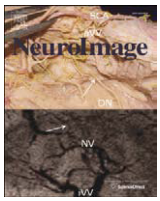
NeuroImage 55 (2011) 688–704

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg



Statistical parametric network analysis of functional connectivity dynamics during a working memory task

Cedric E. Ginestet *, Andrew Simmons

King's College London, Institute of Psychiatry, Centre for Neuroimaging Sciences (CNS), UK
National Institute of Health Research (NIHR) Biomedical Research Centre for Mental Health, UK

ARTICLE INFO

Article history:
Received 21 May 2010
Revised 19 October 2010
Accepted 8 November 2010
Available online 21 November 2010

Keywords:
Statistical parametric network (SPN)
Small-world topology
Network dynamics
N-back
Cognitive load
Working memory
Weighted cost
Cost-integrated metrics

ABSTRACT

Network analysis has become a tool of choice for the study of functional and structural Magnetic Resonance Imaging (MRI) data. Little research, however, has investigated connectivity dynamics in relation to varying cognitive load. In fMRI, correlations among slow (<0.1 Hz) fluctuations of blood oxygen level dependent (BOLD) signal can be used to construct functional connectivity networks. Using an anatomical parcellation scheme, we produced undirected weighted graphs linking 90 regions of the brain representing major cortical gyri and subcortical nuclei, in a population of healthy adults ($n = 43$). Topological changes in these networks were investigated under different conditions of a classical working memory task — the *N*-back paradigm. A mass-univariate approach was adopted to construct statistical parametric networks (SPNs) that reflect significant modifications in functional connectivity between *N*-back conditions. Our proposed method allowed the extraction of 'lost' and 'gained' functional networks, providing concise graphical summaries of whole-brain network topological changes. Robust estimates of functional networks are obtained by pooling information about edges and vertices over subjects. Graph thresholding is therefore here supplanted by inference. The analysis proceeds by firstly considering changes in weighted cost (i.e. mean between-region correlation) over the different *N*-back conditions and secondly comparing small-world topological measures integrated over network cost, thereby controlling for differences in mean correlation between conditions. The results are threefold: (i) functional networks in the four conditions were all found to satisfy the small-world property and cost-integrated global and local efficiency levels were approximately preserved across the different experimental conditions; (ii) weighted cost considerably decreased as working memory load increased; and (iii) subject-specific weighted costs significantly predicted behavioral performances on the *N*-back task ($Wald\ F = 13.39, df_1 = 1, df_2 = 83, p < 0.001$), and therefore conferred predictive validity to functional connectivity strength, as measured by weighted cost. The results were found to be highly sensitive to the frequency band used for the computation of the between-region correlations, with the relationship between weighted cost and behavioral performance being most salient at very low frequencies (0.01–0.03 Hz). These findings are discussed in relation to the integration/specialization functional dichotomy. The pruning of functional networks under increasing cognitive load may permit greater modular specialization, thereby enhancing performance.

Introduction

Network analysis has become a promising framework within which systems neuroscience can address a range of challenging questions (Bassett and Bullmore, 2006; Bullmore and Sporns, 2009). Some of this interest has been motivated by the fact that several topological metrics capturing core aspects of cortical networks have been found to be heritable (see Bassett and Bullmore (2009), for a review). Brain functional and structural networks rely on the assumption that correlation in BOLD signal or correlation in gray matter (GM) density are indicative of neuronal connectivity. These networks have been found to have predictive validity as network architecture appears to be related to medication dosage and to psychometric scores such as the Positive and Negative Syndrome Scale (PANSS) (Rubinov et al., 2009). In addition, whole-brain network topological metrics have also been used as biomarkers for specific neurodegenerative diseases such as Alzheimer's disease (AD) with high levels of specificity and sensitivity (Supekar et al., 2008). Research on the human connectome seems especially important from

* Corresponding author. Centre for Neuroimaging Sciences, NIHR Biomedical Research Centre, Institute of Psychiatry, Box P089, King's College London, De Crespigny Park, London, SE5 8AF, UK.
E-mail address: cedric.ginestet@kcl.ac.uk (C.E. Ginestet).

1053-8119/$ – see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.neuroimage.2010.11.030

a clinical perspective, as several psychopathological disorders can be specifically described as 'disconnection syndromes' (Bassett and Bullmore, 2009; Catani and Mesulam, 2008).

One particular network architecture, which is commonly emphasized in this emerging field, is small-worldness (Milgram, 1967; Watts and Strogatz, 1998). The small-world property is a qualitative description of a network characterized by high levels of local clustering and short path lengths linking all nodes of the network. This constitutes a particularly attractive model of brain network organization, since it can account for the combination of both *specialized* and *distributed* information processing, as well as minimizing wiring cost in brain circuitry (Achard et al., 2006). Small-worldness of functional networks in humans has been validated by several teams of researchers working with different modalities, such as functional MRI (Achard et al., 2006), magnetoencephalography (Bassett et al., 2006) and electroencephalography (Micheloyannis et al., 2009). Although networks can be directed or undirected, weighted or unweighted, most of the research in systems neuroscience thus far has concentrated on the simplest case: undirected and unweighted graphs using a thresholding function. Following this trend, we will similarly focus on unweighted, undirected networks based on functional MRI data.

Only a small subset of human behavior has been investigated from a network analytic perspective. Resting-state functional networks have received most research attention (Achard et al., 2006; Salvador et al., 2005). These investigations have shed light on the intrinsic architecture of spontaneous fluctuations in BOLD signal when subjects are at rest. In addition, the functional connectivity networks underlying finger tapping and listening to music have been summarized by Eguiluz et al. (2005). For these two activities, the brain also appears to display a small-world architecture. Intriguingly, Eguiluz et al. (2005) reported that listening to music was associated with functional networks characterized by higher mean degrees than those subtending finger tapping. That is, listening to music seems to induce, on average, a greater amount of functional connectivity than finger tapping. Other authors have investigated functional connectivity patterns linked with movement execution (De Vico Fallani et al., 2008; Cecchi et al., 2007; Astolfi et al., 2009). By contrast, there have been very few studies exploring the relationship between human cognition and functional networks' topological metrics (Sporns et al., 2004; Bullmore and Sporns, 2009). The exploration of such changes in topology will aid systems neuroscience in probing the dynamics of functional networks.

Several studies have investigated changes in functional network topologies across different cognitive tasks. Some research in this area has been based on the *N*-back task and has used either EEG (Pachou et al., 2008), MEG (Bassett et al., 2009) or fMRI data (Salvador et al., 2008). The *N*-back paradigm is a popular cognitive task, which allows evaluation of changes in working memory load under several conditions of increasing difficulty (Gevins and Cutillo, 1993). The task requires on-line monitoring of a series of stimuli, as well as updating and manipulating remembered information. It is generally thought to place great demands on several key processes involved in memory (Honey et al., 2002; Owen et al., 2005). Some authors, however, have recently argued that *N*-back performance is not a pure measure of working memory, but may be better described as a measure of processing speed (Miller et al., 2009). Research in this area generally seeks to unveil the neural basis of the central executive system of working memory (Baddeley, 1998; D'Esposito et al., 1995; Kane et al., 2007). The *N*-back task has extensively been used in fMRI to identify pathological responses in clinical populations (Harvey et al., 2005; Caseras et al., 2006; Wishart et al., 2006; Van Snellenberg et al., 2006). A recent meta-analysis of 24 primary studies using the *N*-back task demonstrated that the cortical regions that tend to be robustly activated under this paradigm are the lateral premotor cortex, the dorsal cingulate and medial premotor cortices, the dorsolateral and centrolateral prefrontal cortices, the frontal poles, and the medial and lateral posterior parietal cortices (Owen et al., 2005). The pattern of interactions between these different regions, however, is still one of the key questions in the study of working memory. Network analysis may be particularly useful in this respect, by enabling researchers to map the functional connectome subtending different working memory loads. Our first objective in this article is therefore to explore the dynamics of functional connectivity patterns under different levels of cognitive load.

The second contribution of this paper is methodological. We introduce a concise way of summarizing inference on networks, by using a mass-univariate approach. This strategy is similar in spirit to the one adopted in the statistical parametric mapping (SPM) framework (Friston, 1994), and will therefore be referred to as a statistical parametric network (SPN). SPM revolutionized the analysis of PET and later MRI data, by providing summary maps, which facilitate the identification of a set of regions of interest, through the application of specific levels of significance (Friston, 1994). We argue that network analysis urgently needs to adopt a similar approach when combining networks over a population of subject-specific correlation matrices. Currently, networks constructed on the basis of structural and functional data are treated in a different manner. Whereas in fMRI studies, researchers have generally computed inter-regional correlations over time (Achard et al., 2006; Achard and Bullmore, 2007), studies based on structural MRI data have utilized inter-regional correlations over subjects (e.g. Bassett et al., 2008). In this paper, we will focus on the specific problem posed by the study of functional MRI cortical networks, where correlation matrices are subject-specific with regional covariances computed with respect to time. It will therefore be assumed throughout this paper that a population of subject-specific networks is available for the network analysis.

Our proposed SPN framework pursues some of the ideas put forward by Achard et al. (2006), He et al. (2007) and He et al. (2009b) and generalizes them in order to produce summary networks constructed over the different conditions of an experiment, or similarly for distinct populations of interest. Achard et al. (2006) and He et al. (2009b) have presented summary networks, where an edge is only included if a test statistic for that edge computed over all the subjects in the population of interest is significant. We will term these networks *mean* or *summary* SPNs. In addition, we will also construct what we term *differential* or *difference* SPNs, which provide information about which edges have been significantly 'lost' and which edges have been significantly 'gained' from one experimental condition to another, or from one subject population to another. Similar approaches to functional network comparison have been adopted by Zalesky et al. (2010) and Richiardi et al. (in press) using network-based statistics and machine learning methods, respectively, for the comparison of two populations of subjects and controls. Our approach is slightly more general since it accommodates complex experimental designs where information is pooled over several (more than two) experimental conditions. The SPN framework shares with SPM a concise visualization of the data, and is therefore particularly useful for the presentation and reporting of results. Specific to network analysis, differential SPNs have the additional advantage of replacing some of the concerns with the choice of a threshold value to compute the adjacency matrix with the selection of a specific p-value. In this context, thresholding is supplanted by inference (Achard et al., 2006; He et al., 2009b).

A recent review of the literature on complex brain networks has emphasized the paucity of research papers linking network topology with cognitive and behavioral functions (Bullmore and Sporns, 2009). The present investigation should therefore be regarded as exploratory in nature, as only few previous results can be used for comparison. As little research has been done on the dynamics of functional networks under different levels of cognitive demand, no prior hypothesis was

formulated on the nature of the potential relationship between the global and local efficiencies of these functional networks under different levels of working memory load.

Methods

Subjects

The participants were adult healthy controls. Functional MRI data was acquired from $n=43$ (incl. 21 females) volunteers with no previous history of mental disorders. The subjects had a mean age of 68.23 years ($sd=13.17$ and $range=\{30,89\}$). Full ethical approval for this project was gained from the South London and Maudsley (SLaM) ethics committee. The 43 subjects selected for this study were all screened on their ability to respond correctly to the demands of the *N*-back task. In order to verify the subjects' attendance to the task at hand, we selected participants on the basis of their performance on the most challenging experimental condition: 3-back. Only subjects who had responded correctly to more than 75% of the items in the 3-back condition were included in the study. Therefore, the 43 subjects here discussed constitute a subsample of high-performers among an original population of 84 healthy controls. This screening of the participants based on their capacity to successfully complete the task's requirements echoed the perspective of previous *N*-back researchers, who have questioned the validity of fMRI studies based on unsuccessful performances (Callicott et al., 1999).

Working memory task

In the specific variant of the *N*-back paradigm used in the present study, the stimuli were letters and the input modality was visual (see Owen et al. (2005), for a review of variants of the *N*-back task). Subjects were shown one letter every 2 s, and were asked to monitor the stimuli, in order to indicate by the push of a button whether the current letter was identical to the one presented *N* trials previously, where $N=\{1,2,3\}$. A control or null condition was also included – the 0-back task – which consisted of simply indicating whether the current letter was an X. The full experiment consisted of 12 randomized blocks organized in three full experimental cycles including each of the four conditions. Each block lasted 31 s and contained three test stimuli requiring a positive answer from the subjects, and 12 other letters, which did not necessitate an answer. Written instructions on the screen between each block lasted for 3 s, and the entire experiment therefore took 372 s per subject to complete. Response times were recorded for all trials in milliseconds. All subjects were provided with as many training blocks as they requested to ensure that they were confident with the task at hand.

Data acquisition and preprocessing

Functional MRI data were acquired on a GE Signa 1.5 T MRI scanner (General Electric, Milwaukee, WI, USA). 186 $T2^*$-weighted EPI volumes depicting BOLD contrast were acquired in each of 16 non-contiguous near-axial planes (5.0 mm thick with 0.5 mm interval, in-plane resolution 3 mm) parallel to the Anterior Commissure Posterior Commissure line (TE = 40 ms, TR = 2000 ms, flip angle 90°, and number of signal averages was 1). The entire preprocessing of the images was conducted within the FMRIB Software Library (FSL) environment (Smith et al., 2004). The McFLIRT package was used for affine registration of the time series to the first image with 6 degrees of freedom (*df*), including correction for movement artifacts (Jenkinson et al., 2002). Images were normalized to the Montreal Neurological Institute (MNI-152) template based on $2\,mm^3$ voxels. Normalization (6*df*) was based on the minimization of the correlation ratio and used trilinear interpolation (Jenkinson and Smith, 2001). Finally, the 4D volumes were regressed on the movement correction parameters, including the square, first derivatives and square of the first derivatives of these movement correction parameters, with a total of 24 regression parameters (Lund et al., 2006). Images were not smoothed during preprocessing prior to parcellation.

Parcellation and wavelet decomposition

The FSL residual time series for the 43 subjects were parcellated using the Anatomical Automatic Labelling (AAL) templates (Tzourio-Mazoyer et al., 2002). This parcellation scheme produced a set of 45 cortical and subcortical regions for each hemisphere. Each volume was thus parcellated into 90 regions of interest and the BOLD signal was averaged over the voxels in each of these regions at every time point. Note that we did not include cerebellar regions in our investigation. The resulting regional mean time series constituted the main input for the remaining part of the analysis.

These regional mean time series were further decomposed into six different frequency bands using maximal overlap discrete wavelet transform (MODWT), as described in Percival and Walden (2000), and available in the waveslim package in the R language (www.cran.r-project.org). The frequencies corresponding to the different scales of the wavelet bases used in our investigations were computed by calculating the spectral density functions of the wavelet coefficients. For each wavelet scale, we computed the frequency interval as follows. Using the standard rule, $[2^{-k-1}/TR, 2^{-k}/TR]$, to compute the bounds of the frequency interval of the k^{th} scale, we obtained 0.031–0.062 Hz and 0.015–0.031 Hz for wavelet scales 3 and 4, which will be referred to as the 0.03–0.06 Hz and 0.01–0.03 Hz frequency intervals, respectively. Our procedure is consistent with previous research, which has highlighted that low frequency bands (<0.1 Hz) are particularly salient to the study of functional connectivity (Achard et al., 2006; Biswal et al., 1995; Cordes et al., 2000; Fransson, 2005). Wavelet decomposition has been extensively applied to time–frequency analysis of non-stationary signals (Bullmore et al., 2003). Such decompositions allow the detection of small signal changes or correlations against a noisy background, and have proved especially useful in functional MRI, where it tends to increase sensitivity (Brammer, 1998). One of the advantages of using wavelets over other band-pass filters is that they produce an orthogonal decomposition of the signal of interest.

The decomposed mean time series were separated according to the four conditions of the *N*-back paradigm. The first three volumes of each block were discarded, to allow for the lag in hemodynamic response between blocks. The remaining volumes were concatenated to form four distinct time series representing the four different conditions. Although other authors have used the concatenation of time series in the context of functional connectivity (Fair et al., 2007), the combination of wavelet decomposition with block concatenation tends to inflate the between-series correlations. We have explored this issue in detail in the Supplementary Methods. Through the use of simulated time series, we have shown that wavelet decomposition in conjunction with block concatenation (hereafter the WDC procedure) increases both the within-condition variance and the between-condition covariances of the inter-region correlation coefficients. In addition, between-condition differences in correlation coefficients were also increased by WDC procedure. Most importantly, this inflation of between-condition differences in correlations was not found to be systematically biased in favor of one condition. Since our main purpose in this paper was to infer between-condition differences, we therefore concluded that no bias was introduced through out the preprocessing protocol.

Statistical Parametric Networks (SPNs)

Networks derived from both structural and functional neuroimaging data are generally produced by thresholding correlation matrices

(Bullmore and Sporns, 2009). Thresholding is equivalent to the application of an entrywise indicator function. Such a function, however, is not linear. Precisely, an indicator function is a *quasilinear* map (see Appendix for technical details). Therefore, given a thresholding function $T_\tau(\cdot)$, which applies a threshold τ, the following inequation holds,

$$\sum_{i=1}^{n} T_\tau(\mathbf{R}_i) \neq T_\tau\left(\sum_{i=1}^{n} \mathbf{R}_i\right) \tag{1}$$

where $i=1,\ldots,n$ defines the subjects taking part in the experiment, and where $\mathbf{R}_i$'s are subject-specific correlation matrices. Eq. (1) states that averaging over thresholded matrices is not equal to thresholding the mean correlation matrix. There is currently no guidance on which one of these two quantities should be reported when summarizing the results of a particular study. There is therefore a pressing need to formalize and standardize the compilation of summary networks, which adequately represent variability over a population of networks. Some progress in that direction has been made by Achard et al. (2006) and He et al. (2009b) for the case of a single population of networks. Here, we extend these ideas to the case of several populations of networks.

The problem of comparing networks over different experimental conditions centers on two distinct questions: (i) what are the nodes and (ii) what are the edges that significantly vary due to the experimental design. Statistical inference for these two distinct but related questions permits the construction of SPNs, similar in spirit to SPMs, thus justifying our choice of name. One of the key advantages of this method is that it answers the problem posed by the quasilinearity of the thresholding function highlighted in Eq. (1). Since we are drawing inference from the correlation coefficients themselves, we bypass the problem of averaging over a set of networks, while still providing a statistical summary in the form of a graph. In effect, we are trading the choice of a thresholding cut-off point with the use of a specific significance level. Note, however, that this is only true for differential SPNs and not for mean SPNs, which still require the application of a particular cut-off point.

We will employ standard graph-theoretical notation to introduce our approach to the above problem (Bollobas, 1998). A graph G is formally defined as an ordered pair (V,E), where $V:=V(G)$ is the set of *vertices* (or nodes) in the graph and $E:=E(G)$ is the set of *edges* (or connections), with $:=$ indicating the use of a definition. Note that $E\ V\otimes V$, where the Cartesian product $V\otimes V$ indicates the complete graph, containing all possible edges. An element of E will be denoted by $e:=v_a v_b$, which represents a connection between vertices v_a and v_b. For notational convenience, the sets V and E will refer to both the sets of their elements and the sets of the corresponding indices of these elements. The total number of edges and total number of nodes in G can be obtained from the cardinality of sets E and V, and will be denoted by $N_E:=|E|$ and $N_V:=|V|$, respectively, where $|\cdot|$ denotes the cardinality of a set.

The raw data of interest is a $(n\times J)$-matrix of correlation matrices, where the indexes $i=1,\ldots,n$ and $j=1,\ldots,J$ label the individual subjects and the conditions of the experiment of interest, respectively. For the N-back task used in this paper, there are $J=4$ conditions. The entire data set can thus be represented by the following matrix,

$$\mathbf{R} := \begin{vmatrix} \mathbf{R}_{11} & \ldots & \mathbf{R}_{1J} \\ \vdots & \ddots & \vdots \\ \mathbf{R}_{n1} & \ldots & \mathbf{R}_{nJ} \end{vmatrix}, \tag{2}$$

where each element $\mathbf{R}_{ij}$ in Eq. (2) is a correlation matrix of order $N_V\times N_V$. In our specific case, we have used $N_V=90$ vertices corresponding to the 90 regions in the AAL parcellation scheme. These vertices are labelled by $v=1,\ldots,N_V$. For simplicity, an entry in the $\mathbf{R}$, will be denoted by r_{ij}^e, where $e\in E$ denotes an edge from the saturated graph G_S, which has the maximal number of edges — that is, $|E(G_S)|=N_V(N_V-1)/2$. We will systematically use superscripts to denote edges. The $\mathbf{R}$ matrix and the production of mean and differential SPNs from this raw data are illustrated in Fig. 1.

Mean SPNs

A mean SPN provides a concise manner to statistically infer the 'mean' functional network of a population of subjects. These SPNs are constructed using a mass-univariate approach based on a set of N_E z-tests. In the context of our specific experiment, this operation is repeated for each condition $j=1,\ldots,J$, once for each of the columns of array $\mathbf{R}$ from Eq. (2). That is, we consider a vector of correlation matrices,

$$\mathbf{R}_j := \left[\mathbf{R}_{1j},\ldots,\mathbf{R}_{nj}\right], \tag{3}$$

independently for each of the J conditions. Every correlation coefficient was Fisher z-transformed in order to allow the application of the normal theory to the distribution of these quantities. We standardize each condition-specific mean correlation coefficient $\left\{\overline{r}_j^e\right\}$, with respect to the grand sample mean $\overline{\overline{r}}$ and the grand sample standard deviation $sd(\mathbf{r})$ of the population of correlation coefficients in $\mathbf{R}$, where the grand sample mean is defined as

$$\overline{\overline{r}} := \frac{1}{nJN_E}\sum_{i=1}^{n}\sum_{j=1}^{J}\sum_{e=1}^{N_E} r_{ij}^e, \tag{4}$$

and the grand sample standard deviation is

$$sd(\mathbf{r}) := \left(\frac{1}{nJN_E-1}\sum_{i=1}^{n}\sum_{j=1}^{J}\sum_{e=1}^{N_E}\left(r_{ij}^e-\overline{\overline{r}}\right)^2\right)^{1/2}, \tag{5}$$

where the vector $\mathbf{r}$ combines all the upper off-diagonal elements in every matrix contained in $\mathbf{R}$, in Eq. (2). The standardized correlation coefficients are then defined as

$$z_j^e := \frac{|\overline{r}_j^e-\overline{\overline{r}}|}{sd(\mathbf{r})/\sqrt{(n)}}, \tag{6}$$

where $\overline{r}_j^e := n^{-1}\sum_{i=1}^{n} r_{ij}^e$ for the e^{th} edge. These standardized values thus become location test statistics, which are then tested with respect to a normal distribution. For every edge $e\in E$, we test whether the following inequality is verified:

$$1-\Phi\left(-z_j^e\right) \leq \alpha, \tag{7}$$

where $\Phi(\cdot)$ is the standard normal cumulative distribution function (CDF), and α is a given significance level to test edges.

The use of a z-test in this context is justified by our preliminary use of the Fisher's z-transform on every correlation coefficient in the sample. Inference testing is here directional, since we are only interested in identifying the edges, which are significantly larger than the grand sample mean, $\overline{\overline{r}}$. The mean networks will be denoted by $\overline{SPN}_j$ for each of the $j=1,\ldots,J$ conditions.

Differential SPNs

For both edge- and node-specific differential SPNs, we will make extensive use of mixed effect models (see Demidenko (2004), for a comprehensive treatment). The question of whether edges are 'gained' or 'lost' as a result of a particular experimental manipulation can be answered by producing a pair of differential networks: the

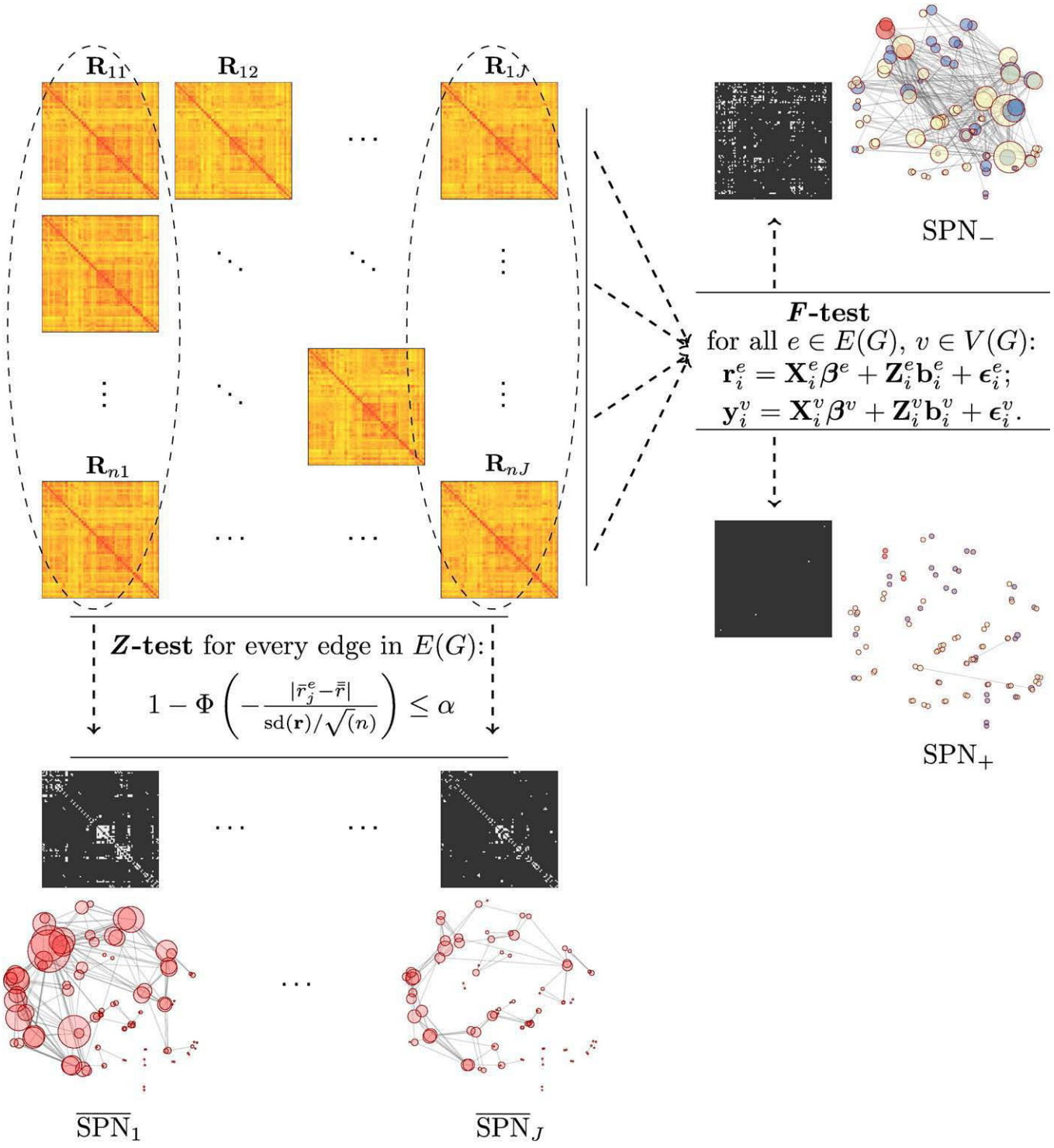


Fig. 1. Statistical Parametric Network (SPN) framework. On the left-hand side, the set of correlation matrices $\mathbf{R}_{ij}$ is indexed per individual subject $i = 1,\ldots,n$ and per level of experimental manipulation $j = 1,\ldots,J$, which correspond to rows and columns, respectively. On the bottom left-hand side of the figure, the mean networks ($\overline{SPN}_j$) for each experimental condition are produced by running z-tests for every individual edge, combining data over all individuals in each column, with respect to the grand sample mean, $\bar{\bar{r}}$. On the right-hand side, we make use of the entire array of correlation matrices $\{\mathbf{R}_{ij}\}$ to compute F-tests for every individual edge and voxel in order to produce downweighting and upweighting networks, denoted SPN_- and SPN_+, respectively. For all SPNs, the adjacency matrices used to construct the graphs are thresholded using the FDR with a base rate of $\alpha_0 = 0.05$.

downweighted or upweighted SPNs, which we will denote as SPN_- and SPN_+, respectively. As for mean SPNs, this problem can also be tackled using a mass-univariate approach. This time, however, such inference needs to be drawn from the entire set of correlation coefficients — that is, all of the data contained in $\mathbf{R}$, as described in Fig. 1. This involves N_E linear mixed effects models, which can be formulated using the formalism introduced by Laird and Ware (1982) as follows,

$$\mathbf{r}_i^e = \mathbf{X}_i^e \boldsymbol{\beta}^e + \mathbf{Z}_i^e \mathbf{b}_i^e + \epsilon_i^e, \qquad i = 1,\ldots,n, \tag{8}$$

where $\mathbf{r}_i^e := [r_{i1}^e, \ldots, r_{ij}^e]^T$ are the correlation coefficients of interest, $\boldsymbol{\beta}^e := [\beta_1^e, \ldots, \beta_j^e]^T$ is a vector of fixed effect, which does not vary over subjects, $\mathbf{b}_i^e := [b_{i1}^e, \ldots, b_{ij}^e]^T$ is the vector of subject-specific random effects and $\epsilon_i^e := [\epsilon_{i1}^e, \ldots, \epsilon_{ij}^e]^T$ are the residuals. Finally, the matrices $\mathbf{X}_i^e$'s and $\mathbf{Z}_i^e$'s are given the following specification,

$$\mathbf{X}_i^e = \mathbf{Z}_i^e = \begin{vmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{vmatrix}, \tag{9}$$

for every $i = 1, \ldots, n$. This linear mixed model is commonly referred to as a *growth curve* model, where subject-specific trajectories through the different levels of the experimental factor – in our case, the four conditions of the *N*-back task – are extracted before considering the experiment-wise effect $\boldsymbol{\beta}^e$ (see Bollen and Curran (2006), for an introduction). An edge was included in a differential SPN when the corresponding *F*-test for the experimental factor was found to be significant. The parametric test, in this context, is the Wald *F*-test, which tests whether the inclusion of $\boldsymbol{\beta}^e$ in the model explains more variance than the nested model without that parameter (Pinheiro and Bates, 2000). Edges were included in SPN_- and SPN_+ if the sign of the largest regression coefficient in $\boldsymbol{\beta}^e$ was negative or positive, respectively; and the corresponding *F*-test was found to be significant.

Similarly, the estimation of the upweighting and downweighting of the wavelet coefficients at single *nodes* was computed by a set of N_V growth curve models. Here, the data is the set of matrices $\mathbf{Y}^v := \{y_{ij}^v\}$, where $v \in V$, which corresponds to the time-averaged wavelet coefficients for the v^{th} region in the j^{th} experimental condition, for subject *i*. One can reformulate the system of equations for evaluating edges in Eq. (8) by using superscripts to denote vertices. This gives

$$\mathbf{y}_i^v = \mathbf{X}_i^v \boldsymbol{\beta}^v + \mathbf{Z}_i^v \mathbf{b}_i^v + \epsilon_i^v \qquad i = 1, \ldots, n, \tag{10}$$

where $\mathbf{y}_i^e := [y_{i1}^e, \ldots, y_{ij}^e]^T$ and the definitions used in Eq. (8) also apply to $\mathbf{X}_i^v$, $\mathbf{Z}_i^v$, $\boldsymbol{\beta}^v$, $\mathbf{b}_i^v$ and ϵ_i^v. As for edge-specific differential SPNs, a node was judged to be either significantly upweighted or significantly downweighted depending on the sign of the largest coefficient in the corresponding vector $\boldsymbol{\beta}^v$, and whether the *F*-test for that model was above a given significance level α.

Although our methodology was here described for a growth curve model, which is particularly suited to the data at hand, any other general linear models (GLMs) could be applied in this context. Our proposed procedure is a mass-univariate approach to the problem of comparing topological networks, and therefore general remarks about SPMs are also applicable to SPNs. All mixed linear models were fitted using restricted maximum likelihood (REML) as developed by Bartlett (1937). Computations were conducted using the lme4 package freely available in the R language (see www.cran.r-project.org and Pinheiro and Bates, 2000). Throughout the rest of this paper, we infer the significance of edges and nodes in both $\overline{SPN}_j$ and $SPN_\pm$ by using the false discovery rate (FDR) with a base rate of $\alpha_0 := 0.05$ (Benjamini and Hochberg, 1995; Nichols and Hayasaka, 2003). The conventional thresholding used in network analysis is therefore here superseded by the application of standard multiple testing correction methods. The advantage of this approach lies in its pooling of information over several subjects, to produce robust edge- and node-specific statistics.

Measures of network topology

Small-worldness has been quantified using a range of different metrics. Metrics of network topology commonly utilized in neuroscience and computational biology have been reviewed by Rubinov and Sporns (2010). A comprehensive introduction to the statistical mechanics of networks can be found in Albert and Barabasi (2002). Here, we favored the use of global and local efficiencies (Latora and Marchiori, 2001), over the classical measures of small-worldness — i.e. characteristic path length and clustering coefficients (see Dijkstra (1959) and Luce and Perry (1949), for the historical references about the concepts). Previous investigators such as Achard and Bullmore (2007) have already noted that such quantities are advantageous from a computational perspective. Global efficiency should therefore be interpreted as inversely related to characteristic path length, whereas local efficiency can be regarded as related to the clustering coefficient (Latora and Marchiori, 2001).

The efficiency of any network – connected or disconnected – is defined by the following formula due to Latora and Marchiori (2001),

$$E(G) := \frac{1}{N_V(N_V - 1)} \sum_{i \in V} \sum_{j \neq i \in V} d_{ij}^{-1}, \tag{11}$$

with $N_V := |V|$, and where d_{ij} is the length of the shortest path between vertices *i* and *j*, and the set $\{j \neq i \in V\}$ is the set of all indices in *V*, which are different from *i*. Efficiency can be seen to be equivalent to the inverse of the harmonic mean of the length of the shortest paths between each pair of nodes in the network *G*. From Eq. (11), we can immediately derive the definition of global efficiency, which is $E^{Glo}(G) := E(G)$. This quantity can be understood as the efficiency of parallel information transfer in the entire system.

Local efficiency, by contrast, measures how fault-tolerant the system is at a local level. Let sub-network G_i, which contains all the neighbors of the i^{th} vertex. That is, $V(G_i) := \{v_j \in G | v_j \sim v_i\}$, where $v_j \sim v_i$ signifies that nodes *i* and *j* are connected. By convention, we have $v_i \notin V(G_i)$ (see Latora and Marchiori, 2001; Latora and Marchiori, 2003). Now, the definitions of global and local efficiencies are directly related to each other in the following fashion:

$$E^{Glo}(G) := E(G), \quad \text{and} \quad E^{Loc}(G) := \frac{1}{N_V} \sum_{i \in V} E(G_i). \tag{12}$$

That is, these two quantities are derived from the general concept of efficiency described in Eq. (11). While global efficiency is the efficiency of the entire graph *G*, local efficiency is the averaged efficiency of all first-order neighborhoods. The computation of two metrics similar to the characteristic path length and the clustering coefficient, using one single measure of efficiency is a remarkable insight. Whereas Watts and Strogatz (1998) considered global information flow and local clustering as two distinct phenomena, Latora and Marchiori (2001) combined these two notions into a single network property by showing that they both stem from a unique concept, which can loosely be interpreted as the efficiency of information transfer.

In the sequel, we will also require the regional efficiency of a node. The region-specific global efficiency, denoted $E^{Glo}(G, v)$ and sometimes simply referred to as *regional* efficiency (Achard and Bullmore, 2007), quantifies the connectivity of each node to all the other nodes in the network. It is defined as follows

$$E^{Glo}(G, v) := \frac{1}{N_V - 1} \sum_{j \neq v \in V} d_{vj}^{-1}, \tag{13}$$

where it should be noted that $E^{Glo}(G) = N_V^{-1} \sum_{v \in V} E^{Glo}(G, v)$.

Global, local and regional efficiencies are dependent on the cost of the underlying networks. Our goal in this paper is to compare network efficiencies over different experimental conditions. Experimental manipulation, however, may influence functional network costs (i.e. overall level of connectivity). Therefore, it is of interest to derive efficiency metrics, which are independent of network cost, thereby allowing cost-independent comparisons across conditions. Let k_i denote the degree of vertex i — i.e. the number of connections of

the i^{th} node, such that $k_i := \sum_{j \in V} a_{ij}$, where matrix $\mathbf{A}$ with entries $\{a_{ij}\}$ is the adjacency matrix of G satisfying

$$a_{ij} := \begin{cases} 1 & \text{if } v_i \sim v_j; \\ 0 & \text{otherwise.} \end{cases} \tag{14}$$

Another measure of a graph's connectivity strength is its cost, denoted $K(G)$. To define this quantity, we need to compute the saturated or complete graph, G_S, corresponding to our graph of interest G. The network G_S has the same number of vertices as G, but the maximal number of edges, such that $|E(G_S)| = N_V(N_V - 1)/2$; where recall that $N_V := |V|$ is the number of vertices in G. We can then define the cost of G by

$$K(G) := \frac{|E(G)|}{|E(G_S)|} = \frac{N_E}{N_V(N_V - 1)/2}, \tag{15}$$

where as before $|\cdot|$ indicates the cardinality of each set. The cost of a graph is inversely related to the sparsity of its adjacency matrix. The sparsity of G is the percentage of non-zero entries among the off-diagonal elements of the adjacency matrix subtending G. Cost is important when one is concerned with evaluating the changes in levels of connectivity across different conditions, or between different populations. In this vein, global and local efficiencies as defined in Eq. (12) can be said to be absolute metrics (He et al., 2009a). In addition, we will make use of the weighted cost, which we define for any correlation matrix R_{ij} as follows,

$$K_W\left(\mathbf{R}_{ij}\right) := \frac{1}{N_E} \sum_{e=1}^{N_E} r_{ij}^e, \tag{16}$$

for every i^{th} subject and every j^{th} condition, where $N_E := N_V(N_V - 1)/2$ since we are here considering weighted networks. An identical measure has been previously defined by Barrat et al. (2004) under the name of *average nodal strength*. A non-standardized version of the cost of a weighted network similar to Eq. (16) was introduced by De Vico Fallani et al. (2008). Note that $K_W(\mathbf{R}_{ij})$ is equivalent to the mean correlation coefficient in $\mathbf{R}_{ij}$.

We will also require *cost-integrated* topological metrics, in order to conduct inference across experimental conditions. By integrating over cost, the topological metrics of interest become independent of the differences in level of connectivity of the networks under comparison (He et al., 2009a). It therefore suffices to integrate the aforementioned measures of global and local efficiencies with respect to $K \in [0,1]$. This gives

$$\begin{aligned} \overline{E}^{Glo}(G) &:= \int_{[0,1]} E^{Glo}(G(K))dK, \\ \overline{E}^{Loc}(G) &:= \int_{[0,1]} E^{Loc}(G(K))dK, \end{aligned} \tag{17}$$

for the global and local efficiencies, respectively; where $G(K)$ is the *unweighted* graph obtained by ranking the entries of the corresponding correlation matrix and giving a value of 1.0 to all entries greater or equal to the $(N_E - 1 + KN_E)^{\text{th}}$ ranked correlation coefficient and a value of 0 otherwise. This procedure produces an adjacency matrix with cost K. One can also compute the integrated version of the regional efficiency, originally described in Eq. (13),

$$\overline{E}^{Glo}(G, v) := \int_{[0,1]} E^{Glo}(G(K), v)dK, \tag{18}$$

for the v^{th} vertex. In the following part of this paper, the cost-integrated quantities in Eqs. (17) and (18) were evaluated by discretizing the cost domains of the relevant efficiency measures over 30 intervals and averaging over them using the length of these intervals (every interval has approximate length of 0.03). A more sophisticated technique for approximating these integrals based on Monte Carlo integration was also implemented and is reported in Section S3 of the Supplementary Methods. Cost-integration in this context can therefore be interpreted as an approximation of the mean topological measure over all possible cost levels.

Finally, we utilized matched random and regular networks for comparison purposes. That is, networks, which have the same number of vertices and edges as the networks of interest. The random network, denoted G^{Rdm}, matches the properties of G, such that it satisfies $V(G^{Rdm}) = V(G)$. Its edges are allocated using the following rule,

$$\mathbb{P}\left[v_i \sim v_j\right] := \frac{|E(G)|}{|E(G_S)|} = K(G), \tag{19}$$

which is equivalent to the cost of graph G, as described in Eq. (15). In effect, we specified a uniform distribution over the off-diagonal elements of the adjacency matrix of G^{Rdm}. This distribution is proportional to the number of edges in G divided by the number of edges in the corresponding saturated graph. We also made use of matched regular networks or regular lattices, which we denoted by G^{Reg}. These were constructed by assigning unit entries to the off-diagonal elements in the adjacency matrix of G^{Reg}, starting from the elements closest to the diagonal. This process was repeated until exhaustion of the number of edges in the original network G. A representation of the adjacency matrices of a small-world graph G and its (N_V, N_E)-matched random and regular networks is given in Fig. 2.

In summary, this paper will focus on the following four measures of subject-specific connectivity strength and cost-integrated topological metrics:

i. Connectivity strength is measured by weighted cost (i.e. mean correlation coefficient), denoted $K_W(\mathbf{R}_{ij})$.
ii. Cost-integrated topology, by contrast, is measured using:
   a. cost-integrated global efficiency, $\overline{E}^{Glo}(G_{ij})$,
   b. cost-integrated local efficiency, $\overline{E}^{Loc}(G_{ij})$,
   c. and cost-integrated regional efficiency, $\overline{E}^{Glo}(G_{ij}, v)$.

All of these four quantities are subject-specific in the sense that they are computed for every i^{th} subject and every j^{th} condition. In addition, graphical visualization of the weighted networks will be provided by using the statistical properties of the mean and differential SPNs, denoted $\overline{SPN}_j$ and $SPN_{\pm}$, respectively. These summary measures are available in the form of object-oriented functions in the R language, within the *NetworkAnalysis* package (see http://cran.r-project.org/web/packages/NetworkAnalysis/index.html).

Results

Our main focus, in this report, was low frequency (0.01–0.03 Hz) BOLD signal correlations. Results for a slightly higher frequency interval (0.03–0.06 Hz) were also computed. Generally, results for the lower frequency band were not replicated for networks built from wavelet coefficients in the 0.03–0.06 Hz frequency interval. As such investigations are equivalent to testing two separate sets of hypotheses for each frequency band, we corrected for such multiple comparisons using the Bonferroni's correction. In the sequel, we present our findings for the lower frequency interval (0.01–0.03 Hz), and discuss the notable differences between the two frequency intervals. We considered functional networks' changes in both connectivity strength (i.e. weighted cost) and cost-integrated efficiency measures (i.e. cost-integrated global, local and regional efficiencies).

Connectivity strength is inversely related to working memory load

Differences in the sparsity of the adjacency matrices subtending the mean SPNs in each condition can be visualized in Fig. 3. These differences can be tested independently of the choice of threshold τ, by

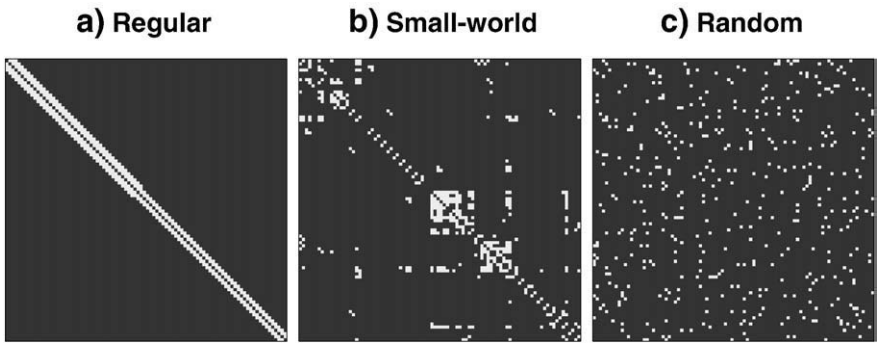


Fig. 2. Adjacency matrices for a small-world network G and its (N_V, N_E)-matched regular and random networks. Panel (b) shows the original network of interest G characterized by a small-world structure, which is used to construct the (N_V, N_E)-matched regular lattice G^{Reg}, in panel (a); and the random graph G^{Rdm}, in panel (c). By convention, the diagonal entries satisfy $a_{ii} = 0$ for every vertex i in these three matrices.

considering the mean correlation coefficients under each condition, computed for each subject. We thus drew inference from the correlation matrices in the first row of Fig. 3, as opposed to the thresholded matrices in the second row, thereby bypassing the arbitrary choice of a threshold value τ. Note that this procedure is similar, in spirit, to the consideration of *weighted* networks, as described by some authors (Rubinov and Sporns, 2010). A growth curve model, as described in Eq. (8), specifying subject-specific random effects demonstrated that the *N*-back factor was a significant predictor of the mean correlation coefficient (Wald $F = 3.59, df_1 = 3, df_2 = 126, p = 0.015$), where $df_2 = nJ-(J-1)-n$ due to the use of n random effects in the mixed linear model. This result remained significant when the 0-back condition was excluded from the analysis (Wald $F = 3.85, df_1 = 2, df_2 = 84, p = 0.025$). A similar analysis was conducted with mean correlation coefficients based on a higher frequency interval (0.03–0.06 Hz), and yielded no significant results (Wald $F = 0.55, df_1 = 3, df_2 = 126, p = 0.64$). Thus, after applying Bonferroni's correction for multiple comparisons over two different frequency bands, connectivity strength at low frequencies (0.01–0.03 Hz) was found to significantly diminish as subjects experienced increasing working memory load, but not in the higher frequency band (0.03–0.06 Hz).

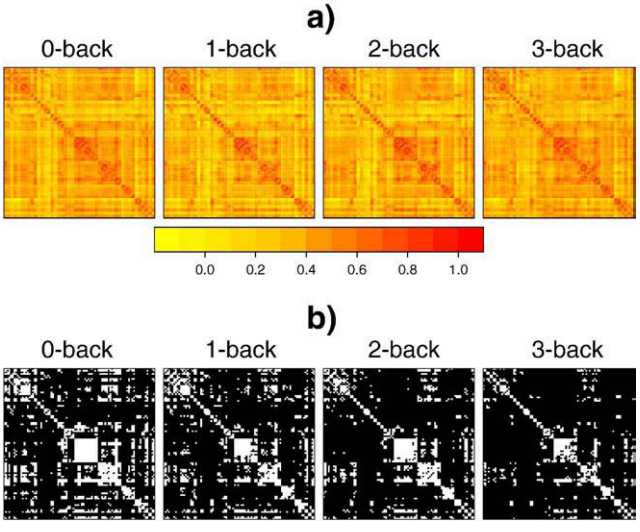


Fig. 3. Adjacency matrices become sparser with increasing working memory load. In panel (a), heatmaps corresponding to the correlation matrices in each of the four *N*-back conditions, for $n = 43$ subjects. In panel (b), the adjacency matrices were obtained by constructing $\overline{SPN}_j$, using a mass-univariate approach based on z-tests with respect to the grand sample mean $\bar{\bar{r}}$ and the grand sample standard deviation $sd(\mathbf{r})$ with FDR correction (base rate $\alpha_0 = 0.05$). The frequency interval was 0.01–0.03 Hz. Zero entries are denoted in black in the adjacency matrices.

Pruning of functional networks

A more detailed analysis of the differences between the functional networks underlying cognitive activity in the four experimental conditions was obtained by directly evaluating mean and differential SPNs (see Figs. 4 and 5, respectively). The $\overline{SPN}_j$'s in Fig. 4 summarize statistical inference for each edge of the functional networks in each condition. In particular, one could observe that an inter-hemispheric connectivity pattern is emerging with symmetric regions in each hemisphere being more likely to be connected. This finding is consistent with the general consensus in brain anatomy, stating that most fibres in the corpus callosum link symmetrical regions in the brain (Zaidel and Iacoboni, 2003). This symmetry was also observable in terms of nodal degrees: Identical regions in each hemisphere appeared to be more likely to exhibit a similar number of connections.

Fig. 5 reports negative and positive differential SPNs, corresponding to significantly 'lost' and 'gained' connections, respectively, as subjects experienced increasing working memory load. Visual inspection of these differential SPNs indicates that while a number of anterior–posterior connections were found to be significantly downweighted, few connections were found to be significantly upweighted, when comparing the 0-back with the 3-back condition. An increase in working memory load seemed therefore to correspond to a considerable *pruning* of a number of connections. Additionally, one could also observe that the vertices, whose wavelet coefficients have been significantly down-weighted (indicated in blue in Fig. 5) approximately mapped onto the set of high-degree nodes — thereby providing some moderate agreement between the pruning of some vertices, and the diminution of the BOLD signal in these regions. That is, the regions that had been most affected by a loss of connections (represented by large circles in Fig. 5) tended to have also witnessed a diminution in level of activation, as measured by the size of the wavelet coefficient in the same frequency band. This effect of cognitive load on functional connectivity was found to be also present when considering pairs of experimental conditions. We report such pairwise comparisons in the Supplementary Methods section. A number of edges are shown to be deprecated between 0-back and 1-back, between 1-back and 2-back and between 2-back and 3-back (see Supplementary Methods S2). These figures support the idea that a gradual decrease in connectivity parallels an increase in cognitive load.

Several high-degree nodes in panel (a), however, did not show a sufficient decrease in BOLD signal to reach significance. The color scheme in Fig. 5 only partially reflect changes in levels of node-specific BOLD signal, as it is restricted to the particular frequency band of interest (0.01–0.03 Hz, in our case). No significant relationships were found between the nodes' degrees and average (over subjects) region-specific wavelet coefficient ($t = -1.43, df = 88, p = 0.15$). The differential SPN framework allowed the computation of the topological properties of the deprecated or 'lost' network. Interestingly, it can be shown that the SPN_- in this experiment displayed the hallmarks of a

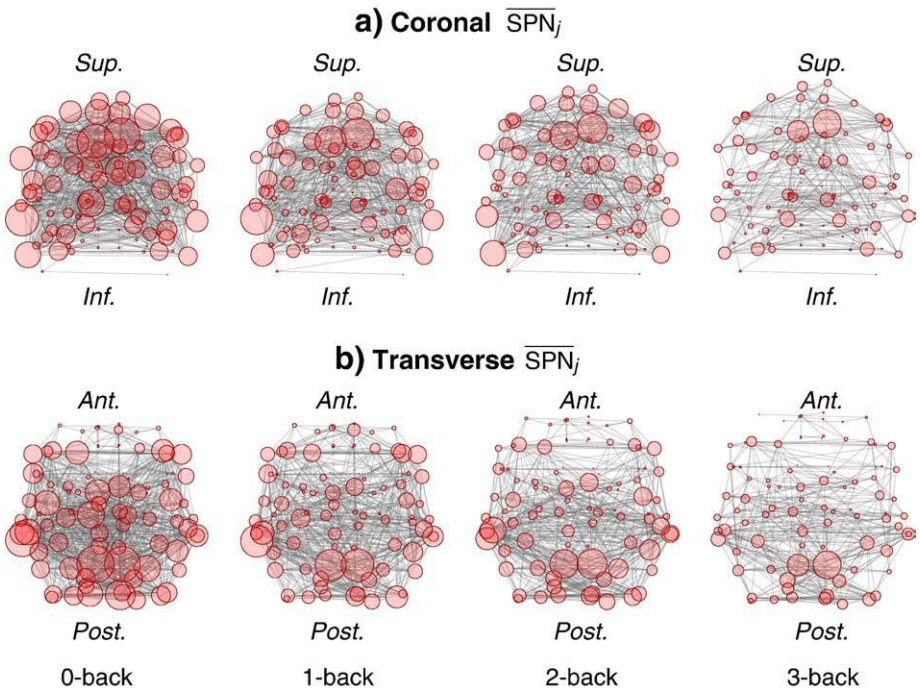


Fig. 4. Mean Statistical Parametric Networks ($\overline{SPN}_j$) over the 4 levels of the *N*-back task, in the coronal (a) and transverse (b) planes, based on wavelet coefficients in the 0.01–0.03 Hz frequency band, with FDR correction (base rate $\alpha_0 = 0.05$). Locations of the nodes correspond to the stereotaxic centroids of the corresponding cortical regions. The orientation axes are indicated in italics: inferior–superior and anterior–posterior for the coronal and transverse sections, respectively. The size of each node is proportional to its degree.

small-world topology. The local efficiency of SPN_- was found to be greater than that of a random network with an identical number of nodes and edges ($E^{Loc}(SPN_-) = 0.23$, whereas $E^{Loc}(G^{Rdm}) = 0.085$). The global efficiency of SPN_- was also larger than that of a (N_V, N_E)-matched regular lattice ($E^{Glo}(SPN_-) = 0.33$, compared to 0.211 for G^{Reg}). Functional networks were therefore found to be considerably pruned as working memory load increased, and the network formed by downweighted connections was itself found to exhibit the small-world property.

Integrated global and local efficiencies were preserved

Fig. 6 provides a description of the variations in global and local efficiencies of the *N*-back functional networks. The evolution of both E^{Glo} and E^{Loc} were evaluated over the full range of graph costs by varying the value of the significance threshold α of the $\overline{SPN}_j$'s. All functional networks under the different conditions of the *N*-back task tended to remain in the small-world regime with generally higher E^{Glo} than an equivalent regular lattice, as well as higher E^{Loc} than an equivalent random network. The exact pattern of changes between functional networks corresponding to different working memory load seemed to vary with the cost of the network. Functional graphs appeared to become more random as working memory load increased, especially when considering high cost ($K \geq 0.60$). However, a different pattern seemed to emerge when comparing *N*-back functional networks under lower cost regimes ($0.2 \geq K \geq 0.50$). We clarified these differences by comparing the cost-integrated global and local efficiencies integrated over the entire cost domain ($K \in [0,1]$) by the four *N*-back conditions.

To assess whether functional networks significantly differed in levels of global and local efficiency, the effect of the *N*-back factor on the topological metrics across the experimental conditions was tested. However, since the number of edges varied from one condition to the next, our chosen measures of small-worldness were standardized by integrating E^{Glo} and E^{Loc} with respect to cost, as described in Eq. (17), and denoted $\overline{E}^{Glo}$ and $\overline{E}^{Loc}$, respectively. These cost-integrated topological measures correspond to the areas under the curves in Fig. 6. Using these standardized metrics, the main effect of the *N*-back factor on global efficiency was not found to be statistically significant (Wald $F = 1.81, df_1 = 3, df_2 = 126, p = 0.15$), using a mixed effect model, controlling for within-subject variance (see Pinheiro and Bates, 2000). Boxplots of the data used in this test can be seen in Fig. 7, panel (a). Similarly, we tested for the effect of the *N*-back factor on integrated local efficiency specifying a random effect for each subject, which was also found to be non-significant (Wald $F = 1.08, df_1 = 3, df_2 = 126, p = 0.36$). Similar results were found for functional networks based on higher frequency bands (0.03–0.06 Hz) with Wald *F*-tests producing *p*-values in the region of 0.94 and 0.70 for cost-integrated global and local efficiencies, respectively. Boxplots for this data are presented in Fig. 7, panel (b). Since other authors have used different domains of cost-integration when comparing functional networks (see Achard and Bullmore (2007), for instance), we have also assessed the effect of the *N*-back experimental factor on partially cost-integrated global and local efficiencies, where $K(G)$ was made to vary over $[0, 1/2]$ (see Section S3 in Supplementary Methods). These subsequent analyses also yielded negative results, in the sense that the *N*-back factor was not found to have a significant effect on either partially cost-integrated global or partially cost-integrated local efficiencies. Taken together, these *F*-tests indicated that the cost-integrated small-world properties of the subjects' functional networks were not influenced by an increase in working memory load.

Mean correlation predicts individual N-back performance

Since connectivity strength was found to be related to the effect of the *N*-back paradigm, it was of interest to evaluate whether such changes in weighted cost (i.e. mean correlation) could predict subject-specific performances on the cognitive task under scrutiny.

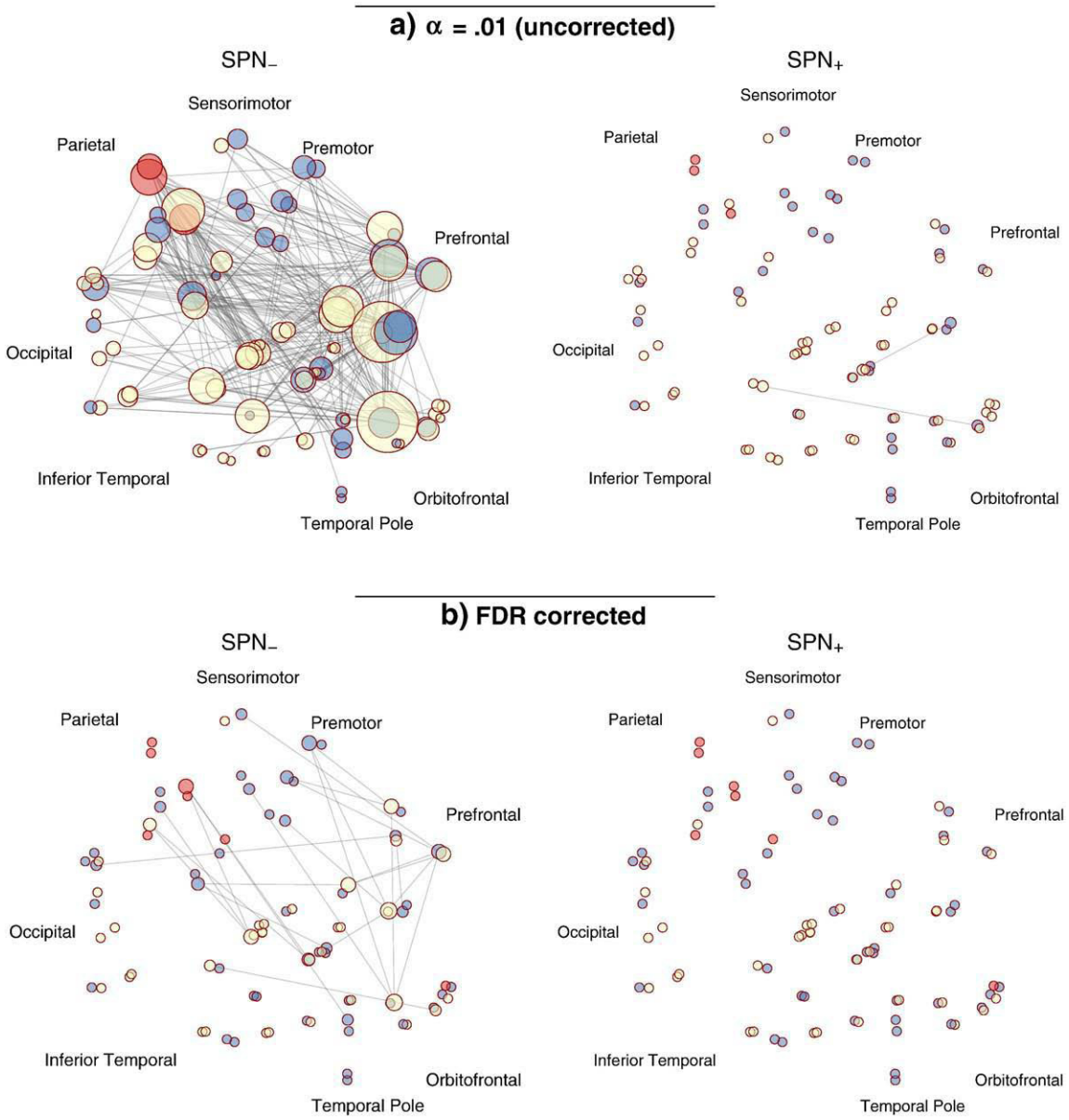


Fig. 5. Sagittal sections of the negative and positive differential SPNs, which represent the significantly 'lost' and 'gained' edges. The presence of edges reflects the thresholding of *p*-values at 0.01 (uncorrected) or using the FDR in panels (a) and (b), respectively. These SPNs are based on 43 subjects, and constructed with wavelet coefficients in the 0.01–0.03 Hz frequency band. The size of each node is proportional to the node's degree, such that the radius of each circle is $\rho_v \propto k_v$ for every vertex $v \in V$. The color scheme of the vertices indicates whether the time-averaged wavelet coefficients have been significantly downweighted (blue) or significantly upweighted (red) from 0-back to 3-back — yellow implying no significant effect for these vertices. The locations of the cortical regions correspond to their stereotaxic coordinates in the sagittal plane.

Our measure of *N*-back performance combined a measure of speed (reaction time (RT) for correct answers) with a measure of accuracy (penalty for failing to answer, when an answer was expected). We termed this measure *penalized* RT, denoted *p*RT. Our adopted scoring system was thus defined as follows,

$$p\mathrm{RT}_{ij} := \sum_{b=1}^{B} \sum_{l=1}^{L} x_{bl}, \tag{20}$$

for every $i = 1,\ldots,n$ subjects, and every $j = 1,\ldots,4$ conditions. The indices $b = 1,\ldots,B$ label the blocks ($B = 3$ blocks per condition), whereas $l = 1,\ldots,L$ denote the test items for which an answer was expected ($L = 3$ expected answers per block). The x_{bl}'s are binary variables taking the value of RT_{bl} for that trial if the subject had pressed the button, and 2000, if the subject had not answered. That is,

$$x_{bl} := \begin{cases} \mathrm{RT}_{bl} & \text{if answered;} \\ 2000 & \text{if no answer.} \end{cases} \tag{21}$$

Each RT_{bl} is the reaction time in ms for the l^{th} expected answer in the b^{th} block. In Eq. (20), the penalty for no answer was therefore equivalent to the maximal RT of 2 s — recall that each trial lasted for 2 s. In effect, *p*RT was therefore equivalent to bounding RT to 2 s, where an absence of answer was scored as the slowest possible answer.

Boxplots in Fig. 8 describe changes in RT and network cost over the four experimental conditions. The cost of each subject- and condition-

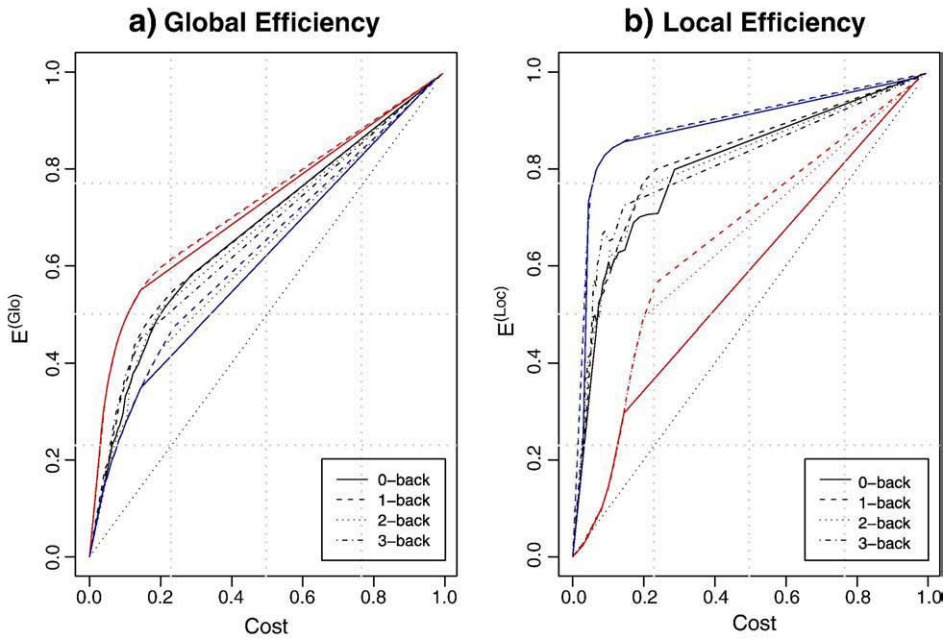


Fig. 6. Functional networks satisfied the small-world property under the four conditions of the *N*-back task. The global and local efficiencies of the mean SPNs are represented in panels (a) and (b), respectively. Corresponding efficiencies for the equivalent random graphs (indicated in red), and equivalent regular lattices (indicated in blue) were derived by matching the number of edges and vertices in the mean SPNs. Cost-specific global and local efficiencies were computed by thresholding the condition-specific correlation matrices at different cost levels. Different choices of *p*-values permitted to span the full gamut of cost, $K \in [0,1]$. Global and local efficiencies of the (N_V, N_E)-matched random graphs were averaged over 50 Monte Carlo realizations.

specific graph was computed in order to produce a set of $n \times 4$ network costs, denoted $K_W(\mathbf{R}_{ij})$. We regressed *p*RT on cost, after controlling for the *N*-back factor, using a mixed linear model with subject-specific random effects (see Pinheiro and Bates, 2000). The effect of network cost on *p*RT was found to be significant after Bonferroni's correction for using two different frequency bands (Wald $F = 13.82, df_1 = 1, df_2 = 125, p<0.001$), with a negative fixed effect ($\beta_{Cost} = -31.29$). This relationship, however, could simply be due to a difference between the no memory (0-back) and memory conditions (*j*-back, for all $j \geq 1$). Thus, we also tested for the effect of weighted cost on behavioral performance when considering only the 1- to 3-back conditions. Controlling for the *N*-back factor and using a subject-specific random effect, this gave a significant result at 0.01–0.03 Hz, after Bonferroni's correction for testing several frequency bands (Wald $F = 13.39, df_1 = 1, df_2 = 83, p<0.001$). That is, weighted cost was a negative predictor of overall *N*-back performance, even when only considering differences between levels of cognitive load. As subject-specific network cost increased, penalized

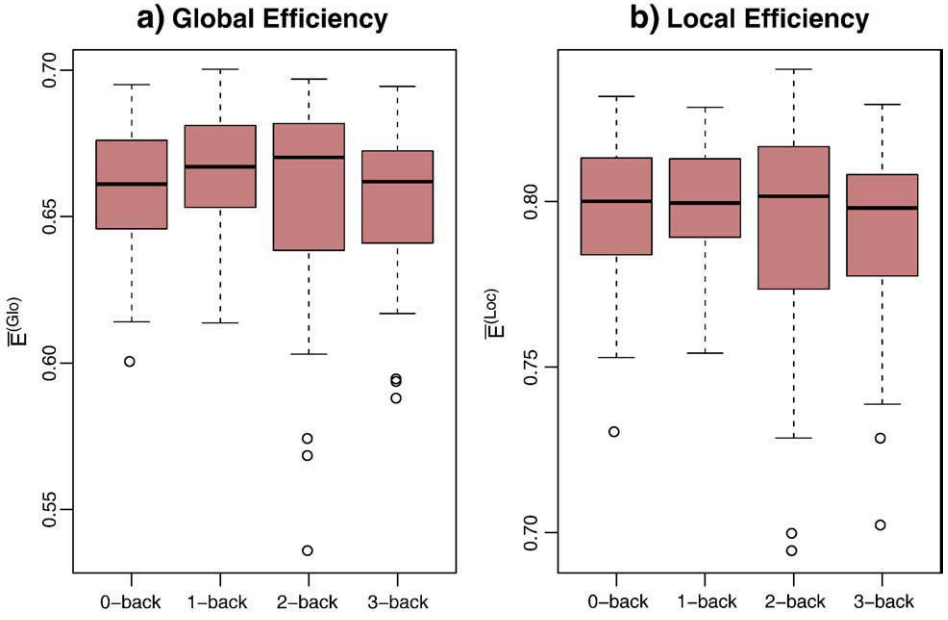


Fig. 7. Boxplot representations ($\{0.025, 0.25, 0.50, 0.75, 0.975\}$ quantiles) of subject-specific integrated global and local efficiencies denoted $\bar{E}^{Glo}(G_{ij})$ and $\bar{E}^{Loc}(G_{ij})$ in panels (a) and (b), respectively, where G_{ij} denotes the functional network for the i^{th} subject in the j^{th} condition. Integrated global and local efficiencies were not found to differ significantly with increasing working memory load. Note that inference is here conducted on weighted graphs, irrespective of any choice of threshold.

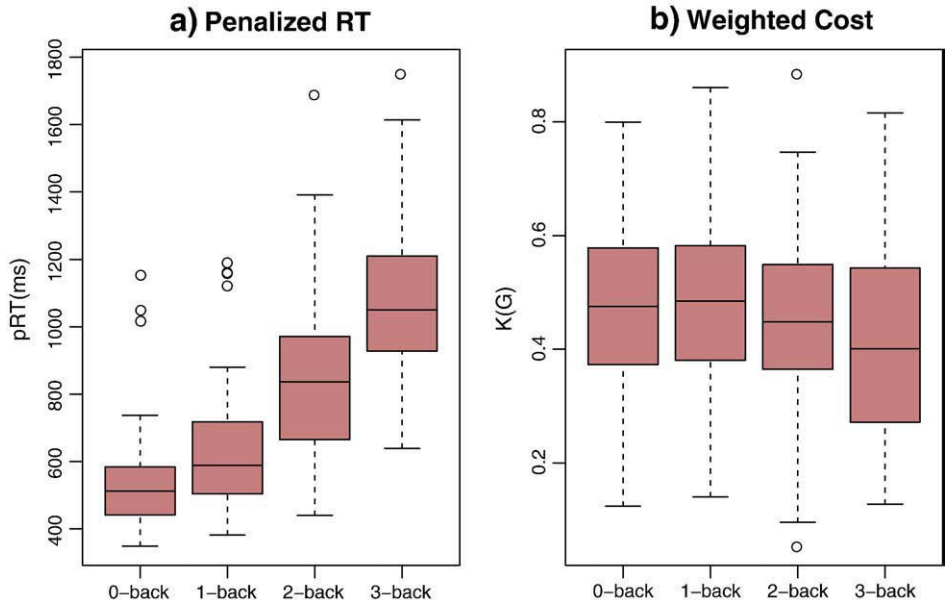


Fig. 8. Boxplot representations ($\{0.025, 0.25, 0.50, 0.75, 0.975\}$ quantiles) of (a) penalized reaction time (denoted *p*RT in ms) and (b) network cost per *N*-back conditions. Regression of *p*RT on subject-specific weighted cost ($K_W(G_{ij})$ for the i^{th} subject under the j^{th} condition) after controlling for the *N*-back factor was found to be significant. Weighted cost (i.e. mean between-region correlation coefficient) contributed negatively to performance on the *N*-back task. Note that inference is here conducted on weighted graphs, irrespective of any choice of threshold.

RT tended to significantly decrease. We note that the effect of K_W did not generalize to a higher frequency band. This effect was found to be non-significant for functional networks based on wavelet coefficients in the 0.03–0.06 Hz frequency interval (results not shown). Additionally, mixed effect models regressing *p*RT on cost-integrated global and local efficiencies ($\overline{E}^{Glo}$ and $\overline{E}^{Loc}$) were found to be non-significant ($p = 0.27$ and $p = 0.192$, respectively). Therefore, behavioral performance on the *N*-back task was significantly predicted by measures of

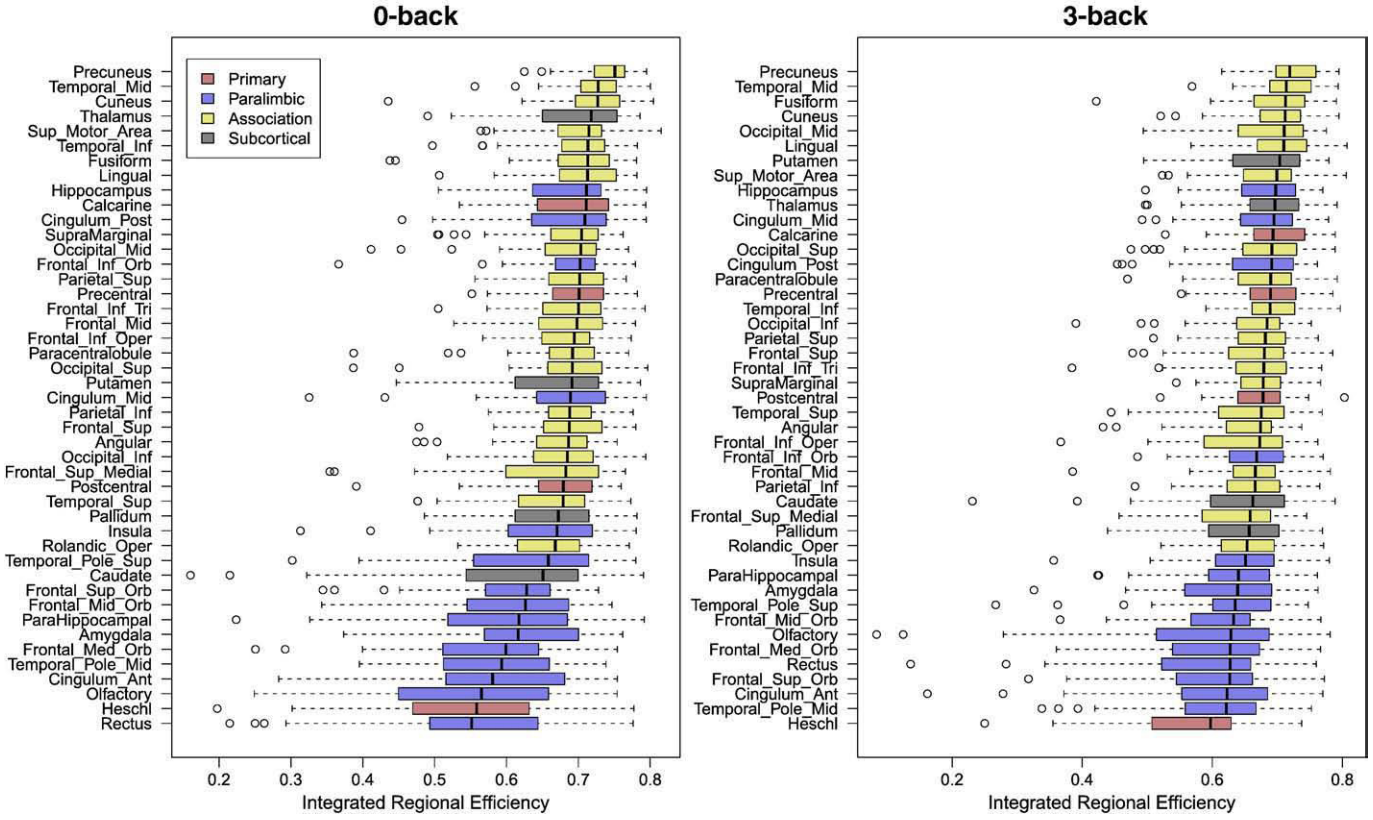


Fig. 9. Cost-integrated regional efficiencies for the 0-back and 3-back conditions. The subject-specific regional efficiencies, $\overline{E}^{Glo}(G_{ij}, v)$, for each symmetric pair of AAL regions (one region in each hemisphere) were averaged to give 45 regions. Each individual boxplot represents the variability of the regional efficiency for a vertex *v* over the population of subject-specific functional networks. Association areas played a prominent role in terms of regional efficiency under both experimental conditions. Paralimbic areas, by contrast, were mostly found to exhibit medium to low levels of regional efficiency. A full description of the AAL areas and the abbreviations used in this figure can be found in Tzourio-Mazoyer et al. (2002).

connectivity strength (i.e. weighted cost), but not by cost-integrated metrics of small-worldness.

Regional efficiencies under different working memory loads

Cost-integrated regional efficiencies (see Eq. (18)) for each of the 90 parcellated regions were computed. Each homologous pair of regions (one in each hemisphere) was averaged in order to produce 45 distributions of cost-integrated regional efficiencies, which are reported in Fig. 9, for the 0-back and 3-back conditions. The association cortices including the precuneus, the cuneus and the middle temporal gyri were found to be among the most efficient regions in both conditions. By contrast, most of the paralimbic regions exhibited cost-integrated low regional efficiencies. The sole paralimbic areas, which contravened this rule were the hippocampi, and the posterior and median cingulate gyri, as well as the inferior frontal orbital cortices.

Comparisons of these results for the 0-back and 3-back conditions indicated that the regions' ordering was well preserved. We tested for differences in rank across the two experimental conditions, using Spearman's rank correlation coefficient between the median values of the cost-integrated regional efficiencies for the 0-back and the 3-back conditions. Rank correlation was high with $\rho = 0.91$, which was highly significant ($S = 1266, df = 43, p < 10^{-15}$, where p-values were computed non-parametrically using algorithm AS 89, as described by Best and Roberts (1975)). Highly significant Spearman's correlation was also found for all other pairwise combination of the experimental conditions with all p-values lower than 10^{-15}. Highly significant rank correlations were also exhibited by cost-integrated regional efficiencies for wavelet coefficients in the 0.03–0.06 Hz frequency band, with $\rho = 0.89$ ($S = 1522, df = 43, p < 10^{-15}$). These results were reproduced for non-aggregated regional efficiencies, where each region was treated independently, as opposed to averaged over the two hemispheres ($S = 16876, df = 88, p < 10^{-16}$). Moreover, the ordering of the regions in terms of their cost-integrated regional efficiencies was also found to be overtly similar to the ordering of the regions with respect to their nodal degree, when considering $\overline{SPN}_j$'s (results not shown). Thus, the ordering of the regions in terms of their relative importance in the functional networks was found to be statistically independent of the cognitive load experienced by the subjects.

Discussion

In this paper, we have used graph-theoretical metrics and a principled framework to extract and compare functional networks between *N*-back conditions. The main findings of this investigation are that: (i) connectivity strength, as measured by weighted cost (i.e. mean correlation coefficient) significantly decreased as working memory load increased; (ii) cost-integrated connectivity, such as cost-integrated global and local efficiencies and ranking of cortical areas in terms of regional efficiencies were preserved under different levels of working memory load; and (iii) subject-specific measures of connectivity strength (i.e. network cost) predicted subject-specific performance on the *N*-back task above chance.

Connectivity strength decreased with working memory load

The main conclusion of this study is the occurrence of a considerable pruning of the functional MRI networks as working memory load increases. Comparisons of condition-specific connectivity strength demonstrated that macro-scale spontaneous correlations tend to diminish when participants are subjected to more challenging cognitive tasks. In particular, the use of differential SPNs to extract significantly deprecated edges showed that a substantial number of edges were 'lost' when comparing the 0-back and 3-back conditions, or when contrasting the 1-back and 3-back conditions. These findings are in line with previous research on working memory. Fransson and Marrelec (2008) in a study investigating changes in the default mode network (DMN) under a continuous verbal *N*-back task also reported a global reduction in functional interactivity between the cortical and subcortical regions of interest that they studied.

Tentatively, one can also draw some links between the present findings and some of the research conducted on differences in functional connectivity between adults and children. Several lines of evidence converge to indicate that gray matter (GM) maturation is associated with a diminution of GM density, generally explained by increased synaptic pruning during adolescence and early adulthood (Sowell et al., 2003; Gogtay et al., 2004). Specifically, the transition from childhood to adulthood seems to be characterized by a dynamic sequence of over-connectivity followed by pruning of the functional networks. These findings have been corroborated at a system level by the investigation of developmental changes in functional connectivity using fMRI data. Supekar et al. (2009) demonstrated that children (aged 7–9 years) tend to exhibit connectivity patterns different from those of young adults (aged 19–22 years). Intriguingly, we found that a similar pattern seemed to be displayed by adults experiencing different levels of working memory load. This may suggest that the anatomical pruning witnessed in adolescence and early adulthood responsible for greater modular specialization of cognitive faculties could be conceptually similar to the task-specific functional trimming that we have reported in this study. These parallels remain naturally very speculative, but we believe that they may pave the way for further research in the interaction between the functional and anatomical connectomes.

Cost-integrated efficiency metrics were preserved

The mean SPNs corresponding to the four experimental conditions of interest were found to all satisfy the small-world property. These functional networks surpassed an equivalent random network in terms of local efficiency, and outperformed an equivalent regular lattice in terms of global efficiency (see Fig. 6). These results are supported by previous studies, which have shown that the small-world topology of functional MRI networks are robust to changes under a range of different conditions. Supekar et al. (2009) demonstrated that children and young adults exhibit functional networks with similar small-world architecture. Our findings corroborate the general consensus in neuroscience, which states that the brain has evolved in order to optimize the balance between specialization and integration constraints, thereby satisfying the small-world property (Sporns et al., 2000, 2004).

Additionally, our analysis of cost-integrated local and global efficiencies also showed that such properties are robust to changes in cognitive load. Various studies in the literature on small-world networks have indicated some weakening of these characteristics in different populations. Comparisons of functional networks between young adults and healthy elderly showed substantial reductions in both local and global efficiencies with age (Achard and Bullmore, 2007). Administration of dopamine receptor antagonist has also been found to reduce both local and global efficiencies in populations of young and elderly subjects (Achard and Bullmore, 2007). Similarly, an EEG study of functional connectivity based on the *N*-back task highlighted an abatement of small-world efficiency in patients with schizophrenia (Pachou et al., 2008). In the study at hand, however, cognitive load did not produce significant changes in levels of cost-integrated global and local efficiencies. Finally, we note that the present results should be contrasted with the findings of Bassett et al. (2009), who reported that 'cost-efficiency' – defined by these authors as efficiency minus cost – was found to increase with performance on the *N*-back task. Two major differences, however, distinguish our results with this previous work: (i) our results are based on fMRI and low frequency (<0.1 Hz) oscillations, whereas Bassett et al. (2009)

have used MEG data and studied high frequency oscillations (15–30 Hz); and (ii) we have simply related weighted cost and performance, whereas Bassett et al. (2009) have considered cost-efficiency. Further research will be needed to clarify the relationship between cost-efficiency and weighted cost.

Collectively, these results indicate that connectivity strength may be at least as important as small-world properties when studying functional networks. The changes undergone by the networks subtending cognitive load in our sample were most salient when observed through the lens of connectivity strength. When controlling for differences in cost across the functional networks, however, we found no significant differences between the experimental conditions either in terms of cost-integrated efficiency or regarding the ordering of regional hubs. Our findings in this paper seem therefore to indicate that consideration of variability in weighted cost may add substantial explanatory power to routine analyses of small-world networks across different populations, or experimental conditions.

Limitations and extensions

Throughout this paper, we have assumed that the *N*-back paradigm allows the detection of changes in working memory load. Miller et al. (2009), however, have shown that behavioral performance on the *N*-back task does not necessarily correlate with performance on established measures of working memory, such as forward and backward digit span (Wechsler, 1997). Thus, although the *N*-back paradigm is generally conceived as an indicator of overall processing speed and memorization, it does not appear to constitute a reliable marker of the working memory construct, as defined by digit spans. For the purpose of this paper, however, we will adhere to the former perspective and consider performances on the *N*-back task to be related to general working memory function.

An important limiting factor of this study was the age range of the subjects. Several studies have investigated changes in functional connectivity with age (Achard and Bullmore, 2007), and with pathologies associated with age (Stam et al., 2007). Future research will thus be needed to replicate these findings with younger adults. Another important limitation of the present study is our concentration on specific frequency bands. Our focus on low (<0.1) frequency oscillations is in line with other network analysis based on fMRI data (e.g. Achard and Bullmore, 2007). However, previous studies using EEG-based functional networks during different cognitive tasks have unveiled a variety of different patterns of synchronization and desynchronization depending on the frequency bands that are considered (Pachou et al., 2008; Micheloyannis et al., 2009). Similar findings have been reported by Salvador et al. (2008), who showed that different regions of the cortex may reach maximal coherence with other brain regions at different frequencies. Moreover, other studies have also demonstrated that the salience of cortical and subcortical modules may differ depending on which frequency band is used (Wu et al., 2008). These issues could be usefully tackled by adopting a multimodal approach to functional networks, perhaps combining EEG and fMRI data.

The generality of our results is also constrained by their dependence on our choice of parcellation scheme. Research has shown that different cortical parcellations may lead to different network topologies (Wang et al., 2009). The present study may therefore benefit from being replicated with other types of anatomical parcellations. Alternatively, one could adopt a parcellation-free approach to functional connectivity, by inferring cortical regions from their statistical properties, such as by using independent component analysis (see Biswal et al. (2010), for a recent utilization of these methods to network analysis). Thus, our analysis will need to be corroborated by replications of the present results using different parcellation schemes. Additionally, as highlighted in the Methods section, another potential bias was introduced during the preprocessing stage by the use of wavelet decomposition in conjunction with concatenation of the block-specific time series. We have shown through random simulations that this particular procedure tends to artificially inflate the magnitude of the between-condition differences and the within-condition variances of the correlation coefficients (see Supplementary Methods S1). Moreover, it should also be noted that the effect of the *N*-back experimental factor on weighted cost was only found to be significant for scale 4 (0.01–0.03 Hz) wavelet coefficients. At that scale, the correlation coefficients are based on 12 wavelet bases, against 23 bases for scale 3 wavelets. The scale 4 correlation coefficients are therefore based on a lower number of observations than the scale 3 ones, and may therefore be considerably noisier (see Achard et al., 2008). Our results may therefore be confounded by a considerable amount of noise. Further work will be needed in order to incorporate such sources of noise in the final analysis perhaps through the utilization of Bayesian technique to propagate uncertainty from the preprocessing stages to the final steps of the analysis.

In addition, we also note that our treatment of the interdependence between differences in cost and differences in topology is somewhat unorthodox. While some authors have used cost-integrated metrics (He et al., 2009b), as we have done in this paper, other researchers have integrated topological metrics over a subset of the cost regimen, which is sometimes referred to as the 'small-world' cost regimen (see Achard and Bullmore (2007), for example). As there is not currently a standard of best practice in this field, we have investigated several approaches to the problem of comparing weighted networks by considering various cost-integration domains and cost-integration procedures (See Section S3 in Supplementary Methods). These different approaches led us to conclude that the present data does not support a significant relationship between the *N*-back factor and cost-integrated global and local efficiencies. However, we have also observed a large amount of variability between the subject-specific functional networks in our sample. Since it is reasonable to expect that changes in functional network topologies due to experimental manipulation are considerably small, it may be conjectured that our sample size was not sufficiently large to detect such subtle effects. It would therefore be unwise to rule out any relationship between functional network topology and cognition or behavior on the sole basis of the negative result at hand. Further research based on larger data sets will be needed to clarify these specific relationships.

Future extensions of these findings may need to deploy more sophisticated measures of synchronization such as partial correlation, mutual information, coherence, or synchronization likelihood (Pachou et al., 2008; Micheloyannis et al., 2009). Additionally, correlation between nodes' degrees could also be used. This latter measure of synchronization has successfully been utilized in a range of different studies evaluating network properties (Bianconi et al., 2008; Bianconi, 2008; Bianconi, 2009). For convenience, we have voluntarily restricted ourselves to the study of unweighted and undirected graphs. However, much could be gained from the consideration of weighted undirected networks. Regarding the experimental paradigm, previous fMRI research on the *N*-back task has shown that distinct frontoparietal activation patterns emerge depending on which variants of the task are implemented (Owen et al., 2005). It will therefore be of interest to investigate whether the changes presented in this report are replicated when one considers other variants of the *N*-back paradigm, such as when comparing process- and content-specific frontoparietal activation associated with working memory. Also note that a separate publication will explore in greater details the role played by the different regions in the functional networks investigated thus far. In particular, we will consider a variety of different means to identify the most salient hubs in these networks, using techniques proposed in the literature (Sporns et al., 2007).

Since we have provided some evidence of the predictive power of connectivity strength, one could also envisage the comparison of the outcomes of such a connectivity analysis with standard mass-univariate approaches to fMRI data. Finally, changes in the degree distributions subtending the networks in the different experimental conditions will need to be ascertained. Further research may be required to investigate the nature of these changes, and which parameters are controlling such changes. A Bayesian statistical analysis, using robust model selection methods may be utilized for this purpose. However, it is surmised that larger data sets will be required to fully identify the parameters of interest.

Our stated aim in this paper was to evaluate differences in *network-wise* topology due to the experimental paradigm. We were not directly interested in identifying *edge-specific* or *node-specific* effects, but rather how overall connectivity patterns may change with the experimental manipulations. Future extensions of this work, however, will require the adoption of a more rigorous approach to multiple testing than the FDR in order to identify whether specific between-regions connections are significantly upweighted or down-weighted. Specifically, the FDR only controls for multiple testing weakly. By contrast, a strong control of the familywise error rate (FWE) would require subset pivotality (Nichols and Hayasaka, 2003; Westfall and Troendle, 2008). However, this assumption can be shown not to hold when testing for the significance of the elements of a correlation matrix. This follows from the fact that the free combination condition is not satisfied in this case (see Nichols and Hayasaka, 2003). Other methods may therefore be required to provide a strong control of the FWE in the context of both mean and differential SPNs.

Conclusions

To our knowledge, the present study constitutes one of the first attempts to systematically evaluate functional network dynamics under different levels of a cognitive task. As expected, small-world topological features were found in low frequency intervals (<0.1 Hz) under the 4 conditions of the *N*-back task. This corroborates previous investigations, which have reported that the small-world properties of functional brain networks based on fMRI are most salient in low frequency bands (Achard et al., 2006; Achard and Bullmore, 2007; Supekar et al., 2009). Functional connectivity became more diffuse and the subtending graphs became sparser as subjects faced greater working memory load. Importantly, we found that this relationship had predictive value – that is, subject-specific network cost predicted performance on the *N*-back task – when considering networks built on very low frequency oscillations (0.01–0.03 Hz), but were not noticeable in the 0.03–0.06 Hz frequency band. These discrepancies imply that subtle changes in frequency bands may have an important impact on the saliency of functional MRI connectivity.

Functional networks were found to satisfy the small-world property in all *N*-back conditions. When controlling for cost, we found no significant differences in cost-integrated global and local efficiencies across the networks, thus suggesting that these properties are robust to an increase in working memory load. The ordering of the regions in terms of regional efficiencies was also consistent across the different experimental conditions. This research therefore contributes to the literature on functional connectivity and demonstrates that while connectivity strength tends to decrease with working memory load, cost-integrated connectivity and other small-world properties are maintained.

The present paper also validated a formal approach to network analysis based on mass-univariate statistical inference. Our proposed methodological framework is exploratory in nature. It permits the extraction of differential graphs, reflecting 'lost' and 'gained' connections over a series of tasks. This approach allows the visualization and comparison of network dynamics over several conditions. Although our proposed methodology was based on a growth curve model, these techniques can readily be extended to any type of GLM. These results have shown that the brain appeared to prune superfluous connections, which may reflect an increase in focus and attendance to the cognitive task at hand. Future research will need to elucidate the precise properties of both the specific networks subtending brain activity during the *N*-back task, as well as the properties of the deprecated network.

Acknowledgments

This work was supported by a fellowship from the UK National Institute for Health Research (NIHR) Biomedical Research Centre in mental health. The authors are greatly indebted to Sophie Achard, Mick Brammer and Tom Nichols for useful discussions and comments. We also would like to thank three anonymous reviewers for their valuable input.

Appendix A. Quasilinearity of thresholding function

The great majority of network analysis in systems neuroscience relies on the construction of undirected unweighted graphs through the thresholding of structural or functional correlation matrices (Bullmore and Sporns, 2009; Rubinov and Sporns, 2010). We here describe how such a thresholding function complicates the production of 'summary' networks over a population of subject-specific correlation matrices.

Let the vectorisation of the upper off-diagonal entries of a correlation matrix $\mathbf{R}$, of order $(N_V \times N_V)$, into a vector of size $N_E = N_V(N_V-1)/2$, be denoted by $\mathrm{vec}(\mathbf{R})$. The thresholding function $T(\cdot)$ can then be formulated as the map, $T : \mathbb{R}^d \mapsto \{0,1\}^d$. Use of the inverse of the vectorisation function produces the adjacency matrix, $\mathbf{A} := \mathrm{vec}^{-1}(T(\mathrm{vec}(\mathbf{R})))$, whose N_V^2 entries satisfy the following rule,

$$a_{ij} := T\left(r_{ij}\right) = \begin{cases} 1 & \text{if } r_{ij} \geq \tau, \\ 0 & \text{otherwise;} \end{cases} \tag{22}$$

for some threshold value τ. The thresholding of a matrix therefore simplifies to the thresholding of every single value in that matrix. It therefore suffices to concentrate on the properties of the indicator function $T(r_{ij}) := 1_{[\tau,1]}(r_{ij})$, for a given r_{ij}, which we will denote by r, for convenience.

It can be shown that the indicator function $T(\cdot)$ is neither convex nor concave. However, the following inequality holds,

$$T(\lambda r + (1-\lambda) r') \leq max\left(T(r), T(r')\right). \tag{23}$$

for some $\lambda \in [0,1]$ and two correlation coefficients $r, r' \in [-1,1]$. Eq. (23) states that the function $T(\cdot)$ is *quasiconvex*. Moreover, since one can also verify that

$$T(\lambda r + (1-\lambda) r') \geq max\left(T(r), T(r')\right), \tag{24}$$

$T(\cdot)$ is additionally *quasiconcave*. Taken together these two arguments imply that $T(\cdot)$ is *quasilinear*. Since $T(\cdot)$ is only quasilinear, but not linear, we have

$$\mathbb{E}[T(X)] \neq T(\mathbb{E}[X]), \tag{25}$$

where the expectation is taken with respect to the probability density function of X. Eq. (25) can be put in relation to Eq. (1). This suggests that simply taking the mean correlation matrix $\sum_{i=1}^{n} \mathbf{R}_i$ over n individuals and thresholding that matrix (RHS of Eq. (23)) will not be equal to the expectation of the thresholded correlation matrices (LHS of Eq. (23)).

In addition, note that the standard version of Jensen's inequality cannot be invoked for quasiconvex functions. In cortical network

analysis, undirected unweighted networks are generally of interest. The metrics used for the quantification of the topology of these networks rely on a thresholding function. In such cases, Eq. (25) will necessarily apply, and the problem of summarizing topological information over a family of networks needs therefore to be addressed.

Appendix B. Supplementary data

Supplementary data to this article can be found online at doi:10.1016/j.neuroimage.2010.11.030.

References

Achard, S., Bassett, D.S., Meyer-Lindenberg, A., Bullmore, E., 2008. Fractal connectivity of long-memory networks. Phys. Rev. E 77 (3), 036104–036114.

Achard, S., Bullmore, E., 2007. Efficiency and cost of economical brain functional networks. PLoS Comput. Biol. 3, 174–182.

Achard, S., Salvador, R., Whitcher, B., Suckling, J., Bullmore, E., 2006. A resilient, low-frequency, small-world human brain functional network with highly connected association cortical hubs. J. Neurosci. 26 (1), 63–72.

Albert, R., Barabasi, A., 2002. Statistical mechanics of complex networks. Rev. Mod. Phys. 74 (1), 47–97.

Astolfi, L., Cincotti, F., Mattia, D., De Vico Fallani, F., Salinari, S., Marciani, M., Witte, H., Babiloni, F., 2009. Study of the time-varying cortical connectivity changes during the attempt of foot movements by spinal cord injured and healthy subjects. Conf. Proc. IEEE Eng. Med. Biol. Soc. 2208–2211.

Baddeley, A., 1998. Working memory. CR Acad. Sci. III Sci. Vie 321 (2–3), 167–173.

Barrat, A., Barthlemy, M., Pastor-Satorras, R., Vespignani, A., 2004. The architecture of complex weighted networks. Proc. Natl Acad. Sci. USA 101 (11), 3747–3752.

Bartlett, M.S., 1937. Properties of sufficiency and statistical tests. Proc. R. Soc. Lond. Math. Phys. Sci. 160 (901), 268–282.

Bassett, D.S., Bullmore, E., Verchinski, B.A., Mattay, V.S., Weinberger, D.R., Meyer-Lindenberg, A., 2008. Hierarchical organization of human cortical networks in health and schizophrenia. J. Neurosci. 28 (37), 9239–9248.

Bassett, D.S., Bullmore, E.T., Meyer-Lindenberg, A., Apud, J.A., Weinberger, D.R., Coppola, R., 2009. Cognitive fitness of cost-efficient brain functional networks. Proc. Natl Acad. Sci. 106 (28), 11747–11752.

Bassett, D.S., Meyer-Lindenberg, A., Achard, S., Duke, T., Bullmore, E., 2006. Adaptive re-configuration of fractal small-world human brain functional networks. Proc. Natl Acad. Sci. USA 103 (51), 19518–19523.

Bassett, D.S., Bullmore, E., 2009. Human brain networks in health and disease. Curr. Opin. Neurol. 22, 340–347.

Bassett, D.S., Bullmore, E., 2006. Small-world brain networks. Neuroscientist 12 (6), 512–523.

Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Stat. Soc. B Methodol. 57 (1), 289–300.

Best, D.J., Roberts, D.E., 1975. Algorithm as 89: the upper tail probabilities of spearman's rho. J. R. Stat. Soc. C Appl. Stat. 24 (3), 377–379.

Bianconi, G., 2008. The entropy of randomized network ensembles. EPL (Europhys. Lett.) 81 (2), 28005–28015.

Bianconi, G., 2009. Entropy of network ensembles. Phys. Rev. E 79 (3), 036114–036124.

Bianconi, G., Coolen, A.C.C., Perez Vicente, C.J., 2008. Entropies of complex networks with hierarchically constrained topologies. Phys. Rev. E 78 (1), 16114–16124.

Biswal, B., Yetkin, F.Z., Haughton, V.M., Hyde, J.S., 1995. Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. Magn. Reson. Med. 34 (4), 537–541.

Biswal, B.B., Mennes, M., Zuo, X.N., Gohel, S., Kelly, C., Smith, S.M., Beckmann, C.F., Adelstein, J.S., Buckner, R.L., Colcombe, S., Dogonowski, A.M., Ernst, M., Fair, D., Hampson, M., Hoptman, M.J., Hyde, J.S., Kiviniemi, V.J., Ktter, R., Li, S.J., Lin, C.P., Lowe, M.J., Mackay, C., Madden, D.J., Madsen, K.H., Margulies, D.S., Mayberg, H.S., McMahon, K., Monk, C.S., Mostofsky, S.H., Nagel, B.J., Pekar, J.J., Peltier, S.J., Petersen, S.E., Riedl, V., Rombouts, S.A.R.B., Rypma, B., Schlaggar, B.L., Schmidt, S., Seidler, R.D., Siegle, G.J., Sorg, C., Teng, G.J., Veijola, J., Villringer, A., Walter, M., Wang, L., Weng, X. C., Whitfield-Gabrieli, S., Williamson, P., Windischberger, C., Zang, Y.F., Zhang, H.Y., Castellanos, F.X., Milham, M.P., 2010. Toward discovery science of human brain function. Proc. Natl Acad. Sci. 107 (10), 4734–4739.

Bollen, K., Curran, P., 2006. Latent Curve Models: A Structural Equation Perspective. Wiley Series in Probability and Statistics. Wiley, London.

Bollobas, B., 1998. Modern Graph Theory. Springer, London.

Brammer, M.J., 1998. Multidimensional wavelet analysis of functional magnetic resonance images. Hum. Brain Mapp. 6 (5–6), 378–382.

Bullmore, E., Sporns, O., 2009. Complex brain networks: graph theoretical analysis of structural and functional systems. Nat. Rev. Neurosci. 10 (1), 1–13.

Bullmore, E., Fadili, J., Breakspear, M., Salvador, R., Suckling, J., Brammer, M., 2003. Wavelets and statistical analysis of functional magnetic resonance images of the human brain. Stat. Meth. Med. Res. 12 (5), 375–399.

Callicott, J.H., Mattay, V.S., Bertolino, A., Finn, K., Coppola, R., Frank, J.A., Goldberg, T.E., Weinberger, D.R., 1999. Physiological characteristics of capacity constraints in working memory as revealed by functional MRI. Cereb. Cortex 9 (1), 20–26.

Caseras, X., Mataix-Cols, D., Giampietro, V., Rimes, K.A., Brammer, M., Zelaya, F., Chalder, T., Godfrey, E.L., 2006. Probing the working memory system in chronic fatigue syndrome: a functional magnetic resonance imaging study using the *n*-back task. Psychosom. Med. 68 (6), 947–955.

Catani, M., Mesulam, M., 2008. What is a disconnection syndrome? Cortex 44 (8), 911–913.

Cecchi, G., Rao, A., Centeno, M., Baliki, M., Apkarian, A., Chialvo, D., 2007. Identifying directed links in large scale functional networks: application to brain fMRI. BMC Cell Biol. 8, 1–10.

Cordes, D., Haughton, V.M., Arfanakis, K., Wendt, G.J., Turski, P.A., Moritz, C.H., Quigley, M.A., Meyerand, M.E., 2000. Mapping functionally related regions of brain with functional connectivity MR imaging. AJNR Am. J. Neuroradiol. 21 (9), 1636–1644.

De Vico Fallani, F., Astolfi, L., Cincotti, F., Mattia, D., Marciani, M., Tocci, A., Salinari, S., Witte, H., Hesse, W., Gao, S., Colosimo, A., Babiloni, F., 2008. Cortical network dynamics during foot movements. Neuroinformatics 6 (1), 23–34.

Demidenko, E., 2004. Mixed Models: Theory and Applications. Wiley, London.

D'Esposito, M., Detre, J., Alsop, D., Shin, R., Atlas, S., Grossman, M., 1995. The neural basis of the central executive system of working memory. Nature 378, 279–281.

Dijkstra, E., 1959. A note on two problems in connexion with graphs. Numer. Math. 1, 269–271.

Eguiluz, V.M., Chialvo, D.R., Cecchi, G.A., Baliki, M., Apkarian, A.V., 2005. Scale-free brain functional networks. Phys. Rev. Lett. 94 (1), 18102–18106.

Fair, D.A., Schlaggar, B.L., Cohen, A.L., Miezin, F.M., Dosenbach, N.U., Wenger, K.K., Fox, M.D., Snyder, A.Z., Raichle, M.E., Petersen, S.E., 2007. A method for using blocked and event-related fmri data to study "resting state" functional connectivity. Neuroimage 35 (1), 396–405.

Fransson, P., 2005. Spontaneous low-frequency bold signal fluctuations: an fMRI investigation of the resting-state default mode of brain function hypothesis. Hum. Brain Mapp. 26 (1), 15–29.

Fransson, P., Marrelec, G., 2008. The precuneus/posterior cingulate cortex plays a pivotal role in the default mode network: evidence from a partial correlation network analysis. Neuroimage 42 (3), 1178–1184.

Friston, K.J., 1994. Functional and effective connectivity in neuroimaging: a synthesis. Hum. Brain Mapp. 2 (1–2), 56–78.

Gevins, A., Cutillo, B., 1993. Neuroelectric evidence for distributed processing in human working memory. Electroencephalogr. Clin. Neurophysiol. 87, 128–143.

Gogtay, N., Giedd, J.N., Lusk, L., Hayashi, K.M., Greenstein, D., Vaituzis, A.C., Nugent, T.F., Herman, D.H., Clasen, L.S., Toga, A.W., Rapoport, J.L., Thompson, P.M., 2004. Dynamic mapping of human cortical development during childhood through early adulthood. Proc. Natl Acad. Sci. USA 101 (21), 8174–8179.

Harvey, P.O., Fossati, P., Pochon, J.B., Levy, R., LeBastard, G., Lehericy, S., Allilaire, J.F., Dubois, B., 2005. Cognitive control and brain resources in major depression: an fMRI study using the *n*-back task. Neuroimage 26 (3), 860–869.

He, Y., Chen, Z.J., Evans, A.C., 2007. Small-world anatomical networks in the human brain revealed by cortical thickness from MRI. Cereb. Cortex 17 (10), 2407–2419.

He, Y., Dagher, A., Chen, Z., Charil, A., Zijdenbos, A., Worsley, K., Evans, A., 2009a. Impaired small-world efficiency in structural cortical networks in multiple sclerosis associated with white matter lesion load. Brain 132 (12), 3366–3379.

He, Y., Wang, J., Wang, L., Chen, Z.J., Yan, C., Yang, H., Tang, H., Zhu, C., Gong, Q., Zang, Y., Evans, A.C., 2009b. Uncovering intrinsic modular organization of spontaneous brain activity in humans. PLoS ONE 4 (4), 1–18.

Honey, G.D., Fu, C.H.Y., Kim, J., Brammer, M.J., Croudace, T.J., Suckling, J., Pich, E.M., Williams, S.C.R., Bullmore, E.T., 2002. Effects of verbal working memory load on corticocortical connectivity modeled by path analysis of functional magnetic resonance imaging data. Neuroimage 17 (2), 573–582.

Jenkinson, M., Bannister, P., Brady, M., Smith, S., 2002. Improved optimization for the robust and accurate linear registration and motion correction of brain images. Neuroimage 17 (2), 825–841.

Jenkinson, M., Smith, S., 2001. A global optimisation method for robust affine registration of brain images. Med. Image Anal. 5 (2), 143–156.

Kane, M., Conway, A., Miura, T., Colesh, G., 2007. Working memory, attention control and the *n*-back task: a question of construct validity. J. Exp. Psychol. 33, 615–622.

Laird, N., Ware, J., 1982. Random-effects models for longitudinal data. Biometrics 38, 963–974.

Latora, V., Marchiori, M., 2003. Economic small-world behavior in weighted networks. Eur. Phys. J. B 32 (2), 249–263.

Latora, V., Marchiori, M., 2001. Efficient behavior of small-world networks. Phys. Rev. Lett. 87 (19), 198701–198705.

Luce, R., Perry, A., 1949. A method of matrix analysis of group structure. Psychometrika 14 (2), 95–116.

Lund, T.E., Madsen, K.H., Sidaros, K., Luo, W.L., Nichols, T.E., 2006. Non-white noise in fMRI: does modelling have an impact? Neuroimage 29 (1), 54–66.

Micheloyannis, S., Vourkas, M., Tsirka, V., Karakonstantaki, E., Kanatsouli, K., Stam, C.J., 2009. The influence of ageing on complex brain networks: a graph theoretical analysis. Hum. Brain Mapp. 30 (1), 200–208.

Milgram, S., 1967. The small world problem. Psychol. Today 2, 60–67.

Miller, K., Price, C., Okun, M., Montijo, H., Bowers, D., 2009. Is the *n*-back task a valid neuropsychological measure for assessing working memory? Arch. Clin. Neuropsychol. 24 (7), 711–717.

Nichols, T., Hayasaka, S., 2003. Controlling the familywise error rate in functional neuroimaging: a comparative review. Stat. Meth. Med. Res. 419–446.

Owen, A., McMillan, K., Laird, A., Bullmore, E., 2005. *N*-back working memory paradigm: a meta-analysis of normative functional neuroimaging studies. Hum. Brain Mapp. 25, 46–59.

Pachou, E., Vourkas, M., Simos, P., Smit, D., Stam, C., Tsirka, V., Micheloyannis, S., 2008. Working memory in schizophrenia: an EEG study using power spectrum and

coherence analysis to estimate cortical activation and network behavior. Brain Topogr. 21 (2), 128–137.
Percival, D., Walden, A., 2000. Wavelet Methods for Time Series Analysis. Cambridge University Press, London.
Pinheiro, J., Bates, D., 2000. Mixed-effects Mode in S and S-Plus. Springer, London.
Richiardi, J., Eryilmaz, H., Schwartz, S., Vuilleumier, P., and Van De Ville, D., in press. Decoding brain states from fMRI connectivity graphs. NeuroImage, Corrected Proof.
Rubinov, M., Knock, S.A., Stam, C.J., Micheloyannis, S., Harris, A.W., Williams, L.M., Breaks-pear, M., 2009. Small-world properties of nonlinear brain activity in schizophrenia. Hum. Brain Mapp. 30 (2), 403–416.
Rubinov, M., Sporns, O., 2010. Complex network measures of brain connectivity: uses and interpretations. Neuroimage 52 (3), 1059–1069.
Salvador, R., Martinez, A., Pomarol-Clotet, E., Gomar, J., Vila, F., Sarro, S., Capdevila, A., Bullmore, E., 2008. A simple view of the brain through a frequency-specific functional connectivity measure. Neuroimage 39 (1), 279–289.
Salvador, R., Suckling, J., Coleman, M.R., Pickard, J.D., Menon, D., Bullmore, E., 2005. Neuro-physiological architecture of functional magnetic resonance images of human brain. Cereb. Cortex 15 (9), 1332–1342.
Smith, S.M., Jenkinson, M., Woolrich, M.W., Beckmann, C.F., Behrens, T.E., Johansen-Berg, H., Bannister, P.R., De Luca, M., Drobnjak, I., Flitney, D.E., Niazy, R.K., Saunders, J., Vickers, J., Zhang, Y., De Stefano, N., Brady, J.M., Matthews, P.M., 2004. Advances in functional and structural MR image analysis and implementation as FSL. Neuroimage 23 (Supplement 1), S208–S219.
Sowell, E.R., Peterson, B.S., Thompson, P.M., Welcome, S.E., Henkenius, A.L., Toga, A.W., 2003. Mapping cortical change across the human life span. Nat. Neurosci. 6 (3), 309–315.
Sporns, O., Tononi, G., Edelman, G., 2000. Theoretical neuroanatomy: relating anatomical and functional connectivity in graphs and cortical connection matrices. Cereb. Cortex 10 (2), 127–141.
Sporns, O., Chialvo, D.R., Kaiser, M., Hilgetag, C.C., 2004. Organization, development and function of complex brain networks. Trends Cogn. Sci. 8 (9), 418–425.
Sporns, O., Honey, C., Kotter, R., 2007. Identification and classification of hubs in brain networks. PLoS ONE 2 (10), 1–14.
Stam, C., Jones, B., Nolte, G., Breakspear, M., Scheltens, P., 2007. Small-world networks and functional connectivity in Alzheimer's disease. Cereb. Cortex 17 (1), 92–99.
Supekar, K., Menon, V., Rubin, D., Musen, M., Greicius, M.D., 2008. Network analysis of intrinsic functional brain connectivity in Alzheimer's disease. PLoS Comput. Biol. 4 (6), 1–11.
Supekar, K., Musen, M., Menon, V., 2009. Development of large-scale functional brain networks in children. PLoS Biol. 7 (7), 1–15.
Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., Joliot, M., 2002. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. Neuroimage 15 (1), 273–289.
Van Snellenberg, J., Torres, I., Thornton, A., 2006. Functional neuroimaging of working memory in schizophrenia: task performance as a moderating variable. Neuropsychology 20, 497–510.
Wang, J., Wang, L., Zang, Y., Yang, H., Tang, H., Gong, Q., Chen, Z., Zhu, C., He, Y., 2009. Parcellation-dependent small-world brain functional networks: a resting-state fMRI study. Hum. Brain Mapp. 30 (5), 1511–1523.
Watts, D.J., Strogatz, S.H., 1998. Collective dynamics of 'small-world' networks. Nature 393 (6684), 440–442.
Wechsler, D., 1997. Wechsler Adult Intelligence Scale — Third Edition. The Psychological Corporation, San Antonio, TX.
Westfall, P.H., Troendle, J.F., 2008. Multiple testing with minimal assumptions. Biom. J. 50 (5), 745–755.
Wishart, H.A., Saykin, A.J., Rabin, L.A., Santulli, R.B., Flashman, L.A., Guerin, S.J., Mamourian, A.C., Belloni, D.R., Rhodes, C.H., McAllister, T.W., 2006. Increased brain activation during working memory in cognitively intact adults with the apoe epsilon4 allele. Am. J. Psychiatry 163 (9), 1603–1610.
Wu, C.W., Gu, H., Lu, H., Stein, E.A., Chen, J.H., Yang, Y., 2008. Frequency specificity of functional connectivity in brain networks. Neuroimage 42 (3), 1047–1055.
Zaidel, E., Iacoboni, M., 2003. The Parallel Brain: The Cognitive Neuroscience of the Corpus Callosum. MIT Press, New York.
Zalesky, A., Fornito, A., Bullmore, E.T., 2010. Network-based statistic: identifying differences in brain networks. Neuroimage 53 (4), 1197–1207.