the

24 Alzheimer's Disease Neuroimaging Initiative protocol (http://adni.loni.usc.edu). The MPRAGE

25 sequence parameters were as follows: 3D acquisition, sagittal slab, 160 slices, shot interval 2300ms,

26 inversion time 900ms, TR=7.2ms, TE=2.91ms, field of view (FoV) 256×240×192mm^3, voxel size

27 1×1×1.2mm^3, flip angle 9°.

28 For *Chicago*, resting-state fMRI data was acquired on a 3T GE Signa HDx scanner. Volumes were

29 acquired using a T_2^*-weighted echo-planar imaging (EPI) sequence (29 slices, TR=1775ms, TE=27ms,

30 field of view (FoV) 224×224×145mm^3, voxel size 3.5×3.5×4mm^3, slice gap 1mm, flip angle 60°). A

31 high-resolution anatomical image was acquired using a T1-weighted MPRAGE sequence (3D

32 acquisition, sagittal slab, 166 slices, shot interval 2300ms, inversion time 700ms, TR=6.988ms,

33 TE=2.848ms, field of view (FoV) 260×260×200mm^3, voxel size 1.016×1.016×1.2mm^3, flip angle 8°).

1   For *Dallas*, resting-state fMRI data was acquired on a 3T Philips Achieva scanner. Volumes were

2   acquired using a $T_2{}^*$-weighted echo-planar imaging (EPI) sequence (29 slices, TR=1500ms, TE=27ms,

3   field of view (FoV) $224\times224\times145mm^3$, voxel size $3.5\times3.5\times4mm^3$, slice gap 1mm, flip angle 60°). A

4   high-resolution anatomical image was acquired using a T1-weighted MPRAGE sequence (3D

5   acquisition, sagittal slab, 170 slices, shot interval 3000ms, inversion time 846ms, TR=6.8ms, TE=3.1ms,

6   field of view (FoV) $256\times256\times204mm^3$, voxel size $1\times1\times1.2mm^3$, flip angle 8°).

7   And finally, for *Hartford*, resting-state fMRI data was acquired on a 3T Siemens Allegra scanner.

8   Volumes were acquired using a $T_2{}^*$-weighted echo-planar imaging (EPI) sequence (29 slices,

9   TR=1500ms, TE=27ms, field of view (FoV) $224\times224\times145mm^3$, voxel size $3.5\times3.5\times4mm^3$, slice gap

10   1mm, flip angle 70°). A high-resolution anatomical image was acquired using a T1-weighted MPRAGE

11   sequence (3D acquisition, sagittal slab, 160 slices, shot interval 2300ms, inversion time 900ms,

12   TR=6.8ms, TE=2.91ms, field of view (FoV) $256\times240\times192mm^3$, voxel size $1\times1\times1.2mm^3$, flip angle 9°).
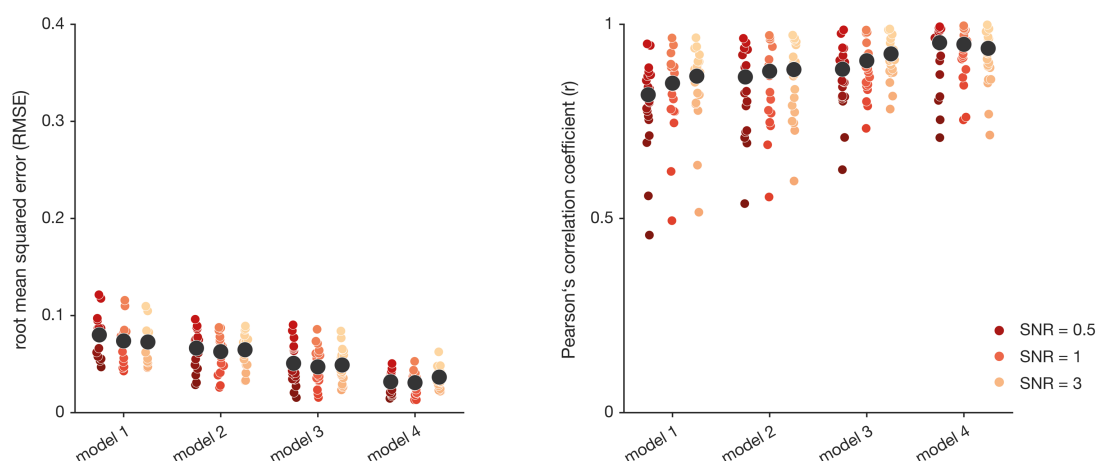
1 **SUPPLEMENTARY FIGURES**

2 **Supplementary Figure S1**
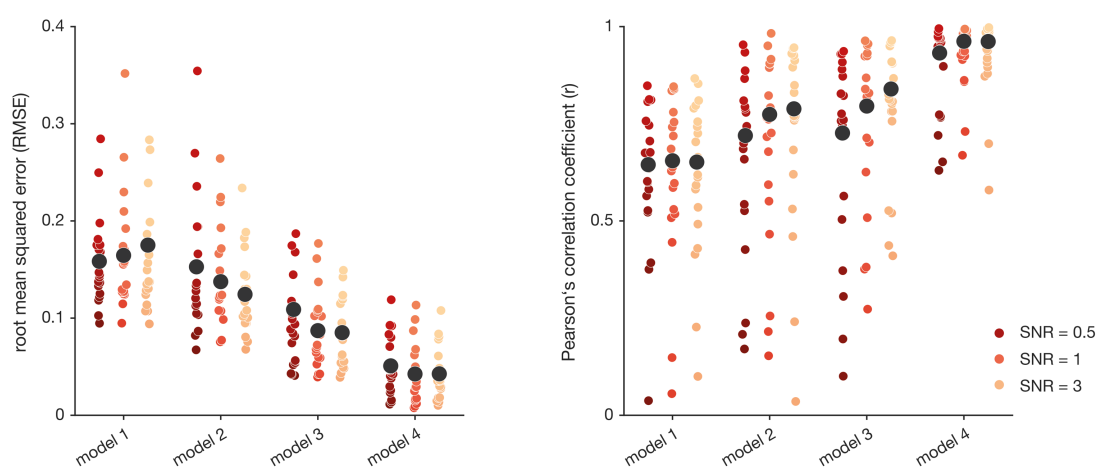

A model parameter recovery (rDCM) | TR = 1s
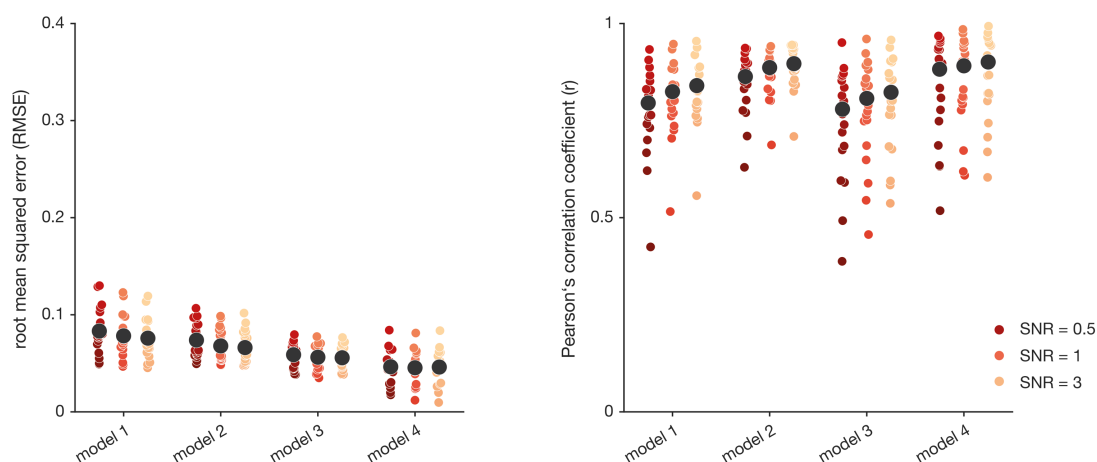
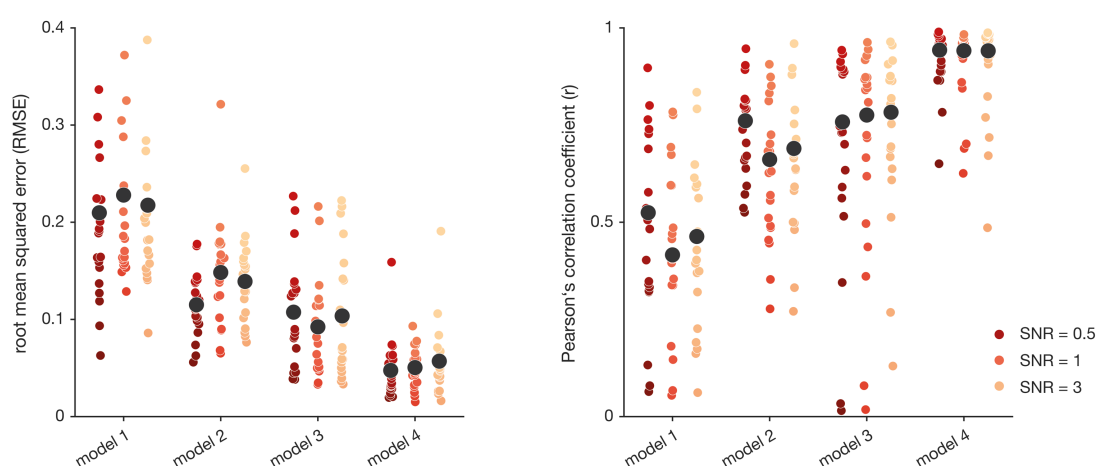**B** model parameter recovery (rDCM) | TR = 0.5s

3

4 **Supplementary Figure S1. Model parameter recovery for regression DCM.** Model parameter recovery for
5 synthetic resting-state fMRI data in terms of the root mean squared error (*left*) and the Pearson's correlation
6 coefficient (*right*) between the inferred and the "true" data-generating parameter values for regression DCM
7 (rDCM) for **(A)** a repetition time (TR) of 1s, and **(B)** a repetition time (TR) of 0.5s. Each colored dot represents
8 the respective RMSE or Pearson's correlation coefficient for an individual synthetic subject. Black dots represent
9 the mean RMSE or mean Pearson's correlation coefficient, averaged across all 20 synthetic subjects. Note that
10 averaging of the individual Pearson's correlation coefficients was performed in z-space (see main text in main
11 manuscript). Results are shown for four different network architectures, where model 1 is the most complex (full)
12 model and model 4 is the sparsest network. Furthermore, results are shown for an exemplary repetition time (TR)
13 of 2s and three different signal-to-noise ratios (SNRs): 0.5 (red), 1 (orange), and 3 (yellow).

1 **Supplementary Figure S2**

A model parameter recovery (spectral DCM) | TR = 1s

B model parameter recovery (spectral DCM) | TR = 0.5s

2

3 **Supplementary Figure S2. Model parameter recovery for spectral DCM.** Model parameter recovery for
4 synthetic resting-state fMRI data in terms of the root mean squared error (*left*) and the Pearson's correlation
5 coefficient (*right*) between the inferred and the "true" data-generating parameter values for spectral DCM for **(A)**
6 a repetition time (TR) of 1s, and **(B)** a repetition time (TR) of 0.5s. Each colored dot represents the respective
7 RMSE or Pearson's correlation coefficient for an individual synthetic subject. Black dots represent the mean
8 RMSE or mean Pearson's correlation coefficient, averaged across all 20 synthetic subjects. Note that averaging of
9 the individual Pearson's correlation coefficients was performed in z-space (see main text in main manuscript).
10 Results are shown for four different network architectures, where model 1 is the most complex (full) model and
11 model 4 is the sparsest network. Furthermore, results are shown for an exemplary repetition time (TR) of 2s and
12 three different signal-to-noise ratios (SNRs): 0.5 (red), 1 (orange), and 3 (yellow).

1   **REFERENCES**

2   Tamminga, C.A., Ivleva, E.I., Keshavan, M.S., Pearlson, G.D., Clementz, B.A., Witte, B., Morris, D.W.,
3   Bishop, J., Thaker, G.K., Sweeney, J.A., 2013. Clinical phenotypes of psychosis in the Bipolar-
4   Schizophrenia Network on Intermediate Phenotypes (B-SNIP). Am J Psychiatry 170, 1263-
5   1274.

6   Tamminga, C.A., Pearlson, G., Keshavan, M., Sweeney, J., Clementz, B., Thaker, G., 2014. Bipolar and
7   schizophrenia network for intermediate phenotypes: outcomes across the psychosis continuum.
8   Schizophr Bull 40 Suppl 2, S131-137.