Dynamic expression of brain functional systems disclosed by fine-scale analysis of edge time series

Olaf Sporns¹, ², ³, ⁴, Joshua Faskowitz¹, ², Andreia Sofia Teixeira³, ⁵, ⁶, Sarah A. Cutts¹, ², and Richard F. Betzel¹, ², ³, ⁴

¹Department of Psychological and Brain Sciences, Indiana University, Bloomington, IN, USA
²Program in Neuroscience, Indiana University, Bloomington, IN, USA
³Network Science Institute, Indiana University, Bloomington, IN, USA
⁴Cognitive Science Program, Indiana University, Bloomington, IN, USA
⁵Center for Social and Biomedical Complexity, School of Informatics, Computing, and Engineering, Indiana University, Bloomington, IN, USA
⁶INESC-ID, Lisboa, Portugal

Keywords: Connectome, Functional MRI, Resting state, Network dynamics, Graph theory

ABSTRACT

Functional connectivity (FC) describes the statistical dependence between neuronal populations or brain regions in resting-state fMRI studies and is commonly estimated as the Pearson correlation of time courses. Clustering or community detection reveals densely coupled sets of regions constituting resting-state networks or functional systems. These systems manifest most clearly when FC is sampled over longer epochs but appear to fluctuate on shorter timescales. Here, we propose a new approach to reveal temporal fluctuations in neuronal time series. Unwrapping FC signal correlations yields pairwise co-fluctuation time series, one for each node pair or edge, and allows tracking of fine-scale dynamics across the network. Co-fluctuations partition the network, at each time step, into exactly two communities. Sampled over time, the overlay of these bipartitions, a binary decomposition of the original time series, very closely approximates functional connectivity. Bipartitions exhibit characteristic spatiotemporal patterns that are reproducible across participants and imaging runs, capture individual differences, and disclose fine-scale temporal expression of functional systems. Our findings document that functional systems appear transiently and intermittently, and that FC results from the overlay of many variable instances of system expression. Potential applications of this decomposition of functional connectivity into a set of binary patterns are discussed.

AUTHOR SUMMARY

Numerous studies of functional connectivity have revealed densely coupled sets of brain regions corresponding to resting-state networks or functional systems. Prior work suggests that functional connectivity fluctuates over time. Here, we extend those studies by suggesting that functional connectivity can be decomposed into a set of momentary network states, with each one partitioning the network into exactly two clusters or communities. We show that these bipartitions exhibit characteristic spatiotemporal patterns that are reproducible across participants and imaging runs, and can capture individual differences. Our decomposition approach discloses fine-scale dynamics of functional systems, and reveals that functional systems coalesce and dissolve at different times and on fast timescales. Numerous applications and extensions of the approach are discussed.
INTRODUCTION

Modern network neuroscience conceptualizes the brain as an interconnected dynamic multiscale system (Bassett & Sporns, 2017; Betzel & Bassett, 2017; Bullmore & Sporns, 2009). At the level of the whole brain, anatomical projections between brain regions shape spontaneous dynamics and constrain the brain’s momentary responses to changes in input, internal state, and environmental demand (Honey et al., 2009; Suárez, Markello, Betzel, & Misis, 2020). The resulting statistical dependencies among regional time courses are generally described as “functional connectivity,” quantified with a variety of bivariate metrics (Buckner, Krienen, & Yeo, 2013; Friston, 2011). In extant fMRI research, the Pearson correlation of blood oxygenation level dependent (BOLD) time courses remains in wide use, generally applied to long epochs of resting or task-evoked responses. The resulting correlation matrix, representing a functional network (Power et al., 2011), or “functional connectome” (Biswal et al., 2010), provides a summary representation of the system’s pairwise dependencies.

Functional connectivity, measured during the resting state, exhibits highly consistent patterns across imaging sessions (Horien, Shen, Scheinost, & Constable, 2019), participant cohorts (Dadi et al., 2019), and parcellations (Arslan et al., 2018), while also expressing individual differences (Marek et al., 2019), state-dependent changes (Betzel et al., 2020), and genetic associations (Demeter et al., 2020). Among its characteristic network features is community structure, the presence of reproducible modules consisting of regions that are internally densely coupled, reflecting their coherent and correlated activity over time. These intrinsic connectivity networks (Damoiseaux et al., 2006), or resting-state networks (RSNs), have become enshrined in the cognitive neuroscience literature, providing a fundamental taxonomy and topographic reference frame for mapping brain/behavior relations (Ito et al., 2017; Uddin, Yeo, & Spreng, 2019). Canonical sets of RSNs have been proposed (Power et al., 2011; Yeo et al., 2011), and their consistent spatial layout has been shown to reflect patterns of coactivation in task-driven fMRI activation studies (Laird et al., 2011). As internally coherent, coactivated, co-fluctuating systems they may be taken to represent building blocks of the brain’s cognitive architecture that supports specialized brain function. RSNs are not, however, sharply delineated. As has been noted in early mapping studies (Fox et al., 2005), and later revealed with data-driven community detection and clustering approaches (Power et al., 2011; Yeo et al., 2011), functional connectivity exhibits communities at multiple spatial scales, arranged in an overlapping nested hierarchy (Akiki & Abdallah, 2019; Doucet et al., 2011). Furthermore, most cognitive processes do not occur within single RSNs, and indeed may require breaking modular boundaries and dynamic reconfiguration of neural resources, including network nodes and edges (Alavash, Tune, & Olesker, 2019; Braun et al., 2015; Petersen & Sporns, 2015).

Functional systems or RSNs manifest in long-time samples of resting brain activity; indeed, their reproducibility across imaging sessions sharply increases with the length of time samples, leveling off at timescales of tens of minutes (Gordon et al., 2017). This raises the question of whether RSNs manifest only on longer timescales or whether they also “exist” at shorter timescales. Recent studies of time-varying functional connectivity (tvFC; Heitmann & Breakspear, 2018; Kucyi, Tambini, Sadaghiani, Keilholz, & Cohen, 2018; Lurie et al., 2020) have addressed the issue, approaching fine temporal structure and dynamics of FC through the use of shorter data samples, such as sliding windows or instantaneous coactivation patterns that result in temporally ordered sequences of functional networks and network states (Allen et al., 2014; Fagheri, Iraj, Damaraju, Turner, & Calhoun, 2020; Liu & Dyun, 2013; Park, Friston, Pae, Park, & Razi, 2018; Preti, Bolton, & Van De Ville, 2017; Shakil, Lee, & Keilholz, 2016). These studies have provided evidence for significant fluctuations of functional connections and network communities on timescales of tens of seconds to minutes (Grandjean et al., 2018).
Dynamic expression of brain functional systems

et al., 2017; Hilger, Fukushima, Sporns, & Fiebach, 2020; Liao, Cao, Xia, & He, 2017; Liégeois et al., 2019; Vohryzek, Deco, Cessac, Krügelbach, & Cabral, 2020). These fluctuations involve a shifting balance between segregated (high modularity) and integrated (low modularity) states (Betzel, Fukushima, He, Zuo, & Sporns, 2016; Zalesky, Fornito, Cocchi, Gollo, & Breakspear, 2014), with episodes of high modularity exhibiting consistent topology across time (Fukushima et al., 2018) and subject to modulation by internal state, task performance, or behavior (Cohen, 2018; Shine et al., 2016).

Recently, we suggested a new approach to functional connectivity, by focusing on the dynamics and networks formed by “edge time series” (Esfahlan et al., 2020; Faskowitz, Esfahlan, Jo, Sporns, & Betzel, 2020; Jo, Faskowitz, Esfahlan, Sporns, & Betzel, 2020). The approach unwraps time-averaged FC into time series of co-fluctuating signals on network edges resolved at the timescale of single frames in MRI acquisition, thus allowing inspection of network dynamics at fine timescales. Here we build on this approach and show that a simple proxy for the resulting framewise community structure, expressed as a set of bipartitions of the network into two positively co-fluctuating ensembles of nodes, represents a compact decomposition of the functional connectivity. Examining the patterns and frequencies of these bipartitions allows addressing several issues related to FC dynamics. How well do bipartitions, sampled at fine-scale temporal resolution, represent “classic” system-level architecture as derived from long-time FC? How does the community structure of single frames combine into the community structure of FC? Do systems, as coherent blocks, manifest continuously or do they appear transiently at short timescales? Do systems differ in their patterns of “functional expression” across time?

RESULTS

Extraction of Bipartitions From Time Series

This expository section introduces the basic constructs employed in this study (Figure 1); for more detail see the Methods section.

Starting from node time series (BOLD activations), the cross-correlation between each pair of nodes defines their linear statistical dependence (Figure 1A). The correlations of all node pairs within a given system are that system’s functional connectivity. Employing a standard definition of the cross-correlation, the average of the products of the standard scores of the two variables, yields scalar correlation estimates. Omitting the averaging step retains the sum-mands, corresponding to a temporal unwrapping of the scalar correlation estimates into vectors (time series) along each edge (corresponding to the link between a node pair). These edge time series represent co-fluctuations of node pairs, which are positive when the sign of the two nodes’ signal amplitude agrees, and negative otherwise. The average of these edge time series is equivalent to the value of the corresponding correlation (functional connectivity) and, when computed across all edges, is equivalent to the FC matrix (Figure 1B). A useful summary metric aggregates the amplitudes of all edge co-fluctuations, computed as the square root of the sum of their squared values (root sum square), here denoted RSS. High RSS values indicate that node signals strongly agree/disagree at a given point in time.

Removing amplitudes and retaining only the sign of co-fluctuation along edges naturally partitions the network into exactly two sets of nodes (Figure 1C), one set comprising nodes with positive z-scores and a complementary set comprising the remaining nodes with negative z-scores. This is equivalent to thresholding each frame’s node vector at $z = 0$. These two sets
Dynamic expression of brain functional systems

Figure 1. Schematic illustration of main constructs related to time series and bipartitions. (A) Two-node time series (BOLD signals, converted to standard scores, for nodes $i$ and $j$) and the corresponding edge time series (BOLD signal co-fluctuation, computed as the product of the two-node time series, for edge $i,j$). Positive (negative) BOLD signals and positive (negative) co-fluctuations indicated in red (blue). (B) Edge $i,j$ time series (same as in panel A) depicted as a matrix row. The set of all $(N^2 - N)/2$ edge time series for a given network composed of $N$ nodes can be folded into $N \times N$ matrix form. Examples of single time steps (frames) of such $N \times N$ edge co-fluctuation matrices are shown at the bottom of the panel. The time average of these single-frame matrices is the network’s functional connectivity. (C) Binarized edge $i,j$ time series, by thresholding co-fluctuations at $z = 0$. Positive elements correspond to time points where nodes $i$ and $j$ exhibited positive co-fluctuations (i.e., the sign of their BOLD signals agreed). Frames below correspond to the frames shown in panel B. Each frame is split into exactly two communities. The time average of these frames is equivalent to the agreement matrix (consensus co-classification) of these communities.

of nodes internally co-fluctuate positively and exhibit negative co-fluctuations between them, thus defining a bipartition of the network. Reversing the sign of BOLD amplitudes will retain the exact same co-fluctuation pattern and bipartition; we will therefore disregard the signs of $z$-scored node amplitudes in further analysis. Note also that applying the $z = 0$ threshold, while inherent to the computation of FC from edge time series, should not imply functional activation of nodes above $z = 0$; it merely indicates that regions are active above or below their own mean.

Bipartitions divide the network into exactly two communities, and, over all time frames, these community assignments can be combined into a co-assignment or agreement matrix. In network science, agreement matrices are often used to represent graded assessments of community affiliation (also called co-classification or co-assignment), for example in consensus clustering (Lancichinetti & Fortunato, 2012) and multiresolution community detection (Jeub, Sporns, & Fortunato, 2018). In general, the agreement matrix expresses the frequency with which each node pair is grouped into the same community across many partitions. Here, we calculate the agreement of many bipartitions across many time points.

Bipartitions, as special cases of partitions that bisect the network into two communities, are described by a binary node vector of community assignments. The similarity between two such vectors can be measured with several distance metrics such as the Jaccard distance, the cosine similarity, the variation of information, or the mutual information. Here, we adopt mutual information ($MI$) as the principal metric used for assessing similarity between bipartitions.

Agreement matrix:
Sometimes also called co-classification or co-assignment matrix. Matrix entries record the frequency, over an ensemble of partitions, with which a node pair is co-assigned to the same network community.
Other metrics are highly correlated with MI, and their application gives qualitatively similar results to those reported in this article.

Variations of the bipartition approach are possible, and a few are explored as part of this study. For example, the zero-threshold dividing each frame into two sets of nodes based on the sign of their z-scored time courses may be modified by adopting an arbitrary threshold \( \theta \). Another approach is to define two thresholds \(+\theta\) and \(-\theta\) that separate highly positive and highly negative activations from activations near the temporal mean, thus yielding tripartitions.

**Bipartitions Are Strongly Related to Functional Connectivity**

All analyses reported in this article have been carried out on resting-state fMRI scans acquired in a cohort of 95 participants, a quality-controlled subset of the “100 unrelated” Human Connectome Project (Glasser et al., 2013) cohort. After preprocessing and nuisance regression, each one of the four separate runs was composed of 1,100 frames (TR = 720 ms, total length 792 s). BOLD time courses from cerebral cortex were parcellated into 200 nodes according to a standard template (Schaefer et al., 2018). Some variations of MRI preprocessing have been explored and are discussed below, including a second parcellation scheme into a finer set of 300 nodes and an alternative nuisance regression strategy that omits global signal regression (referred to as “non-GSR data”). For details on participants, scanning, and fMRI processing see the Methods section. To allow division of brain regions into a set of functional systems, each network node was assigned to one of seven canonical RSNs (Yeo et al., 2011), comprising the visual (VIS), somatomotor (SOM), dorsal attention (DAN), ventral attention (VAN), limbic (LIM), frontoparietal (FP), and default mode (DMN) systems (Figure S1 in the Supporting Information).

Classic FC is equal to the mean over all frames (time points) of the edge time series (Figure 2A). Edge time series are converted to binary form by applying a threshold based on the sign of the momentary co-fluctuation, an operation that results in a series of bipartitions (Figure 2B). The agreement matrix constructed from these bipartitions is highly correlated with the corresponding FC matrix (\( \hat{r} = 0.964 \pm 0.008 \), 95 participants, one run). The strong correlation between this bipartition overlay and traditional that FC is robust against different choices of node parcellation and fMRI preprocessing. Figure S2 shows consistently strong similarity between FC and agreement matrix for a finer nodal parcellation (300 nodes) and for time series data omitting global signal regression. Interestingly, for both variants of preprocessing the agreement matrix contains negative entries, representing node pairs whose co-assignment into the same module was below the level predicted by chance. In global-signal-regressed data, these entries strongly overlap with negative functional connectivity. Variants of the bipartition approach also yield high matches of agreement and FC matrices. Adopting an arbitrary (non-zero) threshold \( \theta \) to create bipartitions, or adopting an approach using tripartitions, results in close approximation of FC over a wide range of the \( \theta \) parameter (Figure S3).

The set of bipartitions is a binary decomposition of functional connectivity. The characteristic patterning of FC is constructed from the specific spatiotemporal patterns of its constituent bipartitions, as shown in Figure 3. Subsampling randomly chosen bipartitions gradually approximates FC, with even modest proportions of frames (around 10%) resulting in a very close match with the full-length FC estimate (Figure S4A). When varying run length and using all frames, the quality of the match between agreement and FC matrices remains high even when runs are short (Figure S4B). Prior work (Esfahlani et al., 2020) noted that selecting edge time series frames based on their rankings in RSS magnitude approximates FC more quickly when frames are ranked from high to low RSS amplitudes, as opposed to ranking them from low.
Dynamic expression of brain functional systems

Figure 2. Example of edge time series, FC, bipartitions, and agreement matrix, for one representative participant, one imaging run, in a 200-node parcellation of the cerebral cortex. (A) Edge time series recording co-fluctuations between node pairs (19,900 unique edges) over 1,100 frames (left). The vector of the means of these time series (middle), when refolded into matrix form (right), is equal to the functional connectivity. Nodes are ordered according to 7 canonical functional systems. (B) Thresholding the edge time series at $z = 0$ yields binary time series that track whether co-fluctuations are positive or negative (left). Their average (middle) records, for each edge (node pair), the frequency of positive co-fluctuation that corresponds to their co-assignment (agreement) to the same bipartite community. The agreement matrix (right) is constructed from the complete set of bipartitions and is very highly correlated with the FC matrix (Pearson’s $r = 0.967$, Spearman’s $\rho = 0.963$, cosine similarity = 0.962; all computed on the upper diagonal 19,900-element vector).

to high. Bipartitions behave very similarly (Figure S4C). This effect persists when accounting for the autocorrelation structure (temporal adjacency) of the selected frames (Figure S4D). The level to which bipartitions approximate FC is unrelated to framewise head motion. “Scrubbing” (removing) high-motion frames (retaining only the frames below the 90th percentile of the framewise displacement, per participant, per run) does not significantly affect the match between agreement and FC matrices (Figure S5).

Representing the fMRI time series as a series of bipartitions allows computing their pairwise similarity (quantified as mutual information) across time. Figure 3A displays an example of such a similarity matrix for a single participant and a single run. Notably, some instances of bipartitions recur throughout the run, as indicated by strongly positive MI between remote time points (off-diagonal entries in the matrix plot). Reordering frames by RSS magnitude on each run, followed by averaging over all participants, reveals that high similarity of bipartitions is largely restricted to episodes when RSS amplitudes are near maximal (Figure 3B). The bipartition similarity between adjacent time points (pairwise MI) is correlated with RSS amplitude (Figure 3C shows data from a representative participant; $\rho = 0.502$, $p = 10^{-71}$; $\hat{\rho} = 0.494 \pm 0.049$, 95 participants, one run). Lower values of pairwise MI occur when RSS amplitudes are small,
Dynamic expression of brain functional systems

Figure 3. Spatiotemporal patterns of bipartitions. (A) Pairwise mutual information between bipartitions on adjacent frames, for a single representative participant and imaging run. Plots displays percentiles of the MI distribution (90th, 95th, and 99th percentiles). Note recurrent MI peaks between remote frames. (B) Mean pairwise MI, over all participants on a single run, computed after ranking each participant’s frames by RSS amplitude. Note high mean MI is predominantly evident on high-amplitude frames. (C) Scatterplot of pairwise MI (adjacent frames) versus RSS amplitude (computed as the mean of the two adjacent frames), in one representative participant on one imaging run. The two measures are significantly correlated (Spearman’s $\rho = 0.502$, $p = 10^{-71}$). (D) Switching rates of brain regions, plotted as the ratio of rates observed when RSS amplitudes are high versus low. To compute rates, the bipartition communities on selected frames (top or bottom 10% RSS amplitude) and their immediate temporal successors were compared to identify those regions that switched their community affiliation. Data were aggregated across all participants and all four imaging runs. The plot shows each region’s number of switches during high RSS amplitude frames divided by the number during low RSS amplitude frames. All regions’ ratios are less than 1, indicating lower switch rates on high-amplitude frames, with lowest rates exhibited by regions in lateral parietal, medial parietal, and medial frontal cortex (light colors).

and higher pairwise MI occurs predominantly when RSS amplitudes are large. This relationship indicates that the community structure expressed in framewise bipartitions is more stable (changes less) when overall co-fluctuations, across the entire network, are large. These time points correspond to moments when BOLD time series (and hence co-fluctuations), on average, exhibit larger amplitudes, that is, are farther from their zero mean. Different nodes switch at different rates (Figure 3D), with several DMN regions (parcels DefA_IPL_1 and DefA_PCC_1, both hemispheres) and VAN regions (parcel SalVentA_ParOper_1, right hemisphere) remaining most stably associated with their host communities during high-amplitude epochs.

Principal component analysis (PCA), applied to the set of bipartitions extracted from each participant’s BOLD time course, yields a small number of principal components (PCs) that account for significant portions of the observed variance and exhibit consistent topography.
Dynamic expression of brain functional systems

Figure 4. Principal components of bipartitions, and relation to RSS, gradients, and Louvain partitions. (A) Largest principal component (PC1) derived from PCA of the complete set of bipartitions, for each participant, single run. All 95 PC1s are shown, rectified and z-scored to facilitate comparison across participants. The component generally captures a mode that bisects the brain into two sets of functional systems (VIS, SOM, DAN, VAN vs. LIM, FP, DMN). (B) Topography of the PC1 mode (averaged over 95 participants). (C) Boxplot of correlations (Spearman's \( \rho \)), across participants, of the PC1 loadings, on 1,100 frames, with the RMS amplitude computed from the edge time series. Note that components are binned by the order in which they appear in each participant’s PCA but may not directly correspond in terms of spatial topography. Higher order PCs, accounting for larger percentages of the variance, are more strongly positively correlated with RSS. (D) Comparison of the node vectors of the PC1 mode (left), the principal gradient computed from the FC matrix (middle), and the principal component of the PCA of the bipartitions derived by modularity maximization of the FC matrix, using the Louvain algorithm (right). All three vectors represent averages over 95 participants, one run, and are z-scored. All pairwise correlations are \( r > 0.98 \).

Across participants (Figure 4A). The largest PC (PC1), on average, accounts for approximately 12% of the variance (12.772 ± 2.272, range 20.393 to 8.970, 95 participants, one run). Averaged across participants and projected onto the cortical surface, the PC1 pattern corresponds to a mode that splits the brain into two co-fluctuating ensembles comprising most regions belonging to the VIS, SOM, DAN, and VAN systems on one side versus most regions belonging to the LIM, FP, and DMN systems on the other (Figure 4B). The temporal expression of PC1 is positively correlated with RSS amplitude (Figure 4C), indicating that the connectivity mode inscribed in PC1 is most strongly expressed at time points with high-amplitude network-wide co-fluctuations. This correlation with RSS is obtained despite discarding amplitude information when creating binary partitions. The PC1 as derived from sets of bipartitions is related to several other more familiar constructs (Figure 4D). It is equivalent to the principal eigenvector of FC (or, more precisely, its corresponding covariance matrix), and thus also exhibits strong resemblance to connectivity “gradients” (Margulies et al., 2016). We derived principal components of the affinity matrix derived from FC (equivalent to the principal FC eigenmode) as in the example shown in Figure 4D. The resulting pattern is very highly correlated with the PC1 derived from bipartitions. Furthermore, the bipartition PC1 pattern closely resembles the first principal component of bipartitions identified by applying the Louvain modularity
maximization algorithm to the long-time-averaged FC matrix (Figure 4D). These strong relationships indicate that the set of bipartitions encapsulates characteristic features of the FC matrix, including its eigenmodes and community structure.

Collectively, these results demonstrate a strong relationship between bipartitions and traditional functional connectivity as expressed in the FC matrix. The set of bipartitions derived from the BOLD time series reflects several important spatial and topographic features of FC, while also disclosing its fine temporal structure.

**Bipartitions Map Onto Basis Sets of Templates**

Bipartitions divide the network, at each point in time, into exactly two communities. These two communities are often approximately equal in size, with only 5% comprising node sets that have fewer than 70 (out of 200) members. This fact begs the question of how these large communities relate to canonical subdivisions of the brain into much more compact functional systems, the largest of which (the DMN in the 200-node parcellation) comprising 46 nodes. One way to address this question is to compare each of the empirically observed bipartitions with a standard or basis set of templates that split the brain into bipartitions defined along the boundaries of canonical functional systems (Figure 5A). The basis set used here comprises 7

---

**Figure 5.** Matching bipartitions to a template basis set. (A) Illustration of the template basis set and examples of templates. Each template is a binary 200 × 200 node mask, defining a bipartition along the boundaries of 7 canonical resting-state networks. The complete set of 63 templates is indicated at the top. For example, template 1 divides the brain’s 200 nodes into those belonging to the VIS network (29 nodes) and the complement, the remaining 171 nodes. Three example templates are shown at the bottom of the panel. (B) Time courses of the mutual information computed between the observed bipartition and three example templates from the basis set, for a single participant on a single imaging run. (C) Templates that best match observed bipartitions are aggregated across each imaging run and each participant. The boxplot shows their median frequency in order of abundance, across all 95 participants, single run. The frequency is stated as the number of frames when a given template provides the best match (out of 1,100 total). (D) Each template’s time course of MI, relative to the observed bipartitions on a given run, was correlated to the same run’s RSS amplitude. The boxplot shows the median correlation (Spearman’s ρ) across all 95 participants.
templates that divide 7 canonical RSNs (Yeo et al., 2011) into 1 versus 6 networks, 21 templates that divide them into 2 versus 5 networks, and 35 templates that divide them into 3 versus 4 networks, for a total of 63 such templates. Since these templates are drawn along RSN boundaries (which themselves are defined based on their coherent co-fluctuations over long timescales), one would expect that bipartitions observed at each frame will at least partially align with the boundaries of the 7 systems as captured in the 63-template basis set.

Comparison of templates with observed bipartitions over time allows tracking of several metrics: (a) the similarity (mutual information) of each bipartition with each basis set template; (b) the identification of the single basis set template that most closely resembles the observed bipartition; and (c) computing which of these best-matched templates occur most frequently and which correlate most strongly with framewise measures such as RSS amplitude. Figure 5B shows examples of three MI time courses for three examples of templates (cf. Figure 5A), one each that divides the network into 1 + 6, 2 + 5, and 3 + 4 systems. The full set of 63 MI time courses represent how well each observed bipartition resembles each of the 63 basis set templates and may be interpreted as an index of how strongly a given template is realized at a given point in time. Selecting, at each time frame, the template for which the MI is maximal allows representing the sequence of highly variable bipartitions as a sequence of integers, each representing the single best match (highest MI) out of the 63 templates. Figure S6 provides examples of observed bipartitions and their best matches in the template set determined by maximal MI, for three example templates.

For each participant and imaging run, templates can be ordered by their median frequency, based on the number of times they were selected as the best match for the observed bipartitions (Figure 5C). Once a single best-matching basis set template is assigned to each frame, their occurrence can be compared against RSS amplitude (Figure 5D). The most frequently observed basis set template (template 63) most strongly correlates with framewise RSS, indicating that it is predominantly expressed when BOLD signals and their co-fluctuation patterns exhibit high amplitudes. Note that the template 63 pattern strongly resembles the PC1 extracted from observed bipartitions (cf. Figure 5A). Qualitatively similar rankings of basis set templates and correlations with RSS amplitude are obtained for a finer node parcellation and for non-GSR data (Figure S7).

Finally, the best-matching template set represents a highly compressed set of features of the framewise decomposition, specific to each imaging session and to each participant. Discarding the temporal ordering of the templates, which is immaterial for computing or reconstructing FC, results in a string of 63 numbers encoding a frequency spectrum. The agreement matrix of the template set, as encoded in the 63-element vector, closely matches the down-sampled system-wise FC (Figure S7). This suggests that the frequency of the template basis set may represent a highly compact encoding of the empirically observed FC.

**Bipartitions Contain Signatures of Individual Differences**

If templates represent an efficient encoding of the framewise decomposition of the full-length FC time course, can features of this encoding be useful for identifying individual differences in FC? One way to approach this question is to compare individual variations in template frequencies across multiple resting-state runs (4 runs yielding 6 possible pairwise comparisons: runs 1 vs. 2, 1 vs. 3, 1 vs. 4, 2 vs. 3, 2 vs. 4, 3 vs. 4). Results show that frequencies of a subset of templates are significantly correlated when comparing pairs of runs within participants.
Dynamic expression of brain functional systems

(Figure 6A). This suggests that, at least for some template classes, the level of expression during resting state is moderately conserved across individuals.

A second way to address the question is to pursue connectional “fingerprinting” (Finn et al., 2015) by comparing patterns of functional connectivity across individuals and across runs. Similarity of FC across the cohort is assessed by computing the differential identifiability and accuracy (see the Methods section), both computed for all 6 pairwise combinations of runs and then averaged. Using all 1,100 frames per run to compute FC yields $I_{diff} = 12.97$ and accuracy $= 68.58$ (mean of 6 comparisons). Substituting the corresponding agreement matrix yields $I_{diff} = 12.32$ and accuracy $= 65.75$. Framewise decomposition of FC time series offers the opportunity to select subsets of frames, derive the means of their edge time series and the agreement of their bipartitions, and then compute $I_{diff}$ and accuracy for these samples.

Frames may be selected by applying a wide range of criteria. Here, we selected a criterion that was designed to return frames with bipartitions that were closest to each of the 63 templates in the basis set (Figure 5A). The criterion was computed as follows. For each participant and run, we previously calculated time series of MI between the observed bipartition and each of 63 templates (see Figure 5B). At each time step, the set of 63 MI values was converted to z-scores. For each participant, each run, and all templates we then selected the top 10% of frames with the highest z-scores, comprising a subset of frames with bipartitions that were highly similar to a given target template. Identifiability and accuracy computed from the selected subset are compared against multiple (20) instantiations of a null model where the selected time points (for each participant, each run) were shifted by a random value. Figures 6B and 6C show $I_{diff}$ and accuracy for each of the 63 target templates. Results suggest that several target templates significantly outperform their respective nulls. Highest levels of identifiability and accuracy are attained when subsets of frames are selected based on their similarity to template 4, representing high co-fluctuation within the VAN system, and template 24, representing high co-fluctuations within the VAN and FP systems.

Figure 6D shows normalized (z-scored) similarity matrices, displayed as the mean of 6 matrices of Pearson correlations of subsampled mean edge time series computed across all pairs of participants. Values on the main diagonal refer to self-similarities of participants across runs, while off-diagonal values refer to similarities between different participants across scans. Normalization suggests that increases in $I_{diff}$ and accuracy for target template 4 relative to target template 63 are due to increased self-similarity (cf. Finn et al., 2015). Figure 6E shows the between-participant correlations of the mean agreement matrix (means over all 4 runs) for selected subsets of frames. Selecting subsets of frames based on target template 4 results in lower between-participant correlations ($\hat{r} = 0.562$) than selecting them based on template 63 ($\hat{r} = 0.835$; corresponding means for edge time series patterns: $\hat{r} = 0.601$ and $\hat{r} = 0.828$; mean correlations when using all 1,100 frames: $\hat{r} = 0.740$ for agreement and $\hat{r} = 0.752$ for edge time series). Figure 6F displays the average agreement matrix (across all participants and runs) for target templates 4 and 63.

Collectively, these findings suggest that framewise decomposition and template classification may exhibit some level of stable individual differences and may be useful for enhancing participant-specific network patterns. Regarding the latter point, it appears that selecting moments in time when VAN and/or FP are strongly co-fluctuating results in network patterns that express enhanced levels of individual differences.

Bipartitions Are Constrained by Functional Connectivity

So far, findings indicate that the set of bipartitions observed during single resting-state fMRI runs closely approximates FC (Figure 2) and exhibits characteristic spatiotemporal patterns.
Dynamic expression of brain functional systems

Figure 6. Individual differences. (A) Correlation (Spearman’s rho) between six pairs of runs, for each of 63 template frequencies, computed across all participants. Correlations were compared against 10,000 random shuffles of participant labels, with data points in red corresponding to instances when the empirical correlation exceeded the null distribution with $p < 0.05/6$. (B) Differential identifiability ($\text{Idiff}$), displayed as the mean of six comparisons of scan pairs. Colored symbols are empirical data; grayed symbols are from 20 instantiations of a null model that preserves frame sequence but shifts frames by a random amount. Data are arranged in order of empirical $\text{Idiff}$ per target template. (C) Same as panel B, but for accuracy. (D) Between-participant similarity matrices, with each matrix entry recording the Pearson correlation of the FC pattern (upper triangle of the FC matrix), averaged over six pairwise run comparisons and converted to z-scores to allow comparison (examples shown are for target templates 4 and 63). Values on the main diagonal record self-similarity between different runs, and off-diagonal values record similarity between individuals on different runs. Distributions of diagonal values are significantly different ($p = 9.2 \times 10^{-16}$), while off-diagonal values are not significantly different ($p = 0.27$, two-sided $t$ test). (E) Between-participant similarity (averaged across all scans) of the agreement matrices shown in panel F. (F) Examples of mean agreement matrices computed from frames selected under two different criteria (target templates 4 and 63).

(Figures 3, 4, 5, and 6). Working backwards from a given FC matrix, we can ask to what extent the long-term pattern constrains the set of underlying fine-scale bipartitions from which it is composed. Obviously, many different sets of bipartitions (many different sets of time courses) can yield identical FC. To what extent are sets of bipartitions free to vary once their final superposition in FC is fixed?

Network Neuroscience

Network Neuroscience
An optimization approach, searching the space of all possible bipartitions, can help address this question (Figure 7). The approach adopts a variant of the Metropolis algorithm (Metropolis, Rosenbluth, Rosenbluth, Teller, & Teller, 1953) by maximizing an objective function, defined as the similarity between an empirically observed agreement matrix (which, as established above, very closely resembles FC) and an agreement matrix derived from a set of bipartitions that are subject to incremental optimization. The initial state consists of a completely random set of bipartitions that give rise to a flat agreement matrix. Then, at each subsequent iteration, a single node’s community affiliation on a single time frame (both chosen uniformly and randomly) is swapped. The objective function is recomputed after each swap, and the swap is retained if similarity is increased, subject to a simulated annealing paradigm (Kirkpatrick, Gelatt, & Vecchi, 1983) applied to ensure that the end state corresponds, as closely as possible, to a global optimum. Three different objective functions are employed, the Pearson correlation, the cosine distance, and the root-mean-square distance (additional data shown in Figures S9A and S9B), with near-identical outcomes. Applying the algorithm to data from single participants and single imaging runs succeeds in identifying sets of bipartitions that closely approximate the agreement matrix derived from the empirical BOLD time series (Figure 7A). Importantly, the optimized set of bipartitions resembles the set of observed
Dynamic expression of brain functional systems

bipartitions, as determined by comparing their respective best-matching basis set templates (Figure 7B; Figure S9C). Optimization also yields closely matching sets of bipartitions when the optimized set of bipartitions is significantly smaller than the length of the original time series (Figure S9D). For example, if the optimized set is limited to one-tenth of the length of the original time series (110 frames), optimization still converges and resulting bipartitions continue to resemble those in the observed set.

These findings suggest that the set of bipartitions encountered in the decomposition of fMRI data is highly constrained by the long-time average functional connectivity. Recall that each bipartition represents a snapshot of how momentary co-fluctuations distribute across the network, and that the total set of these snapshots exhibits significant fluctuations across time. The optimization approach suggests that these fluctuations are necessary for reconstructing long-time averages in FC, as optimized bipartitions strongly resemble the highly variable observed set.

Expression of Canonical Systems Varies Across Time

The findings presented so far suggest that bipartitions offer an opportunity to compress time courses into discrete feature sets that retain long-time characteristics of FC while also disclosing fine-scale dynamics. A complementary approach to extract fine-scale network states is possible, as explored in this final section. The expression of individual functional systems across time can be tracked directly, by examining co-fluctuation patterns at fine-scale temporal resolution. The mean co-fluctuation of functional systems can be computed across all $7 \times 7$ subblocks (each system and each system interaction), yielding 28 unique time series. An example is shown in Figure 8A. The temporal averages of these time series are identical to the corresponding down-sampled $7 \times 7$ functional connectivity matrix (cf. Figure S8). On each time step, mean co-fluctuations are compared with a null distribution derived by randomly shuffling system labels and recomputing co-fluctuations (100 independent shuffles per time step). This comparison yields z-scores for each system and pairwise system interaction, where the z-score expresses how much the signal deviates from the label-reshuffling null. Discretizing these time courses by applying a z-score threshold (Figure 8B) yields discrete “network states,” with systems and between-system interactions either exceeding or failing to exceed the threshold of expression. Visual inspection of the sample time course suggests that each of the seven RSNs is significantly expressed, as indicated by exceeding the co-fluctuation z-score threshold, only intermittently, on a fraction of time points. Recall that co-fluctuations should not be taken as “mean activation time courses” as they take on positive values when participating nodes are either jointly above or jointly below their long-time $z = 0$ means.

Considering the above-threshold expression of each of the seven RSNs (leaving aside their mutual interactions, and noting that strongly negative z-scores do not occur) yields, for each point in time, a binary seven-element vector (a total of 128 such states are possible, with between 0 and 7 RSNs expressed at a given time). Aggregating these states (95 participants, 4 runs, 418,000 frames) provides summary statistics on their frequency (Figure 8C). The most frequent state (occurring in approximately 13% of all frames) is one where no RSN is strongly expressed. Individual participants range between 7.4% and 22.4%, and expression levels are correlated across imaging runs (95 participants, $\hat{\rho} = 0.375$, $p < 10^{-4}$ for five out of six pairwise correlations; Figure S10A). The next most frequent states predominantly include those where single RSNs are significantly expressed, while states that involve simultaneous co-expression of multiple systems are less frequent. Frequency ranks of states remain stable when changing z-thresholds (Figure S10B).
Figure 8. Temporal patterns of RSN expression. (A) Edge time series (cf. Figure 2) were aggregated (averaged) based on edges’ placement within or between 7 canonical functional systems. This is equivalent to down-sampling \(200 \times 200\) (node \(\times\) node) frames into a \(7 \times 7\) (system \(\times\) system) matrix, the latter comprising 28 unique elements. The plot shows the resulting 28 edge time series, for a single participant, single imaging run. Note that within-system time courses exhibit intermittent peaks of high and almost exclusively positive co-fluctuations. Between-system interactions show similar intermittency, with both positive and negative co-fluctuations. (B) Same data as in panel A, after discretizing time courses by applying a threshold after z-scoring against a label permuting null model. The threshold shown here is set at \(z = 6\). (C) Each column (time step) in panel B corresponds to a discrete system state. The plot at the left shows the most frequent states encountered after aggregating all 95 participants, all four runs (comprising a total of 418,000 time steps and states). States are displayed by frequency, ordered top to bottom. States with frequencies less than 1% of total frames are not shown. Frequencies are plotted at the right, in corresponding order. Variants of the plot for different z-thresholds are shown in Figure S10. (D) Relation of system states with best-matching templates from the 63-template basis set. Each row of the matrix is normalized to 1. Note that the most frequent system state (no system strongly co-fluctuating) has no clear correspondence with basis set. Other states correlate strongly with specific basis set templates, establishing a link between bipartitions and system states. (E) Average co-fluctuation patterns computed across frames during which specific system states are encountered (top 10 most frequent states shown).
As discussed above, bipartitions decompose FC into sequences of two-community assignment vectors that can be matched to templates from a basis set. As defined in this section, network states also represent sequences of discrete patterns directly derived from significant excursions of edge time series. How do these two representations relate to each other? Network states derived from system-wise expression levels partially reflect the community structure of bipartitions. Many of the states expressing one or several canonical functional systems have clear counterparts within the bipartition template set, that is, the two representations coincide in time (Figure 8D). Aggregating the bipartitions observed on each time point corresponding to the 10 most frequent network states confirms that most states map onto consistent patterns of co-fluctuation as indexed by the bipartition approach (Figure 8E). This comparison establishes a relationship between network states as defined here, through framewise averaging of co-fluctuations, and the community structure of bipartitions as defined in previous sections. Both represent compact descriptions of the dynamic expression of functional systems on fine timescales.

DISCUSSION

Fine-scale analysis of BOLD signal co-fluctuations (edge time series) demonstrates that canonical functional systems are not expressed uniformly or stably across time. Instead, their levels of expression fluctuate significantly, as individual functional systems coalesce and dissolve, singly or in varying combinations. While found reliably and reproducibly in long-timescale FC, this appearance is the result of the overlap of many transient and fleeting manifestations. The proposed decomposition of FC into bipartitions and network states allows tracking these dynamics at fine timescales, only limited by the acquisition rate of single MRI frames. The approach complements traditional network analysis of FC, estimated on long timescales, and of time-varying FC, estimated on shorter windows or epochs.

We propose that FC can be decomposed into sets of bipartitions that map onto discrete network states. These bipartitions exhibit characteristic spatiotemporal patterns, with systems and combinations of systems expressed at different times, and in varying combinations. The most common patterns are those where none of the systems are expressed, or where systems are expressed singly and in isolation. The statistics of bipartitions and network states are reproducible across imaging runs and participants, and do not appear to depend critically on choices made in fMRI preprocessing (e.g., parcellations and global signal regression). Their patterning reflects the complex multiscale community structure of long-time FC, which has been, to this point, the primary target of functional network analysis.

Our work builds on and extends previous investigations of time-varying functional connectivity that has provided evidence for time-dependent fluctuations in functional connections (Chang & Glover, 2010) and network patterns and states (Allen et al., 2014; Lurie et al., 2020; Zalesky et al., 2014). Consistent with prior studies of tvFC, our approach reveals spatiotemporal patterns of network-wide co-fluctuations. Notably, we detect a dominant (segregated or modular) connectivity mode that covaries with overall signal amplitudes (RSS), appears intermittently over time, and exhibits consistent topography (Betzel et al., 2016; Fukushima et al., 2018; Shine et al., 2016). Unlike most tvFC studies, our approach does not require defining sliding windows and hence allows tracking system dynamics at higher temporal resolution. The decomposition of the edge time series into discrete sets of bipartitions and/or network states offers not only a highly compressed encoding of system dynamics but also potential new targets for analysis and modeling of both resting and task-evoked fMRI time series data. Our approach is related to other methods, including coactivation pattern (CAP) analysis.
Dynamic expression of brain functional systems (Karahanoğlu & Van De Ville, 2015; Liu & Dyun, 2013; Liu, Zhang, Chang, & Duyn, 2018) as well as approaches for detecting transient network states that employ hidden Markov models (Vidaurre et al., 2018; Vidaurre, Smith, & Woolrich, 2017), dynamic effective connectivity (Park et al., 2018), instantaneous phase synchrony (Pedersen, Omidvarnia, Walz, Zalesky, & Jackson, 2017; Pedersen, Omidvarnia, Zalesky, & Jackson, 2018), patterns of instantaneous phase-locked BOLD responses (Vohryzek et al., 2020), or filter banks (Faghiri et al., 2020). In contrast to most of these approaches, our method is distinct in that it does not require the user to specify a seed region or a threshold for an “event” (Allan et al., 2015; Petridou, Gaudes, Dryden, Francis, & Gowland, 2013; Tagliazucchi, Balenzuela, Fraiman, & Chialvo, 2012), and does not require “windowing” or specification of models or parameters for the inference of states and state transitions. Importantly, the temporal average of edge time series generated by our approach is exactly equal to time-averaged FC, making it possible to retrieve the precise contributions of individual frames to the static correlation pattern.

The topography of the dominant principal component of the bipartitions bears strong resemblance to the principal mode of BOLD dynamics observed during high RSS amplitude “events” (Esfahlani et al., 2020), as well as patterns characterized by strong excursions (Betzel et al., 2016) or high modularity (Fukushima et al., 2018) in time-varying functional connectivity. Similar patterns representing a decoupling of mainly task-positive from task-negative regions have been described and interpreted in previous studies as a major intrinsic/extrinsic dichotomy in functional architecture (Doucet et al., 2011; Fox et al., 2005; Golland, Golland, Bentin, & Malach, 2008; Zhang et al., 2019). The pattern reported here is also very highly correlated with cortical gradients (Margulies et al., 2016), specifically those derived from eigen-decompositions of the functional connectivity matrix. Indeed, this strong resemblance is due to a mathematical relationship between sets of framewise bipartitions described here (a compression of the original time series) and the spatial patterns of FC eigenmodes. Going beyond static patterns such as gradients (see also Faghiri, Stephen, Wang, Wilson, & Calhoun, 2019), our approach links these connectivity eigenmodes to fluctuating levels of expression of specific functional systems at fine-scale temporal resolution. Their relation to cognitive processes, such as ongoing thought (Mckeown et al., 2020), is a topic for future study.

Our findings suggest that the decomposition of FC into edge time series and bipartitions may offer new approaches for extracting network markers of individual differences. Prior work has established that long-time averaged FC contains connectional features that allow connectotyping (Miranda-Dominguez et al., 2014) or fingerprinting (Finn et al., 2015) of individuals. Subsequent studies have shown that signatures of individual variability involve specific functional systems (Finn et al., 2015), can be extracted from small sets of nodes and connections (Byrge & Kennedy, 2019), and can be enhanced by extracting subsets of principal components (Amico & Goñi, 2018). Analysis of tvFC patterns suggested that cortical fingerprints are expressed at different levels across time, with FC snapshots that are most unique across individuals carrying the most information regarding individual differences (Peña-Gómez, Avena-Koenigsberger, Sepulcre, & Sporns, 2018). Here, we build on this body of work and suggest that FC and agreement matrices derived from subsets of frames may carry enhanced signatures of individual differences. Our findings suggest that frames characterized by strong co-fluctuations within the ventral attention and frontoparietal networks carry information that is specific to participants. These systems, or their close homologs, have also been implied in previous work on FC fingerprinting (Finn et al., 2015; Peña-Gómez et al., 2018). An intriguing avenue for further study, suggested by findings on participant-specific and heritable state frequencies and transitions (Vidaurre et al., 2017), is whether decomposition of FC and selection of frames can uncover new connectivity traits related to cognition and behavior.
We employed an optimization approach to explore whether a given FC matrix can be decomposed into sets of bipartitions that differ radically from the ones that are empirically observed. Our findings suggest this is not the case. Given an observed pattern of FC, the set of framewise patterns from which it is composed, or into which it can be decomposed, is not free to vary. Instead, the statistics of these patterns appear strongly constrained by the correlation structure inscribed in long-time FC. Optimizations invariably retrieve sets of patterns that resemble those observed empirically, even though no dynamic generative model is employed. This makes it harder to dismiss the observed framewise patterns as artifactual or as massively corrupted by noise or uncertainty. Instead it appears that the fluctuating patterns of observed bipartitions are necessary, in the sense that it is difficult if not impossible to construct FC from a radically different (“stationary,” temporally smoother, less variable) set of frames. Our findings suggest the hypothesis that framewise bipartitions reflect a decomposition of the community structure of the FC matrix, unwrapped in time.

Our work opens new avenues for future research. The decomposition of BOLD time series into a set of bipartitions and/or network states represents a compression or encoding of the system’s dynamics into a much more compact feature set. Such feature sets may provide novel opportunities for mapping individual differences, relations to demographic or behavioral measures, task-rest reconfigurations, or relation to the underlying anatomy. They may also serve as tools for further analysis using machine learning or multivariate statistical mapping, including those probing the relation of brain to behavior. In addition, we note that the proposed scheme may also apply to brain data obtained with other acquisition methods, including more highly resolved recordings of neuronal populations or individual neurons. The decomposition of FC into framewise contributions allows us to selectively recombine subsets of frames to get different patterns of FC. It might be possible to select specific subsets of frames/templates to amplify a brain/behavior correlation. Another possibility for future research lies in exploring alternative basis sets of templates. Here, many of our findings related to bipartitions are expressed by reference to a specific choice involving a 63-template basis set. This choice was motivated by a specific research agenda related to the expression of canonical functional systems for which a nodal partition, and hence basis set, are defined. However, many alternative choices of templates are possible, depending on specific working hypotheses. Templates could also be defined in a more data-driven way, as centroids derived by clustering techniques, or from modular partitions of the empirically measured FC matrix. Finally, alternative decomposition strategies and edge metrics may be pursued. Substituting FC by the covariance matrix, a framewise decomposition could retain information on instantaneous BOLD signal variance, a measure that has proven valuable in prior work (Garrett, Kovacevic, McIntosh, & Grady, 2010). Substituting time-lagged correlation, partial correlation, or covariance, or information-theoretic measures such as mutual information, may provide additional opportunities for decomposing long-time averages for node pair interactions into framewise contributions that disclose temporal patterns and fluctuations.

Limitations of the approach should be noted. As is the case with all studies employing functional neuroimaging, the present work inherits most drawbacks of fMRI methodology including its limited temporal and spatial resolution, the indirect link to underlying neural activity, and measurement noise and statistical biases. Following good practices in data preprocessing, the use of multiple data sources, and cautious interpretation of findings can at least partially guard against these limitations. It should also be noted that the basic methodological framework transcends the limitations of fMRI as its mathematical and algorithmic core applies to all time-dependent data sources regardless of origin. Future work should aim to reproduce and refine the proposed feature sets, spatiotemporal patterns, and statistics on system expression...
by leveraging new data sources and participant cohorts. Interventional studies and multimodal experimentation are needed to identify putative neurobiological mechanisms that underpin or drive temporal fluctuations.

Other limitations relate more directly to the analytic approach. The calculation of cross-correlations as measures of BOLD functional connectivity implies z-scoring of the original time series, which effectively normalizes and centers all time series. If means are nonstationary, for example due to task-induced regional activations, more sophisticated methods that correct for such transient responses are needed (e.g., Cole et al., 2019). Finally, as has been pointed out in the case of “windowing” approaches in tvFC, finer temporal scales (shorter time segments) may increase the noisiness of each estimate of FC (Leonardi & Van De Ville, 2015; Zalesky & Breakspear, 2015). In our approach, no windowing is applied, and each co-fluctuation pattern is directly derived from framewise BOLD responses. It is worth pointing out that long-time FC and noisy BOLD patterns are intertwined. Adding noise to framewise BOLD signals would imply adding noise to the FC as well, and vice versa. Because the decomposition into frames and bipartitions is mathematically exact, it is not subject to noisy estimation.

In conclusion, fine-scale temporal fluctuations in the community structure of resting brain activity suggest that the brain’s functional systems express only transiently, intermittently, and infrequently across time. Their robust manifestation over long timescales results from the superposition of large numbers of spatially distinct and temporally variable patterns. Novel insights and applications may result from the proposed decomposition of brain dynamics into network bipartitions and states.

**MATERIALS AND METHODS**

**Dataset and fMRI Preprocessing**

All analyses reported in this article were carried out on data originally collected by the Human Connectome Project (HCP; Van Essen et al., 2013), specifically resting-state fMRI data from 100 unrelated adult participants (54% female, mean age = 29.11 ± 3.67, age range = 22–36). The study was approved by the Washington University Institutional Review Board and informed consent was obtained from all subjects. Participants underwent four 15-min resting-state fMRI scans (here referred to as four imaging runs) spread out over a 2-day span. For a full description of the imaging parameters and image preprocessing, see Glasser et al. (2013). Briefly, data were acquired with a gradient-echo EPI sequence (run duration = 14:33 min, TR = 720 ms, TE = 33.1 ms, flip angle = 52, 2-mm isotropic voxel resolution, multiband factor = 8). Participants were instructed to keep their eyes open and fixate on a cross. Images were collected on a 3T Siemens Connectome Skyra with a 32-channel head coil. Subjects were considered for data exclusion based on the mean and mean absolute deviation of the relative root-mean-square motion across either four resting-state MRI scans (file: Movement_RelativeRMS.txt) or one diffusion MRI scan (file: eddy_unwarped_images.eddy_movement_rms), resulting in four summary motion measures. If a subject exceeded 1.5 times the interquartile range (in the adverse direction) of the measurement distribution in two or more of these measures, the subject was excluded. These exclusion criteria were established before the current study. Four subjects were excluded based on these criteria. One subject was excluded for software error during diffusion MRI processing. Even though diffusion MRI was not part of the present study, this subset was created to include subjects with adequate resting-state and diffusion data for future analysis. The remaining subset of 95 subjects have the following demographic characteristics: 56% female, mean age = 29.29 ± 3.66, age range = 22–36. Finally, we note here that we defined framewise displacement as the relative root-mean-square motion (file:
Movement_RelativeRMS.txt), which was computed with the FSL function rmsdiff via the HCP pipelines. We used this information to censor the resting-state scans at a frame-by-frame level in a supplementary analysis.

HCP data were minimally preprocessed as described in Glasser et al. (2013). Briefly, data were corrected for gradient distortion, susceptibility distortion, and motion, and then aligned to a corresponding T1-weighted (T1w) image with one spline interpolation step. This volume was further corrected for intensity bias and normalized to a mean of 10,000. This volume was then projected onto the 32k fs LR mesh, excluding outliers, and aligned to a common space using a multimodal surface registration (Robinson et al., 2014).

A functional parcellation designed to optimize both local gradient and global similarity measures of the fMRI signal (Schaefer et al., 2018; Schaefer200) was used to define 200 regions (parcels or nodes) of the cerebral cortex. These nodes can be mapped to a set of canonical functional networks (Schaefer et al., 2018; Yeo et al., 2011); in the current study we adopt a mapping to seven canonical networks that comprise the visual (VIS), somatomotor (SOM), dorsal attention (DAN), ventral attention (VAN), limbic (LIM), frontoparietal (FP), and default mode (DMN) systems (cf. Figure S1). For HCP data, the Schaefer200 is openly available in 32k fs LR space as a cifti file. A second processing variant used a finer parcellation into 300 nodes (Schaefer300) following the same basic procedure.

We employed two variants in preprocessing to explore the robustness of main findings reported in this article. All analyses were first carried out on data processed with the inclusion of global signal regression. For this strategy, the mean BOLD signal for each cortical node was linearly detrended, band-pass filtered (0.008–0.08 Hz; Parkes, Fulcher, Yücel, & Fornito, 2018), confound regressed, and standardized using Nilearn’s signal.clean, which removes confounds orthogonally to the temporal filters (Lindquist, Geuter, Wager, & Cao, 2019). The confound regression employed (Satterthwaite et al., 2013) included six motion estimates, time series of the mean CSF, mean WM, and mean global signal, the derivatives of these nine regressors, and the squares of these 18 terms. Following confound regression and filtering, the first and last 50 frames of the time series were discarded. Furthermore, a spike regressor was added for each fMRI frame exceeding a motion threshold (0.25-mm root-mean-squared displacement). This confound strategy has been shown to be effective in reducing motion-related artifacts (Parkes et al., 2018). For validation, we also preprocessed the data using aCompCor (Behzadi, Restom, Liau, & Liu, 2007). These data were linearly detrended, band-pass filtered, and trimmed identically to the previous strategy. This confound regression included five high-variance signals estimated from the CSF and white matter each (10 total), as well as six motion estimates, their derivatives, and the squares of these 12 terms. This strategy did not incorporate spike regressors. Following preprocessing and nuisance regression, residual mean BOLD time series at each node was recovered using Connectome Workbench. All data were visually inspected.

**Functional Connectivity and Edge Time Series**

Functional connectivity (FC) is generally estimated from fMRI data by computing the Pearson correlation between the BOLD time series recorded from each node pair. Hence, each FC estimate represents a linear similarity between the respective time courses, interpreted as their mutual statistical dependence. It is, by definition, a nondirected, noncausal metric that does not distinguish between node pairs that are structurally (anatomically) coupled or uncoupled. All node pairs maintain nonzero FC, and FC estimates may be negative or positive. In a system composed of N nodes, the system’s FC matrix has dimensions \([N \times N]\), due to symmetry with a total of \(K = (N^2 - N) / 2\) unique entries (all node pairs \(i, j\) with \(i \neq j\)).
One definition of the Pearson correlation coefficient states that it is the mean of the product of the standard scores of the two individual variables. Specifically,

\[ r_{xy} = \frac{1}{n-1} \sum_{i=1}^{n} \left( \frac{x_i - \bar{X}}{s_x} \right) \left( \frac{y_i - \bar{Y}}{s_y} \right), \]

where \( \bar{X} = \frac{1}{n} \sum_{i=1}^{n} x_i \) is the mean of \( x \) (and applied analogously for \( y \)) and \( s_x = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{X})^2} \) is the standard deviation of \( x \) (and applied analogously for \( y \)), and \( n \) is the number of observations (for time series data the number of observations is equal to the number of time points). Thus, there are three steps involved in this computation: the conversion of two node time series to z-scores (note that \( z(x_i) = \left( \frac{x_i - \bar{X}}{s_x} \right) \)), forming their product to create a single time series for node pair \( i, j \), and finally forming the mean of this pairwise time series to yield the FC estimate for the node pair (cf. Figure 1). This procedure, when repeated for all pairs of nodes, results in a node-by-node correlation matrix, that is, an estimate of FC. Following an approach developed in prior work (Esfahlani et al., 2020; Faskowitz et al., 2020; Jo et al., 2020), we may omit the final averaging step and retain the time series for each node pair. Since each node pair subtends a unique network edge, we refer to this construct as edge time series. The mean of each edge time series is equal to the corresponding node pair’s FC. Rather than collapsing all time steps into a single scalar FC estimate, omitting the averaging step effectively unwraps FC into a set of edge time series that track co-fluctuation between each node pair. At each point in time, the edge time series report a product of two z-scored BOLD signals; the product is positive if both signals are above or below their respective zero mean, and negative otherwise. The amplitude of their product varies with the joint amplitudes of the two signals. The full set of edge time series comprises a matrix of dimension \( [K \times T] \), with \( K \) equal to the number of unique edges and \( T \) equal to the number of time points.

At each moment in time, the amplitude of the co-fluctuations along all edges can be computed as the root sum square, denoted \( \text{RSS} \). This metric takes on high amplitude when edge-wise co-fluctuations, on average, are high (either positively or negatively), and it takes on low amplitudes when co-fluctuations are low (again, irrespective of their sign). Prior work has utilized \( \text{RSS} \) to track co-fluctuation amplitudes and stratify or order time points according to their magnitudes. High-amplitude time points coincide with intermittent and recurrent patterns of network activity and connectivity (Esfahlani et al., 2020).

**Bipartitions and Agreement Matrix**

Edge time series can be converted to binary form, by applying a threshold at the zero crossings that retains only if co-fluctuations were positive or negative. Positive co-fluctuations occur if and only if two signals both exhibit above mean (positive z-score) amplitudes or if both exhibit below mean (negative z-score) amplitudes. Negative co-fluctuations occur when the two signals deviate in opposite directions. A simple extension of this fact is that, on each time step, positively co-fluctuating node pairs split the network into exactly two communities that are fluctuating negatively with respect to each other. This obligatory two-community split results in a bipartition of the network. The network’s time evolution may be represented as a sequence or set of such bipartitions. Adopting bipartitions largely removes information on signal amplitudes (noting that the bipartition does depend on standardizing the individual node time series) while creating a compact description of the original FC as a set of finely resolved modular partitions, without the need to perform computational community detection. Communities are directly evident from the binary edge time series.
It is common practice in network science to combine multiple partitions, for example those obtained from multiple runs of a community detection algorithm, into a single co-classification or agreement matrix (Jeub et al., 2018; Lancichinetti & Fortunato, 2012). The elements of this matrix express, for each node pair, the frequency with which the two nodes are assigned to the same network community. To correct for the rate at which this occurs due to chance, one can subtract the expected frequency if community labels are randomly permuted. Under the assumption that for each sampled partition the number and sizes of clusters are fixed but nodes are otherwise assigned randomly to clusters, one obtains a constant null computed as (Jeub et al., 2018)

\[ P_{null} = \frac{1}{l} \sum_{k=1}^{l} \sum_{s=1}^{C_k} \frac{|C_{ks}|}{N} \frac{|C_{ks}| - 1}{N - 1}, \]

where \( l \) is the number of samples, \( C_k \) is the partition of the \( k \)-th sample, \( |C_{ks}| \) is the number of nodes in clusters of partition \( C_k \), and \( N \) is the number of nodes. In applications to bipartitions from time series, \( l \) is equal to the number of time points \( T \). Subtracting the constant null results in agreement matrices that contain negative entries for all node pairs where the observed frequency of co-classification is smaller than that expected under the adopted null model. The subtraction step does not change relative order of frequencies in the agreement matrix and hence has no impact on correlations or similarity metrics computed against FC.

**Bipartition Similarity**

Similarity or distance between two modular partitions can be defined in several ways, including the mutual information computed as

\[ MI(M, M') = \sum_{m \in M} \sum_{m' \in M'} P(m, m') \log \frac{P(m, m')}{P(m)P(m')} , \]

where \( M \) and \( M' \) indicate the two partitions, \( m \) and \( m' \) indicate modules belonging to the two partitions, and \( P(m, m') = \frac{n_{mm'}}{n} \) with \( n_{mm'} \) corresponding to the number of nodes that are members of module \( m \) as well as module \( m' \) (Rubinov & Sporns, 2010). In the case of bipartitions, other metrics such as the cosine similarity or the Jaccard distance are also possible. In practice, all these metrics give highly similar results. We adopt the mutual information as the principal metric for assessing similarity between pairs of bipartitions.

**Modularity Maximization**

Modularity maximization is a commonly used approach for detecting communities in brain networks (Newman & Girvan, 2004; Sporns & Betzel, 2016) that attempts to partition a network into nonoverlapping communities such that the observed density of connections within subnetworks maximally exceeds what would be expected by chance. The choice of null models should reflect the nature of the network data, which in our case is a correlation matrix (MacMahon & Garlaschelli, 2015). We adopt a constant null (the Potts model; Traag, Van Dooren, & Nesterov, 2011) and retain the full FC matrix, including its negative entries, for the purpose of community detection by applying the Louvain algorithm. Louvain bipartitions are identified by first scanning a wide range of the resolution parameter, selecting upper and lower limits within which a two-community structure appears, followed by a finer sampling of the range to retrieve bipartitions (1,000 samples).

**Gradients**

So-called gradients, when computed from FC matrices, represent major connectivity modes (node vectors) that can be mapped back onto the original node set, for example the surface of
Dynamic expression of brain functional systems

the cerebral cortex. Following a standard workflow (de Wael et al., 2020), after starting from an FC matrix, one first derives an affinity matrix that essentially represents a node-wise distance matrix of size \([N \times N]\). Here, we compute the affinity matrix, from the full unthresholded FC, as the cosine similarity for each node pair excluding their mutual connections. The first principal component of the affinity matrix is retained for purposes of analysis and comparison. It is identical to the largest eigenvector of the FC matrix. Other variants for computing the affinity matrix may include additional steps such as thresholding or alternative distance transforms. These variants have little to no impact on the relevant findings reported in this article.

**Individual Differences**

Two measures are employed to assess consistency and variance in patterns of functional connectivity or agreement matrices across participants. FC and agreement are computed as described above, as means of edge time series or from sets of bipartitions (taken over all 1,100 frames or from a subset of selected frames). This results in a square \([N \times N]\) matrix, from which the upper triangle is converted into a vector of \((N^2 - N)/2\) elements. One such vector is derived for all 95 participants and for each of the four resting-state runs. For each pair of runs (six pairs in total), we compute the Pearson correlation of these vectors across all pairs of participants. The resulting matrix carries similarities between pairs of participants on different runs, with the main diagonal of the matrix representing self-similarity (same participant, different run). The differential identifiability (Amico & Goñi, 2018) is computed as

\[
I_{\text{diff}} = 100 \times (\hat{r}_{\text{self}} - \hat{r}_{\text{non-self}}),
\]

with \(\hat{r}_{\text{self}}\) corresponding to the mean of the correlations on the main diagonal and \(\hat{r}_{\text{non-self}}\) corresponding to the mean of the off-diagonal correlations. The accuracy is computed as the number of times (out of 95) that the correlation value on the main diagonal (self-similarity) is maximal along the matrix rows and columns. The accuracy estimate is the mean of these two numbers.

Many different criteria can be employed for selecting specific subsets of frames to compute means of edge time series and agreement matrices, both of size \([N \times N]\). Identifiability and accuracy are computed from such subsets of frames in the same manner as described above. The selection criterion employed here involves three steps: (a) computing the mutual information between each frame’s bipartition and all of the 63 templates in the basis set; (b) normalizing by converting the resulting 63 \(MI\) values (per frame) to z-scores; (c) selecting, for each of the 63 templates, a fraction (here, the top 10%) of frames with the highest z-scores across a single run. Mean edge time series and agreement matrices are derived from this subset, representing 10% of the data. Resulting measures of identifiability and accuracy are compared against a distribution resulting from several independent realizations (here, 20) of a null model. In the null model, the timings of the selected frames are shifted by a random offset (chosen randomly and uniformly between 1 and 1,100). Once the offset is chosen, frames are rotated in time by the chosen amount. This approach largely preserves the spacing (autocorrelation) of the frames in the selected subset.

**Optimization**

The purpose of the optimization approach pursued in this study was to discover sets of bipartitions of \(N\) nodes, comprising \(P\) instances, that approximate the observed bipartitions’ (size \([N \times T]\)) agreement matrix. We adapted a variant of the Metropolis-Hastings algorithm (Metropolis et al., 1953) to generate samples from the very large distribution of possible sets of
Dynamic expression of brain functional systems

bipartitions of size \([N \times P]\), starting from a random sample and then iteratively creating variants that are either rejected or accepted, moving into the next iteration. The rejection or acceptance is governed by simulated annealing (Kirkpatrick et al., 1983) and by an objective function \(D\), taken here to be a measure of the distance between the optimized and observed agreement matrix. New variants are accepted if the objective function improves (lower distance) or if the annealing criterion \(e^{-\Delta D/Temp} > R(0,1)\) is fulfilled, where \(Temp\) refers to a simulated “temperature” and \(R(0,1)\) is a random number uniformly drawn from the \([0,1]\) interval. Essentially, the annealing criterion allows suboptimal variants to pass, as a function of the current temperature. The temperature decays exponentially as a function of the number of iterations, \(h\), as \(Temp = T_0T_{exp}^{-h}\). Temperature parameters \(T_0, T_{exp}\) were selected such that stable solutions near the global minimum (distance of zero) emerged in reasonable time. The initial conditions were chosen as sets of completely random bipartitions, with equal probability (“flipping a coin”) on all node co-assignments. On each step of the optimization, a single element in a single bipartition (both chosen at random) was flipped. Note that this optimization procedure does not implement a true generative process for bipartitions, as they are not derived from time series data and hence contain no information on temporal sequences. While more realistic scenarios for discovering optimally matching sets of bipartitions are conceivable, they were not pursued in the current study.

Three different formulations of the objective function were tested, all computed from the agreement matrix’s \(K\) unique (upper triangle) elements, with highly reproducible results: (a) the Pearson correlation; (b) the rank-order correlation (Spearman’s \(\rho\)); and (c) the cosine similarity. Optimizations were also carried out by substituting the observed agreement matrix with the observed FC matrix in the objective function, with near-identical outcomes.

The number of bipartitions in the optimized set, \(P\), is a free parameter. Different settings of \(P\), varying the number of bipartitions from 1,100 (matching the number of experimental time steps \(T\)) down to 11, were explored. Smaller values of \(P\) yield more compact optimized sets, while also resulting in less accurate matches between the observed and the optimized agreement matrix.

ACKNOWLEDGMENTS

Data were provided, in part, by the Human Connectome Project, WU-Minn Consortium (principal investigators: D. Van Essen and K. Ugurbil; 1U54MH091657), funded by the 16 National Institutes of Health (NIH) institutes and centers that support the NIH Blueprint for Neuroscience Research and by the McDonnell Center for Systems Neuroscience at Washington University.

SUPPORTING INFORMATION

Supporting information for this article is available at https://doi.org/10.1162/netn_a_00182. Matlab code and auxiliary data files demonstrating key concepts introduced in this paper are available at https://www.brainnetworkslab.com/s/bipartitions-code-package.zip (Sporns, Faskowitz, Teixeira, Cutts, & Betzel, 2021a). A movie of edge time series and RSS amplitudes can be downloaded at https://www.brainnetworkslab.com/s/ets_movie_1.mp4 (Sporns, Faskowitz, Teixeira, Cutts, & Betzel, 2021b).

AUTHOR CONTRIBUTIONS

Olaf Sporns: Conceptualization; Formal analysis; Investigation; Methodology; Software; Visualization; Writing – original draft; Writing – review & editing. Joshua Faskowitz: Conceptualization; Formal analysis; Software; Writing – review & editing. Sofia Teixeira: Conceptualization; Methodology; Writing – review & editing. Sarah A. Cutts: Conceptualization;
Dynamic expression of brain functional systems

Formal analysis; Writing – review & editing. Richard Betzel: Conceptualization; Formal analysis; Software; Writing – review & editing.

FUNDING INFORMATION
Richard Betzel, Indiana University Office of the Vice President for Research, Emerging Area of Research Initiative, Award ID: Learning: Brains, Machines, and Children. Richard Betzel, National Science Foundation (http://dx.doi.org/10.13039/100000001), Award ID: 2023985. Joshua Faskowitz, National Science Foundation (http://dx.doi.org/10.13039/100000001), Award ID: 1342962. Sofia Teixeira, Fundação para a Ciência e a Tecnologia, Award ID: UIDB/50021/2020.

REFERENCES


Dynamic expression of brain functional systems


Dynamic expression of brain functional systems

Image, 180, 594–608. DOI: https://doi.org/10.1016/j.neuroimage.2017.11.033, PMID: 29158202, PMCID: PMC6138953


Dynamic expression of brain functional systems


