Title: Robust extraction of spatio-temporal patterns from resting state fMRI

Authors: Katharina Glomb[1,*], Adrián Ponce-Alvarez[1], Matthieu Gilson[1], Petra Ritter[2,3,4,5], and Gustavo Deco[1,6]

Affiliations:

1 Center for Brain and Cognition, Dept. of Technology and Information, Universitat Pompeu Fabra, Carrer Tànger, 122-140, 08018 Barcelona, Spain

2 Minerva Research Group Brain Modes, Max Planck Institute for Human Cognitive and Brain Sciences, Stephanstrasse 1, 04103 Leipzig, Germany

3 Dept. of Neurology, Charité - University Medicine, Charitéplatz 1, 10117 Berlin, Germany

4 Bernstein Focus State Dependencies of Learning & Bernstein Center for Computational Neuro-science, Philippstrasse 12, 10115 Berlin, Germany

5 Berlin School of Mind and Brain & Mind and Brain Institute, Humboldt University, Luisenstrasse 56, 10117 Berlin, Germany

6 Institució Catalana de la Recerca i Estudis Avançats, Universitat Barcelona, Passeig Lluís Companys 23, 08010 Barcelona, Spain

* corresponding author, katharina.glomb@upf.edu

Abstract

It is well-established that patterns of functional connectivity (FC) - measures of correlated activity between pairs of voxels or regions observed in the human brain using neuroimaging - are robustly expressed in spontaneous activity during rest. These patterns are not static, but exhibit complex spatio-temporal dynamics. In this study, we use a whole-brain approach combining data analysis and modelling of FC dynamics between 66 ROIs covering the entire cortex. We simultaneously utilize temporal and spatial information by creating tensors that are subsequently decomposed into sets of brain regions ("communities") that share similar temporal dynamics, and their associated time courses. The tensors contain pairwise FC computed inside of overlapping sliding windows. Communities are discovered by clustering features obtained from 24 healthy subjects, thereby ensuring that they generalize across subjects. First, we determine that at this resolution, four communities that resemble known RSNs can be clearly discerned in the empirical data: DMN, visual network, control networks, and sensorimotor network. Second, we use a noise-driven stationary mean field model which possesses simple node dynamics and realistic anatomical connectivity derived from DTI and fiber tracking. It has been shown to explain resting state FC as averaged over time and multiple subjects, however, this average FC summarizes the spatial distribution of correlations while hiding their temporal dynamics. Thus, it is unclear whether the same type of model can reproduce FC at different points in time. We find that this is the case for all four networks using the spatio-temporal information revealed by tensor decomposition if nodes in the simulation are connected according to model-based effective connectivity. Furthermore, we find that these results require only a small part of the FC values, namely the highest values that occur across time and ROI pair. Our findings suggest that in resting state fMRI, FC patterns that occur over time are mostly derived from the average FC, are shaped by underlying structural connectivity, and that the activation of these patterns is limited to brief periods in time. We provide an innovative method that does not make strong assumptions about the underlying data and is generally applicable to resting state or task data from different subject populations.

1 Introduction

The question of how large-scale cortical function arises from underlying anatomical connectivity has been the object of much investigation since the advent of non-invasive imaging techniques (Vincent et al., 2007; Matsui et al., 2011; Wang et al., 2013), in particular since it was discovered that interareal functional relationships found under task conditions are maintained during rest (Biswal et al., 1995; Cordes et al., 2000; Beckmann and Smith, 2004; Fox et al., 2005). With magnetic resonance imaging (MRI) it is possible to obtain both functional and structural connectivities (FC and SC, respectively). Although there is large variability across subjects and sessions, both in SC (Heiervang et al., 2006) and FC measures (Mueller et al., 2013; Finn et al., 2015), studies using group averages have revealed general principles of information processing in the brain (Raichle et al., 2001; Doucet et al., 2011; Van den Heuvel and Sporns, 2011; Deco and Jirsa, 2012; Haimovici et al., 2013).

In order to connect SC and FC, computational models are an important tool for understanding how activity propagates from one node to another to produce the observed data (Honey et al., 2009; Cabral et al., 2012; Deco et al., 2014). Most models optimize their parameters by fitting the average FC. Only recently, the question of whether and how relevant information can be extracted from the fluctuations in pairwise FC strength, and how to describe the richness of the temporal dynamics,

has received increasing attention in data analysis (Chang and Glover, 2010; Hutchison et al., 2012; Allen et al., 2012; Liu et al., 2013) and modelling (Hansen et al., 2014; Ponce-Alvarez et al., 2015). This has lead to the notion of dynamic functional connectivity (dFC).

Here, we use a dynamic mean field model (Wong and Wang, 2006) of the human cortex which has been shown to reproduce average resting state (RS) FC (Deco et al., 2014). It is our goal to determine whether simulated data exhibit FC patterns over time that resemble those of empirical data. To this end, we analyze RS data from 24 healthy subjects (Schirner et al., 2015) and compare to simulated data. The cortex is modelled by 66 nodes corresponding to 66 brain areas also used to parcellate the empirical data. The nodes are connected according to empirical SC derived from the same subjects.

We opt for tensor decomposition for extracting relevant and general features of the spatio-temporal dynamics. This method has been shown to work well for community detection (Gauvin et al., 2014) and has been applied to brain data (Cichocki, 2013; Leonardi and Van de Ville, 2013; Ponce-Alvarez et al., 2015). Unlike ICA, which has become the standard method for extracting RSNs (McKeown et al., 1998; Beckmann et al., 2005; Mantini et al., 2007), tensor factorization does not assume spatial independence of the underlying components, which is a strong constraint not directly motivated by the data. Here, such a constraint is not required and hence, the space of possible solutions is not unnecessarily restricted. Furthermore, it has the advantage that it can readily be used at our level of coarse spatial resolution.

The modelling approach aims at linking FC and SC. One conceptual problem of SC is that it provides neither directionality information nor the weights of the connections. These two points are addressed by the concept of effective connectivity (EC). SC can be viewed as an approximation to EC, and it is the latter that is genuinely related to the dynamics in network models (Friston, 1994). Reversely, underlying connectivity (SC or EC) can be inferred from FC, or more generally, from the dynamics found in the data, through the same kinds of models. Gilson et al. (2016) developed a method to extract EC from RS fMRI data using a noise diffusion model. They show that the EC that accounts best for empirical FC significantly differs from the SC in a number of points. We use both SC and EC as underlying connectivity in our model and explore how their properties are linked to the spatio-temporal patterns found in empirical and simulated data.

2

2 Methods

2.1 Empirical data

RS fMRI as well as corresponding diffusion weighted (dw) MRI data were collected from 24 healthy participants (11 female) at the Charité Berlin, Germany, by Petra Ritter and co-workers. The original dataset consisted of 49 subjects, but we chose only those aged 18 to 35 years (mean 25.7 years) since it is known that FC changes with age (Meunier et al., 2009). Each dataset amounts to 661 time points recorded at TR=2s, i.e. about 21 minutes. In the same session, EEG was also recorded, but we do not use the data here. RS BOLD was recorded while subjects were asked to stay awake with their eyes closed, using a 3T Siemens Trim Trio scanner and a 12 channel Siemens head coil. Voxel time courses are averaged inside ROIs defined by the Desikan-Killiany atlas (Desikan et al., 2006) as implemented in FreeSurfer. We removed the areas labeled as corpus callosum on both sides since they only contain white matter, amounting to 33 cortical ROIs for each hemisphere. The dwMRI data were subjected to fiber tracking to obtain structural connectivity (SC) matrices for each subject. Details are available in Schirner et al. (2015).

2.2 Model data

A mean field approximation of a network of populations of leaky integrate and fire neurons (Wong and Wang, 2006) was used to simulate RS activity as described in Deco et al. (2014). The excitatory populations are connected using a) an average over the SC matrices derived from dwMRI data (section 2.1), and b) an effective connectivity (EC) matrix derived from these SC matrices as described in section 2.9.

The activity of the populations is computed using a set of coupled nonlinear stochastic differential

Table 1: Parameters of the DMF model.

Excitatory	Inhibitory
$a_E = 310(\mathrm{nC}^{-1})$	$a_I = 615(\mathrm{nC}^{-1})$
$b_E = 125(\mathrm{Hz})$	$b_I = 177(\mathrm{Hz})$
$d_E = 0.16(\mathrm{s})$	$d_I = 0.087(\mathrm{s})$
$\tau_E = \tau_{\mathrm{NMDA}} = 100(\mathrm{ms})$	$\tau_I = \tau_{\mathrm{GABA}} = 10(\mathrm{ms})$
$I_{0,E} = 0.382$	$I_{0,I} = 0.267$
$w_{EE} = 0.21$	$w_{II} = 1,\ w_{IE} = 0.15$

equations:

$$I_i^{(E)} = I_{0,E} + w_{\mathrm{EE}}S_i^{(E)} + G\sum_j C_{ij}S_j^{(E)} - w_{\mathrm{EI},i}S_i^{(I)}$$

$$I_i^{(I)} = I_{0,I} + w_{\mathrm{IE}}S_i^{(E)} - w_{\mathrm{II}}S_i^{(I)}$$

$$r_i^{(E)} = H^{(E)}(I_i^{(E)}) = \frac{a_E I_i^{(E)} - b_E}{1 - \exp(-d_E(a_E I_i^{(E)} - b_E))}$$

$$r_i^{(I)} = H^{(I)}(I_i^{(I)}) = \frac{a_I I_i^{(I)} - b_I}{1 - \exp(-d_I(a_I I_i^{(I)} - b_I))}$$

$$\frac{\mathrm{d}S_i^{(E)}(t)}{\mathrm{d}t} = -\frac{S_i^{(E)}}{\tau_E} + (1 - S_i^{(E)})\gamma r_i^{(E)} + \sigma\nu_i(t)$$

$$\frac{\mathrm{d}S_i^{(I)}(t)}{\mathrm{d}t} = -\frac{S_i^{(I)}}{\tau_I} + r_i^{(I)} + \sigma\nu_i(t)$$
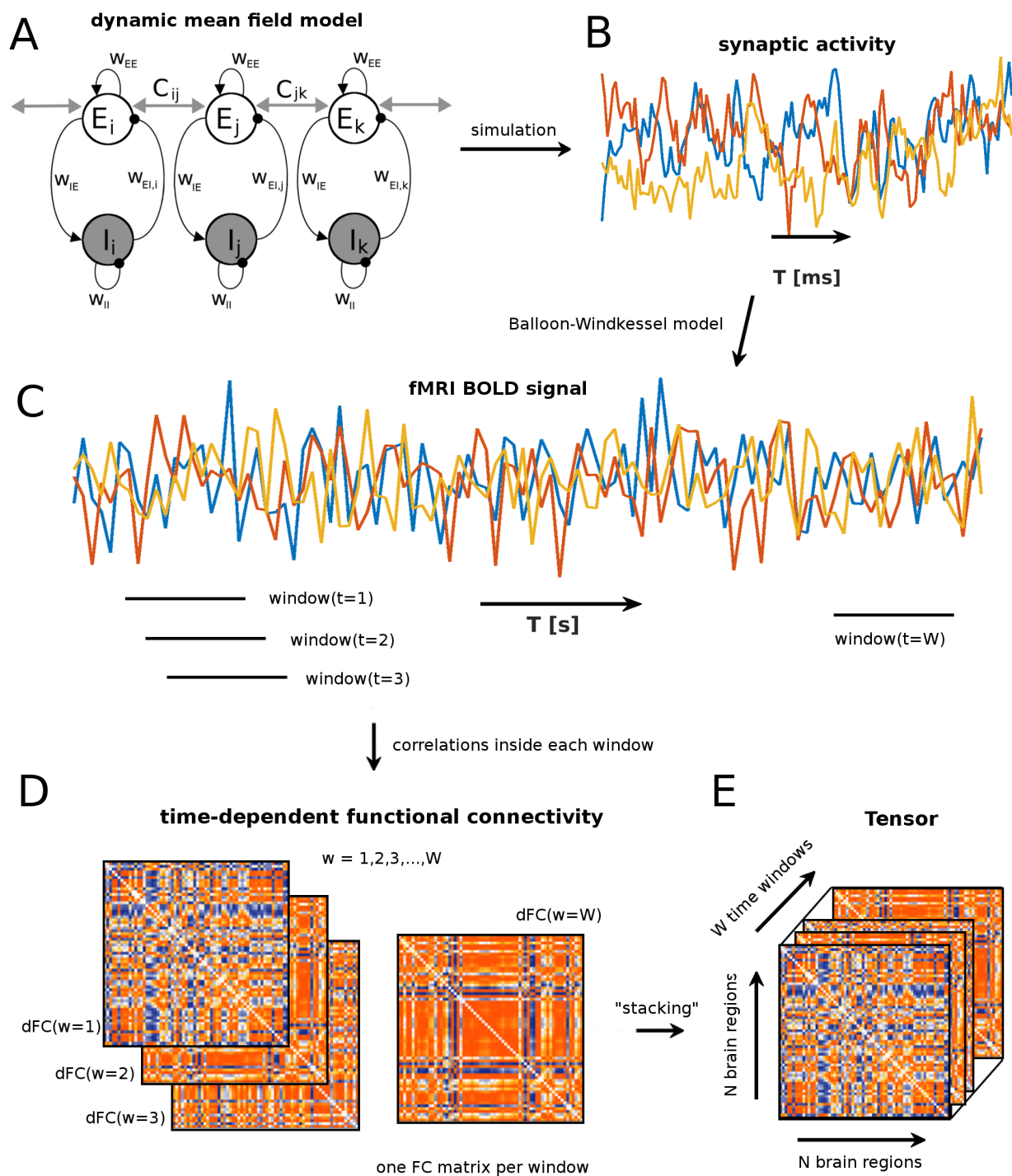
(1)

Constants are listed in table 1; see figure 1A for an illustration. Super-/subscripts E and I denote the excitatory and inhibitory pools of population i, respectively. I_i denotes synaptic input currents, which are turned into population firing rates r_i via sigmoid transfer functions $H(\cdot)$. S_i denotes the average synaptic gating variable, or activity, and $\nu(t)$ is Gaussian noise with amplitude $\sigma = 0.01$. The kinetic parameter $\gamma = 0.641$.

Synaptic currents I_i are a result of of inputs from the local network, i.e. $S_i^{(E)}$ and $S_i^{(I)}$, and inputs from other network nodes j, i.e. $S_j^{(E)}$. Local inputs are governed by four weights, w_{EE}, $w_{\mathrm{EI},i}$, w_{IE}, and w_{II}. Additionally, there are constant inputs to each population, denoted by $I_{0,E}$ and $I_{0,I}$. Inputs from other parts of the network are provided by the excitatory populations and weighted by the entries C_{ij} for the connectivity from region i to region j, noted in the SC or EC. The diagonal of $\mathbf{C}$ is set to 0. Weights are scaled by the global coupling parameter G.

The feedback inhibition, $w_{\mathrm{EI},i}$, is adjusted before the simulation to ensure that the network is in its asynchronous state where we only have one stable fixed point with firing rates between 3 and 10 Hz for all regions (Deco and Jirsa, 2012). We can determine the stability of the system by taking advantage of a semi-analytical solution (Deco et al., 2014). We calculate the Jacobian matrix and confirm that all eigenvalues are negative with zero imaginary parts. Simulations were only performed for values of G that warranted stability of the system.

Number and length of simulations are matched to the empirical data. BOLD time courses are obtained from the synaptic activities of the excitatory pools via the Balloon-Windkessel model (Friston et al., 2003; Deco et al., 2013). Time courses are downsampled to match the TR of the experimental data (figure 1B,C).

Figure 1: (see next page) Overview of analysis pipeline that leads to tensorized data. For simulated data, panels A to E apply, for empirical data, panels C to E. **A** Schematic view of dynamic mean field (DMF) model used to simulate synaptic activity. Each brain area is modelled by a pair of excitatory (E) and inhibitory (I) pools. The local connectivity is governed by the four weights w_{EE}, $w_{\mathrm{EI},i}$, w_{IE}, and w_{II}, whereby $w_{\mathrm{EI},i}$ is adjusted for each population individually so as to keep the firing rates at a low level (3-10 Hz). Black lines with spheres signify GABA connections, black arrows, NMDA. Gray arrows are long range connections mediated by AMPA synapses and whose weights are set by the SC or EC matrix. **B** Synaptic activity is simulated on a millisecond scale. **C** BOLD activity on the scale of seconds is obtained from simulated activity by applying the Balloon-Windkessel hemodynamic model. **D** Starting from these BOLD time courses, we compute time dependent dFC(w)-matrices from overlapping time windows of width 120s (60 frames) starting at time points $t = 1, 2, ..., W$. Each dFC(w)-matrix has dimensions $N \times N$, N being the total number of ROIs. **E** This results in an $N \times N \times W$ tensor. The entry at $dFC_{ij}(w)$ is the functional connectivity value of ROIs i and j in window t. dFC(w)-matrices are symmetric.

A — **dynamic mean field model**

B — **synaptic activity**

simulation

T [ms]

Balloon-Windkessel model

C — **fMRI BOLD signal**

window(t=1)

window(t=2)

window(t=3)

window(t=W)

T [s]

correlations inside each window

D — **time-dependent functional connectivity**

w = 1,2,3,...,W

dFC(w=1)

dFC(w=2)

dFC(w=3)

dFC(w=W)

one FC matrix per window

"stacking"

E — **Tensor**

W time windows

N brain regions

N brain regions

2.3 Tensorization of the data

In order to take advantage of the temporal information, we adopt the widely used approach of sliding windows to compute time-dependent dynamic FC (dFC). We use overlapping windows w of width 120s (60 data points, TR=2s) that we advance along the time course in increments of 2s (1 frame), which results in $W = T - 60$ windows for each dataset (subject or simulation), T being the number of frames. For each window, we compute an $N \times N$ matrix of pairwise connectivity values, $dFC(w)$. By concatenating these matrices along the temporal dimension, we represent each dataset as a tensor of dimensions $N \times N \times W$ (see figures 1D and E).

We use two measures of FC: on the one hand the most widely used one, i.e. Pearson correlation, on the other, mutual information (MI), a non-negative and nonlinear measure that allows us to constrain the tensor decomposition. The calculation of MI follows Kraskov et al. (2004) and is based on nearest neighbor distances, thus being adaptive and continuous.

In short, when estimating the MI between the 60 data points inside our window, we determine the distance between any two points $[x_i, y_i]$ and $[x_j, y_j]$ with $i \neq j$ and take the maximum norm, i.e.

$$d_{ij} = \parallel z_i - z_j \parallel = \max\{\parallel x_i - x_j \parallel, \parallel y_i - y_j \parallel\} \tag{2}$$

The nearest neighbor to each point $[x_i, y_i]$ is the point with the minimum d_{ij}, and we term this distance ϵ_i. For each point $[x_i, y_i]$, we count how many points are within this minimum distance ϵ_i, separately for the x- and y-directions, resulting in two numbers $n_x(i)$ and $n_y(i)$. We estimate MI as

$$I(X, Y) = \psi(k) - \langle \psi(n_x + 1) + \psi(n_y + 1) \rangle + \psi(N) \tag{3}$$

where X and Y are the two time series, and $\psi(\cdot)$ denotes the digamma function

$$\psi(x) = \frac{d}{dx} \ln\left(\Gamma(x)\right) \tag{4}$$

$k = 1$ because we consider the nearest neighbor; N is the number of data points, i.e. 60. Since this is a continuous measure, it is possible for $I(\cdot, \cdot)$ to become negative whenever the following inequality is satisfied:

$$\psi(k) + \psi(N) < \langle \psi(n_x + 1) + \psi(n_y + 1) \rangle$$

This happens when there is very little MI and many points are closer than the nearest neighbor. In these cases, we simply set MI to zero.

2.4 Extracting communities from low resolution fMRI data

We apply tensor factorization to both the empirical and the model data (figure 2). The problem for the three-dimensional case treated here can be formulated as follows (Cichocki et al., 2009):

$$\underline{\mathbf{Y}} = \sum_{f=1}^{F} a_f \circ b_f \circ c_f + \underline{\mathbf{E}} = \hat{\underline{\mathbf{Y}}} + \underline{\mathbf{E}} \tag{5}$$

where $\underline{\mathbf{Y}}$ is the data tensor of dimensions $N \times N \times W$, defined as in section 2.3. F is the number of features we wish to extract. $\circ$ denotes the outer product. $\mathbf{A} = [a_1, a_2, ..., a_f]$, $\mathbf{B} = [b_1, b_2, ..., b_f]$ and $\mathbf{C} = [c_1, c_2, ..., c_f]$ are the factor matrices that contain the features of each of the three dimensions, respectively. Here, $\mathbf{A} = \mathbf{B}$ due to the symmetry of the dFC matrices. They contain F vectors of length N with weights that are interpreted as membership values for a community. The third set $\mathbf{C}$ contains their associated time courses and is of dimensions $F \times W$.

$\hat{\underline{\mathbf{Y}}}$ is an approximation of the data based on the features, and $\underline{\mathbf{E}}$ is the error/noise not described by the features. Hence, the distance between $\underline{\mathbf{Y}}$ and $\hat{\underline{\mathbf{Y}}}$ can be used to assess how well the extracted features approximate the data. We use the Frobenius norm in the case of continuous, non-thresholded tensors and the Hamming distance for thresholded, binarized tensors (section 2.5). Note that in the latter case, the result of the decomposition is continuous although the input is binary, so we threshold and binarize the reconstructed tensor such that the number of ones is preserved.

This technique is based on the very general Tucker model which can be viewed as a generalization of SVD to tensors. Unlike for SVD, though, convergence to a unique and optimal solution is not guaranteed. Consequently, it is impossible to determine the true rank of the tensors and thus, the appropriate number of features F. This problem is mitigated by the inclusion of further constraints, in this case, non-negativity when using absolute value of correlation or MI to construct the tensors (section 2.3).

To decompose tensors constructed using correlation, we use the algorithm described in Phan et al. (2013). For the non-negative measures (absolute value of correlation, mutual information), we apply non-negativity of the resulting features as an additional constraint and use the algorithm described in Kim and Park (2012). Both algorithms are implemented in Matlab, requiring the tensor toolbox (Bader et al., 2015; Acar et al., 2011), and available on-line.

2.5 Thresholding

We reduce noise by thresholding and binarizing the tensors. This approach was also chosen in Ponce-Alvarez et al. (2015). It would perhaps be most desirable to use only the dFC values that are significant. However, it is too time demanding to generate the appropriate number of surrogate datasets to achieve the desired significance level to account for the high number of ROI pairs, windows and subjects/simulations. Hence, we simply use the x-th percentile as a threshold θ, where $x = \{0, 75, 80, 90, 91, ..., 99\}$ and, for $x > 0$, set all elements Y_{ijt} of tensor $\hat{\underline{\mathbf{Y}}}$ which are bigger than or equal to that percentile to one and everything else to zero. Note that we do not make any claims about non-stationarity of the FC time courses.
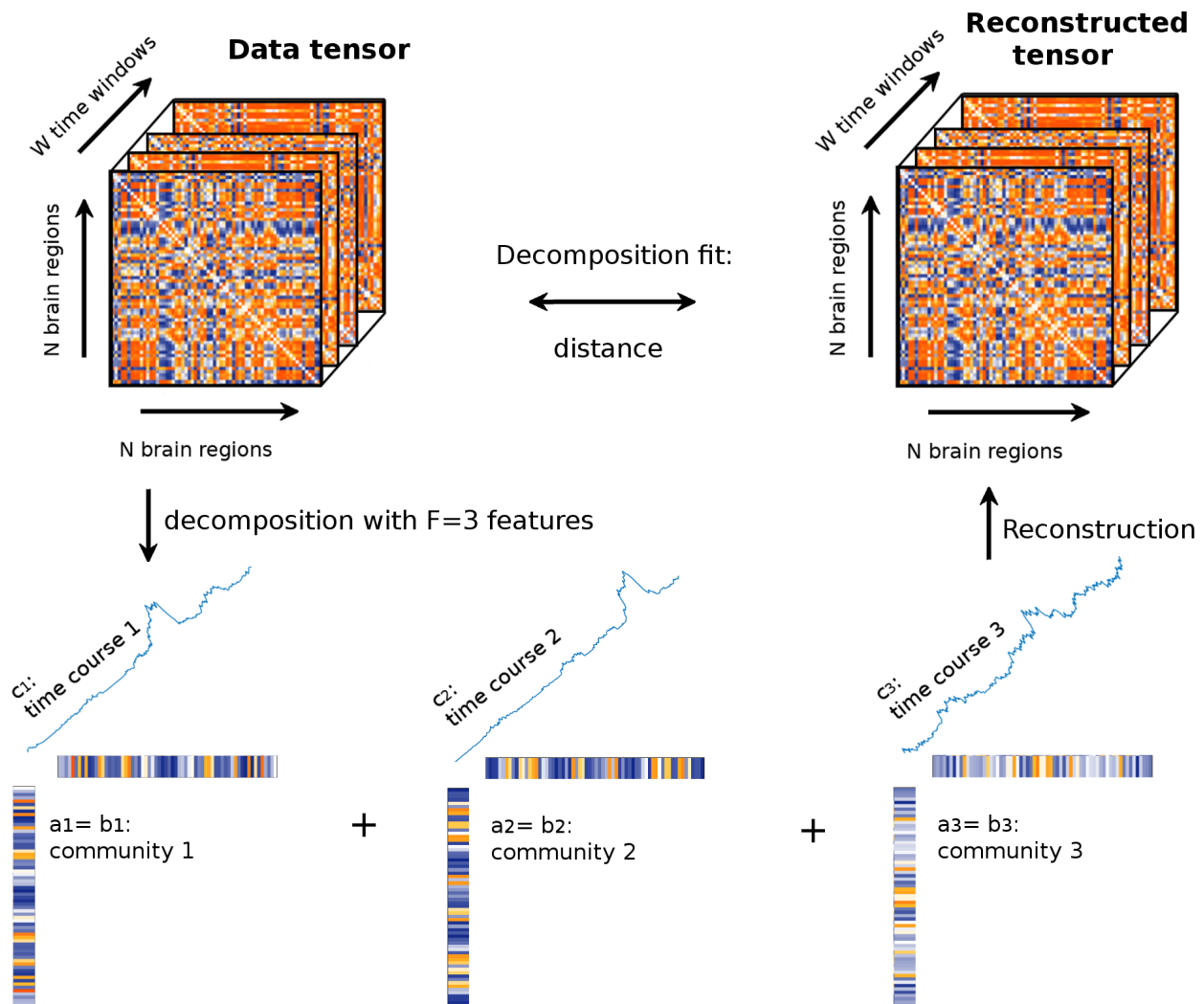
8

Figure 2: Illustration of tensor factorization with $F = 3$ features. The original data tensor (top left) is decomposed into three components, i.e. three communities with their associated time courses are extracted. As shown in the bottom row, each set of vectors, consisting of two identical "community vectors" a_f and b_f and a time course, c_f, constitutes a rank-1 tensor, which can be explicitly computed by taking the outer product as in equation 5 (taking the outer product of a_f and b_f, one would obtain an FC matrix like the "layers" of the tensor). By adding these rank-1 tensors, we obtain an approximation of the data tensor, i.e. a reconstructed tensor (top right), and can evaluate the fit between the two by taking the distance.

2.6 Surrogate data

To further validate our results, we conduct analyses with surrogate data alongside those for real data. Surrogate data are constructed by removing the pairwise correlation structure of the original time series while keeping the Fourier spectrum constant. More specifically, we Fourier transform our original time courses $x_i(t)$ of region i for time points $t = 1, 2, ..., T$ using Fast Fourier Transform and add random phases φ_r to each frequency bin before transforming back.

$$\tilde{x}_i(k) = \sum_{t=1}^{T} x_i(t) e^{-j\frac{2\pi kt}{T}}$$

$$x_i^{\text{surr}}(t) = \frac{1}{T} \sum_{k=1}^{T} |\tilde{x}_i(k)| e^{-j\frac{2\pi kt}{T} + \varphi_r}$$

(6)

$k = 1, 2, ...T$ and φ_r is uniformly distributed between $-\pi$ and π. This results in time courses that have the same spectral properties and autocorrelations as the original data, but are uncorrelated.

2.7 Constructing templates from empirical data

We construct spatial templates directly from the empirical data. First, we decompose each subject's data tensor as described above, using a range of numbers F of features. The true value of F is impossible to determine due to the aforementioned problems with uniqueness and optimality of the solution. In our case, the goal is to extract spatial maps that are common across subjects, so we apply a simple clustering algorithm (K-means clustering) to the set of all $F \cdot S$ features (i.e. pooled from all subjects) with different numbers K of clusters. We test the quality of the clustering by evaluating the mean silhouette value (de Amorim and Hennig, 2015). The resulting cluster centers of the parameter combination with the maximum silhouette value are used as templates.

For each point i (here, an N-dimensional feature) the average distance (in terms of correlation) to points assigned to the same cluster is evaluated, denoted by a_i, as well as the smallest average distance to all points assigned to a different cluster (i.e., the closest cluster), denoted by b_i. Then the silhouette value is calculated as

$$s(i) = \frac{b_i - a_i}{max\{a_i, b_i\}}$$

(7)

Obviously, $s(i)$ will lie between -1 if the point is entirely in the wrong cluster, and 1 if the assignment is perfect. Taking the mean over all features gives an estimate of how well the data points are clustered.

2.8 Calculating the overlap between templates and features

In order to determine how well, overall, communities extracted from simulated data match those from empirical data, we compute the average overlap between all simulated spatial features with any of the templates, comparing connectivities (EC vs. SC) and using the full range of global coupling parameter values (G), i.e. $G = 0.5, 0.6, ...4$ for simulations with the SC and $G = 1, 1.1, ...6$ for simulations with the EC.

Due to the high thresholds, the features are somewhat sparse. Therefore, correlation is not a good choice to measure the distance between features and templates. Instead, we use confusion matrices and Cohen's kappa. Briefly, we quantize the features on three levels and compute the overlap between any two vectors (one template, one feature extracted from simulated data) by determining the overlap for each level, creating a confusion matrix. Cohen's kappa is a summary of the confusion matrix:

$$\kappa = \frac{P_a - P_e}{1 - P_e} \tag{8}$$

P_a is the overlap, P_e is the expected overlap.

As the overall match between a feature extracted from simulated data and the templates of the empirical data, we consider the mean over the maximum κ values for each feature in each simulation. In other words, we assume that each simulated feature corresponds to only one template. We compute an overall match for each value of G by averaging this value over the features and simulations.

2.9 Effective connectivity

In this study, we compare simulation results from two different underlying connectivities. On the one hand, we have dwMRI-derived SC containing estimates of fiber densities from the same 24 subjects whose BOLD signals are analyzed, on the other, model-based EC. In the following, be briefly describe the method developed by Gilson et al. (2016) for constructing EC matrices by combining SC and FC. The key is to extract information about cortical interactions from BOLD covariances with non-zero time shifts:

$$Q_{ij}^{\tau} := \langle (x_i^t - \bar{x}_i)(x_j^{t+\tau} - \bar{x}_j) \rangle \tag{9}$$

Q_{ij}^{τ} is the covariance between BOLD time courses x of ROIs i and j, with x_j shifted by τ against x_i. The angular brackets denote averaging over randomness due to noise in the model, such that the mean BOLD for node i is $\bar{x}_i = \langle x_i^t \rangle$. For $\tau \neq 0$, this matrix is non-symmetric. The goal is to estimate the underlying connectivity such that the model minimizes the error between model covariances (Q^0 and Q^{τ}) and their empirical counterparts ($\hat{Q}^0$ and $\hat{Q}^{\tau}$), for a given τ equal to 1 TR.

EC is model based, meaning that there is an underlying assumption of how activity propagates through the brain using the present connections to activate the nodes. We use a noise diffusion model with a static nonlinearity, Φ:

$$\mathrm{d}x_i^t = \left[-\frac{x_i^t}{\tau_x} + \Phi\left(\sum_{k \neq i} C_{ik} x_k^t + e \right) \right] \mathrm{d}t + \mathrm{d}B_i^t \tag{10}$$

Time course x_i of region i is subject to an exponential decay with time constant τ_x at each time point t. C is the connectivity matrix that contains weights linking regions i and j, and the sum is over all regions k from which i receives input. This means that activation is only provided by the input from other nodes, weights for which are defined in C. The background input e is shared by

11

all i. Fluctuations of the activities are driven by Gaussian noise dB_i^t. The model directly simulates BOLD activity, hence the Jacobian of the system is simply:

$$J_{ij} = -\frac{\delta_{ij}}{\tau_x} + C_{ij} \, \Phi'\Big(\sum_{k \neq i} C_{ik}\bar{x}_k + e\Big), \tag{11}$$

where δ_{ij} is the Kronecker delta and Phi' denotes the first time derivative of Φ. Therefore, the model is solely constrained by C.

We want to estimate J and therefore C such that it satisfies the steady state of the second order fluctuations:

$$\begin{aligned} JQ^0 + Q^0 J^\dagger + \Sigma &= 0 \\ Q^\tau &= Q^0 \exp(J^\dagger \tau) \end{aligned} \tag{12}$$

Σ is the noise matrix with diagonal terms $(\sigma_i)^2 = \langle dB_t^t \, dB_j^t \rangle$; $\dagger$ denotes the matrix transpose and **exp** the matrix exponential. Since we are using empirical covariances estimated from fMRI data, the objective covariances $\hat{Q}^0$ and $\hat{Q}^\tau$ are very noisy and make a direct estimation via an analytical approach unfeasible. Therefore, we use the iterative Lyapunov optimization (LO) procedure described in Gilson et al. (2016).

The update works by simulating the BOLD activity using the model defined in equation 10 without noise and the current connectivity C, so as to evaluate the mean activity $\bar{x}_i$ for all regions, yielding the Jacobian J in equation 11. Then, the model covariances Q^0 and Q^τ are given by equation 12, using the Bartels-Stewart algorithm for the first line and using the current values for J and Σ. The model covariance matrices are compared to the objective covariances $\hat{Q}^0$ and $\hat{Q}^\tau$. Then, the Jacobian update is evaluated according to:

$$\Delta J = \frac{1}{\tau}[(Q^0)^{-1}(\Delta Q^0 + \Delta Q^\tau \exp(-J^\dagger \tau))]^\dagger \tag{13}$$

Finally, we obtain the connectivity update $\Delta C_{ij} = \Delta J_{ij}/\Phi(\sum_k C_{ik}\bar{x}_k + e)$.

We apply a mask when updating C, thus only tuning connections that are present in the SC as well and are above a certain threshold. The only exception are the elements on the secondary diagonal, which are added regardless of whether they are present in the SC or not, in order to account for homotopic connections.

In addition, the noise matrix Σ (see equation 12) is optimized at the same time as C. We assume that each node receives independent noise, meaning that Σ is diagonal. The update is performed such that the model variances coincide with the empirical values:

$$\Delta \Sigma_{ii} = \epsilon_\Sigma(\hat{Q}_{ii}^0 - Q_{ii}^0) \tag{14}$$

It was shown that a time shift τ equal to 1 or 2 TR gives a good estimation performance (Gilson et al., 2016). In fact, τ has to roughly match the time scale on which the neural activity decays, i.e. τ_x in equation 11. The latter is estimated from the slope of the autocorrelation of each region (the slope is close to $1/\tau_x$), and results in $\tau = 5.3s$ which leads us to set the time shift to 1 TR=2s.

2.10 Analysis pipeline

To summarize the methodology of the paper, we give a brief overview of the steps involved in obtaining the results described in the next section. Illustrations are shown in figures 1 and 2. We have $S = 24$ sets of empirical resting state data (from 24 subjects) of length $T = 661$ frames (TR=2s). Voxel time courses are averaged inside of $N = 66$ ROIs that cover the entire cortex. To match this number, we simulate 24 times for each one of evenly spaced values of the global coupling parameter G from a suitable range, using the DMF model. We obtain 2 sets of simulations, one using the SC matrix and one using the EC matrix to set the underlying connectivity between the 66 regions. Each dataset is tensorized by computing dFC matrices inside of rectangular sliding windows w of width 2 minutes (60 frames) which are moved by one frame, resulting in W $dFC(w)$ matrices for each subject/simulation. These matrices are concatenated into tensors of dimensions $N \times N \times W$. Entries of $dFC(w)$ are calculated using correlation (positive and negative values), the absolute value of correlation (non-negative values) and mutual information (non-negative values). Each tensor is then decomposed into sets of spatial (communities) and temporal (time courses) features.

In a first step, in order to obtain community templates, we decompose the empirical tensors using different thresholds θ, binarizing the tensors in the case of non-zero thresholds. Since the rank of the tensors is unknown, we use different numbers F of features, where F ranges from 3 to 9. Hence, we obtain $F \cdot S$ spatial (and temporal) features for each F. We do not expect the temporal features to have anything in common except some general dynamic properties, so we continue with only the spatial features and cluster them, calculating the silhouette value as a quality measure for each instance of clustering. We choose the combination of F, K and θ that yields the most well-defined clusters, i.e. the highest silhouette value. The means of those clusters are the templates (i.e., K is the number of templates). In a second step we extract features from simulated data ($S = 24$ runs, different values of G) using the same F and θ, and compare each feature to the templates. We calculate an overall match for each value of G by taking the mean over the maximum match of each feature with any of the templates.

3 Results

3.1 A method for extracting RSNs from single subject data

Our general goal is to understand the spatio-temporal dynamics of human resting state (RS) fMRI, using time-dependent, or dynamic, FC (Chang and Glover, 2010; Hutchison et al., 2012; Allen et al., 2012; Liu et al., 2013). To this end, we combine data analysis of long-range connectivity and and a whole-brain modelling approach (Deco et al., 2014) to investigate whether the dynamics of the model can reproduce the empirical data. We apply tensor decomposition (Cichocki et al., 2014), a method that allows us to simultaneously consider spatial and temporal structure of FC. We compute pair-wise dynamic FC (dFC) using mutual information (MI) (Kraskov et al., 2004) inside of overlapping sliding time windows (Chang and Glover, 2010; Kiviniemi et al., 2011; Hutchison et al., 2012; Allen et al., 2012; Leonardi and Van de Ville, 2013), obtaining a dFC matrix, $dFC(w)$, for each window w. These matrices are then concatenated along the temporal dimension into a 3-way-tensor (figure 1C-E). Additionally, we apply a binarization threshold in order to reduce noise. This is done for both empirical as well as simulated data.

Tensor factorization allows us to decompose dFC of a single subject (or simulation run) into spatial patterns, so-called "communities", and associated time courses. The spatial patterns are expected
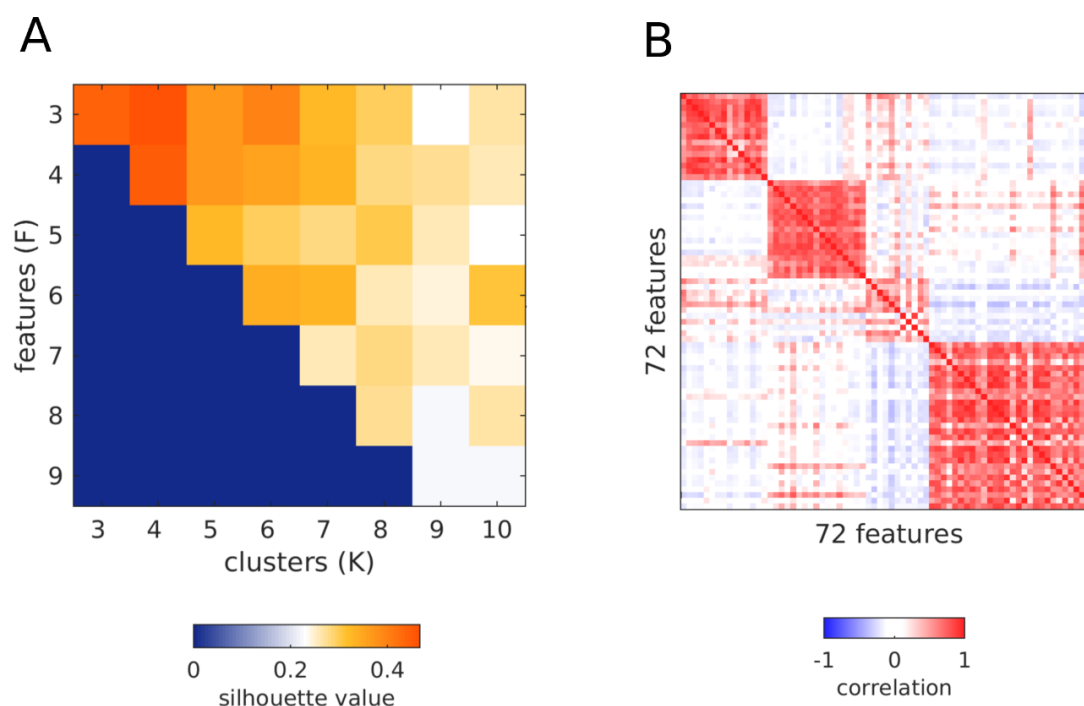
Figure 3: Result of template extraction with optimal parameter settings. **A** Silhouette values for clustering all features extracted from MI tensors for the best threshold, i.e. 98th percentile. **B** Correlation between the same features, ordered by assignment to $K = 4$ clusters. The clusters are in the same order as the templates in figure 4.

to be common among subjects, while the temporal evolution of each community is specific to a subject. We extract $F = 3, 4, ...9$ communities/time courses (features) (Beckmann and Smith, 2004; Mantini et al., 2007) from each of $S = 24$ subjects which gives a pool of $F \cdot S$ communities. These are then grouped into $K = F + 1, F + 2, ...10$ clusters using K-means clustering in order to find common patterns. Goodness of clustering is assessed using the silhouette value (section 2.7); a high value indicates that the clusters represent the data well, i.e. cluster centers can be seen as "prototype communities", or templates, that can be used on the whole group of subjects. In order to validate our results, we compare them against surrogate data and consider the difference in clustering performance. The surrogate data do no exhibit any long-term correlations, so any clustering structure is due to their Fourier spectrum, autocorrelations, and due to the method itself. We find that $F = 3$ and $K = 4$ results in the maximum value of 0.54. Phase randomized surrogate data (section 2.6), on the other hand, only reach 0.08. This indicates that, while there are clearly still a lot of inter-individual differences, assuming a cluster structure is supported by the data.

Figure 3A shows the difference between silhouette values for real and surrogate data for all combinations of F and K. Panel B visualizes the clusters by showing the correlations between communities. The clusters are quite well separated and correspond to four common patterns that are similar to previously described resting state networks, namely the default mode network, somatomotor network, right and left control networks, and visual network (figure 4).
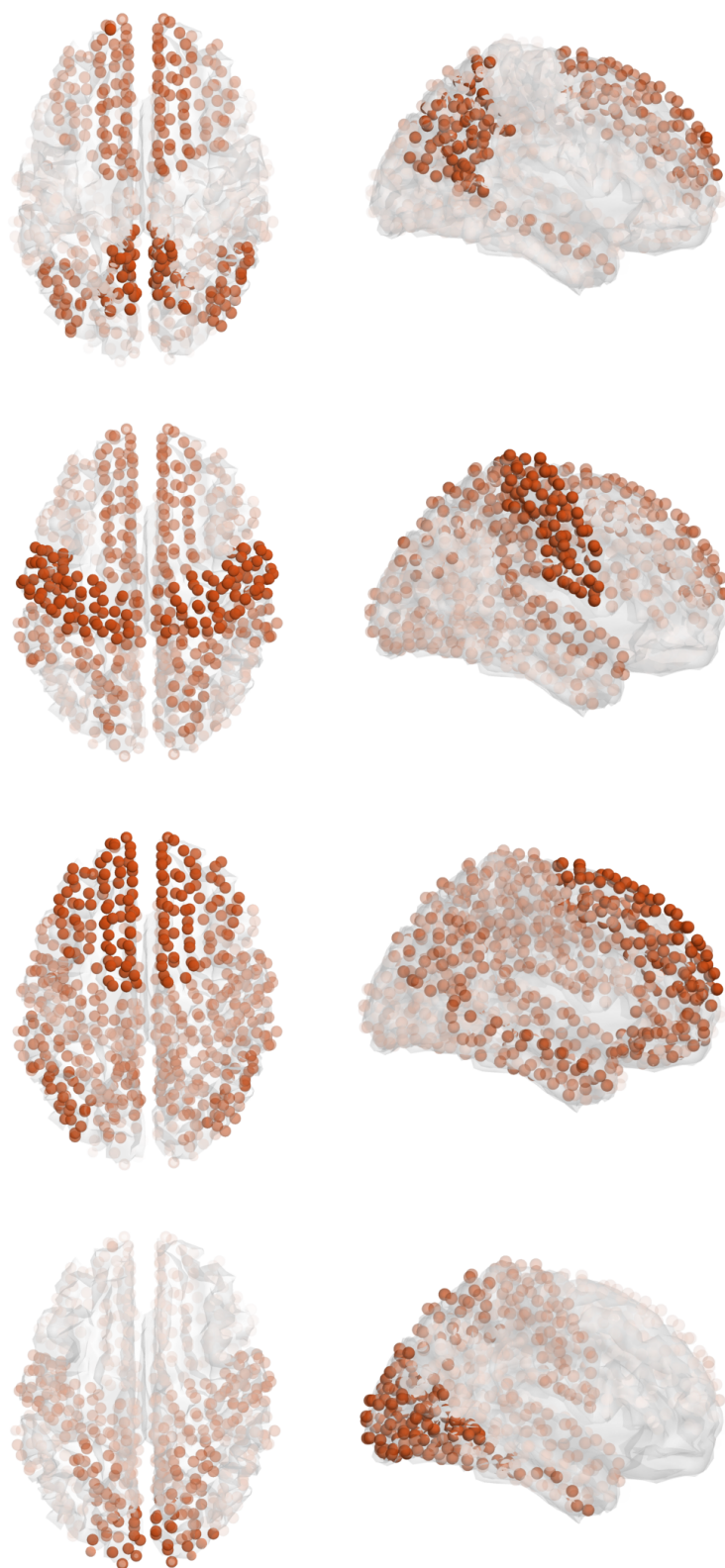
14

Figure 4: Templates plotted in 3D. From top to bottom: default mode network, somatomotor network, control networks, visual network. Although the templates are obtained at a resolution of 66 cortical regions, for better visualization we use 998 centers of mass, each of which is clearly assigned to one of the 66 areas. The opacity of each sphere is proportional to the weight assigned to it in the spatial map of the template.

3.2 Mutual information and a high threshold produce communities that generalize well

We use MI to compute dFC (Kraskov et al., 2004) because we find that MI is a better choice than correlation, based on clustering performance as well as reconstruction fits. Additionally, we apply a binarization threshold to the MI values, keeping only the highest ones, in order to reduce noise, which is in the following demonstrated to be necessary in order to obtain the desired generalized communities.

We start out without applying a binarization threshold, computing pairwise dFC values with Pearson correlation. Decomposing these tensors and clustering the resulting communities, we obtain the best silhouette value at $F = K = 3$ with a value of 0.31 for real data, and 0.19 for surrogates. This means that most of the cluster structure is due to spectral properties and autocorrelation of the data rather than real dFC. In comparison, when using MI without a threshold, we obtain the best value at $F = K = 3$ with 0.45 for real and 0.10 for surrogate data.

One advantage of using MI is that it allows us to apply an additional constraint when decomposing the data, i.e. non-negativity. In order to exclude that the poor result obtained with correlation is due to this, we repeat the procedure replacing the negative correlations with their absolute values. Again, $F = K = 3$, and we obtain 0.44 for real data, i.e. just as high as for MI, but again, 0.19 for surrogate data.

Taken together, clustering can be improved by applying the non-negativity constraint for both correlation and MI, but when using correlation, also the surrogate data exhibit more of a cluster structure, which makes MI the better choice in this application.

Another way to improve clustering is to reduce noise and thereby mitigate the effect of inter-individual differences. In principle, it would be best to apply significance testing and keep only the significant dFC pairs/windows, however due to the large number of windows and pairs, this is computationally not feasible since we would have to generate thousands of surrogate tensors to reach a satisfying significance level, which would prove prohibitive in terms of computation time and storage space. Therefore, we just apply different thresholds and decompose the resulting binary tensors. We use different percentiles as thresholds, $\theta = 0, 75, 80, 90, 91, 92, ...99$, and compare the results, using the silhouette value.

For correlation, the absolute value of correlation, and MI, the thresholds are determined to be the 98th, the 97th and the 98th percentile, respectively. For all measures, $F = 3$ and $K = 4$. Correlation reaches a maximum mean silhouette value of 0.54 with a corresponding surrogate value of 0.13. The two non-negative measures produce better and equally good results: using the absolute value of correlation, we obtain silhouette values of 0.59 and 0.13 for real and surrogate data, respectively; for MI, the values are 0.54 and 0.08 (these templates are the ones shown in figure 4).

Apart from the clustering performance, we consider the reconstruction fits that quantify how well the extracted features describe the original tensor. At their best thresholds, MI and absolute value of correlation reach average fits of 0.39 and 0.48, respectively. The corresponding surrogate tensors yielded 0.09 and 0.34, confirming that while using the absolute value of correlation results in a good decomposition performance, the surrogate tensors constructed in this way also exhibit a lot of structure. Taken together, MI is the better choice.

3.3 Dynamic mean field model reproduces RSNs

In the next step, we use the templates shown in figure 4 to determine whether the dynamic mean field (DMF) model can produce data that exhibit spatio-temporal patterns similar to those found in

the empirical data. We run $S = 24$ simulations of the same length as the empirical data and use the parameters previously determined with the empirical data to the resulting tensors, i.e. $F = 3$ and $\theta = 98$. The only free parameter is the global coupling, G (figure 1A) which is a factor by which the connectivity matrix is multiplied and is hence related to the overall amount of activation in the system. We determine for each of the $F \cdot S = 72$ extracted features the maximum correspondence to any of the $K = 4$ empirical templates and use the mean across all features and all runs as a measure for the match between features (simulated) and templates (empirical) and thus of how well the model reproduces the empirical data. As before, we consider the difference to matches obtained using surrogates, i.e. of features extracted from tensors that are calculated from phase randomized simulated data to empirical templates.

We use the parameters obtained from the empirical data instead of running the same parameter selection procedure on the simulated data. This is because the simulations run on an average connectivity matrix and therefore, clustering across simulation runs does not make sense; the variance in FC structure is small across runs for a given value of G.

Figure 5A shows the overall overlap between simulated features and templates depending on the global coupling parameter. The 95% confidence intervals of real and surrogate data overlap for most values of G, except for $G = 1.9, 2.5 - 2.7, 2.9$, and $3.1 - 3.7$, the latter values matching the region in which the fit of the average FC is best. For these values, the simulated data can be shown to contain FC patterns that match empirical data to a higher degree than surrogate features. However, even the best match at $G = 3.7$ is moderate. Figure 6 shows the simulated features next to the empirical template it best agrees with. We display the communities in vector form, and the ROIs are ordered as indicated in table 2, i.e. symmetrically. Translated to the communities on the cortex, it is obvious that the main problem is that the features are not symmetrical, but rather, lateralized. For example, all features that match with the somato-motor template have their members in the right hemisphere. Surprisingly, the visual network, which has proven to be the most clearly pronounced one in the empirical data, is not found at all.

3.4 Effective connectivity is crucial for modelling realistic communities

SC is derived by applying fiber-tracking algorithms to diffusion tensors obtained via dwMRI. We use the method developed by Gilson et al. (2016), briefly described in section 2.9, to obtain from this SC an effective connectivity (EC) for our dataset. This EC contains meaningful weights as well as directionality information. Importantly, only the weights that are also present in the SC are tuned by the procedure, plus the weights on the secondary diagonal to account for homotopic connections that are not represented well by fiber tracking.

Figure 5B shows the overall overlap between simulated features and templates when EC is used (note that the scale on the y-axis is different). The non-overlapping region of the 95% confidence intervals of real and surrogate data covers a wider range of G (2.9 to 6.0), the maximum being at $G = 4.2$. Figure 7 shows the features extracted at this point next to the templates in vector form. Clearly, they are now symmetrical and have much more in common with the templates although the overlap in figure 5B may still seem somewhat modest.

We explain this better match by considering panels B and C of figure 8. Several differences between SC and EC are immediately obvious, although we first note that both of them differ greatly from the FC shown in panel A. First of all, homotopic connections that are missing in SC are prominent in EC. This can explain the symmetry in the features obtained from EC-based simulations. Furthermore, the weights are more uniformly distributed in EC, making it appear more dense. Panel D of figure 8 reveals that the node degrees are largely equalized, such that most of them range in the middle of
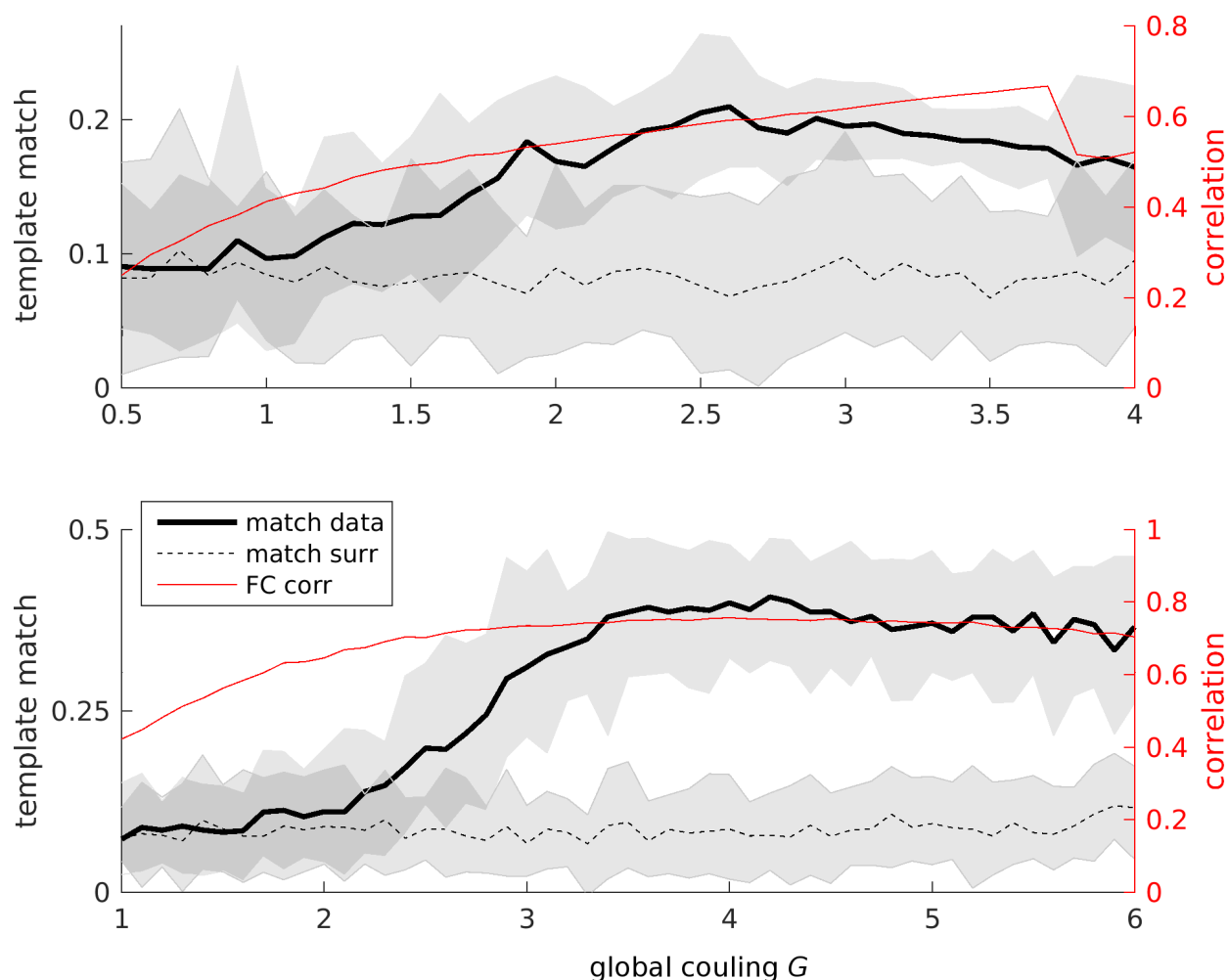
Figure 5: Overall template matches in comparison to correlation fits between FC matrices, results for simulations using SC (top) and EC (bottom). Black curves: template matches (left axes), thick black curve is real data, dashed line surrogate data. Shaded areas indicate the 95% confidence interval. Red curves: correlation of average FC matrices (right axes).
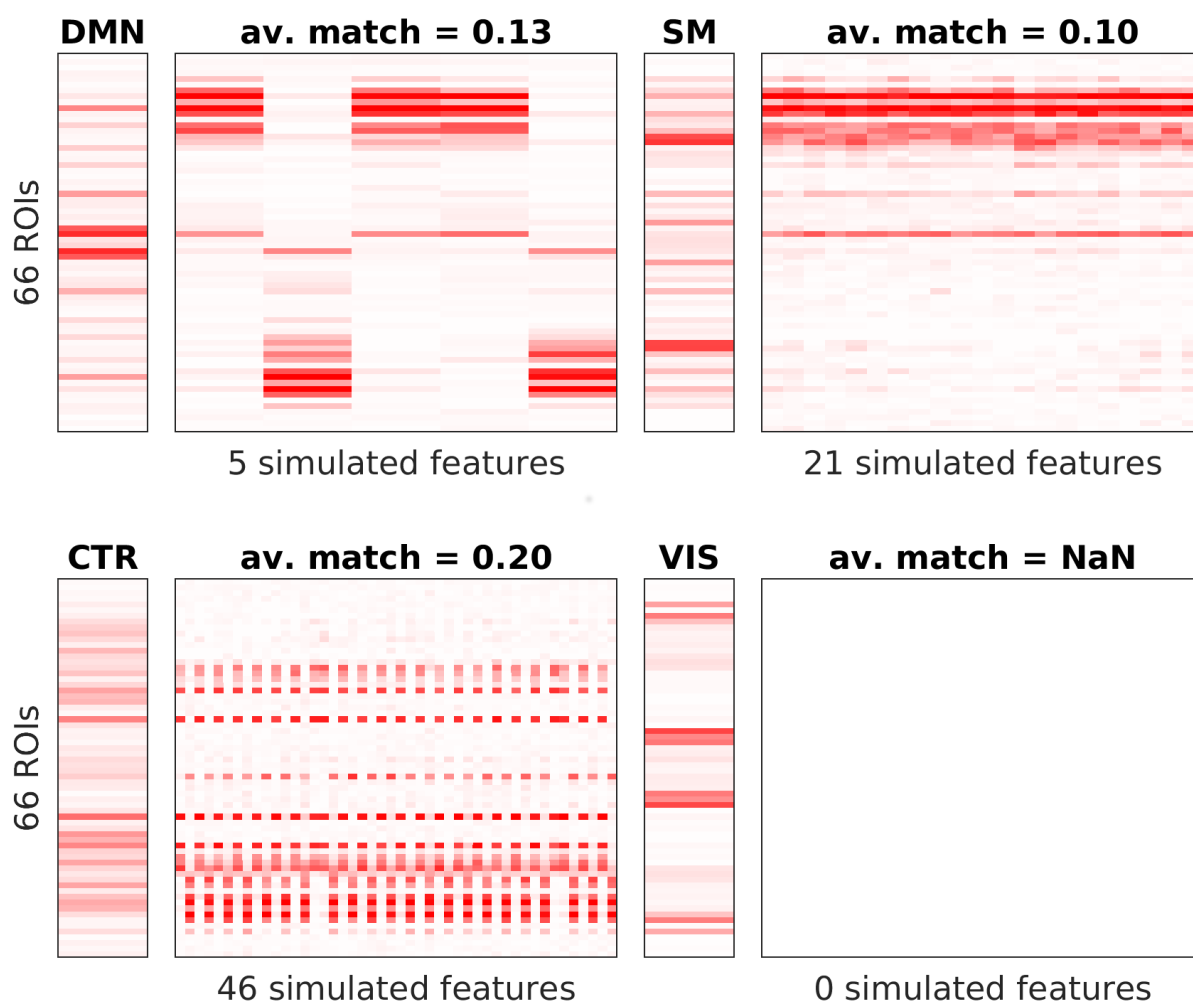
Figure 6: Spatial features extracted from 24 simulation runs using SC as underlying connectivity and the G with the best match (figure 5: $G = 3.7$) with $\theta = 98$th percentile and $F = 3$, plotted next to the templates with which they exhibit the best match. Templates correspond to DMN - default mode network, SM - sensorimotor network, CTR - left and right control networks, VIS - visual network. For each set of features, the number of matched features and the average overlap for only this set is given.

Table 2: ROIs in the same order in which they appear in the vectors in figures 6 and 7, as well as in the matrices in figure 8.

	name		name
1	R entorhinal cortex	34	L posterior-cing cortex
2	R parahippocampal gyrus	35	L precuneus cortex
3	R temp pole	36	L isthmuscing cortex
4	R front pole	37	L paracentral lobule
5	R fusiform gyrus	38	L cuneus cortex
6	R transverse temp cortex	39	L pericalcarine cortex
7	R lateral occipital cortex	40	L lingual gyrus
8	R sup parietal cortex	41	L medial orbital front cortex
9	R inf temp gyrus	42	L sup front gyrus
10	R inf parietal cortex	43	L rostral ant cing cortex
11	R supramarginal gyrus	44	L caudal ant-cing cortex
12	R banks sup temp sulcus	45	L lateral orbital front cortex
13	R middle temp gyrus	46	L pars orbitalis
14	R sup temp gyrus	47	L rostral middle front gyrus
15	R postcentral gyrus	48	L pars triangularis
16	R precentral gyrus	49	L pars opercularis
17	R caudal middle front gyrus	50	L caudal middle front gyrus
18	R pars opercularis	51	L precentral gyrus
19	R pars triangularis	52	L postcentral gyrus
20	R rostral middle front gyrus	53	L sup temp gyrus
21	R pars orbitalis	54	L middle temp gyrus
22	R lateral orbital front cortex	55	L banks sup temp sulcus
23	R caudal ant-cing cortex	56	L supramarginal gyrus
24	R rostral ant cing cortex	57	L inf parietal cortex
25	R sup front gyrus	58	L inf temp gyrus
26	R medial orbital front cortex	59	L sup parietal cortex
27	R lingual gyrus	60	L lateral occipital cortex
28	R pericalcarine cortex	61	L transverse temp cortex
29	R cuneus cortex	62	L fusiform gyrus
30	R paracentral lobule	63	L front pole
31	R isthmuscing cortex	64	L temp pole
32	R precuneus cortex	65	L parahippocampal gyrus
33	R posterior-cing cortex	66	L entorhinal cortex
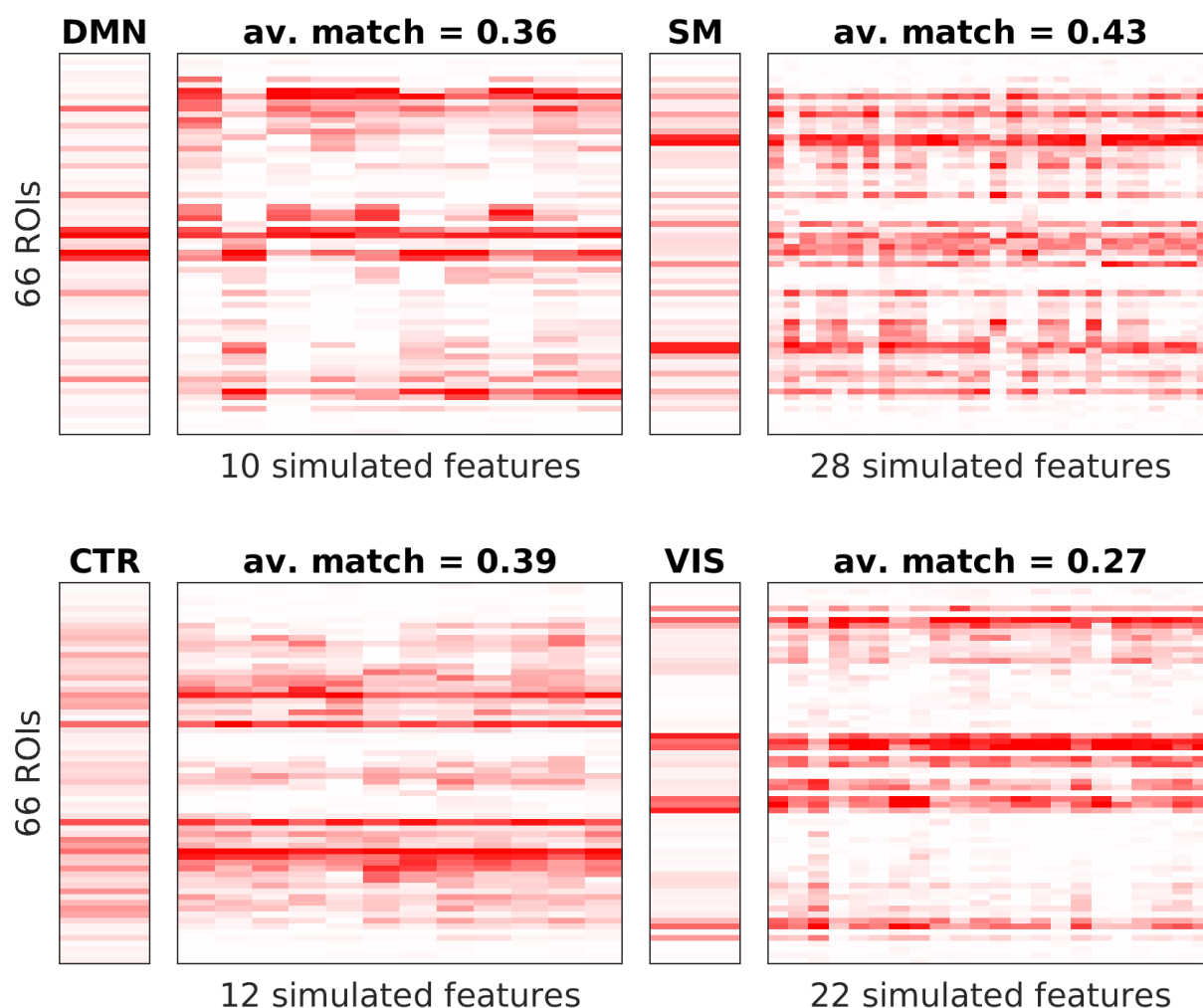
Figure 7: Spatial features extracted from 24 simulation runs using EC as underlying connectivity and the G with the best match (figure 5: $G = 4.2$) with $\theta = 98$th percentile and $F = 3$, plotted next to the templates with which they exhibit the highest match. Templates correspond to DMN - default mode network, SM - sensorimotor network, CTR - left and right control networks, VIS - visual network. For each set of features, the number of matched features and the average overlap for only this set is given.

21
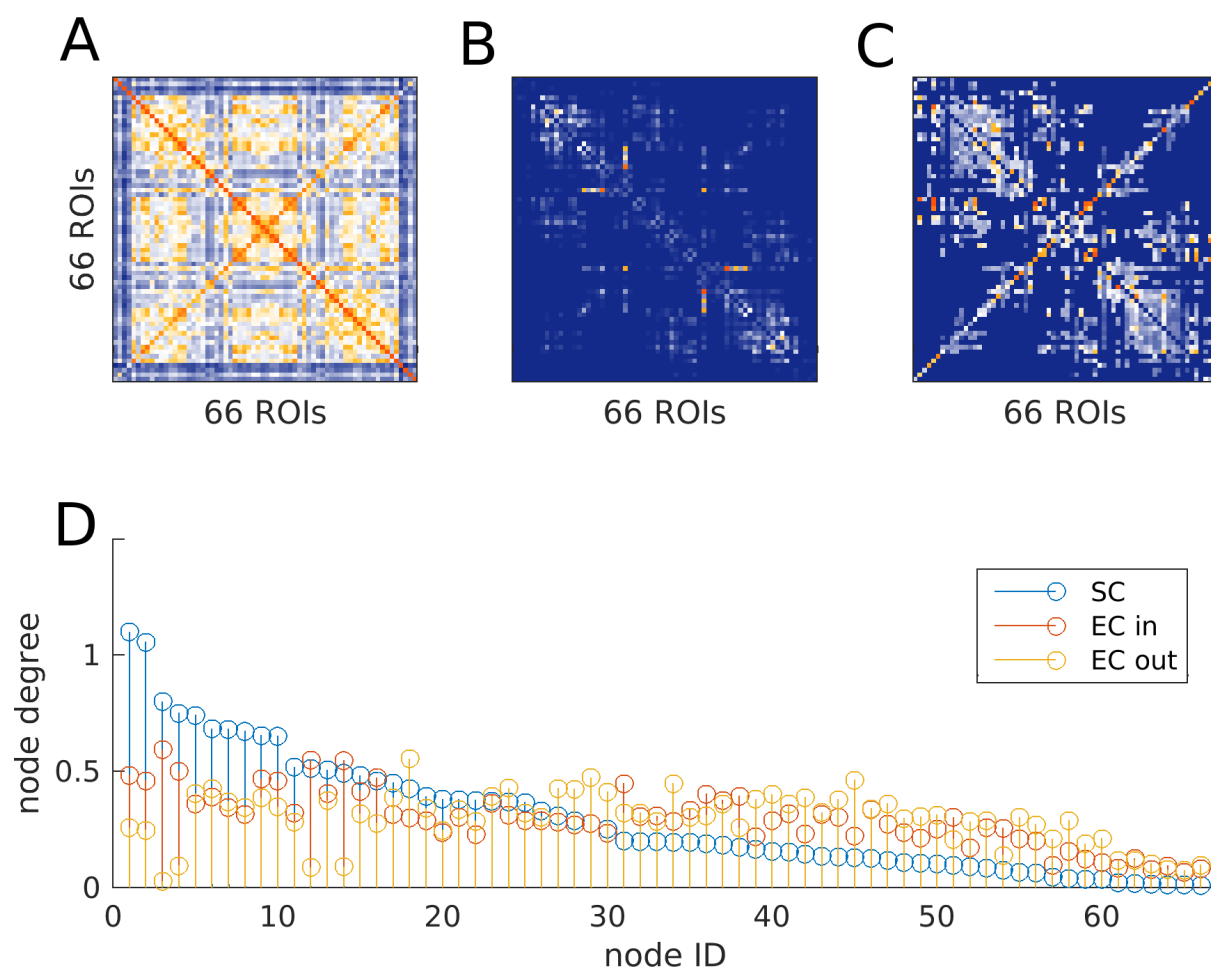
Figure 8: Connectivity matrices. **A** Functional connectivity matrix averaged over 24 subjects. Each entry is the correlation over the entire time course of a pair of ROIs. **B** Average structural connectivity matrix from the same 24 subjects. **C** Effective connectivity matrix derived from the FC, using the method in section 2.9. **D** Degrees of the nodes. For SC (blue), in and out degree are identical. Here, nodes are ordered according to their degree. EC nodes (in-degree: orange, out-degree: yellow) are shown in the same order as for SC.

the distribution. This explains the different ranges of G that can be used in the simulations. For SC, the nodes with the largest degree cause the firing rates in the corresponding excitatory pools to rise to the point where the inhibitory pools cannot compensate for them any longer, and the asynchronous, low-activity regime of the system becomes unstable. The more uniform node degree distribution in the EC matrix allows for higher values of G and thus, the communities become more pronounced due to improved signal to noise-ratio. The EC matrix is also non-symmetric. This property likely further contributes to the stability of the simulations and more generally to the more realistic shape of extracted communities, both on the spatial and temporal level because it allows for a more diverse propagation of the activity through the entire network.

4 Discussion

Our goal was to characterize spatio-temporal features of human resting state (RS) fMRI and to quantify to what extent a noise-driven stationary mean field model (Wong and Wang, 2006; Deco et al., 2014) can reproduce them. Using tensor decomposition (Cichocki et al., 2009), we identify four communities that generalize across subjects and resemble known RSNs (Fox et al., 2005; Beckmann et al., 2005; Mantini et al., 2007). We utilize temporal information by computing pair-wise dynamic FC (dFC) in sliding windows of 2 minutes to build our tensors. We compare three dFC measures: correlation, absolute value of correlation, and mutual information (MI). We determine the dFC measure, number of extracted features F, number of templates K, and binarization threshold θ, that yield the best clustering performance, and take cluster centers as templates. We find that using a low number of features ($F = 3$) and of clusters ($K = 4$) combined with a high threshold ($\theta = 98$th percentile) applied to dFC calculated from MI works best (figures 3, 4).
We determine the range of global coupling for which the match of communities extracted from the model data to the templates is maximal (figure 5). We compare two underlying connectivities of the model: dwMRI derived structural connectivity (SC) and model-based effective connectivity (EC), which estimates directionality and weights of connections using time-shifted covariances(Gilson et al., 2016). EC produces more realistic FC patterns than SC alone.

4.1 Few data points are sufficient to recover RSNs

We find that applying a threshold and binarizing the tensors much improves the clustering performance. For MI, with $\theta = 0$, the maximum silhouette value is 0.45 for real data and 0.10 for surrogates, finding only three templates that do not match known functional networks. For the best threshold of $\theta = 98$, we find a silhouette value of 0.54 (surrogates: 0.09) and we succeed in extracting FC patterns that resemble RSNs (figures 3, 4). $\theta = 98$ translates to using only the 2% biggest dFC values in the decomposition.
Tagliazucchi et al. (2012) transformed RS fMRI time courses into a point process, reducing the data by 94% and keeping only extreme events. Even so, the authors were able to recover RSNs, suggesting that "avalanching events which involve short and long range cortical co-activations" explained their results. Here, we find that only the very largest dFC values are necessary to recover RSNs, suggesting that periods of highly structured FC are equivalent to periods of high variance in the BOLD signal. Indeed, figure 9 shows a strongly fluctuating time course of information content (as measured by supra-threshold MI pairs). We conclude that the peaks of these fluctuations represent periods of high modularity which are detected by the factorization algorithm. Modularity has been shown to dynamically fluctuate by Betzel et al. (2016).
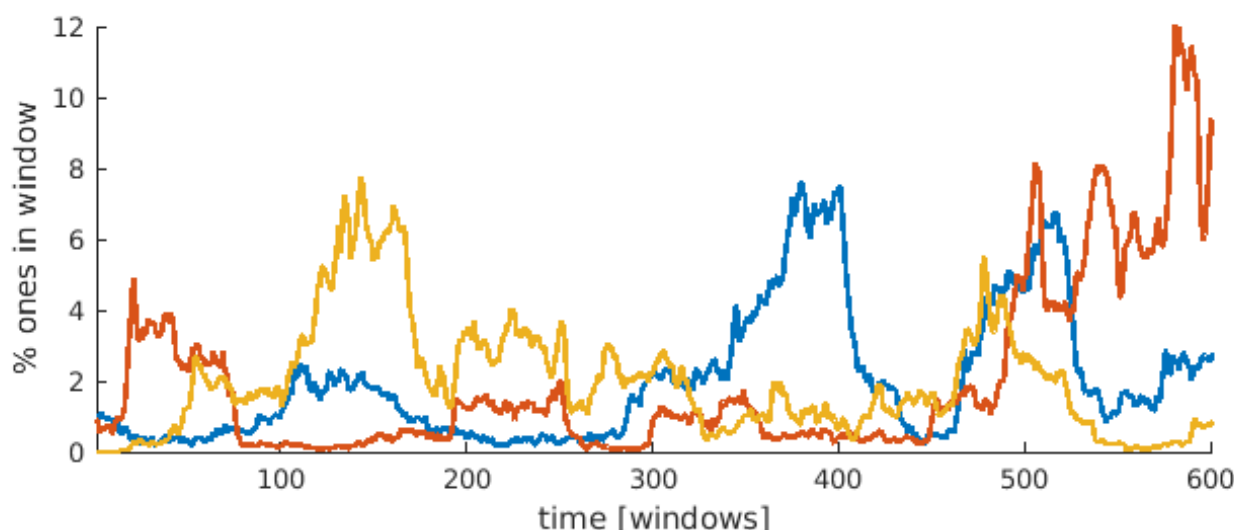
Figure 9: Time courses of the number of supra-threshold FC pairs per window for three subjects. The tensors were created using MI and $\theta = 98$th percentile.

Considering these large fluctuations, our findings are compatible with results by Mitra et al. (2014). The authors explain observed data in terms of waves of activity that propagate from regions acting as sources to others acting as sinks, involving the entire cortex. This results in stereotypical dFC patterns derived from average FC. Since so few data points are necessary to obtain our findings, and since they are concentrated around a very small number of peaks in each subject, it seems plausible to assume that in any given peak, *all* encountered communities are present. This concurs with the notion of a stereotypical global event. Furthermore, Messé et al. (2014) showed that a stationary model assuming fluctuations around a single fixed point, i.e., the average FC, predicts spatial structure of FC more accurately than more complex models, even when dFC is included. This is in parallel to what we show here: even a model that possesses only one attractor can produce time-varying FC patterns that are picked up by a decomposition algorithm.

4.2 Why we use tensor decomposition

In this study, we use a method that is relatively unknown in neuroscience, i.e. tensor decomposition (Cichocki, 2013; Leonardi and Van de Ville, 2013; Gauvin et al., 2014; Ponce-Alvarez et al., 2015). There are several reasons for our choice. First of all, with a resolution of 661 time frames and 66 ROIs, ICA cannot be readily applied. We would have to use temporal ICA (Calhoun et al., 2001) instead of the more widely used spatial ICA (McKeown et al., 1998; Beckmann and Smith, 2004; Calhoun et al., 2009), making it difficult to compare to studies that identify communities (Beckmann et al., 2005; Damoiseaux et al., 2006; De Luca et al., 2006; Mantini et al., 2007; Van den Heuvel and Hulshoff Pol, 2010). However, the low spatial resolution was necessary in order to assess whether our specific model - the dynamic mean field model (Deco et al., 2014) - does indeed have the capacity to reproduce RSNs. Apart from that, investigating specifically long-range connections in a whole-brain approach has its own merit. In any case, tensor decomposition is nothing more but a generalization of PCA or SVD and is thus a fairly generic data analysis technique.

Our second reason is that this method allows us to use temporal fluctuations of FC explicitly, but without making any assumptions. Indeed, non-negativity is the only constraint we apply, and it is completely natural to do so when using a non-negative FC measure like MI. In other words, while

24

ICA starts with the BOLD time courses, we use pair-wise FC, investigating time-evolving community structure explicitly. The result is that we get spatial and temporal features at the same time, without having to assume independence in either dimension.

Third of all, and importantly, this method allows us to extract features from single subjects or simulations, making it unnecessary to map group results back to the individual like with dual regression in group ICA. It is straightforward to match extracted features to a group template and estimate how well a subject agrees with the group average, which is summarized in the silhouette value used here to validate clustering results and tune parameters.

4.3 Correlation versus mutual information

Another perhaps slightly unusual choice is to favor MI over correlation as a measure of FC. It has, however, several advantages. First, it is a non-negative measure which allows us to further constrain the tensor decomposition, leading to more reliable results. Second, spurious dFC structure in correlation-based tensors is reduced, as consistently shown by our results.

The reason for this difference has to lie in the spectral properties of the data, because they are preserved in the surrogate data. We found many time windows that contain outliers, i.e. a handful of time points with much higher activity than the rest. MI as computed here (Kraskov et al., 2004) is robust against these outliers, but correlation overestimates dFC in such windows. We hypothesize that these outliers are a result of global slow fluctuations in the signal that are preserved in the surrogates. MI as a non-parametric measure is robust against these outliers while correlation assumes normality and therefore is sensitive to them.

This observation calls into question the usefulness of Pearson correlation for investigating dFC despite its popularity (Allen et al., 2012; Hutchison et al., 2013; Hansen et al., 2014). Concerns have been addressed in some publications (Lindquist et al., 2014; Hindriks et al., 2015), and cross-validation with appropriate surrogate data is highly commendable.

4.4 Using state-of-the-art connectivity matrices

In SC, each connection is symmetric and its weight is determined by the number of fibers detected by the tracking algorithm. However, it is known that many long-range connections are missed by these algorithms because of crossing fibers, notably in the region of the corpus callosum but also connections between frontal and occipital regions. Therefore, a lot of the interhemispheric connections between homotopic areas are missing. Furthermore, the results of fiber tracking do *not* allow direct inference of the weight of a connection. Lastly, fiber tracking results exhibit a high variability across subjects and sessions (Jones et al., 2013; Jeurissen et al., 2013).

On top of these shortcomings, we know that, *in vivo*, the asymmetry of the underlying connectivity shapes the dynamics of the system. To estimate weights and directionality, we need suitable observables and a dynamical model. We use the conceptual framework of EC (Friston, 1994), which describes the influence that one cortical area has over another. It depends, in principle, on synaptic plasticity, previous brain activity, and neuromodulation. Therefore, *in vivo*, EC changes dynamically depending on the experimental condition, individual differences, etc.

The method developed in Gilson et al. (2016) allows us to extract a single (group level) EC from our data that remains unchanged for all simulations. Note that we constrain the optimization to direct connections that are present in the SC, plus homotopic connections. Directionality and weights of connections are estimated using time-shifted, i.e., asymmetric, covariances as observables, as opposed to correlations, which allows us to also estimate the noise matrix, i.e. the amount of

variability in each node. We apply a simple noise diffusion model, the type of model which has been found to reproduce RS FC quite well (Messé et al., 2014).

The use of EC leads to more realistic communities that are impossible to obtain with SC alone. One of the main benefits of using EC is that homotopic connections are strengthened, enabling realistic symmetric communities. Messé et al. (2014) found that just adding homotopic connections by setting them to a fixed value greatly improved predictive power. It would be interesting to see how model performances play out when using EC, because the shortcomings of SC are likely to impair modelling results very strongly. Perhaps, once this limitation is mitigated, we would favor a more complex model, as was the authors' initial hypothesis.

We emphasize that the communities themselves do not contribute in any way to the EC optimization procedure. They are extracted using temporal information in the shape of $dFC(w)$ matrices that are obtained on time scales that are far greater than those used in the optimization procedure. Thus, the two methodologies make completely different use of dFC. Furthermore, the template extraction which drives the adjustment of parameter setting (choice of F, K, dFC measure, and θ) is done on the single subject basis and is independent of any simulations.

4.5 Conclusion

We have shown that a dynamic mean field model with a single attractor that is explored through noise is sufficient to explain a lot of the spatio-temporal structure found in large-scale resting state fMRI. Noise-driven fluctuations around the average functional connectivity structure are shaped by the underlying connectivity and the simple dynamics of the model in such a way that over time, known functional networks are expressed. We emphasize that we added to previous findings according to which the average FC is reproduced by the model by demonstrating that dFC patterns occurring *over time* are contained in the simulated data. We achieved this by decomposing tensors, or "stacks" of time-dependent dFC matrices. We also showed that applying a high binarization threshold to these tensors, keeping only the overall largest pair-wise dFC values, is necessary to obtain communities that generalize well across subjects.

Future studies should try to further characterize the temporal structure, for example determining how separate in time different networks are. This would enable us to pinpoint differences between experimental groups (different ages, gender, patient groups) and regarding tasks (e.g. attention, decision-making, motor) by describing and modelling how large-scale networks interact and how these characteristics are related to performance or clinical markers.

References

Acar E, Dunlavy DM, Kolda TG (2011) A scalable optimization approach for fitting canonical tensor decompositions. *Journal of Chemometrics* 25:67–86.

Allen EA, Damaraju E, Plis SM, Erhardt EB, Eichele T, Calhoun VD (2012) Tracking whole-brain connectivity dynamics in the resting state. *Cerebral Cortex* p. bhs352.

Bader BW, Kolda TG et al. (2015) Matlab tensor toolbox version 2.6, Available online.

Beckmann CF, DeLuca M, Devlin JT, Smith SM (2005) Investigations into resting-state connectivity using independent component analysis. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* 360:1001–13.

Beckmann CF, Smith SM (2004) Probabilistic independent component analysis for functional magnetic resonance imaging. *IEEE transactions on medical imaging* 23:137–52.

Betzel RF, Fukushima M, He Y, Zuo XN, Sporns O (2016) Dynamic fluctuations coincide with periods of high and low modularity in resting-state functional brain networks. *NeuroImage* 127:287–297.

Biswal B, Zerrin Yetkin F, Haughton V, Hyde J (1995) Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. *Magnetic resonance in medicine* 34:537–541.

Cabral J, Hugues E, Kringelbach ML, Deco G (2012) Modeling the outcome of structural disconnection on resting-state functional connectivity. *NeuroImage* 62:1342–1353.

Calhoun V, Adali T, Pearlson G, Pekar J (2001) Spatial and temporal independent component analysis of functional MRI data containing a pair of task-related waveforms. *Human Brain Mapping* 13:43–53.

Calhoun VD, Liu J, Adali T (2009) A review of group ICA for fMRI data and ICA for joint inference of imaging, genetic, and ERP data. *NeuroImage* 45:S163–72.

Chang C, Glover GH (2010) Time–frequency dynamics of resting-state brain connectivity measured with fMRI. *NeuroImage* 50:81–98.

Cichocki A, Mandic D, Phan A, Caiafa C, Zhou G, Zhao Q, De Lathauwer L (2014) Tensor decompositions for signal processing applications from two-way to multiway component analysis. *arXiv preprint* .

Cichocki A (2013) Tensor decompositions: A new concept in brain data analysis? *arXiv* pp. 507–517.

Cichocki A, Zdunek R, Phan AH, Amari S (2009) *Nonnegative matrix and tensor factorizations: applications to exploratory multi-way data analysis and blind source separation* John Wiley & Sons.

Cordes D, Haughton VM, Arfanakis K, Wendt GJ, Turski PA, Moritz CH, Quigley MA, Meyerand ME (2000) Mapping functionally related regions of brain with functional connectivity MR imaging. *American Journal of Neuroradiology* 21:1636–1644.

Damoiseaux J, Rombouts S, Barkhof F, Scheltens P, Stam C, Smith S, Beckmann C (2006) Consistent resting-state networks across healthy subjects. *Proceedings of the national academy of sciences* 103:13848–13853.

de Amorim RC, Hennig C (2015) Recovering the number of clusters in data sets with noise features using feature rescaling factors. *Information Sciences* 324:126–145.

De Luca M, Beckmann C, De Stefano N, Matthews P, Smith S et al. (2006) FMRI resting state networks define distinct modes of long-distance interactions in the human brain. *NeuroImage* 29:1359–1367.

Deco G, Ponce-Alvarez A, Hagmann P, Romani G, Mantini D, Corbetta M (2014) How local excitation-inhibition ratio impacts the whole brain dynamics. *The Journal of Neuroscience* 34:7886–7898.

Deco G, Jirsa VK (2012) Ongoing cortical activity at rest: criticality, multistability, and ghost attractors. *The Journal of Neuroscience* 32:3366–3375.

Deco G, McIntosh AR, Shen K, Hutchison RM, Menon RS, Everling S, Hagmann P, Jirsa VK (2014) Identification of optimal structural connectivity using functional connectivity and neural modeling. *The Journal of Neuroscience* 34:7910–6.

Deco G, Ponce-Alvarez A, Mantini D, Romani GL, Hagmann P, Corbetta M (2013) Resting-state functional connectivity emerges from structurally and dynamically shaped slow linear fluctuations. *The Journal of Neuroscience* 33:11239–11252.

Desikan RS, Segonne F, Fischl B, Quinn BT, Dickerson BC, Blacker D, Buckner RL, Dale AM, Maguire RP, Hyman BT, Albert MS, Killiany RJ (2006) An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *NeuroImage* 31:968–980.

Doucet G, Naveau M, Petit L, Delcroix N, Zago L, Crivello F, Jobard G, Tzourio-Mazoyer N, Mazoyer B, Mellet E, Joliot M (2011) Brain activity at rest: a multiscale hierarchical functional organization. *Journal of Neurophysiology* 105:2753–2763.

Finn ES, Shen X, Scheinost D, Rosenberg MD, Huang J, Chun MM, Papademetris X, Constable RT (2015) Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity. *Nature Neuroscience* 18:1664–1671.

Fox M, Snyder A, Vincent J, Corbetta M, Van Essen D, Raichle M (2005) The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proceedings of the National Academy of Sciences* 102:9673–9678.

Friston KJ (1994) Functional and effective connectivity in neuroimaging: A synthesis. *Human Brain Mapping* 2:56–78.

Friston K, Harrison L, Penny W (2003) Dynamic causal modelling. *NeuroImage* 19:1273–1302.

Gauvin L, Panisson A, Cattuto C (2014) Detecting the community structure and activity patterns of temporal networks: a non-negative tensor factorization approach. *PLoS ONE* 9:e86028.

Gilson M, Moreno-Bote R, Ponce-Alvarez A, Ritter P, Deco G (2016) Estimation of directed effective connectivity from fMRI functional connectivity hints at asymmetries in cortical connectome. *PLoS Computational Biology* .

Haimovici A, Tagliazucchi E, Balenzuela P, Chialvo DR (2013) Brain organization into resting state networks emerges at criticality on a model of the human connectome. *Physical Review Letters* 110:178101.

Hansen EC, Battaglia D, Spiegler A, Deco G, Jirsa VK (2014) Functional Connectivity Dynamics: Modeling the switching behavior of the resting state. *NeuroImage* 105:525–535.

Heiervang E, Behrens TEJ, Mackay CE, Robson MD, Johansen-Berg H (2006) Between session reproducibility and between subject variability of diffusion MR and tractography measures. *NeuroImage* 33:867–77.

Hindriks R, Adhikari M, Murayama Y, Ganzetti M, Mantini D, Logothetis N, Deco G (2015) Can sliding-window correlations reveal dynamic functional connectivity in resting-state fMRI? *NeuroImage* 127:242–256.

Honey CJ, Sporns O, Cammoun L, Gigandet X, Thiran JP, Meuli R, Hagmann P (2009) Predicting human resting-state functional connectivity from structural connectivity. *Proceedings of the National Academy of Sciences* 106:2035–40.

Hutchison RM, Womelsdorf T, Allen EA, Bandettini PA, Calhoun VD, Corbetta M, Della Penna S, Duyn JH, Glover GH, Gonzalez-Castillo J, Handwerker DA, Keilholz S, Kiviniemi V, Leopold DA, de Pasquale F, Sporns O, Walter M, Chang C (2013) Dynamic functional connectivity: promise, issues, and interpretations. *NeuroImage* 80:360–78.

Hutchison RM, Womelsdorf T, Gati JS, Everling S, Menon RS (2012) Resting-state networks show dynamic functional connectivity in awake humans and anesthetized macaques. *Human Brain Mapping* 34:2154–2177.

Jeurissen B, Leemans A, Tournier JD, Jones DK, Sijbers J (2013) Investigating the prevalence of complex fiber configurations in white matter tissue with diffusion magnetic resonance imaging. *Human Brain Mapping* 34:2747–66.

Jones DK, Knösche TR, Turner R (2013) White matter integrity, fiber count, and other fallacies: the do's and don'ts of diffusion MRI. *NeuroImage* 73:239–54.

Kim J, Park H (2012) Fast nonnegative tensor factorization with an active-set-like method In *High-Performance Scientific Computing*, pp. 311–326. Springer.

Kiviniemi V, Vire T, Remes J, Elseoud AA, Starck T, Tervonen O, Nikkinen J (2011) A sliding time-window ICA reveals spatial variability of the default mode network in time. *Brain connectivity* 1:339–347.

Kraskov A, Stögbauer H, Grassberger P (2004) Estimating mutual information. *Physical Review E* 69:066138.

Leonardi N, Van de Ville D (2013) Identifying network correlates of brain states using tensor decompositions of whole-brain dynamic functional connectivity In *2013 International Workshop on Pattern Recognition in Neuroimaging*, pp. 74–77. IEEE.

Lindquist MA, Xu Y, Nebel MB, Caffo BS (2014) Evaluating dynamic bivariate correlations in resting-state fMRI: a comparison study and a new approach. *NeuroImage* 101:531–46.

Liu X, Chang C, Duyn JH (2013) Decomposition of spontaneous brain activity into distinct fMRI co-activation patterns. *Frontiers in Systems Neuroscience* 7:101.

Mantini D, Perrucci M, Del Gratta C, Romani G, Corbetta M (2007) Electrophysiological signatures of resting state networks in the human brain. *Proceedings of the National Academy of Sciences* 104:13170.

Matsui T, Tamura K, Koyano KW, Takeuchi D, Adachi Y, Osada T, Miyashita Y, Tamura K (2011) Direct comparison of spontaneous functional connectivity and effective connectivity measured by intracortical microstimulation: An fMRI study in macaque monkeys. *Cerebral Cortex* 21:2348–2356.

McKeown MJ, Makeig S, Brown GG, Jung TP, Kindermann SS, Bell AJ, Sejnowski TJ (1998) Analysis of fMRI data by blind separation into independent spatial components. *Human Brain Mapping* 6:160–88.

Messé A, Rudrauf D, Benali H, Marrelec G (2014) Relating structure and function in the human brain: relative contributions of anatomy, stationary dynamics, and non-stationarities. *PLoS Computational Biology* 10:e1003530.

Meunier D, Achard S, Morcom A, Bullmore E (2009) Age-related changes in modular organization of human brain functional networks. *NeuroImage* 44:715–723.

Mitra A, Snyder AZ, Hacker CD, Raichle ME (2014) Lag structure in resting-state fMRI. *Journal of Neurophysiology* 111:2374–91.

Mueller S, Wang D, Fox MD, Yeo BTT, Sepulcre J, Sabuncu MR, Shafee R, Lu J, Liu H (2013) Individual variability in functional connectivity architecture of the human brain. *Neuron* 77:586–95.

Phan AH, Tichavský P, Cichocki A (2013) Low complexity damped Gauss–Newton algorithms for CANDECOMP/PARAFAC. *SIAM Journal on Matrix Analysis and Applications* 34:126–147.

Ponce-Alvarez A, Deco G, Hagmann P, Romani GL, Mantini D, Corbetta M (2015) Resting-state temporal synchronization networks emerge from connectivity topology and heterogeneity. *PLoS Computational Biology* 11:e1004100.

Raichle M, MacLeod A, Snyder A, Powers W, Gusnard D, Shulman G (2001) A default mode of brain function. *Proceedings of the National Academy of Sciences* 98:676–682.

Schirner M, Rothmeier S, Jirsa VK, McIntosh AR, Ritter P (2015) An automated pipeline for constructing personalized virtual brains from multimodal neuroimaging data. *NeuroImage* 117:343–57.

Tagliazucchi E, Balenzuela P, Fraiman D, Chialvo DR (2012) Criticality in large-scale brain fMRI dynamics unveiled by a novel point process analysis. *Frontiers in Physiology* 3:15.

Van den Heuvel MP, Hulshoff Pol HE (2010) Exploring the brain network: a review on resting-state fMRI functional connectivity. *European neuropsychopharmacology* 20:519–34.

Van den Heuvel MP, Sporns O (2011) Rich-club organization of the human connectome. *Journal of Neuroscience* 31:15775–15786.

Vincent J, Patel G, Fox M, Snyder A, Baker J, Van Essen D, Zempel J, Snyder L, Corbetta M, Raichle M (2007) Intrinsic functional architecture in the anaesthetized monkey brain. *Nature* 447:83–86.

Wang Z, Chen LM, Négyessy L, Friedman RM, Mishra A, Gore JC, Roe AW (2013) The relationship of anatomical and functional connectivity to resting-state connectivity in primate somatosensory cortex. *Neuron* 78:1116–1126.

Wong KF, Wang XJ (2006) A recurrent network mechanism of time integration in perceptual decisions. *The Journal of Neuroscience* 26:1314–1328.