

# Dynamic functional connectivity using state-based dynamic community structure: Method and application to opioid analgesia



Lucy F. Robinson<sup>a,1,\*</sup>, Lauren Y. Atlas<sup>b,1</sup>, Tor D. Wager<sup>c</sup>

<sup>a</sup> Department Epidemiology and Biostatistics, Drexel University, 3215 Market St, Philadelphia, PA 19104, USA

<sup>b</sup> Department of Psychology, New York University, 6 Washington Place, New York, NY 10003, USA

<sup>c</sup> Department of Psychology and Neuroscience, University of Colorado, Boulder, Muenzinger D244, 345 UCB, Boulder, CO 80309, USA

## ARTICLE INFO

### Article history:

Accepted 11 December 2014

Available online 20 December 2014

### Keywords:

fMRI

Functional connectivity

## ABSTRACT

We present a new method, State-based Dynamic Community Structure, that detects time-dependent community structure in networks of brain regions. Most analyses of functional connectivity assume that network behavior is static in time, or differs between task conditions with known timing. Our goal is to determine whether brain network topology remains stationary over time, or if changes in network organization occur at unknown time points. Changes in network organization may be related to shifts in neurological state, such as those associated with learning, drug uptake or experimental conditions. Using a hidden Markov stochastic blockmodel, we define a time-dependent community structure. We apply this approach to data from a functional magnetic resonance imaging experiment examining how contextual factors influence drug-induced analgesia. Results reveal that networks involved in pain, working memory, and emotion show distinct profiles of time-varying connectivity.

© 2014 Elsevier Inc. All rights reserved.

## Introduction

The explosion of several new topics in neuroimaging, including large-scale resting state studies (Van Dijk et al., 2012; Zuo et al., 2012; Smith et al., 2013), studies of psychopathology (Jafri et al., 2008; Greicius, 2008), and pharmacological functional magnetic resonance imaging (phfMRI) (Schwarz et al., 2007; Honey and Bullmore, 2004) has led to exciting innovations in methods for functional connectivity. Two emerging themes are the need for dynamic connectivity methods that allow for changes in connectivity over time, and network models that capture the complex structure of connectivity. In this paper we introduce a new method, State-based dynamic community structure (SDCS), that combines these approaches to identify distinct functional connectivity states with respect to the organization of networks of brain regions into communities of brain regions. This method is particularly well-suited for experimental contexts in which the timing of shifts between states cannot necessarily be specified *a priori*.

A growing body of work (Beckmann et al., 2005; De Luca et al., 2006; Damoiseaux et al., 2006; Ferrarini et al., 2009; Fair et al., 2009) has suggested that brain functional connectivity networks exhibit complex structure such that the whole-brain network can be represented as a collection of subnetworks, where each subnetwork is comprised of a set of spatially distributed brain regions. Recent work (Bassett et al., 2011; Bowman et al., 2012; Schwatz et al., 2009; Shen et al., 2010;

Meunier et al., 2009) has shown that these subnetworks can be illustrated as “communities” or modules of “nodes”. Nodes within a community are highly connected and/or share some properties with respect to connectivity. Methods that estimate this network structure by partitioning a collection of brain regions into subgroups can provide complementary information to descriptive measures of the global properties of the functional connectivity network (Bullmore and Sporns, 2009) such as degree distributions (number of connections per node), clustering coefficients, and small world properties (Bullmore and Sporns, 2009; Rubinov and Sporns, 2010; Sporns et al., 2004; Simpson et al., 2013).

Several recent developments in fMRI analysis methods are motivated by the recognition that functional connectivity characteristics are dynamic, fluctuating over time with changes in mental states and other physiological processes (Hutchison et al., 2013). Recent work has incorporated temporal dynamics into the varied descriptions of functional connectivity properties, providing a flexible alternative to standard approaches that assume network behavior is static over time, or differs between task conditions with known timing. Changes in network characteristics include shifts in the strength of connection between a given pair or group of nodes, the degree distribution or clustering coefficient of the network as a whole, or other global network properties. Dynamic connectivity regression (Cribben et al., 2012) provides a method for detecting change points in an fMRI time series based on the connectivity graph of a group of selected regions. Dynamics of spatial activation components have been estimated using dynamic ICA (Allen et al., 2012) and dynamic principal components analysis (Leonardi et al., 2013). The characteristics of the functional connectivity network have been shown to differ according to various task demands (Hutchison et al.,

\* Corresponding author.

E-mail address: [lucy.f.robinson@drexel.edu](mailto:lucy.f.robinson@drexel.edu) (L.F. Robinson).

<sup>1</sup> These authors contributed equally to this work.

2013), and psychoactive drug conditions (Greicius, 2008; Boveroux et al., 2010). Of particular relevance to this paper, the dynamics of the modular structure of the functional connectivity network have been explored in resting state (Jones et al., 2012), learning (Bassett et al., 2011), and studies of emotion (Kinnison et al., 2012).

Recent developments in techniques for dynamic networks in neuroimaging data have occurred alongside new approaches to general dynamic network analysis problems, and in particular, work on dynamic community structure. Popular approaches include extensions of network partitioning algorithms based on modularity maximization or related “quality functions” to time-dependent networks, often by incorporating temporal smoothness into the quality function (Mucha et al., 2010; Kawadia and Sreenivasan, 2012; Chakrabarti et al., 2006; Chen et al., 2013). Other algorithms are based on incremental updates to estimates of community structure (Cazabet et al., 2010; Nguyen et al., 2011). Applications of Dynamic network models to fMRI data must address its particular characteristics, specifically low signal to noise ratio, relative lack of temporal resolution, and the related lack of observable instantaneous, or at least temporally identified, connections (*cf* EEG or email data). As will be discussed, we pursue an approach based on a stochastic generative model for network data, rather than an algorithmic method, and focus on identifying distinct states with respect to community structure.

In this paper, we introduce a dynamic network analysis technique, State-based dynamic community structure (SDCS), that identifies distinct temporal states with respect to community structure over time without *a priori* assumptions on the timing or type of changes in structure. Previous approaches to dynamic connectivity have largely used simple measures of connectivity, or have been confined to a small number of predefined regions and do not scale up easily to larger networks. As discussed in Hutchison et al. (2013), the sliding window approaches employed in many dynamic connectivity studies are difficult to interpret and induce spurious fluctuations under stable connectivity conditions. By contrast, the SDCS approach is able to capture complex changes, estimating the temporal and graphical structure of the network in an integrated manner. We characterize the organization of the network using the stochastic blockmodel (Nowicki and Snijders, 2001), a statistical model for describing community structure. The goal of this approach is to estimate whether and how this network structure shifts over time in experiments with possible state-related changes in psychological activity. A change in the organization of the connectivity network could consist of, for example, a shift from a state in which connectivity is highly modular, *i.e.* connectivity is concentrated within subnetworks of regions, to a state in which the pattern of connectivity is less modular and more integrated throughout the network.

SDCS identifies state-based changes using a Hidden Markov model, a widely used method for identifying latent properties of time series data. Under this framework, we assume that there is an underlying, unobserved state property associated with the network at each point in time, and these states may shift at unknown times. The Hidden Markov model allows us to estimate the properties of the underlying states, and identify where likely shifts between states have occurred. Thus, we can assess whether the functional connectivity network is stationary in time with respect to structural organization, or whether there are shifts in structure, and identify the distinct structural patterns of the functional connectivity network. State-space models have been used previously in related fMRI problems, including studies of changes in effective connectivity for EEG/MEG data under a Dynamic Causal Model (Olier et al., 2013), and in multivariate state-space models for brain function (Janoos et al., 2013).

We apply the SDCS functional connectivity method to fMRI data acquired during a study of opioid- and expectancy-based pain modulation (Atlas et al., 2012). Our previous paper used voxel-wise event-related analyses to test whether opioid drug effects on pain-evoked responses differ as a function of belief (*i.e.* during Open, relative to Hidden, administration). In the present analysis, we use SDCS to examine the

timecourse of connectivity within regions associated with psychological processes thought to underlie placebo effects and pain modulation: pain processing, emotion, and executive function/ working memory (Atlas and Wager, 2013). We build on our previous work by testing whether different networks have different timecourses of connectivity during Open Drug administration, irrespective of task design. Because our task involved changes on the order of minutes (due to the pharmacokinetics of the opioid analgesic) as well as shifts in context (baseline, instructions signaling onset and offset of drug infusion, post-infusion washout period), the present analysis can determine whether distinct functional networks and their within-network organization are differentially influenced by these factors.

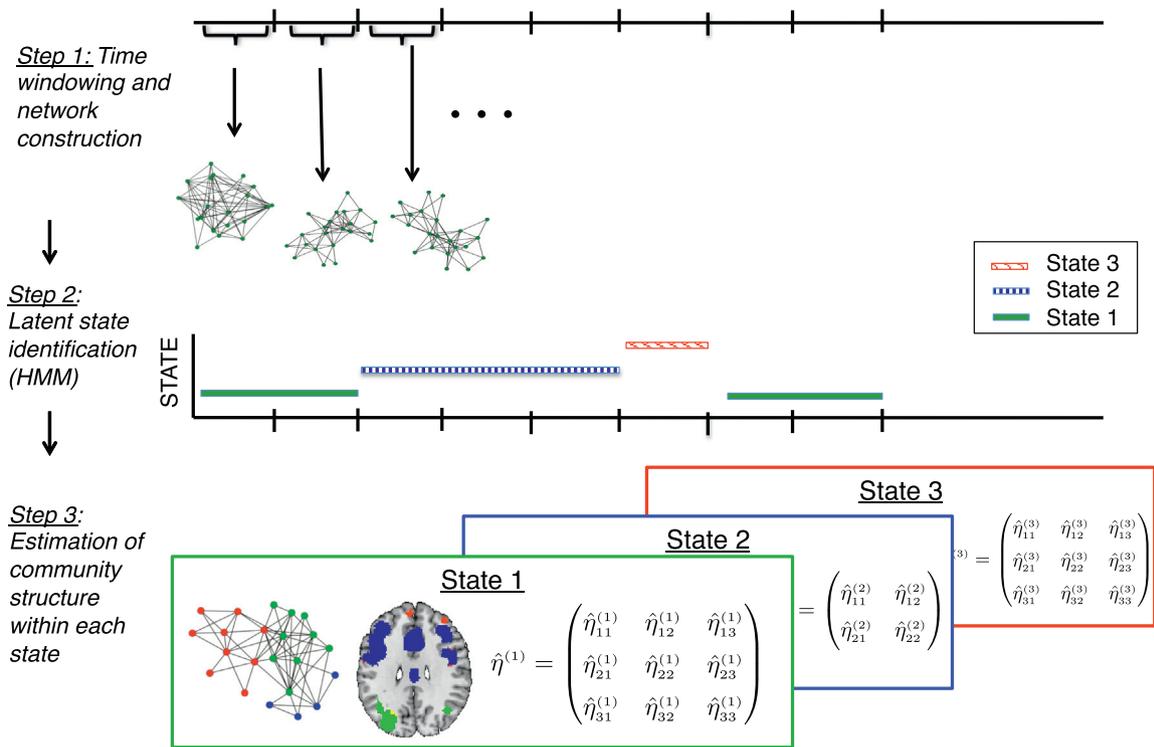
## Methods

The method for estimating time-dependent functional network structure consists of three stages. First, the total time interval is divided into a series of non-overlapping windows and the functional connectivity network is estimated independently in each window, as described in the section [Network representation](#). Next, a hidden Markov model is fit to the sequence of estimated networks to identify possible distinct connectivity states and the most likely locations in time of meaningful shifts in community structure. The procedure for identifying connectivity states and the change points between them is described in the section [Changes over time](#). Finally, the estimates of functional connectivity structure for each of the identified states are refined using Markov Chain Monte Carlo (MCMC). A schematic is shown in [Fig. 1](#).

### Network representation

Before estimating dynamic community structure, the brain regions comprising the network and the measure of connectivity must be defined, as well as the time windows on which the SDCS algorithm will operate. In this paper, we focused on connectivity within parcels of multi-voxel regions defined based on *a priori* functional and anatomical boundaries, examined 150-second windows (75 TRs), and used thresholded coherence as a measure of connectivity. We discuss each of these decisions and associated methods in more detail below. Though the application of our method involves making choices about how to define regions and the length of time windows to integrate over, the SDCS method is general and does not depend on particular choices of region definition and time windowing. Different choices will likely be optimal for different applications.

In network modeling of fMRI data, we have a choice in the definition of the nodes and edges. Voxels themselves can be used as nodes, as in ICA or seed region analysis, or voxels can be segmented into spatially contiguous groups, which are then treated as nodes in the functional connectivity network. While a single voxel approach has the advantage of introducing fewer assumptions on the network, a higher-level representation of connectivity can provide complementary information about patterns of connectivity which can't be attributed to spatial adjacency. In this work, we were primarily interested in brain networks involved in pain processing, working memory, or emotion, as each of these processes has been implicated in expectancy-based pain modulation (Atlas and Wager, 2013). We therefore used meta analyses to identify brain regions consistently activated by working memory (Wager and Smith, 2003) and emotion (Kober et al., 2008) as well as a mega-analysis of five previous studies of thermal pain to identify regions involved in pain processing (Atlas et al., 2010). We note that meta-analysis and mega-analysis provide ways to pool across previous studies to identify regions that are most likely to be activated by a given process. Thus they provide a principled way to identify functional networks *a priori* irrespective of specific experimental details such as fMRI field strength, study population, study location, *etc.* Details of these methods have been described in detail in previous work (Wager et al., 2007; Costafreda, 2009; Salimi-Khorshidi et al., 2009; Wager and Smith,



**Fig. 1.** A schematic for state base dynamic community analysis. Step 1, top: The network is estimated in non-overlapping time windows. Step 2, middle: A hidden Markov model is used to identify distinct connectivity states and where potential state shifts may occur. Step 3, bottom: Estimates of community structure in each state are updated using MCMC.

2003; Kober et al., 2008; Atlas et al., 2010). Voxels identified as being involved in the functional networks of interest were then divided into parcels based on anatomical boundaries, with boundaries defined by the SPM anatomy toolbox (Eickhoff et al., 2005). We regressed out nuisance parameters (spikes, global signal, motion, motion-squared) from voxel-wise timeseries data within the networks for each participant, then averaged across spatially contiguous voxels within parcels. This generated timeseries of discrete regions within networks, which were then used in connectivity analyses.

The next stage of the SDCS estimation algorithm requires the user to specify the length of the windows. We were interested in characterizing functional connectivity and network estimation during noxious thermal stimulation and opioid analgesia. Our task included periods of five minute blocks of stimulation (150 2-second TRs) followed by 30-second periods in which the thermode was moved to a different skin site (Atlas et al., 2012). To apply SDCS and examine connectivity within networks of interest, we excluded these 30-second time periods and constructed a continuous timeseries that reflected brain activation during the thermal stimulation periods. Windows of 150-s (75 TRs) were then used for SDCS analyses. Each corresponds to half of the length of a block, and there were six blocks across the entire run of Open remifentanil administration (Atlas et al., 2012), leading to 12 windows analyzed here (see Fig. 1). Our windows are defined relative to task structure, as instructions were delivered when the thermode was moved between skin sites, but do not incorporate prior information about drug administration timing.

We note that the choice of window length in SDCS depends on balancing two competing factors: accurate estimation of the functional connectivity network in each time window, and the ability to precisely identify potential change points. Model performance using different window lengths under various conditions is analyzed in a simulation study in the section [Changes over time](#). Within each window, the functional connectivity between a given pair of nodes is estimated based on coherence (although correlation or other measures of connectivity could be used). Unlike correlation, reliable estimation of sample

coherence within a given window requires averaging across multiple subwindows, which increases the number of data points required for a stable estimate. More generally, increasing the window length will decrease the variance of the connectivity estimates, providing a less noisy estimate of the network. However, because the estimates of change points are based on the discretized time scale, longer windows will lead to a relative loss of precision in the identification of change points, since there may be considerable heterogeneity within a given window.

Under a few assumptions, we can calculate the standard deviation of errors in estimated locations of state transitions associated with boundary effects from non-overlapping windows as a function of window length. Assuming the locations of transitions between states, in terms of discrete TRs, are distributed uniformly within windows, and that state transitions will be correctly identified at the nearest window boundary, the windowing error will have a discrete uniform distribution  $U(0, L/2)$ , where  $L$  is the window length. The standard deviation of these errors is then given by

$$\frac{L^2 - 4}{48}. \quad (1)$$

As an illustration, the SDs associated with windows of length 30, 50 and 75 TRs are 4.3, 7.2, and 10.8 TRs respectively. While fluctuations in connectivity may occur on several time scales, the SDCS method will perform best in detecting complex changes in the network organization that occur on time scales similar to the window length or greater.

We define connections between nodes based on the degree of linear dependence in the BOLD time series for each node. The time series for each node are created by averaging the voxel-wise time series within each region (with appropriate nuisance variables regressed out, as described above). Because the BOLD signal may have a different shape and duration in different part of the brain, we expect that linear dependence between a pair of regions may be expressed by correlation between their time series at some temporal lag. Rather than specify the

most informative lag(s) between each possible pair of regions and use the associated cross-correlation value as a measure of dependence in the time domain, we characterize linear dependence in the frequency domain. This approach has the advantage of flexibility in discovering relationships at undetermined lags, which is useful in measuring connections between disparate BOLD signals across the brain, and in situations where the BOLD signal may be altered by drug uptake or other neurophysiological state changes. Additionally, we assume that BOLD-related activity occurs at frequencies between .01 and .15 Hz (Sun et al., 2004) and that correlation between regions at higher frequencies is largely spatially correlated noise. The cross spectrum  $f_{xy}$  at frequency  $\omega$  is the Fourier transform of the cross covariance  $\gamma_{xy}(h) = \text{cov}(X_t, Y_{t-h})$ :

$$f_{xy}(\omega) = \sum_{h=-\infty}^{\infty} \gamma_{xy}(h) e^{(i\omega 2\pi h)} \quad (2)$$

which gives the covariance at a given frequency integrated over all possible lags. The coherence (or magnitude-squared coherence)  $\kappa_{xy}(\omega)$  is a normalized version of the cross spectrum which is the spectral-domain analog to the correlation:

$$\kappa_{xy}(\omega) = \frac{|f_{xy}(\omega)|^2}{f_{xx}(\omega)f_{yy}(\omega)}. \quad (3)$$

We use  $\bar{\kappa}_{ij}$ , the coherence averaged over the estimated frequencies in the [.01 Hz, 15 Hz] frequency band, as the measure of connectivity between regions  $i$  and  $j$ .

Coherence matrices can vary greatly between subjects. This between-subject variation is likely due in part to true individual differences in connectivity, but also to artifacts of registration, segmentation, and other scanning and processing factors. In order to estimate connectivity which exhibits consistency across subjects, we average the subject-specific coherence values for each pair of regions to create an overall coherence estimate.

If the number of regions in the network is  $S$ , we can then represent the network using  $W$ , an  $S \times S$  symmetric matrix such that the  $i, j$ th

element  $w_{ij}$  is the coherence between region  $i$  and region  $j$  averaged over subjects,

$$w_{ij} = \frac{1}{n} \sum_{l=1}^n \bar{\kappa}_{ij}^{(l)}, \quad (4)$$

where  $n$  is the number of subjects and  $\bar{\kappa}_{ij}^{(l)}$  is the coherence measure for subject  $l$ .

$W$  is known as a weighted adjacency matrix. As in other functional connectivity studies of network structure (Bowman et al., 2012; Schwatz et al., 2009; Craddock et al., 2012) we will summarize the relationship between a given pair of regions using a binary indicator of connection rather than a continuously valued measure. The binary adjacency matrix  $A$  is created by thresholding  $W$  such that

$$a_{ij} = \begin{cases} 1 & \text{if } w_{ij} > r \\ 0 & \text{otherwise} \end{cases}, \quad (5)$$

where  $a_{ij}$  is the  $(i, j)$ th element of  $A$ . Under this model, connections between regions are symmetric and undirected, and thus  $a_{ij} = a_{ji}$  for  $i, j = 1, \dots, S$ . A schematic for the process used to construct the functional connectivity network is shown in Fig. 2. We choose the threshold  $r$  to be the 95th percentile of the sample coherence when the true coherence is zero (Wang and Tang, 2004), where the sampling distribution depends on the number of windows used to compute the coherence using Welch's averaged, modified periodogram method.

An advantage of using a thresholded rather than weighted network representation is the availability of existing network and graph theoretic models for binary networks, and the relative ease (in some cases) of computation associated with them. Additionally, if the coherence matrix  $W$  can be interpreted as a noisy representation of a "true" network in which each pair of regions is either significantly connected or not, a binary estimate of the network provides a neuroscientifically meaningful representation. However, if the continuous variation of coherences is in fact an informative description of the relative strength of connection, then the thresholding will result in a loss of potentially useful information. As an alternative, a model for network behavior based on weighted

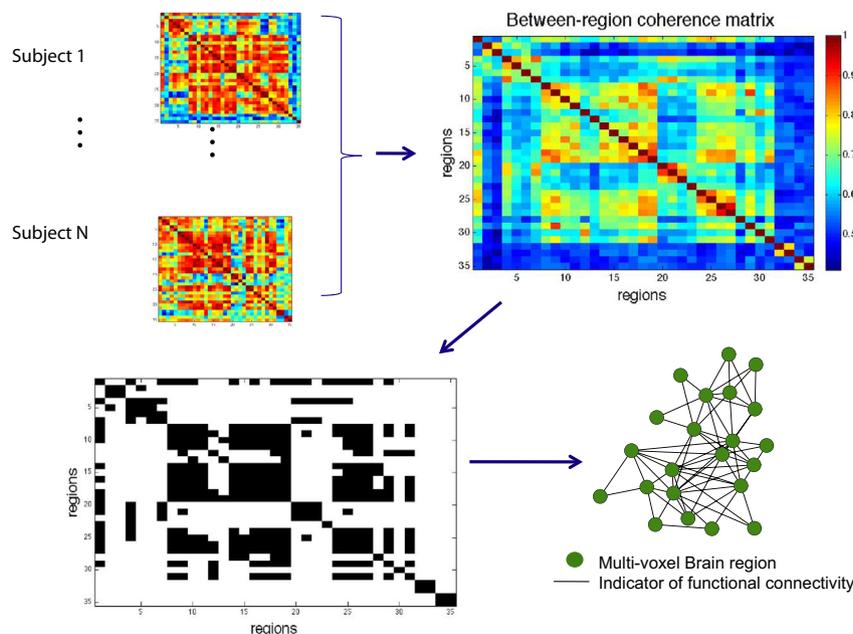


Fig. 2. A schematic for the construction of networks in each time interval. Coherence matrices are averaged over subjects, then thresholded to create a binary adjacency matrix. Regions that have no connections exceeding the threshold are removed from the network.

adjacency matrices could be employed to capture this additional information, for example the mixture model for weighted random graphs described in [Ambrose and Matias \(2012\)](#), although this model would substantially increase the computational burden.

### Stochastic blockmodel

The functional connectivity network can exhibit state-based changes with respect to overall connectivity (averaged across pairs of nodes) and in terms of network structure. We use a stochastic blockmodel ([Fienberg and Wasserman, 1981](#); [Holland et al., 1983](#); [Snijders and Nowicki, 1997](#)) to estimate a partition of the network into a finite number of subnetworks, or 'blocks', each consisting of a group of regions which have similar connectivity characteristics

Description and estimation of block structure are well-studied problems in network analysis (see [Fortunato \(2010\)](#) for a review). Estimation methods include algorithmic approaches such as Newman–Girvan modularity maximization ([Newman, 2006](#)) and spectral clustering ([Von Luxburg, 2007](#)), as well as methods based on statistical network models ([Hoff et al., 2002](#); [Nowicki and Snijders, 2001](#)). Recent theoretical work ([Bickel and Chen, 2009](#)) has shown that model-based partitioning techniques have some advantages as compared to algorithmic methods such as spectral clustering or modularity maximization in terms of statistically consistent estimation. Probabilistic models generate an integrated likelihood of the data under a given community structure, which is useful for making model selection decisions such as how many communities are appropriate, and in assessing whether statistically meaningful shifts in community structure have occurred. We adopt a statistical approach, describing the community structure of the connectivity network using a stochastic blockmodel, described below (see [Fig. 3](#)).

Under the stochastic blockmodel, each brain region  $i$  is associated with an unobserved label  $x_i$ , where  $x_i = k$  indicates that region  $i$  is in block  $k$ , for some  $k = 1 \dots K$ , and  $K$  is the number of blocks. The collection of labels for all regions is  $X = (x_1, \dots, x_S)$ . Prior probabilities for each block are given by  $\alpha = (\alpha_1, \dots, \alpha_K)$ , where  $\alpha_k = P(x_i = k)$  for all  $i = 1 \dots S$ . The vector  $\alpha$  is subject to the condition that

$$\sum_{k=1}^K \alpha_k = 1. \quad (6)$$

Under the stochastic blockmodel, the probability of connection between any two regions, given their block labels, is described by a  $K \times K$  matrix of probabilities  $\eta$  where

$$\eta_{k,l} = P(a_{ij} = 1 | x_i = k, x_j = l). \quad (7)$$

The diagonal values for this matrix,  $\eta_{k,k}$  for  $k = 1 \dots K$ , give the within block connection probabilities, ie the probabilities that two regions which are both in block  $k$  will be connected with one another. The off-diagonal elements  $\eta_{k,l}$ ,  $k \neq l$  give the probability of connection between a pair of regions such that one region is in block  $k$  and the other is in block  $l$ . Because connections are assumed to be symmetric, the matrix  $\eta$  is also symmetric ( $\eta_{kl} = \eta_{lk}$ ). The probability of an adjacency matrix  $A$  given the class labels  $X$  is

$$P(A|X, \eta) = \prod_{i < j} \left( \eta_{x_i, x_j} \right)^{a_{ij}} \left( 1 - \eta_{x_i, x_j} \right)^{(1 - a_{ij})}. \quad (8)$$

Extensions of the stochastic blockmodel exist for more complex network descriptions, such as directed connections ([Nowicki and Snijders, 2001](#)) and degree-corrected models ([Karrer and Newman, 2011](#)).

### Changes over time

We are interested in characterizing changes in network behavior over time. Two types of changes are of interest: changes in the overall level of connectivity across the network, and changes in the community structure. Changes in overall connectivity over the scanning session are assessed using a permutation-based test. Distinct states in the structural organization of the network and the likely locations of shifts between states are estimated using a hidden Markov model. In this paper, we used this approach to identify the dynamics of connectivity separately for three functional networks (pain processing, working memory, and emotion).

Overall connectivity is defined as the average coherence across the network, where the average is taken over all pairs of regions. The complete time interval  $(0, T_{max})$  is divided into  $T$  non-overlapping intervals,  $(0, T_{max}/T), (T_{max}/T + 1, 2T_{max}/T), \dots, (T_{max}(T-1)/T + 1, T_{max})$ . For each interval  $t = 1, \dots, T$  we calculate  $c_t$  where

$$c_t = \frac{1}{\binom{S}{2}} \sum_{i < j} \bar{K}_{ij}, \quad (9)$$

where  $S$  is the total number of regions.

To assess whether fluctuations in average coherence over time indicate meaningful differences rather than merely random sampling variability, we use a permutation analysis to generate a 95% confidence band under the null hypothesis of non-changing coherence. The null distribution of the average coherence over subjects and regions is sampled by permuting the time-ordering of the  $T$  observed coherence matrices within subjects, and then creating averages over subjects in each time interval. This permutation procedure is repeated 1000 times, resulting in 1000 samples of  $(c_1, \dots, c_T)$  from a distribution with time-homogeneous coherence. The .025th and .975th percentiles of the extreme values (maxima and minima respectively) of the samples are computed, providing a 95% confidence for the range of the average coherence over  $T$  time windows. This confidence band reflects the coherence sampling variability over  $T$  estimates, which is more conservative than a confidence band for a single sample coherence value. Because the distribution of the coherence is asymmetric, this confidence band is also asymmetric.

Assessing whether there are significant differences between two time intervals in the structure of the functional connectivity network is a more complex problem. We would like to detect changes in network organization which cannot be completely explained by differences in overall connectivity as defined above.

Computing coherence over  $T$  non-overlapping time intervals produces a sequence of observed functional connectivity networks described by adjacency matrices  $A_1, A_2, \dots, A_T$ . Changes in community structure between neighboring intervals can include changes in  $x = (x_1, \dots, x_S)$ , the partition of regions into blocks, changes in the block connectivity matrix  $\eta$ , and/or changes in the estimated number of communities. We are interested in assessing whether a common block structure  $(X, \alpha, \eta)$  over all time intervals is an adequate description of the functional connectivity network, or if the data reflect two or more distinct block structures  $(x^{(1)}, \alpha^{(1)}, \eta^{(1)}), (x^{(2)}, \alpha^{(2)}, \eta^{(2)}) \dots (x^{(C)}, \alpha^{(C)}, \eta^{(C)})$ , where  $C \geq 2$ . In the latter case, the functional connectivity network moves between  $C$  states over time, following an unknown path.

We model the evolution of block structure over time by imposing a hidden Markov Model (HMM, [Rabiner, 1989](#)) on the observed sequence of functional connectivity networks  $A_1, A_2, \dots, A_T$ . Under a discrete HMM, there is an unobserved latent process  $S_1, S_2, \dots, S_T$ , that determines which of a finite number of states the system is in at any given time. The unobserved state  $S_t$  governs the parameters of the probability distribution that generates the observed data at time  $t$ . In our case, the data of interest are the adjacency matrices,  $A_t$  and the latent state determines which of the community structures  $(x^{(1)}, \alpha^{(1)}, \eta^{(1)}), (x^{(2)}, \alpha^{(2)}, \eta^{(2)}) \dots, (x^{(C)}, \alpha^{(C)}, \eta^{(C)})$

C)) describe the network at time  $t$ . The sequence of observed data is assumed to be conditionally independent in time, given the unobserved sequence  $S_t$ . The temporal dependence in the sequence  $S_1, S_2, \dots, S_T$  is first-order Markovian, meaning that  $P(S_t = c | S_1, S_2, \dots, S_{t-1}) = P(S_t = c | S_{t-1})$ . A schematic for the hidden Markov model is given in Fig. 4.

The dynamic properties of the state sequence are described by a transition matrix  $P$ , where the matrix element  $p_{b,c}$  gives  $P(S_t = b | S_{t-1} = c)$ , the probability the network will be in state  $b$  at interval  $t$ , given that the network was in state  $c$  at time  $t - 1$ , and a vector of initial probabilities  $\delta = (\delta_1, \dots, \delta_C)$ , where  $\delta = P(S_1 = c)$ . The joint probability of sequence of observed adjacency matrices  $A = (A_1, \dots, A_T)$  given parameters  $\theta = (\delta, P, \eta^{(1)}, \dots, \eta^{(C)}, \alpha^{(1)}, \dots, \alpha^{(C)})$  is then given by

$$P(A|\theta) = \sum_{S \in \mathcal{S}} P(A, S|\theta) = \sum_{S \in \mathcal{S}} \delta_{S_1} \prod_{t=1}^T p_{S_t, S_{t-1}} P(A_t | \eta^{(S_t)}, \alpha^{(S_t)}) \quad (10)$$

where  $\mathcal{S}$  is the set of all possible state sequences.

The Baum–Welch algorithm (Rabiner, 1989), a special case of the EM algorithm, is a commonly used method for estimating the parameters of the hidden Markov model. Using the estimated parameters for each latent state, the most likely state sequence can then be determined. The incorporation of the stochastic blockmodel into the hidden Markov is novel (to our knowledge), but in other respects our implementation of the HMM is standard. Therefore, we will not present all of the computational details of the estimation of the HMM but refer the reader to Elliott et al. (1995).

**Model fitting**

If the labels  $x_i, i = 1, \dots, S$  are known *a priori*, the parameters  $\eta$  and  $\alpha$  can be estimated in a straightforward manner. If the community structure is unknown, as in our application, fitting the model from the data requires a more computationally intensive estimation algorithm. The basic steps of our model fitting algorithm are

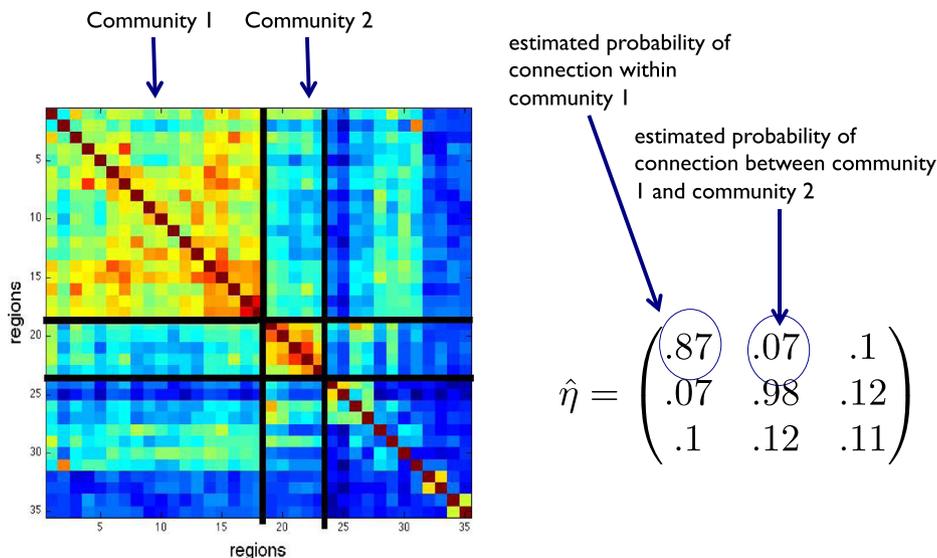
1. Divide the data into a sequence of non-overlapping time intervals and construct an adjacency matrix  $A$  for each interval.
2. Select  $C$  time intervals at random to create initial values for each of the hidden Markov states. In each of the  $C$  intervals, use the Gibbs sampler to estimate the parameters of the stochastic blockmodel, creating initial parameter estimates for each state.

3. Use the hidden Markov model to identify the most likely state sequence, updating the block structure parameters *via* variational EM at each stage of the Baum–Welch algorithm. Using the most likely state sequence (over time), combine the time intervals within each state and use the Gibbs sampler to estimate the block structure for each state.

The algorithm for fitting the hidden Markov model in step 2 can be characterized as an expectation conditional maximization (ECM) algorithm (Meng and Rubin, 1993), in which the variational estimation of the stochastic blockmodel parameters at each state of the Baum–Welch algorithm generates updated estimates of the stochastic block model which increase the likelihood. Consistency properties of the variational EM for the stochastic blockmodel are not fully understood (Daudin et al., 2008; Celisse et al., 2012), but empirical results point to good performance (Gazal et al., 2012). Celisse et al. (2012) has studied the consistency of the variational EM for the stochastic blockmodel, and finds it asymptotically equivalent to the maximum likelihood estimate. Because the variational EM's properties are unclear, we estimate the block structure in each state with a fully Bayesian estimate in step 3. Details for the Gibbs Sampler and variational EM algorithm for estimating the stochastic block structure are given below.

Although the stochastic blockmodel is similar conceptually to the general finite mixture problem, estimation methods such as the EM algorithm cannot be applied because the conditional probability  $P(X|A)$  is intractable. The Gibbs sampler or variational EM approximations (Jordan et al., 1999) can however be applied in a fairly straightforward manner. Gibbs sampling, as described in this context by Nowicki and Snijders (2001), allows us to evaluate the maximum *a posteriori* (MAP) estimates and explore the full posterior distributions, but becomes computationally impractical when applied to many time intervals and within the hidden Markov model (HMM). Full MCMC methods also become difficult to implement as the number of nodes in the network increases. Variational EM methods are much less computationally expensive and give an approximation to the local MAP estimates.

Our approach is to the Gibbs sampler for initial and final estimates of the block structure within each state, as described above, and the variational EM to update the parameter estimates within the HMM fitting algorithm. A brief description of each algorithm is given below; for more details see Nowicki and Snijders (2001) and Daudin et al. (2008).



**Fig. 3.** An example of estimated block structure. The rows and columns of the adjacency matrix are permuted so that regions in the same block appear next to one another. Divisions between blocks are shown by black lines. Within-block connections appear in blocks along the diagonal of the matrix, and between-block connections are in the off-diagonal blocks.

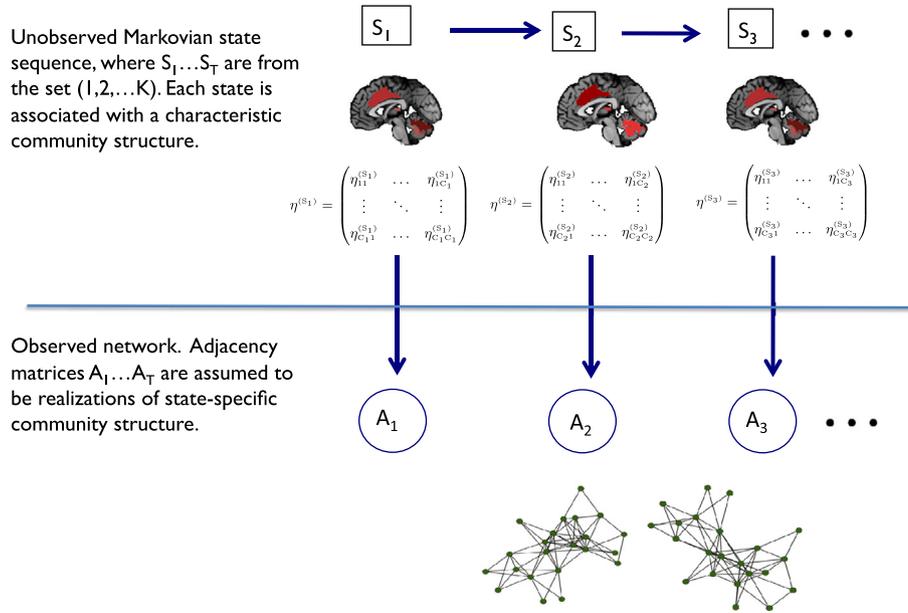


Fig. 4. Hidden Markov model.

Gibbs sampling is a simulation method which can be used to approximate the posterior distribution of  $P(X, \alpha, \eta | A)$ , where  $X$  is the vector of unknown block labels for each region,  $\alpha$  and  $\eta$  are the prior block probability and matrix of connection probabilities respectively, and  $A$  is our observed network for a given time interval. As in other types of mixture models, the block labels  $X$  can be viewed as *missing data*, and  $(A, X)$  is referred to as the *complete data*. A full discussion of the choice of prior distributions for  $\alpha$  and  $\eta$  is available in [Nowicki and Snijders \(2001\)](#). In summary, a  $K$ -dimensional Dirichlet conjugate prior  $D(T, T, \dots, T)$  is used for alpha, with  $T = 100K$ . For  $\eta$ , Dirichlet prior distributions are also used for each column, with some additional constraints to enforce symmetry. Given  $X^{(p)}$ ,  $\alpha^{(p)}$ , and  $\eta^{(p)}$ , the current values of  $X$ ,  $\alpha$  and  $\eta$  respectively, the basic sampling algorithm for generating the next values  $X^{(p+1)}$ ,  $\alpha^{(p+1)}$ , and  $\eta^{(p+1)}$  is

1. Draw values  $\alpha^{(p+1)}$  and  $\eta^{(p+1)}$  from the posterior distribution  $P(\alpha, \eta | A, X)$ . Given the prior distributions described above, the posterior distribution for  $\alpha$  is Dirichlet with parameter vector  $(m_1 + T, m_2 + T, \dots, m_K + T)$ , where

$$m_k = \sum_{i=1}^S I_{(x_i^{(p)}=k)} \quad (11)$$

is the number of regions assigned to block  $k$  in according to the current value of  $X$ , for  $k = 1 \dots K$ . The posteriors for the elements  $\eta_{k,h}$ ,  $h = 1, \dots, K$  of the matrix  $\eta$  are Beta (1-dimensional Dirichlet) with parameter values

$$e_{k,h} = \frac{\sum_{i,j} I_{(a_{ij}=1)} I_{(x_i^{(p)}=k)} I_{(x_j^{(p)}=h)}}{1 + I_{(k=h)}} + 1. \quad (12)$$

2. For each value  $i = 1 \dots S$  in turn, draw  $x_i$  from its conditional distribution given  $A$ , and the current values of  $\eta$ ,  $\alpha$  and  $\{x_j\}_{j \neq i}$ . This distribution is determined by

$$P(x_i = k | A, \alpha, \eta, \{x_j\}_{j \neq i}) = Q \alpha_k \prod_{h=1}^K (\eta_{k,h})^{d_{i,h}} (1 - \eta_{k,h})^{m'_h - d_{i,h}} \quad (13)$$

where  $Q$  is a constant not depending on  $k$ ,

$$d_{i,k} = \sum_{j=1}^{i-1} a_{ij} I_{(x_j^{(p+1)}=k)} + \sum_{j=i+1}^S a_{ij} I_{(x_j^{(p)}=k)}, \quad (14)$$

and

$$m'_k = \sum_{j=1}^{i-1} I_{(x_j^{(p+1)}=k)} + \sum_{j=i+1}^S I_{(x_j^{(p)}=k)}. \quad (15)$$

The value of  $Q$  can be easily calculated from the condition that

$$\sum_{k=1}^K P(x_i = k | A, \alpha, \eta, \{x_j\}_{j \neq i}) = 1. \quad (16)$$

The alternation of the two steps above will create a collection of samples from the posterior distribution of  $P(X, \alpha, \eta | A)$  which can be used to do Bayes estimation of the stochastic block structure.

The variational EM algorithm begins with initial estimates  $\hat{X}^{(0)}$ ,  $\hat{\alpha}^{(0)}$  and  $\hat{\eta}^{(0)}$ . At the  $(q+1)$ <sup>th</sup> iteration, the estimates are updated using the formulas

$$\hat{\alpha}^{(q+1)} = \frac{1}{S} \sum_{i=1}^S \hat{\tau}_{ik}, \quad \hat{\eta}_{hk}^{(q+1)} = \frac{\sum_{i \neq j} \hat{\tau}_{ih} \hat{\tau}_{jk} a_{ij}}{\sum_{i \neq j} \hat{\tau}_{ih} \hat{\tau}_{jk}}, \quad (17)$$

where  $\tau_{ik}$  is a variational parameter which approximates  $P(x_i = k | A, \alpha, \eta)$  based on a simplified conditional likelihood function which provides a lower bound on the true likelihood. The updates for  $\hat{\tau}_{ik}$  are given by the fixed point relation

$$\hat{\tau}_{ik} \propto \prod_{j \neq i, h=1}^K [\eta_{hk}^{a_{ij}} (1 - \eta_{hk})^{1 - a_{ij}}]^{\hat{\tau}_{jh}}. \quad (18)$$

Estimation of the number of blocks,  $K$  is discussed in the section [Model selection](#).

In our application, the variational EM estimation of the stochastic typically converges in a small number of iterations ( $< 10$ , generally). On a machine with eight 2.66 GHz Intel Xeon processors, the Gibbs

Sampler with 5 chains running for 10,000 iterations finishes in under 10 min.

### Model selection

The dynamic community structure model requires specifying  $C$ , the number of temporal states in the hidden Markov model (HMM). Given  $C, K_1, K_2, \dots, K_C$ , the number of blocks in each state, also must be determined. Choosing these values requires identifying the appropriate complexity to describe meaningful features in the data without overfitting. Such model selection problems are typical in clustering procedures, and how to best resolve them is a topic of some debate (Rissanen, 2007). Likelihood ratio tests are generally not valid in this context, which precludes the use of standard statistical hypothesis testing methods, although there is active research in creating tests of model dimension using penalized likelihood ratios (Li and Chen, 2010; Chen et al., 2012). Model selection criteria such as BIC, AIC, MDL and integrated classification likelihood (ICL, Biernacki et al. (2000)) are often used as an alternative, although their properties with respect to consistent estimation of the dimension of the model are generally not completely understood. Hidden Markov models and other change point detection frameworks present a more complex selection problem than standard mixtures, as sequential observed data cannot typically be assumed to be independent. The consistency of BIC and other penalized likelihood criteria for the hidden Markov model order selection is a topic of open research. Large sample properties, however, may not be relevant to our application. Based on empirical simulation results, Scott (2002) find that for small to moderate samples, BIC may underestimate the true number of states and advocate for estimating the number of states using a fully Bayesian variable dimension Monte Carlo (VDMC) method such as reversible jump MCMC or jump diffusion MCMC. While full MCMC methods might have some model selection advantages in our case, we find that the hidden Markov stochastic block model presents considerable computational difficulties in implementation of VDMC, specifically in navigating a parameter space whose dimension is variable both in terms of the number of states and in the number of communities within states, and opt for a more tractable approach.

Estimating the dimension (number of communities) of the stochastic block model presents a similar model selection problem. Although the community estimation problem strongly resembles a standard finite mixture model, the nodes in the network again are not independent of one another. ICL has been proposed as a model selection criterion (Daudin et al., 2008), (Côme and Latouche, 2013), as well as an approximation to BIC (Airoldi et al., 2008).

Our approach is to use an approximate BIC for the entire model to evaluate possible values of  $C$ , and combinations of  $K_1, K_2, \dots, K_C$  for a given  $C$ . Using principles described in Airoldi et al. (2008) and Volinsky and Raftery (2000), the complexity term of this BIC is  $(D_1 + \dots + D_C) \log(R_1 + \dots + R_C)$ , where  $D_c = K_c(1 + (K_c + 1)/2)$  is the number of parameters in state  $c$  and  $R_c$  is the number of observed edges in state  $c$ ,  $c = 1 \dots C$ .

Because fitting the HMM involves comparing the (unpenalized) likelihood of an observation under various potential values of the hidden states, it's not possible to directly fit models for which the number of communities varies across different states. We therefore use a two stage procedure. The HMM is fit with values of  $C$  from 1 to 3, producing for each observation (*i. e.* each network in a given time interval) the probabilities of being in different states. Using these fuzzy estimated state sequences, we fit stochastic block models for each state for  $K = 1, \dots, 4$  communities. The model with the lowest BIC is selected from this suite of candidate models for different combinations of  $C$  and  $K_1, K_2, \dots, K_C$ . Model selection is performed independently for the pain processing, emotion, and working memory networks. While the properties of the selection procedure are not fully understood and are a topic for future research, it provides scientifically reasonable results in our application and in limited simulation studies.

In addition to selecting the complexity of the model through  $C$  and  $K$ , the window length and frequency band for the coherence measure must also be specified prior to analysis. We rely on previous work by Sun et al. (2004) to inform our choice of frequency band. The choice of window length is somewhat subjective and will depend on consideration of the apparent SNR of the data and relative importance of the expected error in estimating state transition locations as discussed in the section Network representation. The impact of window length on the performance of this method under different conditions is explored in the following simulation study.

### Simulations

Synthetic data was used to assess the performance of the method under known conditions, and to assess the sensitivity of results to the choice of window length, number of states, and number of transitions between states.

For each simulation, a series of simulated coherence matrices  $W_1, W_2, \dots, W_T$  are generated according to a state-space mode with 1, 2, or 3 states. The data are simulated with  $S = 70$  nodes in the network and 900 time points. Transitions between states are located at randomly generated time points. Each state has a distinct community structure with respect to the partition of regions into communities given by  $x$ , and the average values of coherence for pairs of regions within and between communities. These average coherence values are given by a matrix  $G$  where  $g_{kl}$  = average coherence between a pair of regions in communities  $k$  and  $l$  respectively. The elements of  $G$  for each state are chosen to reflect typical values of coherence from the data. The values of  $G$  for states 1, 2, and 3 are

$$G^{(1)} = \begin{pmatrix} .6 & .35 & .4 \\ .35 & .55 & .3 \\ .4 & .3 & .45 \end{pmatrix}, G^{(2)} = \begin{pmatrix} .66 & .21 & .24 \\ .21 & .57 & .18 \\ .24 & .18 & .45 \end{pmatrix}, \text{ and } G^{(3)} = \begin{pmatrix} .48 & .25 & .19 \\ .25 & .71 & .30 \\ .19 & .30 & .38 \end{pmatrix}. \quad (19)$$

Given these values for average coherence across subjects within and between communities,  $S \times S$  coherence matrices are generated randomly based on the bootstrapped distribution of the sample coherences at a given window length, where  $S$  is the number of simulated brain regions. For a given pair of regions, the dependence of the cross spectra on window length in the [.01 Hz–.15 Hz] frequency band involves features of the data which may not accurately be simulated by an ARMA or other standard time series model. Rather than simulate time series from a simple generative model, we generate the distribution of the sample coherence at different window lengths using a bootstrap procedure. For a given value of average coherences, pairs of brain regions with similar average values are selected and coherences over 30, 50, 75, and 100 TR windows are computed for each subject. Sample coherences are generated by bootstrapping across regions within a community and across subjects for a given pair. We believe the bootstrapped distributions are more relevant to performance in real applications than purely simulated data, but they unfortunately constrain us to looking at results for our current sample size, 21 subjects. We expect that performance will improve for large numbers of subjects.

Transitions between states are located at randomly generated time points. The location of the transition is independent of the edges of the time windows, so some time windows contain data from two states (or three states, possibly). For these windows, the simulated coherence matrix  $W_i$  is a weighted average of the sample coherence generated by the two state models, where the weight depends on the location of the transition point, *i. e.* what fraction of the time window is in state 1 vs state 2. One goal of the simulation study is to assess how well the characteristics of each state can be recovered under different window lengths.

Community structure for the first state is randomly generated. The magnitude of the difference between states is determined by the number of nodes which are assigned different communities in state 2 or state 3 than in state 1. As the number of nodes which are reassigned decreases, the states become more similar and the changes are more difficult to detect, resulting in errors in the identification of the number of states, location of transitions between states, and community structure in each state. For each time window length, we simulate data under three different magnitudes of change: 15%, 25% and 35% of nodes reassigned. The reassigned communities are selected randomly. We note that random reassignments may be more difficult to detect than more organized changes (which are possibly more plausible in real data), e.g. a connected group of nodes that moves from one community to another, or becomes its own community.

The success of the algorithm is measured using two criteria: identification of the correct number of states using BIC, and recovery of the correct the partition of regions into communities for each state. The latter is measured using the adjusted rand index (ARI), which compares the estimated partition for each state to the true partition for each state, in terms of the numbers of pairs in the network which are correctly identified as belonging to the same or different communities. The ARI takes on values between  $-1$  and  $1$ , where  $ARI = 1$  indicates a perfect recovery of the true partition. For example, in a network with 70 nodes (as in our simulation) an estimated partition with 1 misclassified node produces  $ARI = .95$ , 2 misclassified nodes gives  $ARI = .91$ , 5 misclassified nodes gives  $ARI = .79$ . Heuristically, we consider  $ARI > .75$  to be 'good' performance, indicating a partition relatively close to the true partition. The ARI for each estimated state are averaged

to produce an overall ARI. Model selection is measured by the fraction of simulations in which the true number of states is selected using the BIC criteria, after fitting the model with  $C = 1, 2, 3$  and 4 states.

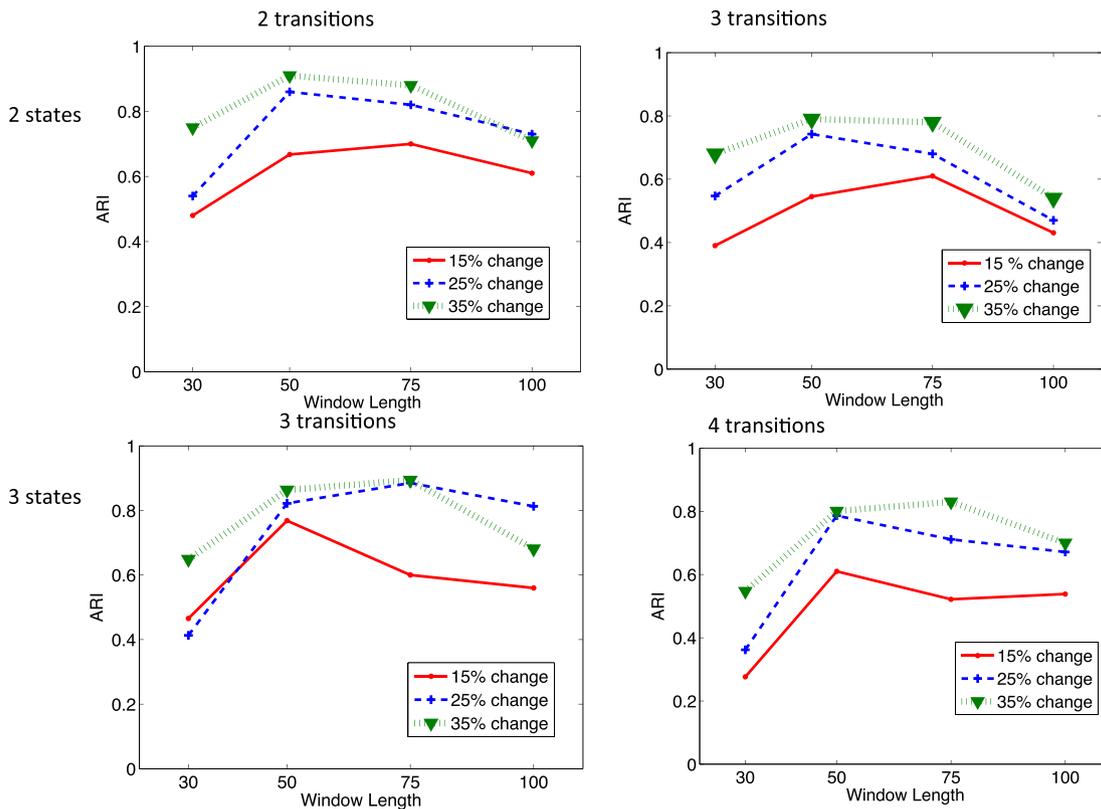
The results of the simulation study for different window lengths, numbers of states (2 or 3), change magnitudes, and number of transitions between states are shown in Figs. 5 and 6. We find that the method is able to produce ARIs  $> .75$  under moderate changes (25–35% nodes reassigned) under a range of conditions. As expected, performance improves with the magnitude of the difference between states. The effect of window length varies somewhat for different numbers of states and numbers of transitions, but overall, for the given sample size and run length, 50 or 75 TRs perform better than 30 or 100 TRs. The disadvantage to window lengths of 100 TRs increases with the number of transitions. The probability of recovery of the true number states is between 75% and 95% for most simulated conditions. As discussed in the section [Model selection](#), when there is an error in the estimation of the number of states, it is typically an underestimation rather than an overestimation (e.g. selecting 2 states rather than 3). In this sense, if model selection is interpreted as a test of homogeneity, for the sample size at hand the test is conservative.

## Results

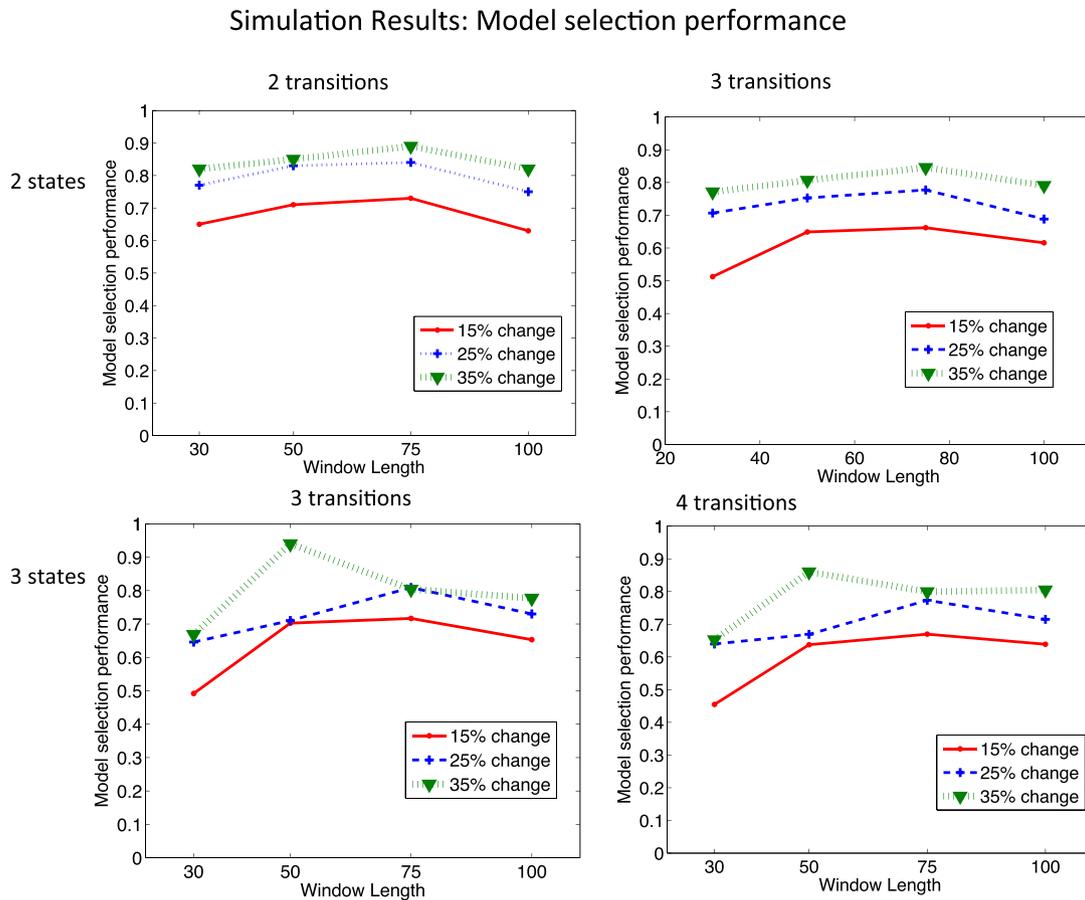
### Overall coherence

The pain processing, emotion and working memory networks data were analyzed separately with respect to changes in overall coherence and network community structure. The 30 minute scanning interval

### Simulation Results: Adjusted Rand Index for Community Estimation



**Fig. 5.** Simulation results for recovery of true community structure. Recovery of true community structure is measured by the adjusted rand Index (ARI) comparing the true community structure to the estimated community structure. The ARI takes on values between  $-1$  and  $1$ , where  $ARI = 1$  indicates a perfect recovery of the true partition, and  $ARI = 0$  is the expected ARI for two random partitions. For example, a network with 70 nodes (as in our simulation) an estimated partition with 1 misclassified node produces  $ARI = .95$ , 2 misclassified nodes gives  $ARI = .91$ , 5 misclassified nodes gives  $ARI = .79$ . Performance is shown for 2 states (top row) and 3 states (bottom row), 15%, 25% and 35% reassigned nodes (shown in different line types), and window lengths of 30, 50, 75 and 100 TRs.



**Fig. 6.** Simulation results for recovery of true number of states. Performance is measured by the fraction of simulations in which the true number of states is recovered using BIC. Performance is shown for 2 states (top row) and 3 states (bottom row), 15%, 25% and 35% reassigned nodes (shown in different line types), and window lengths of 30, 50, 75 and 100 TRs.

(900 TRs) was divided into  $T = 12$  non-overlapping time windows of 2.5 min (75 TRs) each. The average coherences over all pairs of regions and subjects for each time window are shown in Fig. 7, with 95% confidence bounds based on the empirical distributions of the maximum and minimum average coherence over 12 time windows under the hypothesis of time-homogeneous coherence, as discussed in the section [Changes over time](#). The pain processing network exhibits coherences that appear roughly constant in time, with some evidence of increased coherences in the 5–10 and 20–25 minute time intervals, where the average coherence exceeds (by small margins) the 97th percentile of the distribution for the maximum over 12 intervals under time-homogeneous coherence. The working memory network exhibits decreased connectivity over the 12–18 minute time interval, when the drug concentration is expected to be highest, relative to the rest of the interval. There also appears to be increased connectivity in the 0 to 5 minute interval. Connectivity in the emotion network is lower than average in the 5–10 minute interval and higher than average in the 20–25 minute interval. These periods coincide with the delivery of instructions about drug infusion onset and offset, respectively.

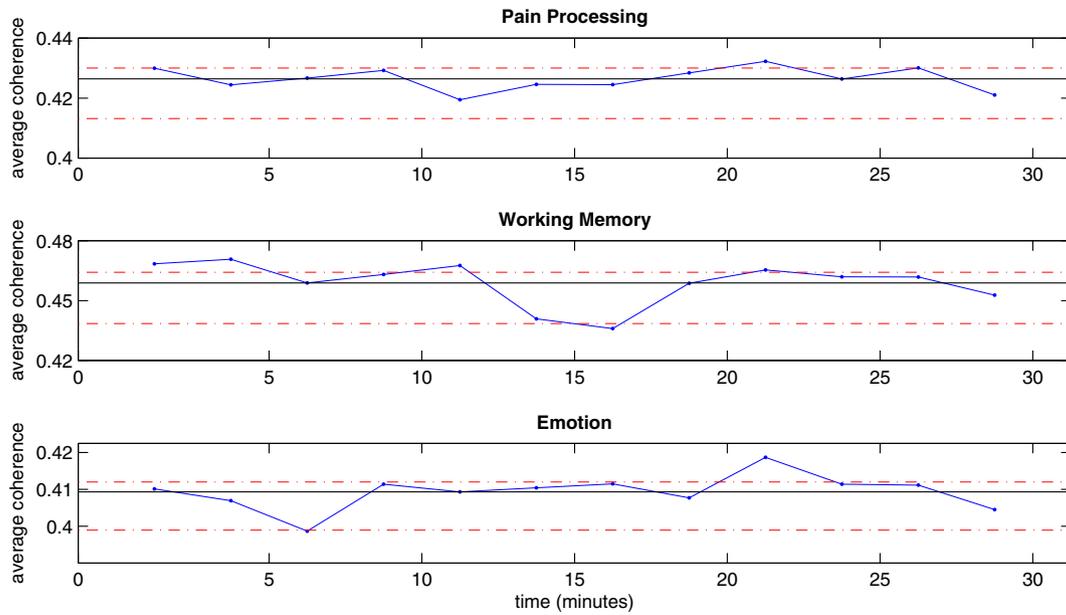
#### Dynamic community structure

The three networks were also evaluated separately with respect to dynamic community organization. As discussed in the section [Model selection](#), dynamic community structure is estimated for each network under models of varying complexity. Here, complexity refers both to temporal structure in terms of  $C$ , the number of latent states, and to

the complexity of the community structure within each state, in terms of  $K_c$ , the number of communities within state  $c$ ,  $c = 1 \dots C$ . Each possible model is evaluated in terms of BIC. Detailed results for each estimated community are shown in the [Appendix A](#).

For the pain processing network, the optimal model based on BIC has two states which differ with respect to the degree of modularity in the functional connectivity network. The most likely (Viterbi) path between states 1 and 2 is shown in Fig. 8. The estimated community structure and connectivity matrix for each State are illustrated in Fig. 8 and reported in Table 1. State 1, which is estimated to be the most likely state before and after peak drug concentration (*i.e.* during standard pain processing), has three communities with very little between-network connectivity. In state 1, we see high within-community coherence in a network composed of dorsal anterior cingulate and bilateral insula (purple in Fig. 8), moderate coherence in thalamus and striatum, and low connectivity in cerebellar clusters. The pain processing network is estimated to most likely be in state 2 during the periods of highest drug concentration. State 2 is characterized by four estimated communities with stronger between-community connectivity than observed in state 1. One interesting shift is that bilateral insula continues to show high within-community coherence, whereas anterior portions of the cingulate cortex join a separate network in state 2, characterized by low within-community connectivity (see Table 1)

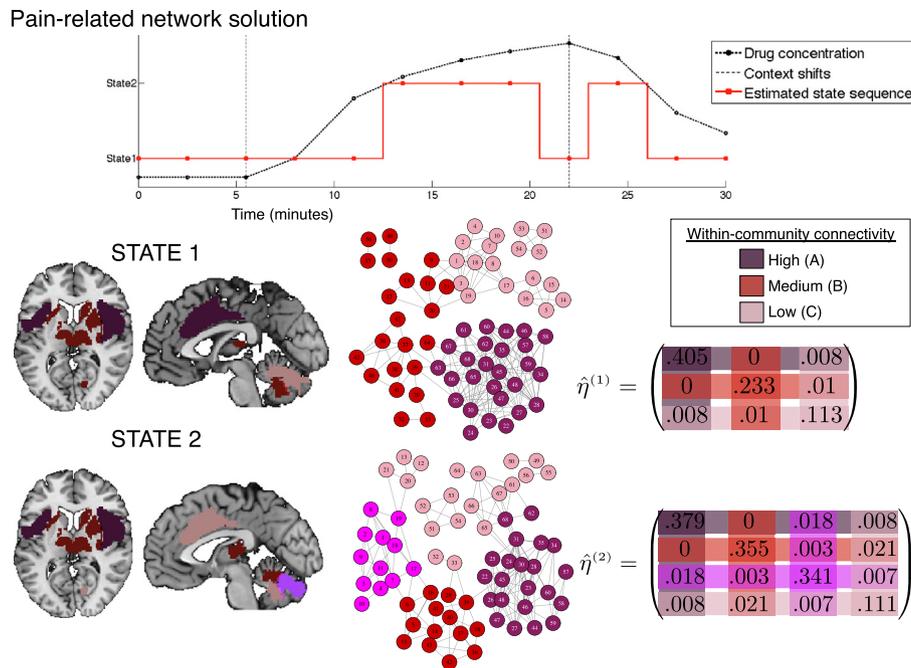
In the working memory network, the model that minimized the BIC also has two states, each with three estimated communities. State 1 roughly corresponds to time periods before and after peak drug intensity, and state 2 has highest probability during peak drug intensity. The two



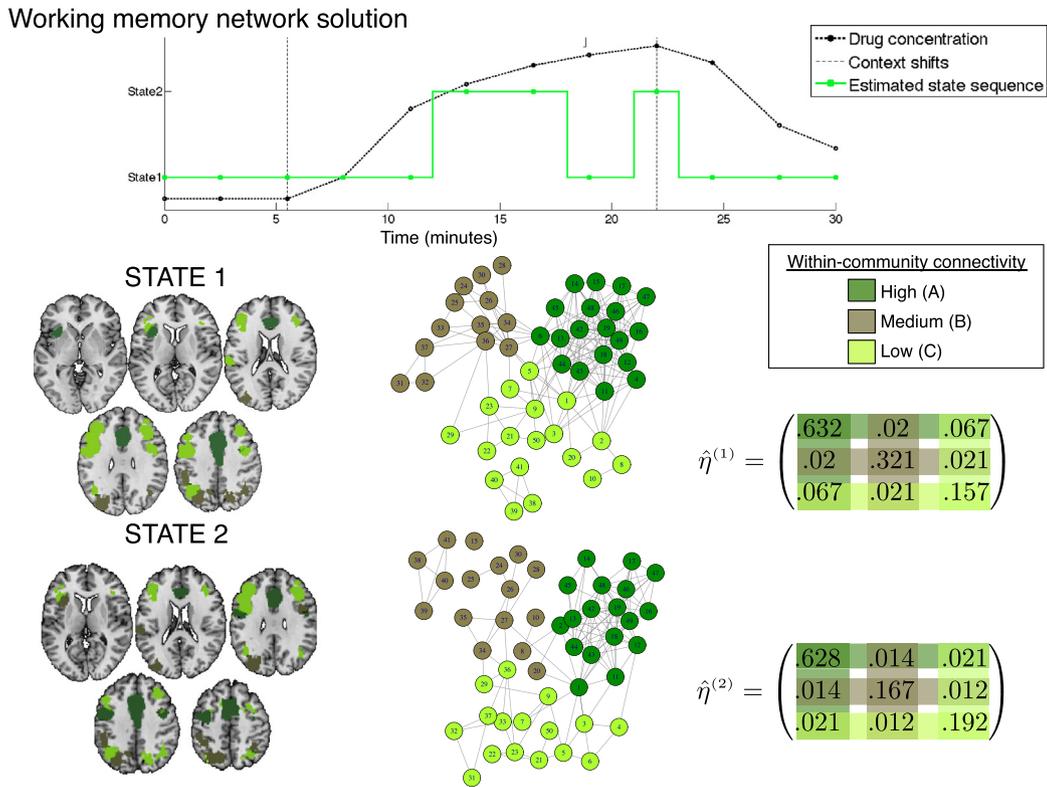
**Fig. 7.** Coherences over time for the three functional networks, averaged across all pairs of regions. The solid horizontal line indicates the time-integrated average coherence. Dashed lines indicate the bootstrapped 95% confidence band for the extreme values (maximum and minimum) over 12 time windows under a hypothesis of time-homogeneous coherence.

states are similar with respect to the organization of regions into three communities (see Table 2). One difference in the estimated network structure between states is the increased within-community connection probabilities in state 1 vs state 2 (see Fig. 9). In both states, the between-community connection probabilities have similar low values. Based on

this result, we see that the decreased overall connectivity displayed in Fig. 7 can largely be attributed to a decrease in connectivity within communities, rather than between communities. Both states show strong within-community coherence in the anterior cingulate (dark green), which shows connectivity with left insula in state 1 and left precentral



**Fig. 8.** For the pain network, the estimated state sequence (top) is shown with the expected drug concentration and context shifts, when participants received information about drug delivery. The network is most likely to be in state 1 during baseline and after drug offset (e.g. during standard pain processing), and state 2 during peak drug concentrations. For each state (middle and bottom rows), the estimated communities, network, and connectivity matrix are shown (see also Table 1). Colors indicate community membership and are color coded by intensity of within-community connectivity (dark red = highest, light pink = lowest), but colors do not indicate belonging to the same community across states. The community with the highest within-community coherence is depicted in dark red (ACC and insula in state 1, insula and cerebellum in state 2). Moderate within-community connectivity is seen in regions in red, which primarily includes thalamus and striatum in both states. Finally, regions with low within-community connectivity are illustrated in pink (cerebellum in state 1, ACC and cerebellum in state 2). Portions of the cerebellum comprised a fourth network with moderate within-community connectivity (depicted in magenta) in state 2.



**Fig. 9.** For the working memory network, the estimated state sequence (top) is shown with the expected drug concentration and context shifts. Similar to the pain processing network, we found a two-state solution where state 1 captured connectivity prior to drug infusion and after offset, whereas state 2 was most likely active during higher drug concentrations. Community membership is reported in detail in Table 2. Regions with high within-community connectivity are depicted in dark green, and include cingulate and insula in State 1, and cingulate and left premotor cortex in state 2. Regions with moderate within-community connectivity are depicted in olive, and regions with low connectivity are illustrated in lime green. The two states differed in organization of frontoparietal networks, such that in State 1, dorsolateral prefrontal cortex was characterized by moderate within-community coherence and parietal regions formed a separate network with lower connectivity, whereas in State 2, we saw reorganization with connectivity between frontal and parietal regions, and lower levels of connectivity in each community.

gyrus in state 2. In addition, there is reorganization of frontoparietal networks between the two states. We see segregation between frontal regions and parietal regions in state 1, with stronger connectivity within the parietal community, whereas these regions become more interconnected in state 2, with the lime green and olive communities consisting of both frontal and parietal regions, and we see moderate to low within-community coherence in each of these networks.

Finally, the optimal model for the emotion network revealed a single state, meaning that emotion network connectivity remained stable despite shifts in drug concentration and context (see Fig. 10). The stable state solution revealed four distinct communities that corresponded to contiguous bilateral regions and showed no between-community connectivity (see Table 3). In order of within-community coherence, the model separately identified the fusiform gyrus (dark blue in Fig. 10), the insula/inferior frontal gyrus (cyan), the amygdala (medium blue), and the thalamus, rostral anterior cingulate cortex, and superior temporal sulcus (light blue).

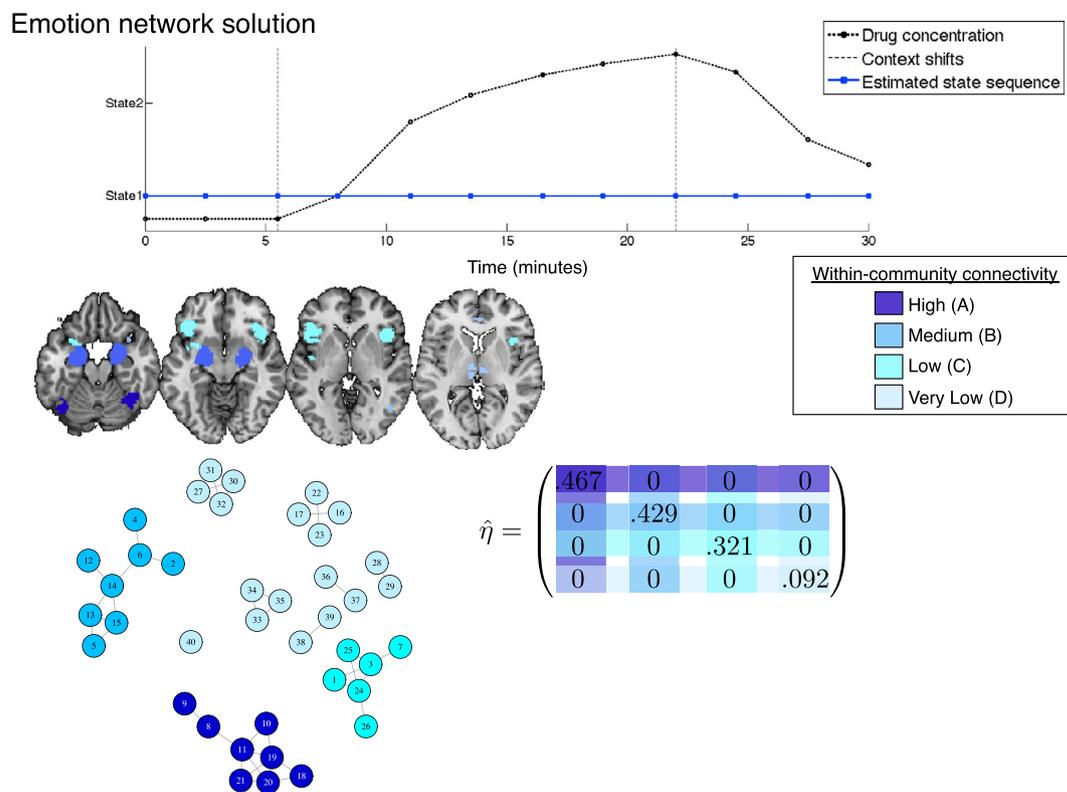
**Discussion**

We have introduced a flexible method that estimates dynamic community structure without *a priori* assumptions on the temporal locations of structural changes. This approach will be useful in situations where changes in the global organization of connectivity in the brain or within functional networks are of interest, rather than changes with respect to a predefined seed region or group of regions. The organization of the network is defined using stochastic block model, which partitions the network into subnetworks and defines connectivity characteristics within and between subnetworks. Temporal dynamics are

modeled using a hidden Markov model for the parameters of the stochastic block model.

SCDS provides a way to detect and describe complex changes in network behavior. In addition to identifying changes in overall connectivity, this method can detect changes that are specific to a subgroup of regions that are not defined *a priori*. As our understanding of the interconnectedness of the brain grows, network modularity and community structure are increasingly important in understanding functional connectivity characteristics. We have applied SDSCS to an experiment with dynamic conditions (drug concentration), but it may fruitfully be applied with some modification to resting state data, or to other task-based studies involving habituation or learning. For resting state data, a modification to the algorithm in which each subject has a unique path through states (and each subject may not enter all states) would be appropriate. In comparison to task-based studies, resting state studies present a potentially more difficult model selection problem as a range of possible values for the number of communities and number of states would be difficult to determine *a priori*. In this application, we use multi-voxel regions defined through meta-analysis, but it could be used in networks defined by PCA or ICA components. For example, a single spatial component which resembles the default mode network could be divided into spatially disjoint regions, either through prior anatomical information or data-driven image segmentation, and these regions could constitute nodes of the network.

In our application, we found evidence for reliable community structure in networks associated with pain processing, working memory, and emotion. Organization of the pain processing and working memory networks was affected by drug concentration and context, while connectivity within the emotion network remained stable across time. This does not



**Fig. 10.** For the emotion network, the estimated state sequence (top) is shown with the expected drug concentration and context shifts. Below, the estimated communities, network, and connectivity matrix are shown (see also Table 3). The emotion network was best characterized by a single state solution, suggesting that connectivity among emotion-related regions did not change substantially in the presence of changing contexts or drug concentrations. In this state, emotion-related regions were organized into four communities that corresponded to bilateral anatomical regions, which ranged from high within-community connectivity (dark blue, fusiform gyrus) to low connectivity (light blue, superior temporal sulcus). Bilateral amygdala and bilateral insula/IFG exhibited moderate levels of within-community connectivity. Unlike the pain processing and working memory networks, we found no evidence of between-community connectivity within the emotion network.

mean that the emotion network was unaffected by the task. Critically, we do know that the magnitude of pain-evoked responses in a number of these regions was influenced by context and drug concentration, based on event-related analyses using the general linear model (Atlas et al., 2012). Our SDCS results show that despite these changes, the overall connectivity among these regions remained stable. In addition, the network solution within the emotion network illustrates that SDCS yields meaningful and sensible solutions, as SDCS identified four communities that were symmetrical bilaterally, segregated based on contiguity, and map onto regions thought to subservise distinct functions (Kober et al., 2008; Lindquist et al., 2012; Phan et al., 2002).

Results in other networks also shed light on the functional significance of this approach. Pain and working memory networks were both characterized by two states of functional connectivity, wherein some connections within communities remained stable across states, and other connections varied. These changes were found in spite of no overall changes in global connectivity, implying that time-varying models like ours are necessary for identifying many of these communities. In both the pain-related network and the working memory network, state 1 captured the organization of connectivity at baseline and as drug concentration decreased after offset. In the pain network, we observed high connectivity in state 1 in a community that consisted primarily of bilateral insula and anterior cingulate, regions with reciprocal connections that are frequently described as part of the same network (Seeley et al., 2007; Menon and Uddin, 2010; Dosenbach et al., 2006). However, during drug administration, when the pain network was in state 2, we saw a reorganization of connectivity structure where insula and anterior cingulate became decoupled. Connectivity within cingulate subregions (parcels) was reduced, whereas connectivity within insula

parcels remained substantial. Previous work suggests that remifentanyl effects are most robust in the insula (Wise et al., 2004; Wise et al., 2002), and the high within-region synchrony we observed may be consistent with these findings. Interestingly, right middle cingulate remained highly connected to insula. Thus in allowing the structure of connectivity to differ over time, we can identify more sophisticated constraints and considerations that relate to connectivity among and between known networks. Similarly, in the working memory network, we observed variations in both the strength and organization of connectivity within regions in the frontoparietal network (Vincent et al., 2008; Corbetta and Shulman, 2002; Wager and Smith, 2003) thought to direct attention and maintain goals. Frontal and parietal regions segregated on the basis of their within-community connectivity in state 1, whereas networks reorganized in state 2. Future work can investigate the functional significance of these changes in connectivity and modularity. One hypothesis is that frontal and parietal regions subservise distinct functions at baseline, and that as the opioid drug influences alertness and arousal and impacts executive function, the network might become more integrated.

As is typical in fMRI studies on functional connectivity, this method reflects several modeling and analysis decisions on the definition of changes and time scales of interest, and the characterization of network nodes and edges. The interpretation of the estimated communities depends of how networks nodes are defined. At the voxel level, network partitioning algorithms based on modularity maximization (Schwartz et al., 2009) or spectral clustering (Craddock et al., 2012) have been used to identify parcels of voxels whose activity is highly correlated. Schwartz et al. (2008) discuss a biological interpretation of voxel-wise community structure, where the degree to which correlation is

concentrated within communities of voxels rather than between communities of voxels reflects the balance of functional segregation vs functional integration in the brain. Communities of voxels may reflect functionally segregated regions of the brain, whereas communities of multi-regions, which are the focus of this paper, can describe groups of regions which tend to activate in concert. In using multi-voxel regions, the choice of parcellation may have a substantive effect on the analysis. Ideally, multi-voxel regions represent spatially contiguous, functionally segregated areas of the brain, with high correlation among voxels within regions under a condition of interest. However, anatomical maps may not necessarily represent the optimal parcellation in terms of functional segregation. Craddock et al. (2012) have shown that parcellations of voxels based on functional connectivity in resting state fMRI bear limited resemblance to commonly used anatomical maps, raising questions about the optimality of standard anatomical atlases for functional connectivity studies. Data-driven techniques also offer ways to identify nodes without potential problems associated with *a priori* specified atlases (Valdés-Sosa et al., 2005; Olier et al., 2013).

The choice of time windows depends on a combination of practical considerations in estimability, experimental conditions, and analysis goals. Choosing different time windows has the potential to affect the results, reflecting both noise in the data and the existence of fluctuations in connectivity on multiple time scales. The results in our application are roughly consistent with analysis using shorter time windows of 50 TRs. Because the sample coherence is unreliable for small samples, we prefer to use somewhat longer time windows, but a similar analysis using correlation rather than coherence as a measure of connectivity could incorporate shorter time windows. Non-overlapping windows are proposed to avoid analytical difficulties inherent in overlapping sliding windows. However, because transitions between states can only be identified at the boundaries of windows, non-overlapping windows necessarily add errors to the estimated location of transition. The expected magnitude of these errors can be estimated based on the window length, as discussed in the section [Network representation](#).

Introducing temporal dynamics into network modeling requires a method for distinguishing random fluctuations in network behavior from meaningful changes. Changes may be complex, operating differently on different subsets of the network and involving multiple model parameters, which makes it difficult to define parametric statistical tests. Sliding window approaches, in which connectivity characteristics are estimated in a series of overlapping time windows, have been used in many studies to estimate functional connectivity dynamics

(refs). Hutchison et al. (2013) discuss the difficulties in interpretation and statistical analysis of sliding window analyses. Essentially, overlapping windows will exhibit smooth fluctuations even under stationary white noise conditions (Robinson et al., 2008), and the dependence structure induced by shared observations between windows makes evaluation of the statistical significance of the observed fluctuations difficult. Rather than estimating smooth fluctuations, Cribben et al. (2012) estimate a partition of the scanning time interval by randomly generating a large number of potential partitions and evaluating their significance based on permutation-based test of BIC values. As an alternative approach to estimating dynamics of community structure, one could use a change point detection framework rather than a state-space model. For example, one might use a stochastic block model to modify a method such as dynamic connectivity regression (Cribben et al., 2013). We use a hidden Markov model that allows for principled model comparison based on model selection criteria, but not explicit hypothesis tests of homogeneity. The hidden Markov stochastic block model is a novel framework for dynamic networks, and its properties will be further explored in future work.

## Conclusion

In summary, SDCS offers a functional connectivity method that can characterize the community structure of brain connectivity networks dynamically, without constraints on the timing of shifts in connectivity. It yields statistically interpretable insights about brain network organization that would not be observed under standard assumptions of static connectivity. Models with varying complexity, both in terms of the number of temporal states and the number of communities within each state, can be evaluated. Finally, our application to opioid analgesia reveals that SDCS provides neuroscientifically meaningful results that complement and extend inferences made based on standard fMRI analyses and lead hypotheses that can guide future work.

## Acknowledgments

This work was supported in part by the National Institute of Mental Health, award number RO1MH076136, and the National Institute of Drug Abuse, award number R01DA027794, of the National Institutes of Health.

## Appendix A

**Table 1**  
Notation.

Parameter	Description
$L$	Length of time window
$\omega$	Frequency argument to coherence
$\gamma$	Cross covariance
$\kappa$	Magnitude-squared coherence
$\bar{\kappa}$	Coherence averaged over frequencies in the [.01 Hz, .15 Hz] band
$S$	Number of nodes in network
$W$	Matrix of coherences between pairs of nodes, averaged over subjects
$A$	Thresholded coherence matrix (network adjacency matrix)
$K$	Number of communities in stochastic block model (SBM)
$x$	Vector of unobserved community labels for each node in SBM
$\alpha$	Vector of prior probabilities of community assignment in SBM
$\eta$	Matrix of probabilities of within and between-community connections in SBM
$T$	Number of time windows
$C$	Number of states in HMM
$\delta$	Initial state probabilities in HMM
$P$	Transition matrix in HMM

Here, we present more detailed information on the state-specific estimated communities for each of the three networks.

**Table 2**  
Community estimation results for the pain processing network.

Region number	Region name	x	y	z	State 1 community	State 2 community
1	Left cerebellum, lobule VIIa Crus I	-31	-66	-29	C	C
2	Right cerebellum, lobule VIIa Crus I	32	-65	-31	C	C
3	Left cerebellum, lobule VIIa Crus I	-22	-76	-32	C	C
4	Right cerebellum, lobule VIIa Crus I	30	-71	-35	C	C
5	Left cerebellum	-6	-48	-18	C	B
6	Left cerebellum, Lobule V	-6	-58	-15	C	B
7	Left cerebellum, lobule VI	-22	-62	-24	C	C
8	Right cerebellum, lobule VI	25	-64	-27	C	C
9	Left cerebellum, lobule VIIa Crus I	-27	-68	-37	B	C
10	Right cerebellum, lobule VIIa Crus I	38	-57	-38	C	C
11	Left cerebellum	-19	-63	-35	B	C
12	Cerebellar vermis	-5	-57	-32	B	D
13	Cerebellar vermis	7	-54	-33	B	D
14	Cerebellar vermis, lobules 1-IV	1	-48	-18	C	B
15	Cerebellar Vermis, Lobules 1-IV	0	-49	-14	C	B
16	Cerebellar vermis, lobule V	-1	-57	-15	C	B
17	Cerebellar Vermis, Lobule VI	1	-67	-17	C	C
18	Cerebellar vermis, lobule VI	2	-74	-22	C	C
19	Cerebellar vermis, lobule VI	2	-68	-27	C	C
20	Cerebellar vermis, lobule VIIa	1	-59	-30	B	D
21	Cerebellar vermis, lobules 1-IV	1	-54	-25	B	D
22	Left inferior frontal gyrus p. opercularis (BA44)	-45	5	10	A	A
23	Right inferior frontal gyrus p. opercularis (BA44)	50	10	12	A	A
24	Right inferior frontal gyrus p. opercularis	46	19	8	A	A
25	Left insula lobe	-35	20	-4	A	A
26	Right insula lobe	38	23	-6	A	A
27	Left rolandic operculum (OP4)	-52	-2	9	A	A
28	Right rolandic operculum (OP3)	51	-13	18	A	A
29	Right ventral striatum	23	5	-10	B	B
30	Left insula lobe	-36	8	4	A	A
31	Right insula lobe	40	3	5	A	A
32	Right hippocampus	17	-5	-14	B	D
33	Right amygdala	22	-2	-11	B	D
34	Right supramarginal gyrus (IPC (PFop))	64	-19	23	A	A
35	Right supramarginal gyrus (IPC (PF))	58	-32	31	A	A
36	Left caudate nucleus	-14	8	15	B	B
37	Right caudate nucleus	15	9	11	B	B
38	Left putamen	-22	7	12	B	B
39	Right putamen	27	3	5	B	B
40	Left thalamus	-13	-2	5	B	B
41	Right thalamus	18	-1	5	B	B
42	Left thalamus	-7	-17	10	B	B
43	Right thalamus	12	-18	10	B	B
44	Right rolandic operculum (OP4)	51	-16	13	A	A
45	Left superior temporal gyrus (OP4)	-49	-5	1	A	A
46	Right superior temporal gyrus (OP1)	56	-22	15	A	A
47	Left temporal pole	-52	7	-1	A	A
48	Right temporal pole	54	6	-3	A	A
49	Left subthalamic nucleus/midbrain	-13	-11	-10	B	D
50	Left midbrain	-11	-9	-13	B	D
51	Right middle frontal gyrus	37	41	4	C	D
52	Right middle frontal gyrus	37	42	1	C	D
53	Right middle frontal gyrus	42	40	4	C	D
54	Right middle frontal gyrus	40	41	2	C	D
55	Right calcarine gyrus (BA17)	14	-74	9	B	D
56	Right calcarine gyrus (BA17)	13	-74	6	B	D
57	Left supramarginal gyrus (IPC (PFop))	-61	-26	19	A	A
58	Left supramarginal gyrus (OP1)	-63	-24	22	A	A
59	Left supramarginal gyrus (IPC (PFop))	-61	-28	23	A	A
60	Left superior temporal gyrus (IPC (PFop))	-59	-29	19	A	A
61	Left SMA (BA6)	0	7	51	A	D
62	Right SMA (BA6)	8	5	54	A	A
63	Left superior medial gyrus	-1	20	43	A	D
64	Right SMA	7	19	48	B	D
65	Left anterior cingulate cortex	-3	22	28	A	D
66	Right anterior cingulate cortex	7	24	27	A	D
67	Left middle cingulate cortex	-3	7	41	A	D
68	Right middle cingulate cortex	6	7	41	A	A

**Table 3**  
Community estimation results for the working memory network.

Region number	Region name	x	y	z	State 1 community1	State 2 community2
1	Left precentral gyrus	-43	-3	42	C	A
2	Right precentral gyrus	42	-3	41	C	A
3	Left superior frontal gyrus	-18	1	59	C	C
4	Right superior frontal gyrus	21	2	58	A	C
5	Left middle frontal gyrus	-38	21	40	C	C
6	Right middle frontal gyrus	32	5	52	A	C
7	Left precentral gyrus	-43	6	30	C	C
8	Right precentral gyrus	42	5	32	C	B
9	Left inferior frontal gyrus p. triangularis	-43	23	22	C	C
10	Right inferior frontal gyrus p. opercularis	41	10	29	C	B
11	Left SMA (BA6)	-5	5	56	A	A
12	Right SMA (BA6)	8	3	56	A	A
13	Left SMA	-4	18	46	A	A
14	Right SMA (BA6)	4	18	48	A	A
15	Left insula lobe	-33	16	4	A	B
16	Left middle cingulate cortex	-3	16	33	A	A
17	Right middle cingulate cortex	4	14	32	A	A
18	Left SMA	-4	6	44	A	A
19	Right middle cingulate cortex (BA6)	6	10	43	A	A
20	Left postcentral gyrus (BA 4a)	-52	-9	41	C	B
21	Right middle frontal gyrus	36	31	32	C	C
22	Right middle frontal gyrus	40	20	35	C	C
23	Right inferior frontal gyrus p. triangularis	42	25	25	C	C
24	Left superior occipital gyrus	-22	-76	35	B	B
25	Left middle occipital gyrus	-32	-78	30	B	B
26	Left superior parietal lobule	-19	-70	51	B	B
27	Left inferior parietal lobule (hIP1)	-40	-48	45	B	B
28	Left inferior parietal lobule (hIP2)	-48	-42	37	B	B
29	Left angular gyrus (IPC (PGa))	-39	-62	38	C	C
30	Left precuneus	-11	-66	47	B	B
31	Right precuneus	19	-70	39	B	C
32	Right superior occipital gyrus	24	-71	41	B	C
33	Right middle occipital gyrus	36	-67	37	B	C
34	Right inferior parietal lobule (hIP3)	38	-53	48	B	B
35	Right inferior parietal lobule (hIP2)	42	-44	45	B	B
36	Right angular gyrus	39	-61	41	B	C
37	Right precuneus	16	-68	43	B	C
38	Left supramarginal gyrus (IPC (Pfp))	-61	-26	19	C	B
39	Left supramarginal gyrus (OP1)	-63	-24	22	C	B
40	Left supramarginal gyrus (IPC (PFop))	-61	-28	23	C	B
41	Left superior temporal gyrus (IPC (PFop))	-59	-29	19	C	B
42	Left SMA (BA6)	0	7	51	A	A
43	Right SMA (BA6)	8	5	54	A	A
44	Left superior medial gyrus	-1	20	43	A	A
45	Right SMA	7	19	48	A	A
46	Left anterior cingulate cortex	-3	22	28	A	A
47	Right anterior cingulate cortex	7	24	27	A	A
48	Left middle cingulate cortex	-3	7	41	A	A
49	Right middle cingulate cortex	6	7	41	A	A
50	Right middle frontal gyrus	37	43	28	C	C

**Table 4**  
Community estimation results for the emotion network.

Region number	Region name	x	y	z	State 1 community1
1	Left Inferior frontal gyrus p. orbitalis	-39	23	-8	C
2	Left extended amygdala/ventral striatum	-20	2	-11	B
3	Left insula lobe	-35	6	-8	C
4	Left hippocampus	-22	-12	-14	B
5	Left amygdala (superficial)	-19	-5	-23	B
6	Left amygdala (basolateral)	-23	-6	-13	B
7	Left insula lobe	-41	-7	-5	C
8	Right fusiform gyrus	41	-54	-17	A
9	Right fusiform gyrus	46	-52	-20	A
10	Right cerebellum	41	-55	-26	A
11	Right fusiform gyrus	37	-55	-22	A
12	Right hippocampus	24	-12	-13	B
13	Right amygdala (superficial)	22	-4	-21	B
14	Right amygdala (basolateral)	26	-4	-13	B
15	Right amygdala	28	3	-19	B
16	Right temporal pole	38	12	-23	D
17	Right temporal pole	40	7	-27	D
18	Left fusiform gyrus	-41	-62	-16	A

(continued on next page)

Table 4 (continued)

Region number	Region name	x	y	z	State 1 community1
19	Left cerebellum (lobule VIIIa Crus I)	−40	−66	−22	A
20	Left fusiform gyrus	−37	−63	−19	A
21	Left cerebellum	−33	−73	−19	A
22	Right insula lobe	41	11	−17	D
23	Right insula lobe	43	15	−17	D
24	Right insula lobe	42	22	−8	C
25	Right insula lobe	42	14	−1	C
26	Right insula lobe	46	17	−13	C
27	Right anterior cingulate cortex	3	35	−5	D
28	Right middle temporal gyrus (hOC5 (V5))	47	−64	2	D
29	Right middle temporal gyrus (hOC5 (V5))	47	−67	−1	D
30	Right anterior cingulate cortex	1	34	−3	D
31	Left anterior cingulate cortex	1	32	−2	D
32	Right anterior cingulate cortex	2	32	−3	D
33	Right anterior cingulate cortex	4	40	2	D
34	Left anterior cingulate cortex	0	40	3	D
35	Right anterior cingulate cortex	4	38	5	D
36	Left thalamus	−5	−16	6	D
37	Right thalamus	8	−19	4	D
38	Left middle occipital gyrus	−48	−74	8	D
39	Left middle temporal gyrus	−49	−70	9	D

## References

- Airoldi, E.M., Blei, D.M., Fienberg, S.E., Xing, E.P., 2008. Mixed membership stochastic blockmodels. *J. Mach. Learn. Res.* 9, 1981–2014.
- Allen, E.A., Damaraju, E., Plis, S.M., Erhardt, E.B., Eichele, T., Calhoun, V.D., 2012. Tracking whole-brain connectivity dynamics in the resting state. *Cereb. Cortex: bhs352*.
- Ambrose, C., Matias, C., 2012. New consistent and asymptotically normal parameter estimates for random-graph mixture models. *J. R. Stat. Soc. Ser. B (Stat. Methodol.)* 74, 3–35.
- Atlas, L.Y., Wager, T.D., 2013. Expectancies and beliefs: insights from cognitive neuroscience. *The Oxford Handbook of Cognitive Neuroscience* vol. 2. The Cutting Edges, pp. 359–381.
- Atlas, L.Y., Bolger, N., Lindquist, M.A., Wager, T.D., 2010. Brain mediators of predictive cue effects on perceived pain. *J. Neurosci.* 30, 12964–12977.
- Atlas, L.Y., Whittington, R.A., Lindquist, M.A., Wielgosz, J., Sonty, N., Wager, T.D., 2012. Dissociable influences of opiates and expectations on pain. *J. Neurosci.* 32, 8053–8064.
- Bassett, D.S., Wymbs, N.F., Porter, M.A., Mucha, P.J., Carlson, J.M., Grafton, S.T., 2011. Dynamic reconfiguration of human brain networks during learning. *Proc. Natl. Acad. Sci. U. S. A.* 108, 7641–7646.
- Beckmann, C.F., DeLuca, M., Devlin, J.T., Smith, S.M., 2005. Investigations into resting-state connectivity using independent component analysis. *Philos. Trans. R. Soc. B Biol. Sci.* 360, 1001–1013.
- Bickel, P.J., Chen, A., 2009. A nonparametric view of network models and Newman–Girvan and other modularities. *Proc. Natl. Acad. Sci.* 106, 21068–21073.
- Biernacki, C., Celeux, G., Govaert, G., 2000. Assessing a mixture model for clustering with the integrated completed likelihood. *IEEE Trans. Pattern Anal. Mach. Intell.* 22, 719–725.
- Boveroux, P., Vanhauwenhuysse, A., Bruno, M.-A., Noirhomme, Q., Lauwrick, S., Luxen, A., Degueldre, C., Plenevaux, A., Schnakers, C., Phillips, C., et al., 2010. Breakdown of within- and between-network resting state functional magnetic resonance imaging connectivity during propofol-induced loss of consciousness. *Anesthesiology* 113, 1038–1053.
- Bowman, F.D., Zhang, L., Derado, G., Chen, S., 2012. Determining functional connectivity using fMRI data with diffusion-based anatomical weighting. *NeuroImage* 62, 1769–1779.
- Bullmore, E., Sporns, O., 2009. Complex brain networks: graph theoretical analysis of structural and functional systems. *Nat. Rev. Neurosci.* 10, 186–198.
- Cazabet, R., Amblard, F., Hanachi, C., 2010. Detection of overlapping communities in dynamical social networks. *Social Computing (SocialCom)*, 2010 IEEE Second International Conference on. IEEE, pp. 309–314.
- Celisse, A., Daudin, J.-J., Pierre, L., et al., 2012. Consistency of maximum-likelihood and variational estimators in the stochastic block model. *Electron. J. Stat.* 6, 1847–1899.
- Chakrabarti, D., Kumar, R., Tomkins, A., 2006. Evolutionary clustering. *Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, pp. 554–560.
- Chen, J., Li, P., Fu, Y., 2012. Inference on the order of a normal mixture. *J. Am. Stat. Assoc.* 107, 1096–1105.
- Chen, Y., Kawadia, V., Urgaonkar, R., 2013. Detecting overlapping temporal community structure in time-evolving networks (arXiv preprint arXiv:1303.7226).
- Côme, E., Latouche, P., 2013. Model selection and clustering in stochastic block models with the exact integrated complete data likelihood (ArXiv e-prints).
- Corbetta, M., Shulman, G.L., 2002. Control of goal-directed and stimulus-driven attention in the brain. *Nat. Rev. Neurosci.* 3, 201–215.
- Costafreda, S.G., 2009. Pooling fMRI data: meta-analysis, mega-analysis and multi-center studies. *Front. Neuroinformatics* 3.
- Craddock, R., James, A., Holtzenheimer III, P., Hu, X., Mayberg, H., 2012. Community structure in networks of functional connectivity: resolving functional organization in the rat brain with pharmacological MRI. *Hum. Brain Mapp.* 33, 1914–1928.
- Cribben, I., Haraldsdottir, R., Atlas, L.Y., Wager, T.D., Lindquist, M.A., 2012. Dynamic connectivity regression: determining state-related changes in brain connectivity. *NeuroImage* 61, 907–920.
- Cribben, I., Wager, T., Lindquist, M., 2013. Detecting functional connectivity change points for single-subject fMRI data. *Front. Comput. Neurosci.* 7.
- Damoiseaux, J., Rombouts, S., Barkhof, F., Scheltens, P., Stam, C., Smith, S.M., Beckmann, C., 2006. Consistent resting-state networks across healthy subjects. *Proc. Natl. Acad. Sci.* 103, 13848–13853.
- Daudin, J.J., Picard, F., Robin, S., 2008. A mixture model for random graphs. *Stat. Comput.* 18, 173–183.
- De Luca, M., Beckmann, C., De Stefano, N., Matthews, P., Smith, S.M., 2006. fMRI resting state networks define distinct modes of long-distance interactions in the human brain. *NeuroImage* 29, 1359–1367.
- Dosenbach, N.U., Visscher, K.M., Palmer, E.D., Miezin, F.M., Wenger, K.K., Kang, H.C., Burgund, E.D., Grimes, A.L., Schlaggar, B.L., Petersen, S.E., 2006. A core system for the implementation of task sets. *Neuron* 50, 799–812.
- Eickhoff, S.B., Stephan, K.E., Mohlberg, H., Grefkes, C., Fink, G.R., Amunts, K., Zilles, K., 2005. A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. *NeuroImage* 25, 1325–1335.
- Elliott, R.J., Aggoun, L., Moore, J.B., 1995. *Hidden Markov Models*. Springer.
- Fair, D.A., Cohen, A.L., Power, J.D., Dosenbach, N.U.F., Church, J.A., Miezin, F.M., Schlaggar, B.L., Petersen, S.E., 2009. Functional brain networks develop from a “local to distributed” organization. *PLoS Comput. Biol.* 5.
- Ferrari, L., Veer, I.M., Baerends, E., van Tol, M.J., Renken, R.J., van der Wee, N.J.A., Veltman, D.J., Aleman, A., Zitman, F.G., Penninx, B., van Buchem, M.A., Reiber, J.H.C., Rombouts, S., Milles, J., 2009. Hierarchical functional modularity in the resting-state human brain. *Hum. Brain Mapp.* 30, 2220–2231.
- Fienberg, S.E., Wasserman, S.S., 1981. Categorical data analysis of single sociometric relations. *Sociol. Methodol.* 156–192.
- Fortunato, S., 2010. Community detection in graphs. *Phys. Rep.* 486, 75–174.
- Gazal, S., Daudin, J.-J., Robin, S., 2012. Accuracy of variational estimates for random graph mixture models. *J. Stat. Comput. Simul.* 82, 849–862.
- Greicius, M., 2008. Resting-state functional connectivity in neuropsychiatric disorders. *Curr. Opin. Neurol.* 21, 424–430.
- Hoff, P., Raftery, A., Handcock, M., 2002. Latent space approaches to social network analysis. *J. Am. Stat. Assoc.* 79.
- Holland, P.W., Laskey, K.B., Leinhardt, S., 1983. Stochastic blockmodels: first steps. *Soc. Networks* 5, 109–137.
- Honey, G., Bullmore, E., 2004. Human pharmacological MRI. *Trends Pharmacol. Sci.* 25, 366–374.
- Hutchison, R.M., Womelsdorf, T., Allen, E.A., Bandettini, P.A., Calhoun, V.D., Corbetta, M., Penna, S.D., Duyn, J., Glover, G., Gonzalez-Castillo, J., et al., 2013. Dynamic functional connectivity: promises, issues, and interpretations. *NeuroImage* 80, 360–378.
- Jafri, M.J., Pearlson, G.D., Stevens, M., Calhoun, V.D., 2008. A method for functional network connectivity among spatially independent resting-state components in schizophrenia. *NeuroImage* 39, 1666–1681.
- Janoos, F., Brown, G., Mórocz, I.A., Wells III, W.M., 2013. State-space analysis of working memory in schizophrenia: an fBIRN study. *Psychometrika* 78, 279–307.
- Jones, D.T., Vemuri, P., Murphy, M.C., Gunter, J.L., Senjem, M.L., Machulda, M.M., Przybelski, S.A., Gregg, B.E., Kantarci, K., Knopman, D.S., Boeve, B.F., Petersen, R.C., Jack Jr., C.R., 2012. Non-stationarity in the resting brains modular architecture. *PLoS ONE* 7, e39731.
- Jordan, M.I., Ghahramani, Z., Jaakkola, T.S., Saul, L.K., 1999. An introduction to variational methods for graphical models. *Mach. Learn.* 37, 183–233.

- Karrer, B., Newman, M.E., 2011. Stochastic blockmodels and community structure in networks. *Phys. Rev. E* 83, 016107.
- Kawadia, V., Sreenivasan, S., 2012. Sequential detection of temporal communities by entanglement confinement. *Sci. Rep.* 2.
- Kinnison, J., Padmala, S., Choi, J.-M., Pessoa, L., 2012. Network analysis reveals increased integration during emotional and motivational processing. *J. Neurosci.* 32, 8361–8372.
- Kober, H., Barrett, L.F., Joseph, J., Bliss-Moreau, E., Lindquist, K., Wager, T.D., 2008. Functional grouping and cortical–subcortical interactions in emotion: a meta-analysis of neuroimaging studies. *NeuroImage* 42, 998–1031.
- Leonardi, N., Richiardi, J., Gschwind, M., Simioni, S., Annoni, J.-M., Schlupe, M., Vuilleumier, P., Ville, D.V.D., 2013. Principal components of functional connectivity: a new approach to study dynamic brain connectivity during rest. *NeuroImage* 83, 937–950.
- Li, P., Chen, J., 2010. Testing the order of a finite mixture. *J. Am. Stat. Assoc.* 105.
- Lindquist, K.A., Wager, T.D., Kober, H., Bliss-Moreau, E., Barrett, L.F., 2012. The brain basis of emotion: a meta-analytic review. *Behav. Brain Sci.* 35, 121–143.
- Meng, X.-L., Rubin, D.B., 1993. Maximum likelihood estimation via the ECM algorithm: a general framework. *Biometrika* 80, 267–278.
- Menon, V., Uddin, L.Q., 2010. Saliency, switching, attention and control: a network model of insula function. *Brain Struct. Funct.* 214, 655–667.
- Meunier, D., Achard, S., Morcom, A., Bullmore, E., 2009. Age-related changes in modular organization of human brain functional networks. *NeuroImage* 44, 715–723.
- Mucha, P.J., Richardson, T., Macon, K., Porter, M.A., Onnela, J.-P., 2010. Community structure in time-dependent, multiscale, and multiplex networks. *Science* 328, 876–878.
- Newman, M., 2006. Modularity and community structure in networks. *Proc. Natl. Acad. Sci.* 103, 8577–8582.
- Nguyen, N.P., Dinh, T.N., Xuan, Y., Thai, M.T., 2011. Adaptive algorithms for detecting community structure in dynamic social networks. *INFOCOM, 2011 Proceedings IEEE. IEEE*, pp. 2282–2290.
- Nowicki, K., Snijders, T.A.B., 2001. Estimation and prediction for stochastic blockstructures. *J. Am. Stat. Assoc.* 96, 1077–1087.
- Olier, I., Trujillo-Barreto, N.J., El-Dereby, W., 2013. A switching multi-scale dynamical network model of EEG/MEG. *NeuroImage* 83, 262–287.
- Phan, K.L., Wager, T., Taylor, S.F., Liberzon, I., 2002. Functional neuroanatomy of emotion: a meta-analysis of emotion activation studies in PET and fMRI. *NeuroImage* 16, 331–348.
- Rabiner, L., 1989. A tutorial on hidden Markov models and selected applications in speech recognition. *Proc. IEEE* 77, 257–286.
- Rissanen, J., 2007. *Information and complexity in statistical modeling*. Springer.
- Robinson, L.F., de la Peña, V.H., Kushnir, Y., 2008. Detecting shifts in correlation and variability with application to ENSO-monsoon rainfall relationships. *Theor. Appl. Climatol.* 94, 215–224.
- Rubinov, M., Sporns, O., 2010. Complex network measures of brain connectivity: uses and interpretations. *NeuroImage* 52, 1059–1069.
- Salimi-Khorshidi, G., Smith, S.M., Keltner, J.R., Wager, T.D., Nichols, T.E., 2009. Meta-analysis of neuroimaging data: a comparison of image-based and coordinate-based pooling of studies. *NeuroImage* 45, 810–823.
- Schwarz, A.J., Gozzi, A., Reese, T., Bifone, A., 2007. In vivo mapping of functional connectivity in neurotransmitter systems using pharmacological MRI. *NeuroImage* 34, 1627–1636.
- Schwartz, A., Gozzi, A., Bifone, A., 2008. Community structure and modularity in networks of correlated brain activity. *Magn. Reson. Imaging* 26, 914–920.
- Schwartz, A., Gozzi, A., Bifone, A., 2009. Community structure in networks of functional connectivity: resolving functional organization in the rat brain with pharmacological MRI. *NeuroImage* 47, 302–311.
- Scott, S.L., 2002. Bayesian methods for hidden Markov models. *J. Am. Stat. Assoc.* 97.
- Seeley, W.W., Menon, V., Schatzberg, A.F., Keller, J., Glover, G.H., Kenna, H., Reiss, A.L., Greicius, M.D., 2007. Dissociable intrinsic connectivity networks for salience processing and executive control. *J. Neurosci.* 27, 2349–2356.
- Shen, X., Papademetris, X., Constable, R.T., 2010. Graph-theory based parcellation of functional subunits in the brain from resting-state fMRI data. *NeuroImage* 50, 1027–1035 (times Cited: 27 Shen, X., Papademetris, X., Constable, R. T. Shen, Xilin /F-4349-2011 27).
- Simpson, S.L., Bowman, F.D., Laurienti, P.J., et al., 2013. Analyzing complex functional brain networks: fusing statistics and network science to understand the brain. *Stat. Surv.* 7, 1–36.
- Smith, S.M., Vidaurre, D., Beckmann, C.F., Glasser, M.F., Jenkinson, M., Miller, K.L., Nichols, T.E., Robinson, E.C., Salimi-Khorshidi, G., Woolrich, M.W., Barch, D.M., UÄYurbil, K., Essen, D.C.V., Smith, S.M., Vidaurre, D., Beckmann, C.F., Glasser, M.F., Jenkinson, M., Miller, K.L., Nichols, T.E., Robinson, E.C., Salimi-Khorshidi, G., Woolrich, M.W., Barch, D.M., UÄYurbil, K., Essen, D.C.V., 2013. Functional connectomics from resting-state fMRI. *Trends Cogn. Sci.* 17, 666–682 (special Issue: The Connectome).
- Snijders, T., Nowicki, K., 1997. Estimation and prediction for stochastic blockmodels for graphs with latent block structure. *J. Classif.* 75.
- Sporns, O., Chialvo, D.R., Kaiser, M., Hilgetag, C.C., 2004. Organization, development and function of complex brain networks. *Trends Cogn. Sci.* 8, 418–425.
- Sun, F.T., Miller, L.M., D'Esposito, M., 2004. Measuring interregional functional connectivity using coherence and partial coherence analyses of fMRI data. *NeuroImage* 21, 647–658.
- Valdés-Sosa, P.A., Sánchez-Bornot, J.M., Lage-Castellanos, A., Vega-Hernández, M., Bosch-Bayard, J., Melie-García, L., Canales-Rodríguez, E., 2005. Estimating brain functional connectivity with sparse multivariate autoregression. *Philos. Trans. R. Soc. B Biol. Sci.* 360, 969–981.
- Van Dijk, K.R., Sabuncu, M.R., Buckner, R.L., 2012. The influence of head motion on intrinsic functional connectivity MRI. *NeuroImage* 59, 431–438.
- Vincent, J.L., Kahn, I., Snyder, A.Z., Raichle, M.E., Buckner, R.L., 2008. Evidence for a frontoparietal control system revealed by intrinsic functional connectivity. *J. Neurophysiol.* 100, 3328–3342.
- Volinsky, C.T., Raftery, A.E., 2000. Bayesian information criterion for censored survival models. *Biometrics* 56, 256–262.
- Von Luxburg, U., 2007. A tutorial on spectral clustering. *Stat. Comput.* 17, 395–416.
- Wager, T.D., Smith, E.E., 2003. Neuroimaging studies of working memory. *Cogn. Affect. Behav. Neurosci.* 3, 255–274.
- Wager, T.D., Lindquist, M., Kaplan, L., 2007. Meta-analysis of functional neuroimaging data: current and future directions. *Soc. Cogn. Affect. Neurosci.* 2, 150–158.
- Wang, S., Tang, M., 2004. Exact confidence interval for magnitude-squared coherence estimates. *IEEE Signal Process. Lett.* 11, 326–329.
- Wise, R.G., Rogers, R., Painter, D., Bantick, S., Ploghaus, A., Williams, P., Rapeport, G., Tracey, I., 2002. Combining fMRI with a pharmacokinetic model to determine which brain areas activated by painful stimulation are specifically modulated by remifentanyl. *NeuroImage* 16, 999–1014.
- Wise, R.G., Williams, P., Tracey, I., 2004. Using fMRI to quantify the time dependence of remifentanyl analgesia in humans. *Neuropsychopharmacology* 29, 626–635.
- Zuo, X.N., Ehmke, R., Mennes, M., Imperati, D., Castellanos, F.X., Sporns, O., Milham, M.P., 2012. Network centrality in the human functional connectome. *Cereb. Cortex* 22, 1862–1875.