Hyperedge bundling: A practical solution to spurious interactions in MEG/EEG source connectivity analyses

Sheng H. Wang[1,2,3], Muriel Lobier[1], Felix Siebenhühner[1,2], Satu Palva[1,3], J. Matias Palva[1]

1. Neuroscience Center, University of Helsinki, Finland

2. Doctoral Programme Brain & Mind, University of Helsinki, Finland

3 BioMag laboratory, HUS Medical Imaging Center, Helsinki, Finland

* The correspondence should be addressed to:

Sheng H., Wang, Neuroscience Center, University of Helsinki

wang.sheng.h@gmail.com

or

J. Matias Palva, Neuroscience Center, University of Helsinki

matias.palva@helsinki.fi

Abstract

Inter-areal functional connectivity (FC), neuronal synchronization in particular, is thought to constitute a key systems-level mechanism for coordination of neuronal processing and communication between brain regions. Evidence to support this hypothesis has been gained largely using invasive electrophysiological approaches. In humans, neuronal activity can be non-invasively recorded only with magneto- and electroencephalography (MEG/EEG), which have been used to assess FC networks with high temporal resolution and whole-scalp coverage. However, even in source-reconstructed MEG/EEG data, signal mixing, or "source leakage", is a significant confounder for FC analyses and network localization.

Signal mixing leads to two distinct kinds of false-positive observations: artificial interactions (AI) caused directly by mixing and spurious interactions (SI) arising indirectly from the spread of signals from true interacting sources to nearby false loci. To date, several interaction metrics have been developed to solve the AI problem, but the SI problem has remained largely intractable in MEG/EEG all-to-all source connectivity studies. Here, we advance a novel approach for correcting SIs in FC analyses using source-reconstructed MEG/EEG data.

Our approach is to bundle observed FC connections into hyperedges by their adjacency in signal mixing. Using realistic simulations, we show here that bundling yields hyperedges with good separability of true positives and little loss in the true positive rate. Hyperedge bundling thus significantly decreases graph noise by minimizing the false-positive to true-positive ratio. Finally, we demonstrate the advantage of edge bundling in the visualization of large-scale cortical networks with real MEG data. We propose that hypergraphs yielded by bundling represent well the set of true cortical interactions that are detectable and dissociable in MEG/EEG connectivity analysis.

Keywords Signal leakage, spurious correlation, artificial correlation, volume conduction, signal mixing, point spread, graph theory, MEG, EEG

Highlights

- A true interaction often is "ghosted" into a multitude of spurious edges (SI)
- Effective in controlling and illustrating SI
- Hyperedges have much improved TPR and graph quality
- Advantages in visualizing connectivity

1 Introduction

Large-scale neuronal networks, *e.g.*, manifested by functional, directed, and effective connectivity(Karl J. 2011), are thought to be critical for healthy brain functions while their abnormalities are thought to underlie many brain diseases (Brookes et al., 2016; Bullmore and Sporns 2009; Bullmore and Sporns 2012; Fornito et al., 2015; Papo et al., 2014; Petersen and Sporns 2015; Rubinov 2015; Sporns 2014; Uhlhaas and Singer 2010; Uhlhaas and Singer 2006). Currently, magneto- and electro-encephalography (MEG/EEG) are the only non-invasive electrophysiological tools for studying connectivity networks with millisecond-range temporal resolution and good coverage of the cortical surface (Kujala et al., 2008; Palva and Palva 2012; S. Baillet et al., 2001; Salmelin and Baillet 2009). Accurately identifying interaction dynamics from MEG/EEG data is of crucial importance for understanding their role in human cognition and its deficits.

To date, numerous interaction metrics have been developed and utilized to assess functional connectivity (FC) in terms of amplitude-, phase-, and phase-amplitude correlations within or across frequency bands for pairs of electrophysiological signals (Bastos and Schoffelen 2016; Kreuz 2011; O'Neill et al., 2015). These pairwise metrics are typically applied to estimate FC among all brain regions, *i.e.*, to obtain “all-to-all” FC connectomes (Sporns et al., 2005). Networks of inter-areal FC are often represented as graphs where brain areas constitute the *nodes* (or vertices) and observed inter-areal connections the *edges* (Bullmore and Sporns 2009; Rubinov and Sporns 2010).

FC graphs estimated from MEG/EEG sensor space data are neuroanatomically uninformative and severely confounded by signal mixing. Signal mixing has two facets: first, any focal neuronal signal is picked up by several sensors. Conversely, one sensor detects a mixture of signals from several distinct sources. Source reconstruction can be used to reduce signal mixing and, importantly, elucidate the likely neuroanatomical sources of the MEG/EEG signals (Buzsaki et al., 2012; Gross et al., 2013; Hamalainen et al., 1993; Palva and Palva 2012; Schoffelen and Gross 2009). Yet, because of ill-posed nature of the inverse problem, no source reconstruction approach can yield an unambiguous estimate of the source topography. Residual signal mixing in source space, signal leakage, is quantitatively dependent on the source-reconstruction method of choice but qualitatively characteristic to all such methods.

Because of signal leakage, FC measures exhibit two distinct types of false positive observations: *artificial interactions* (AI) and *spurious interactions* (SI) (see Box 2, (Palva and Palva 2012)). AIs arise directly from the signal mixing by one true signal being smeared to multiple sensors or sources, regardless of whether true interactions are present. SIs are “ghost” interactions caused by the leakage of the signals from two true connected nodes to their surroundings nodes that in turn become falsely connected like the truly connected nodes. AIs can be suppressed by a number of bivariate metrics that typically aim to remove linear coupling terms, and therefore removing artificial and true interactions with zero- and anti-phase-lag coupling. However, the problem of SIs is much less acknowledged and more difficult to solve because SIs stem from multivariate mixing effects. With typical distributed source modeling approaches, signal leakage causes a large number of SIs that render both the network localization and graph property estimates inaccurate (Drakesmith et al., 2015).

To date, one solution has been proposed for correcting SIs in oscillation amplitude correlation estimates, which simultaneously orthogonalizes all source time series through the Löwdin procedure (Colclough et al., 2015; Colclough et al., 2016). Despite this promising advance, no solutions have yet been advanced to suppress SIs for other interaction metrics.

Here we advance a novel approach, hyperedge-bundling, to alleviate the problem of SIs problem in all-to-all connectivity analyses performed with any interaction metric. Instead of correcting the mixing effects in source signals *per se*, the approach is based on a quantification of the extent of mixing between all sources, evaluation of mixing similarity among all edges, and then clustering the edges into *hyperedge* bundles. This procedure aims to yield a hypergraph where each hyperedge represents a true interaction and its spurious reflections.

In this study, we performed a large set of connectivity simulations and realistic all-to-all MEG source space analyses, in which we estimated phase synchrony as a measure of FC with an AI-insensitive metric. We show that in simulated graphs, hyperedge bundling greatly decreases the number of false positives, *i.e.*, SIs. We illustrated on real MEG data how bundling can support an informative visualization of FC graphs. We suggest that such hypergraphs constitute accurate and unbiased representations of neuronal interactions observable in MEG/EEG source space.

2 Theory

This section covers general topics as follows: signal mixing in MEG/EEG, how spurious interactions (SI) arise from mixing between sources; and bundling of raw edges into hyperedges. The implementations specific to this study are described in the *Methods* section. Throughout the report, we denote a FC graph estimated from reconstructed source time series as raw graph $G_{raw} = (V, E)$, where brain regions are nodes $v_i \in V$ and interactions between nodes are "raw" edges, $e_k = \{(v_i,v_j) \in E | v_i,v_i \in V\}$.

2.1 Signal mixing results in false positive artificial (AI) and spurious interactions (SI)

Let us consider a scenario where a true phase correlation is present between two distant (unmixed) sources V_1 and V_2 (Fig 1A top). The signals from V_1 and V_2 are mixed with signals of their nearby and mutually uncorrelated neighbours V_3 and V_4. Estimating phase FC among all four nodes with the phase-locking value (*PLV*) will reveal both the true edge $E(V_1,V_2)$ and false positive "short-range" AIs between the nearby nodes $E(V_1,V_3)$ and $E(V_2,V_4)$, because *PLV* is inflated by mixing (thick gray edges, Fig 1A bottom). However, due to leakage of the signal from V_1 and V_2 to their neighbors V_3 and V_4, false positive "long-range" SIs $E(V_3,V_4)$, $E(V_2,V_3)$, and $E(V_1,V_4)$ will also be observed (thin dashed edges). These SIs are thus only indirectly caused by mixing and, unlike the zero-phase-lag AIs (see 2.2), SIs inherit the phase-lag of the true interaction. Mixing-insensitive bivariate metrics such as the imaginary part of *PLV* (*iPLV*) can remove AIs but do not eliminate SIs if the true coupling has non-zero phase lag.

2.2 Quantifying the mixing between reconstructed sources

Signal mixing/leakage between two sources is instantaneous and therefore always leads to inflated zero-phase-lag correlations between the sources. Mixing does not vary over time or across frequency bands (Brookes et al., 2012; Brookes et al., 2014; Drakesmith et al., 2013; Nolte et al., 2004; Palva and Palva 2012).

2.2.1 Source-reconstruction

Suppose we have a data matrix $X = \{x^{(1)}, x^{(2)}, \ldots, x^{(n)}\} \in \mathbb{R}^{n\times t}$ representing narrow-band time series of t samples from n neuronal sources. Simulating a MEG/EEG recording, X can be linearly projected to sensor-space:

$$Y = \Gamma X + \varepsilon \tag{1}$$

where $Y \in \mathbb{R}^{s\times t}$ represents the forward-modeled time series from s sensors ($n > s$). Here, $\Gamma \in \mathbb{R}^{s\times n}$ is the forward operator (or the lead field) and $\varepsilon \in \mathbb{R}^{s\times t}$ is the model prediction error derived from measurement noise. Next, Y can be projected back into the source-space, *e.g.*, by minimum-norm estimation (MNE) based inverse modeling:

$$\hat{X} = WY = R\Gamma^T(\Gamma R \Gamma^T + \lambda^2\chi)^{-1}Y \tag{2}$$

where $W \in \mathbb{R}^{n\times s}$ is the inverse operator (sources × sensors), λ^2 is a regularization parameter, R is the source covariance matrix, and χ is the noise covariance matrix. Usually, source vertices are then collapsed onto a number (50-400) of cortical parcels.

2.2.2 Cross-talk function and resolution matrix

In MEG/EEG source connectivity studies, a resolution matrix $R = W\Gamma$ ($R \in \mathbb{R}^{n\times n}$) is often used to describe the relationship between true signals and modeled signals from n sources in the absence of noise (Farahibozorg et al., ; Hauk and Stenroos 2014; Hauk et al., 2011; Liu et al., 2002). In R, each diagonal element quantifies the sensitivity for estimating signals from that source. Each row of R is the "cross-talk" function (CTF) that describes the amount of mixing between one source and all other sources. Each column of R is a "point-spread" function (PSFs) that describes how the modeled signal from any one source is spread across all other sources.

2.2.3 The mixing function

We approximate R numerically by a *mixing function* f_{mix} that describes the mixing between reconstructed sources. We rationalize that if the true source signals are uncorrelated, the amount of correlation at zero-lag between reconstructed signals can only be explained by mixing between the sources. Thus, f_{mix} can be quantified by the zero-lag correlation between parcel time series estimated using a simulated MEG/EEG measurement of uncorrelated source noise.

We first generate uncorrelated signals $X(t) \in \mathbb{R}^{n\times t}$, t samples for n parcels, and forward transform them to obtain sensor signals Y (eq. 1). We next inverse transform Y to obtain $\hat{X}$ (eq. 2). In this process, the reconstructed signals $\hat{X}_0^{vi}$, $\hat{X}_0^{vj}$ of any two nearby sources v_i and v_j become correlated to a certain degree due to mixing. Thus, the mixing from the *true* signal of v_i to the *reconstructed* signal of v_j can be quantified as:

$$f_{mix}(v_i, v_j) = \left| re(cPLV(X_0^{(vi)}, \hat{X}_0^{(vj)})) \right| \quad (3)$$

where $re()$ denotes the real part of a complex number and $cPLV$ is the complex-valued phase locking value:

$$cPLV(\mathrm{A}, \mathrm{B}) = \frac{1}{T}\sum_{t=1}^{T}\left[e^{i(\theta_A(t)-\theta_B(t))}\right] = \frac{1}{T}\sum_{t=1}^{T}\left[\frac{S_A S_B^*}{|S_A||S_B|}\right] \quad , \quad (4)$$

where T denotes the number of samples, θ_A and θ_B are the instantaneous phases of signal A and B; S_A and S_A are complex-valued narrow-band signals from A and B, and * is complex conjugate. Because mixing is instantaneous, re($cPLV(A,B)$) captures all correlations caused by mixing. For parcel pairs that do not become correlated by signal mixing, f_{mix} is near zero. For parcel pairs influenced by signal mixing, $f_{mix} >> 0$ and reaches 1 for complete mixing.

2.3 Signal mixing smears a true interaction into multiple spurious interactions

For a simplified illustration of how signal mixing / source leakage produces SIs, we used model with a 13 × 13 grid of point sources. The mixing function f_{mix} (169 × 169) was defined so that mixing between any two sources was a 2D Gaussian distribution decreasing with distance between the two sources (inset, Fig 1B, methods see *Supplementary*).

We simulated one true edge by setting two sources V_1 and V_2 to have perfect phase coupling with non-zero phase lag and keeping the remaining 167 sources uncorrelated. Next, we introduced mixing between reconstructed sources and mapped all-to-all phase FC with an AI-free metric, the imaginary part of the phase-locking-value (*iPLV*) (Palva and Palva 2012)

$$iPLV = |im(cPLV)| \quad , \quad (5)$$

where $im()$ denotes the imaginary part of a complex number. The *iPLV*, like the imaginary coherency (Nolte et al., 2004), removes zero-lag couplings by excluding the real part of *cPLV*. Therefore, *iPLV* yields only the true phase-lagged interactions and their false positive ghosts (SIs). In this simulation, visualization of the strongest 0.1% of *iPLV* edges revealed the true edge and several SIs, all of which connected sources within the mixing neighbourhoods of the true sources V_1 and V_2 (Fig 1B).

2.4 Raw edges can be bundled into hyperedge by their mixing similarity (S_E)

The *mixing similarity* can next be derived with the known f_{mix} to describe how close these edges are with each other in signal mixing. A bivariate similarity estimation yields a mixing similarity matrix S_E, where each element $S_E(i, j)$ quantifies the similarity between two edges E_i, E_j (for how-to, see 2.5).

Our objective is to classify raw edges by mixing similarity into “hyperedges”, where each hyperedge is a “bundle” of raw edges (including true and SI edges): $e_\kappa^H = \{\{e_k=(v_i,v_j)\}\in E|v_i,v_i\in V\}$. The raw graph is thereby transformed into a hypergraph $G_h = (V, E^H)$. Within any one hyperedge, all raw edges are mixing-wise close to each other but distant from the raw edges of other hyperedges, and thus collectively representing a "community" of raw edges that we hypothesize to include the underlying true interaction and its ghosting SIs.

This classification can be done by partitioning the S_E matrix into clusters with an appropriate clustering method (Fig 1C). In the simplified 13 × 13 toy model, bundling transformed the raw graph with a multitude of false positives into a hypergraph with one hyperedge that captured the true interaction with zero false positives.

For visualizing hyperedges, we utilized a "force directed edge bundling" method that both indicates the adjacency of the constituent raw edges and illustrates the loci where the SIs originated (Holten and Wijk 2009).

2.5 Hyperedge bundling for multiple true interactions

To demonstrate that bundling could be extended to separate multiple true interactions, we expanded the simulation and modeled interactions with three degrees of adjacency: "kin", "nearby", and "far". The estimated raw graph yielded the true-positive (TP) edges surrounded by numerous false positive (FP) SIs (Fig 2D). Estimating and partitioning the edge similarity matrix S_E revealed that: 1) two "kin" edges were inseparable and together with their SIs they merged into the largest hyperedge HE_1 (Fig 2E); 2) the "far" pair was clustered into two clearly separable hyperedges HE_2 and HE_5; 3) the "nearby" pair and their SIs were also clustered into two distinct hyperedges HE_3 and HE_4 with greater inter-hyperedge similarity as measured by mean-linkage (green box) than the "far" pair (magenta box); 4) a few scattered random false positive edges were also clustered into hyperedges (gray box), but they were much smaller in size than any of the hyperedges containing a true edge.

If a hyperedge containing at least one true raw edge is considered as a TP observation, bundling greatly decreased graph noise in terms of the FP/TP ratio. FP/TP in raw graph was 239/6 and 4/5 in the hypergraph, which marks a reduction in the fraction of FPs by a factor of 50. Visualizing these bundles showed that the hypergraph had less visual clutter and facilitated identification of the true interactions compared to the raw graph (Fig 2F).

2.6 Estimation of the edge similarity matrix S_E

Hyperedge bundling is based on the raw connectivity graph A_{FC} (a sparse matrix containing only significant edges), and the mixing function f_{mix} (Fig 2A, C). We first parsed the edges in A_{FC} into a list of node pairs (Fig 2B). We next find the mixing between all the involved nodes from f_{mix} (Fig 2C, and colour-coded and illustrated geometrically in Fig 2D) to compute the edge-to-edge adjacency in signal mixing.

2.6.1 The edge adjacency matrix (A_E)

For a raw graph with m edges, the edge-to-edge adjacency matrix $A_E \in \mathbb{R}^{m \times m}$ represents the pairwise mixing adjacency among all raw edges and is necessary for computing the similarity matrix S_E. The adjacency between two edges $E_i(V_1,V_2)$ and $E_j(V_3,V_4)$} was defined as follows (Fig 2D):

if $V_1 \sim V_4$ *are distinct nodes*

$$A_E(i,j) = max\ [\ f_{mix}(V_1,V_3)\ f_{mix}(V_2,V_4),\ \ f_{mix}(V_1,V_4)\ f_{mix}(V_2,V_3)\ ]$$

elseif $V_1 == V_3 :\ A_E(i,j) = f_{mix}(V_2, V_4)^2$

elseif $V_2 == V_4$: $A_E(i,j) = f_{mix}(V_1,V_3)^2$

elseif $V_1 == V_4$: $A_E(i,j) = f_{mix}(V_2,V_3)^2$

elseif $V_2 == V_3$: $A_E(i,j) = f_{mix}(V_1,V_4)^2$

elseif $i==j$: $A_E(i,j)=0$ % *diagonal of* A_E (6)

here "==" is assertion, "=" is assignment. This algorithm is applied for all pairs of edges in the raw graph to populate the A_E matrix (Fig 2E).

2.6.2 *Evaluation of Edge Similarity (S_E) with correlation of edge mixing profiles in A_E*

We denote rows of the A_E matrix as the *signal mixing profiles* so that $A_E(i)$ and $A_E(j)$ are the mixing profiles of edges E_i and E_j, respectively, and thus indicate their mixing adjacency to all the other raw edges in the graph. If E_i and E_j are similar to each other, *i.e.*, a high correlation between $A_E(i)$ and $A_E(j)$, edge E_i will be similar to all the edges in the raw graph that E_j is similar to, and vice versa (Fig 2F&2G). Such pattern can be already observed in the simplified models (Fig1) where SIs of any given true edge are all close to each other and adjacent to the true interaction.

Conversely, if two edges are far apart in mixing, their mixing profiles exhibit little to no correlation. Using correlation estimates of mixing profiles, it is thus possible to assess the significant similarity of all pairs of edges in A_E and populate the similarity matrix $S_E \in \mathbb{R}^{m\times m}$ (Fig 2H). The significance level of the correlation can be determined through a Fisher z-transformation. Hyperedge bundling is based on the notion that a S_E can be partitioned into clusters of raw edges that are similar to each other in mixing within each cluster and therefore to collectively reflect a shared true underlying interaction.

2.7 The resolution of hyperedge bundling is defined by the cutoff limit

We partition the edge similarity matrix S_E into clusters of "hyperedges" so that within any one hyperedge, the raw edges are mixing-wise close (large S_E values) to each other and distant (small S_E values) from raw edges of other hyperedges.

We now introduce a control parameter, the *cutoff limit* (CL) that dictates the "resolution" of a hypergraph. CL is defined as the ratio of desired number of clusters to the number of available raw edges to be clustered. For example, for a graph of 1000 edges, a CL of 0.1 causes the clustering method to partition the S_E matrix into 100 hyperedges. We chose to control clustering using the CL for better comparability of clustering methods or graphs of different sizes. The similarity matrix $S_E \in \mathbb{R}^{m\times m}$ can be partitioned into arbitrary number of clusters from 1 to $m-1$, *i.e.*, CL ranging from $1/m$ to $(m-1)/m$ (Fig 2I, for details, see *Supplementary*).

2.8 Validate the stability of hyperedge clustering

To ensure that the hyperedges are not random outcomes of partitioning the similarity matrix, the "stability" of partitioning solutions must be evaluated. We ask, at any resolution (CL=c), if the differences between the partitioning solutions of n randomly perturbed versions of a similarity matrix S_E is statistically smaller than their surrogate counterparts, the partitioning solution can be considered as stable (*Supplementary*). The distance between two partitioning solution can be estimated with the *variation of information* (VI,(Meilă

2007)). The independent perturbations to a similarity matrix can be acquired by randomly deleting a small subset, *e.g.*, 10~20%, of the elements in the similarity matrix (Ben-Hur et al., 2002; Williams et al., 2015). The surrogates can be obtained by randomly rewiring the original similarity matrix.

3 Methods

The goal of this study was to assess the performance and applicability of hyperedge bundling in suppressing spurious interactions (SI) in MEG/EEG source connectivity studies. To this end, we obtained large numbers of functional connectivity (FC) graph estimates from simulated data with realistic sources and inverse modeling. We next evaluated the efficacy of hyperedge bundling in capturing true positive (TP) interactions and rejecting false positive (FP) SIs. Finally, we demonstrated the bundling of FC graphs estimated from MEG data recorded in a visual working memory (VWM) experiment.

This section includes the procedural outlines of the simulations and evaluation of bundling efficacy. The preprocessing pipeline, technical details of the simulations and preprocessing of the VWM experiment are described in *Supplementary*.

3.1 Simulating "truth" time series of varying coupling strengths

In real electrophysiological data, mixing is inhomogeneous across source loci and subjects (Brookes et al., 2014) and coupling strengths of neuronal interactions also exhibit great spatiotemporal and inter-subject variability (Preti et al., 2016; Zalesky et al., 2014). To account for such variability, we created 1000 distinct *truth* graphs each containing 200 randomly generated true interactions between 400 cortical parcels in a standard cortical source space. Each node thus connected only to a single other node, which allows an unbiased survey of the whole cortical surface in every graph realization. We did not simulate structured networks to exclude the impact of higher order SI here. These higher order SI can arise from common drive, third-party sources, and cascade effects, although identifying them is of equal importance (Mannino and Bressler 2015; Wollstadt et al., 2015).

For every truth graph, we simulated ten sets of time series, representing two different modes of coupling (gamma distribution or uniform distribution) at 5 different levels of coupling strength each (Supplementary). A set of uncorrelated time series was also simulated as null hypothesis H_0 for each truth graph which would be used for estimating the source-space mixing properties and as the baseline condition in group analysis.

3.2 Estimation of mixing properties using the H_0 time series

Mixing in source reconstructed MEG/EEG data is essentially captured in the forward and inverse operators used in source reconstruction. These operators are determined by the data acquisition system and specifics of the individual source model (Wens 2015). In addition to the mixing function f_{mix} (see 2.2.3), we characterized the source model used here with a set of additional mixing metrics obtained from the 12 subjects from the VWM experiment:

1) Parcel fidelity, f_p, that quantifies the reconstruction accuracy and is defined as the phase correlation between the simulated H_0, and reconstructed $\widehat{H}_0$, time series of each parcel v_i

$$f_p(v) = \left|re(cPLV(x_0^{(vi)}, \hat{x}_0^{(vi)})\right|. \quad (7)$$

2) Edge fidelity, $f_e(v_i,v_j) = f_p(v_i)f_p(v_j)$, that indexes the reconstruction accuracy of raw edges connecting nodes v_i and v_j.

3) Residual spread function, $PLV_0(v_i,v_j)$, that is the phase correlation between two parcels' reconstructed ($\widehat{H}_0$) time series.

$$PLV_0(v_i, v_j) = \left|re(cPLV(\hat{x}_0^{(vi)}, \hat{x}_0^{(vj)}))\right| \quad (8)$$

The definition of PLV_0 appears similar to that of f_{mix}, but they are conceptually different. The f_{mix} measures how much of each source's true signals are picked up in other sources' reconstructed signals. PLV_0, on the other hand, is the correlation between any two sources' modeled $\widehat{H}_0$ time series where both are contaminated by mixing with numerous other sources. Because the *iPLV* estimates can be biased by mixing, we used PLV_0 to exclude edges connecting sources with large mixing.

3.3 Elimination of poorly measurable edges with the intractable-edge-mask (IEM)

We applied an intractable-edge-mask (*IEM*) to exclude edges that connect sources with poor reconstruction accuracy. True interactions between these sources may exist, but cannot be reliably detected because estimations of connectivity between them are unreliable due to the limitations of the source model. We utilized the mixing properties (see 3.2) and construct a group-level *IEM* in two steps:

1) With average edge-fidelity $<f_e>$ and the residual spread $<PLV_0>$, we create two Boolean masks:
   i. The edge-fidelity mask (M_{fe}) to exclude edges with low fidelity, thereby removing edges connecting poorly reconstructed sources.
   ii. The residual spread mask (M_{PLV0}) to exclude edges with large PLV_0, thereby removing edges whose FC estimates likely are much distorted by mixing between these loci.
2) The *IEM* is the union of these two masks.

In this study, we set 0.1 as the threshold for M_{fe}, which removed the 40% most poorly reconstructed edges from all 79,800 ($N(N{-}1)/2$, $N = 400$) possible edges in raw graphs. The M_{PLV0} was acquired by deleting edges whose PLV_0 was greater than the 95th percentile of the PLV_0 matrix.

3.4 Estimation of group-level FC of simulated graphs

The group-level significant *iPLV* estimates thresholded with the *IEM* were used as raw graphs for hyperedge bundling. The group-level analysis for the simulated graphs and for real MEG/EEG data in the VWM experiment were carried out in the same manner. For simulated graphs, we forward- and inverse-modeled the coupled truth time series into 12 subjects' individual source space, thereby introducing mixing into reconstructed signals (Schoffelen and Gross 2009). We next estimated *iPLV* connectivity for these subjects. We then tested across subjects, for each edge in every estimated FC graph, whether there was a significant difference (one-tailed t-test) in the *iPLV* estimate between the coupled and the H_0 condition. Those edges that showed a significant difference were identified as raw edges (corrected for multiple comparisons within each FC graph). We acquired FC graphs with three significance levels $p < 0.05$, 0.01, and 0.001 for each of the ten coupled time series.

3.5 Hyperedge bundling with two clustering methods

After applying the *IEM* to all group-level FC matrices, we followed the procedures described in *Theory* to obtain the similarity matrix S_E for each FC. We next partitioned each S_E into clusters of “hyperedges” with two clustering methods. The unweighted pair group method with arithmetic mean (UPGMA) is an agglomerative hierarchical clustering method that builds a rooted hierarchical tree to represent the distance in signal mixing between all raw edges (Jain et al., 1999). The Louvain method for community detection extracts communities by optimizing the modularity of clusters through a gradient descent procedure (Blondel et al., 2008)(Blondel et al., 2008).

3.6 Comparing hypergraphs with raw graphs

We denoted the TPs as the edges from truth graphs that were identified as significant edges in the group-level FC matrix, and FPs as significant edges in the group-level FC matrix but absent in the truth graph. Thus, the true positive rate (TPR, sensitivity) is given by $TPR = TP/N_{true*}$, where N_{true*} is the number of “detectable true edges” referring to the number of simulated true edge that passed the intractable-edge-mask. We further defined the *noise* as the FP to TP ratio. An ideal group-level FC should capture as many of the true interactions as possible while rejecting other edges, *i.e.*, high TPR and low FP/TP.

We used TPR and FP/TP as the main criteria to characterize raw graphs instead of the commonly used receiver operating characteristic curve (ROC) for two reasons. First, the ROC is derived from the TPR and false positive rate (FPR) which are not directly comparable between raw graphs and hypergraphs, as these are different constructs; second, because the number of FP is disproportionally larger than that of TP (as shown later with an example), the shape of the ROC is misleadingly optimal when limiting the number of raw edges with varying edge weight threshold.

We defined a TP hyperedge (TP_{HE}) as a hyperedge capturing at least one TP raw edge, whereas a FP hyperedge (FP_{HE}) contained only FP raw edges. Hyperedges may also contain multiple TP raw edges. To quantify this, we defined *separability* as the fraction of true positive hyperedges that contain only one TP raw edge out of all true positive hyperedges. An ideal hypergraph should balance high TPR and separability against low FP/TP.

4 Results

This section includes three parts: 1) Demographics of group-level FC of the simulated graphs; 2) Efficacy of hyperedge bundling; 3) Application of hyperedge bundling to real MEG data.

4.1 Group-level FC as raw graphs

In individual subjects, mixing introduced by the virtual MEG experiment distorted *PLV*, *iPLV* and the phase-lag of all measured graphs of varying coupling strength including the H_0 time series (*Supplementary*). To find group-level significant edges, we tested for each edge whether there was a difference in *iPLV* value between the coupled condition and the H_0 condition (Fig 3A, see 3.4). Edges that showed a significant difference were reported as raw edges (corrected for multiple comparisons). Thus, we obtained FC graphs for each of the ten sets of coupled graphs at 3 significance levels of $p < 0.05$, 0.01 and 0.001.

4.1.1 Raw graphs of iPLV edges are noisy

Overall, the number of significant *iPLV* edges increased as coupling strengths increased (Fig 3B). The group-level graphs at all 3 significance levels captured over 75% of all detectable TP edges, except in the case of weak uniform coupling, $C_c(0.1)$ (Fig 3C). We simulated 200 random edges in each ground truth graph and computed the true positive rate (TPR) for each measured group graph as the number of significant edges divided by the number of all simulated true edges that passed through the intractable-edge-mask (IEM). Despite the high TPR, there was a large variability in the ratio of false and true positives, FP/TP, across these graphs (Fig 3D).

4.1.2 Is strict statistical thresholding a good solution for pruning FPs?

We chose the graphs of gamma coupling $C_\lambda(r{=}15)$ and uniform coupling $C_C(0.5)$ to test statistical thresholding (below) and hyperedge bundling (4.2) because they had comparable TPR (Fig 3C) and equivalent true edge strengths (see distribution in Supplementary 1). Moreover, both contained only ~750 edges, which mitigated computational overhead in later clustering analyses.

One sensible way to identify key structures in FC graphs is to apply a statistical threshold to *iPLV* values. We found that by increasing the significance *iPLV* threshold, the number of FP edges decreased at a faster rate than the number of TP edges in both graphs (Fig 3E). Around 120 of the 640 strongest edges were TP, giving a TPR > 90% for 125 detectable true edges, but a FP/TP ratio of 4. When retaining the 20 strongest edges reduced the FP/TP to 0.1 (Fig 3F) but at the cost of reduced TPR, (TPR = 0.15). Overall we found that the mean *iPLV* of TP edges was larger than that of FP edges' (Fig 3G), which suggests that strict thresholding is an applicable solution for reducing FP/TP but comes at a price of an elevated false negative rate, although the shape of ROC curve appeared to be optimal (inset Fig 3E).

4.2 Hypergraphs yields better FP/TP than raw graphs with reasonable TPR cost

4.2.1 The stability of clusters

Evaluating the stability of clustering was a necessary step prior to further analysis of the properties of hyperedge clusters. The resolution of clustering and thereby of the hypergraphs was controlled by the *cutoff limit* (CL, see 2.6). We used bootstrapping to identify the CL range that yielded stable partitioning of the raw graphs (see Methods and *Supplementary*). We found that at CL < 0.4, both UPGMA and Louvain clustering yielded significantly more stable partitions for simulated graphs than their randomly rewired counterparts (Fig 4A). For the 640 raw edge graphs, this CL upper bound corresponded to ~250 hyperedges. In the following analysis, we thus tested bundling with CL ranging from 0.05 to 0.45.

4.2.2 Cluster-size distribution

We next quantified the distributions of hyperedge sizes (numbers of raw edges per hyperedge, Fig 4B) by pooling hyperedges from 500 clustered graphs with CL ranging from 0.05 to 0.45. As expected, we found a systematic shift towards smaller hyperedges with increasing resolution/CL. The Louvain method consistently yielded more small hyperedges than UPGMA.

4.2.3 Hyperedge-bundling performance: trade-offs between separability, TPR and graph noise

Hyperedge bundling aims to detect and separate as many TP interactions as possible while rejecting as many FP as possible. We tried to find an optimal balance among these competing outcomes by taking into account two aspects of hyperedge bundling: separability and noise. We defined *separability* as the ratio between singleton TP hyperedges (containing only one TP raw edge) and all TP hyperedges, and *noise* as the FP/TP ratio of the hyperedges. An ideal hyperedge partitioning would thus have *separability* = 1, FP/TP ~0, and a TPR equal to the TPR of raw edges.

We observed that by increasing the hyperedge resolution (CL from 0.05 to 0.45), the *separability* increased but noise also increased with both clustering methods (Fig 4C, 4D). Thus at coarse resolutions (low CL), multiple TP raw edges were partitioned into one hyperedge but there were very few FP hyperedges, likely because there were less small-sized hyperedges. Conversely, at fine resolutions (high CL), separability was improved but at the cost of having greater numbers of FPs.

Knowing that small hyperedges are more likely to be FPs than large hyperedges (Fig 1E), we further tested whether excluding hyperedges by size would decrease noise. At each resolution level, excluding small hyperedges lead to a decrease in noise (FP/TP decreased with increasing θ_{HEsize}, Fig 4C, D). Nevertheless, this was accompanied by reduced separability (y axis, Fig 4C, D) and a reduced TPR (Fig 4E, F) caused by the removal of small-sized TP hyperedges together with FP hyperedges.

To summarize, at all graph resolutions, hypergraphs were less noisy than raw edge graphs. In the least noisy hypergraph (*e.g.*, Louvain, CL = 0.05 and $\theta_{HEsize} > 8$), 87% of the 125 TP raw edges were retained while achieving a 10^3-fold decrease in noise compared to the underlying raw graphs, *i.e.*, FP/TP decreased from (640-125)/125 = 4.1 (C_γ raw graphs in Fig3E) to 3.8×10^{-3} (leftmost filled box on the cyan curve, Fig 4F). Nevertheless, this improvement came at the cost of poor separability, meaning many hyperedges in CL = 0.05 graphs contained several true edges. To balance an optimal trade-off, we decided to use CL ≥ 0.15 and $\theta_{HEsize} > 2$, expecting to achieve a reduction of FP/TP to 0.1 (from 4.1 in raw edges) with negligible reduction in TPR and adequate separability (0.5).

4.2.4 *Louvain clustering yields less noisy hypergraphs but lower separability than UPGMA clustering*

The Louvain method produced more small hyperedges than the UPGMA method (Fig 4B). Although the Louvain hypergraphs had higher level of noise when retaining singleton hyperedges ($\theta_{HEsize} = 0$), this relation was inverted when singleton hyperedges were screened (Fig 5A). This indicates that the majority of the singleton hyperedges yielded by Louvain were FPs. Moreover, the Louvain hypergraphs had greater TPR when CL values were between 0.15 and 0.25 (Fig 5B). These advantages, however, came at the cost of separability, which was better with UPGMA throughout the tested range (Fig 5C).

4.3 Visual working memory networks: real MEG data

To assess the feasibility of using hyperedge bundling with real MEG/EEG data, we applied bundling to raw graphs that reflected significant strengthening of inter-areal phase synchronization during memory retention compared to pre-stimulus baseline during a visual working memory task (see *Supplementary* and Honkanen et al., 2015).

We found that the *iPLV* estimates in alpha- and gamma-frequency band were greater during memory retention than in pre-stimulus baseline. Here, we picked the 1000 strongest *iPLV* edges and drew them as lines linking the synchronized parcels on a flattened cortical surface (Fig 6A, 6B). We also illustrated a randomly picked graph from our simulations as a comparison (Fig 6C). We applied hyperedge bundling (UPGMA with CL=0.15, θ_{HEsize}>6) to these raw graphs. The resulting hypergraphs, the real MEG and simulated FC graphs alike, offer better visualization of large-scale FC than raw graphs, emphasizing the long-range synchronizations between brain regions(Fig 6D, 6E, 6F).

5 Discussion

MEG and EEG have great potential for yielding insight into the spatio-temporal structure of brain connectivity. Nonetheless, due to the ill-posed nature of the inverse problem, linear mixing and inaccurate source localization complicate MEG/EEG connectivity analyses both by distorting phase and amplitude estimates and by leading to false positive observations of artificial (AIs) and spurious interactions (SIs). We advance here a novel methodological framework, hyperedge bundling, to suppress SIs in brain connectivity graphs. We found that hyperedge bundling can be used to reduce the false positive rate with moderate to little decrease on the true positive rate.

Hyperedge bundling has several features that are advantageous and facilitate its application. First, since it is done only after interaction analyses, it does not require sophisticated preprocessing to suppress mixing effects in the original source time series. Hyperedge bundling only requires the forward and inverse operators and a mixing function estimated analytically or from simulations. Accordingly, it also inherently takes the source-model heterogeneity appropriately into account. Hyperedge bundling is also independent of the interaction metric and can be applied to connectomes estimated with any bivariate interaction metric. Finally, the nodal groups in the hypergraph obtained from hyperedge bundling constitute data-driven coarsening of originally high-resolution source parcellations. We suggest that these nodal groups be more representative of the true co-active local areas than *a priori* constructed low-resolution parcellations. This can be an aspect for future work.

In summary, hyperedge bundling can be used to suppress SIs and identify putative true edges in brain connectivity data and thereby to improve the localization of true interacting neuronal networks.

Hyperedge bundling vs. edge thresholding: reducing FP/TP while maintaining acceptable true positive rate

Some connectivity studies have reduced the amount of edges by applying strict criteria on edge selection. However, biases and instability of graph properties can be introduced when using arbitrary threshold criteria on raw edges (Drakesmith et al., 2015; van Wijk, Bernadette C. M. et al., 2010) and weak connections may also play an important role in cognitive functions (Santarnecchi et al., 2014). Nevertheless, imposing strict criteria for thresholding is an attractive option for increasing the fraction of true positives among all observations, *i.e.*, decreasing the FP/TP ratio (see Fig 3E and F) and for focusing the outcome on most robust effects. However, this approach, while effective in excluding FPs (SIs), also excludes a large fraction of true positives. For example, we found that in raw graphs when we applied a threshold strict to decrease noise (FP/TP ratio dropped from 4 to 0.1), but the TPR was reduced to 0.15. In contrast, with hyperedge bundling we could obtain the same noise level (FP/TP of 0.1) while preserving a TPR of up to 0.88 (see brown line, Fig 4F). Hyperedge bundling is thus superior to strict thresholding in attenuating FP/TP with little decrease in TPR.

Importantly, our simulations show that the raw edges with largest correlational estimates might not correspond to the strongest or most important neurophysiological connections because these estimates appear to be correlated with reconstruction accuracy (*Supplementary*). The reconstruction accuracy is heterogeneous across source space, meaning high accuracy of sources may positively bias the *iPLV* estimates. This bias is another reason for including weak observations in FC graphs.

Control parameters of hyperedge determine resolution and the balance among FP/TP, TPR, separability

In the current implementation, hyperedge bundling is controlled by the cutoff limit (CL) and the hyperedge size threshold (θ_{HEsize}). CL determines the resolution of the hypergraph and the balance between noise (FP/TP) and *separability* of true hyperedges. Low CL values lead to low noise in hypergraphs but poor separation of true raw edges into distinct hyperedges. θ_{HEsize} can be used to prune the smallest hyperedges to further reduce noise, albeit at a cost of pruning TP hyperedges.

We compared two clustering methods, UPGMA and Louvain. While the results showed clearly that by and large both clustering methods yielded similar performance, each method had interesting advantages. Louvain yielded better TPR than UPGMA for CL values between 0.05 and 0.25 (see Fig. 5B), and lower noise when singleton hyperedges were excluded (see Fig. 5A). UPGMA, on the other hand, yielded better separability of TP hyperedges throughout the control parameter ranges. Overall, using either clustering method with CL = 0.15–0.25 and θ_{HEsize} = 1–2 will yield a large reduction in FP/TP (from 4 to 0.1–0.2) with good separability and negligible reduction in TPR.

In applications to real data where the truth graph is unknown, choosing parameters, *i.e.*, to control the trade-off between suppressing noise and maintaining high TPR and separability, can be based on both our simulation results and objectives of the research. If the objective of the hypothesis requires good separability (*e.g.*, connectivity between specific visual areas to inferior parietal region), one should create high resolution hypergraphs, but this will be accompanied by sub-optimal noise reduction. Conversely, if the objective is to establish connectivity between the visual and parietal regions, a low resolution hypergraph (with low noise) is pertinent.

Comparison of hyperedge bundling and symmetric orthogonalization

Symmetric orthogonalization is a pioneering solution to the overall problem of SIs in the context of amplitude correlation estimation (Colclough et al., 2015). Its predecessor, pairwise orthogonalization (Brookes et al., 2012; Hipp et al., 2012) excluded instantaneous mixing and evaluated amplitude correlations for each time-series pair at a time. It is thus applicable to the estimation all-to-all amplitude correlations similarly to any other bivariate AI-free metric for phase or other forms of coupling, and also suffers from SIs in the same manner.

Symmetric orthogonalization overcomes the problem of SIs by simultaneously removing zero phase-lag components from all source time series through a gradient descent procedure known as the Löwdin orthogonalization (Everson 1999; Löwdin 1950). Next, all-to-all amplitude correlations are estimated with partial correlation of amplitude envelopes to keep direct and remove indirect interactions (Marrelec et al., 2006). Because the partial correlation matrix is expected to be sparse, a graphical lasso regularisation of the inverse covariance matrix is applied to penalize near-zero elements (Banerjee et al., 2008; Friedman et al., 2008), which reduces noise in the partial correlation graph.

Symmetric orthogonalization effectively attenuates SIs caused both by signal leakage and by indirect true couplings (*i.e.*, A<->C correlation, when true correlations are A<->B<->C). The two limitations of this method are: *i)* it is applicable only to the estimation of amplitude correlations, *ii)* it is limited by the rank of the data due to its dependence on singular value decomposition. For MEG/EEG data that are preprocessed with signal space separation (SSS) and temporal SSS methods, the rank of the data (~degrees of freedom) is often limited to 60–70 (Haumann et al., 2016). Thus, symmetric orthogonalization should be applied to cortical networks with less than 60–70 independent sources, such as the 19 regions per hemisphere used in (Colclough et al., 2015). For studying FC with greater parcellation resolutions (>>70) or with interaction metrics other than amplitude correlations, hyperedge bundling thus provides an alternative method for SI suppression. The similarities and differences between symmetric orthogonalization and hyperedge bundling are summarized in Table 1.

Optimal source space for brain connectivity analyses

There are numerous source reconstruction methods for MEG/EEG and the choice of method may have profound impacts on source connectivity analysis due to their difference in sensitivity to various synchronization profiles of the interacting sources (Hincapié et al., 2017). Although in the present study we used linear inverse operators (Hamalainen and Sarvas 1989; Hamalainen and Ilmoniemi 1994; Lin et al., 2006), hyperedge bundling can also be used with other source reconstruction methods as long as the amount of mixing among the sources/parcels can be quantified.

Parcel numbers in current MEG/EEG source connectivity studies range from tens of parcels, *e.g.*, 38 in ((Colclough et al., 2015) and around 70 in (Farahibozorg et al., ; Hillebrand et al., 2012), up to 200-400 parcels(Lobier et al., 2017; Siebenhuhner et al., 2016; Zhigalov et al., 2017). We propose that the source-space for FC studies should have a fine spatial resolution that enables the separation of nearby independent signals to an extent allowed by the source reconstruction approach. Neither the neuronal source constellations nor the degrees of freedom in the data are likely to match any *a priori* chosen parcellation scheme and hence coarse parcellations can misrepresent or miss source areas that fall in between the parcels or are much smaller than the parcels.

Our approach to use 400 parcels aims to eliminate the possibility of such pitfall. Moreover, with fine-grained parcellations, hyperedge bundling can well measure the mixing among raw connectivity edges and produce hypergraphs with high confidence of capturing and separating true interactions. Furthermore, the nodal groups connecting hyperedges can be utilized to coarsen a fine-grained source space in a data-driven manner and with consideration of the constraints posed by the source model. On the other hand, hyperedge bundling will likely to fail in a source-space of low spatial sampling, where the mixing similarity between observed edges is likely to be low due to initial low mixing among neighboring parcels.

6 Acknowledgements

We thank Prof. Karim Jerbi for comments to the work. We thank Dr. Giles Colclough for sharing source code and his insights to the symmetric orthogonalization method. This research was funded by the Academy of Finland grants 266745 and 281414. The funding bodies had no role in the design, data acquisition and analysis, decision to publish, or preparation of the manuscript.

7 Table

	Symmetric orthogonalization (Colclough et al., 2015)	Hyperedge bundling
Type of FC	• Amplitude-correlation	• Any form of FC or EC
Data for FC estimation *	• A symmetric multivariate correction is first operated on ***Z*** (narrow-band amplitude-envelope time series) to obtain ***P*** that is the best approximation of ***Z*** in multivariate linear regression sense, and in ***P***, each source's time series is simultaneously orthogonal to each other thus zero-phase-lag free between all source-pairs. ***P***'s amplitude-envelope is next down-sampled to obtain $\widetilde{\boldsymbol{P}}$	• Any narrow- or broad-band time series • Estimation of the mixing functions: $f_{mix}(v_i,v_j)$ and edge-fidelity $f_e(v_i,v_j)$ is a prerequisite
Procedure of FC estimation	• 2 steps: 1) regularize Ω (the inverse of covariance matrix of $\widetilde{\boldsymbol{P}}$), imposing sparsity (Freidman et al., 2008) on the graph to maximise the log-likelihood of a multivariate Gaussian graph model; 2) Compute partial correlation based on the regularized Ω	• 2 steps: 1) estimate pairwise FC, 2) bundle raw edges obtained in (1) using $f_{mix}(v_i,v_j)$
Advantages	• Effective in removing SIs • Regularized partial correlation reports direct FC between $\widetilde{\boldsymbol{P}}$ and excludes indirect FC • Computational effecitient	• Requires no alteration to source-reconstructed time series • Not limited by the rank of the source time series, which allows a high spatial resolution source estimates • Hyperedge bundling can be used with any directed or undirected interaction metric
Limitations	• Blind to true zero-phase-lag interactions • Likely insensitive to low SNR sources • Limited by the rank of the time series • Risk of reporting false positives if the time series are non-Gaussian	• Blind to true zero-phase-lag interactions dependent on the choice of the interaction metric • The resolution of the hypergraph is a free parameter

Figure captions

Fig 1

Spurious edges are indirect products of mixing and they can be bundled. ***A)*** Top: signal mixing causes the detection of artificial (AI) and spurious interactions (SI). Bottom: AIs are always zero-lag connections (solid gray edge) whereas SIs (dashed gray edges) are "ghosts" of the phase-lag of the true interaction (dashed black edge) and thus can be either zero-lag or, more often, non-zero-lag interactions. ***B)*** Toy model 1: one single true interaction $E(V_1, V_2)$ on a grid of 13 x 13 point sources. Inset shows the simulated mixing neighbourhood of V_1 and V_2. FC was estimated with *iPLV*, and the true edge (black) was discovered with multiple SIs (grey) originating from both sources' mixing neighbourhoods. ***C)*** The similarity in signal mixing between all edges (true and SI) can be quantified and all these edges can be bundled into one hyperedge. ***D)*** Toy model 2: three pairs of true edges of varying spatial distance were simulated. ***E)*** Partitioned similarity matrix S_E, for toy model 2, where each row represents one edge and one cluster represents a hyperedge. The grey box indicate false-positive hyperedges; the magenta and green boxes indicate the inter-hyperedge similarity between the "far" and "nearby" pair. ***F)*** Visualization of the hyperedges defined in ***E***.

Fig 2

Bundling of raw edges into hyperedges. ***A)*** The true interaction E_1 and one of its SIs E_2 from Fig 1B schematically shown in matrix form. ***B)*** The raw graph A_{FC} (a sparse matrix containing only significant edges) is parsed to a list node pairs, each pair representing one edge. ***C)*** For E_1 and E_2, the mixing between all of their constituent nodes can be found in the mixing function f_{mix}. ***D***) The edge adjacency (A_E) between E_1 and E_2 is the maximum product of constituent nodes' mixing. ***E)*** A_E is computed for all the pairs of edges found in A_{FC}. Data taken from a randomly selected simulation. ***F)*** Examples of edges that are similar (blue) and not similar (red) in their mixing profiles. ***G)*** Similarity between two edges is the correlation between two edges' mixing profiles. ***H)*** Mixing similarity matrix $S_{E.}$ ***I)*** The partitioning of this S_E at low, medium and high resolutions.

Fig 3

The demographics of group-level FC of simulated graphs ***A)*** Significant edges were determined with a paired one-tailed t-test between a coupled-edge condition (k1) and the H_0 condition for simulated graphs. ***B)*** For initial evaluation of bundling, we chose one set of gamma-distribution-coupled (Cγ) and one set of uniform-distribution-coupled (Cc) graphs, which are indicated by the markers. *C)* True positive rate TPR (see methods) of the two chosen graphs was above 90%. ***D)*** The true positive rate (TRP) as a function of noise (FP/TP) for all coupling strengths. ***E)*** In the chosen sets of graphs, the number of FP decreases exponentially, while the number of TP decreases linearly. Inset shows the ROC of Cγ edge weights threshold. ***F)*** Noise (FP/TP ratio) as a function of TPR. ***G)*** The mean *iPLV* of TP or FP edges alone, and all edges.

Fig 4

Hyperedge bundling outperformed raw edges. ***A)*** The hypergraphs created with both clustering methods were stable below CL of 0.4. ***B)*** The cumulative distribution function (cdf) of hyperedge size at different levels of CL, computed with hyperedges pooled from 500 graphs with 100 iterations within each graph. For both clustering methods *C)* and *D)*, increasingly strict hyperedge size threshold (θ_{HEsize} varying from 0 to 8) caused separability and noise level (FP/TP) to decrease. ***E,F***) The retained true positive raw edges also decreased as hyperedge size threshold increased.

Fig 5

Louvain clustering method yielded hypergraphs with lower noise but also lower separability than UPGMA. A) For CL values 0.15 – 0.45, Louvain hypergraphs had lower noise after singleton hyperedges were deleted. *B*) True positive rate TPR was larger in Louvain hypergraphs for CL values 0.15 and 0.25 and larger in UPGMA hypergraphs for CL values 0.35 and 0.45. *C*) Separability was higher for UPGMA method.

Fig 6

Hypergraphs improve visualization of real and simulated data. Visual crowding of numerous group-level *iPLV* edges of 1:1 phase synchronization in ***A*)** alpha and ***B*)** gamma frequency band during VWM retention (real MEG data), ***C***) a simulated graph overlaid on a flattened 2D map of cortical regions. ***D, E, F*)** Hypergraphs of ***A,B,C***). ***D*)** In alpha band, bundles of long-range hyperedges connect occipital and parietal areas. Hyperedges were created with CL=0.15, θ_{HEsize}>6. ***E)*** In gamma band, long-range hyperedges were observed in the frontal and central regions. On these 2D map, different parcel colours indicate functional sub-systems defined by (Yeo et al. 2011) and in hypergraphs, edge colours are obtained by mixing of the colours of connected parcels. CN: cuneus; CS: central sulcus; iPGsup: supramarginal gyrus; mFG: middle frontal gyrus; mOG: middle occipital gyrus; mOS: middle occipital sulcus and lunatus sulcus; laSp: posterior ramus; prCG: precental gyrus; pCIm: middle posterior cingulate; prCN: precuneus; sPG:superior parietal lobule; sOG: superior occipital gyrus.

8 Figures

Fig1

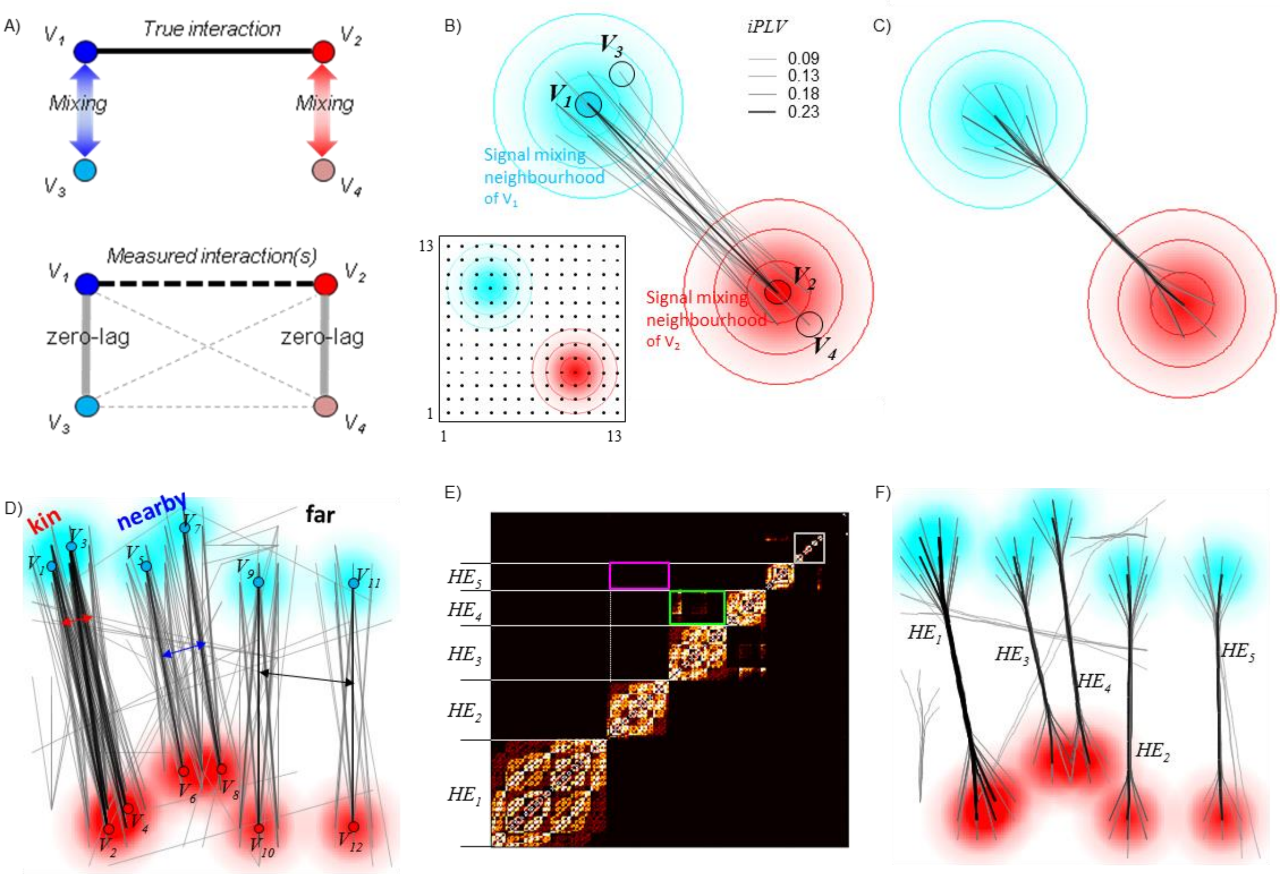
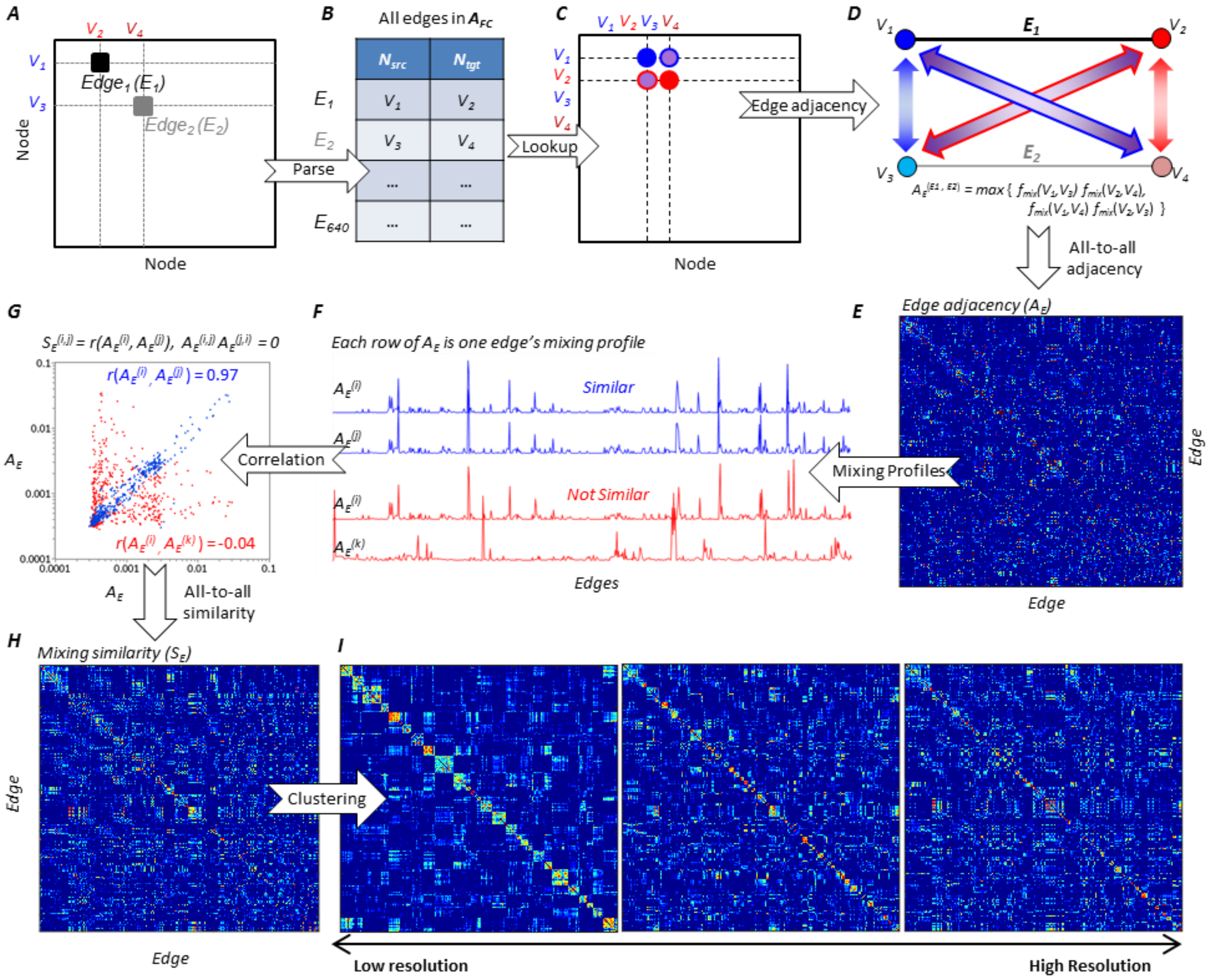
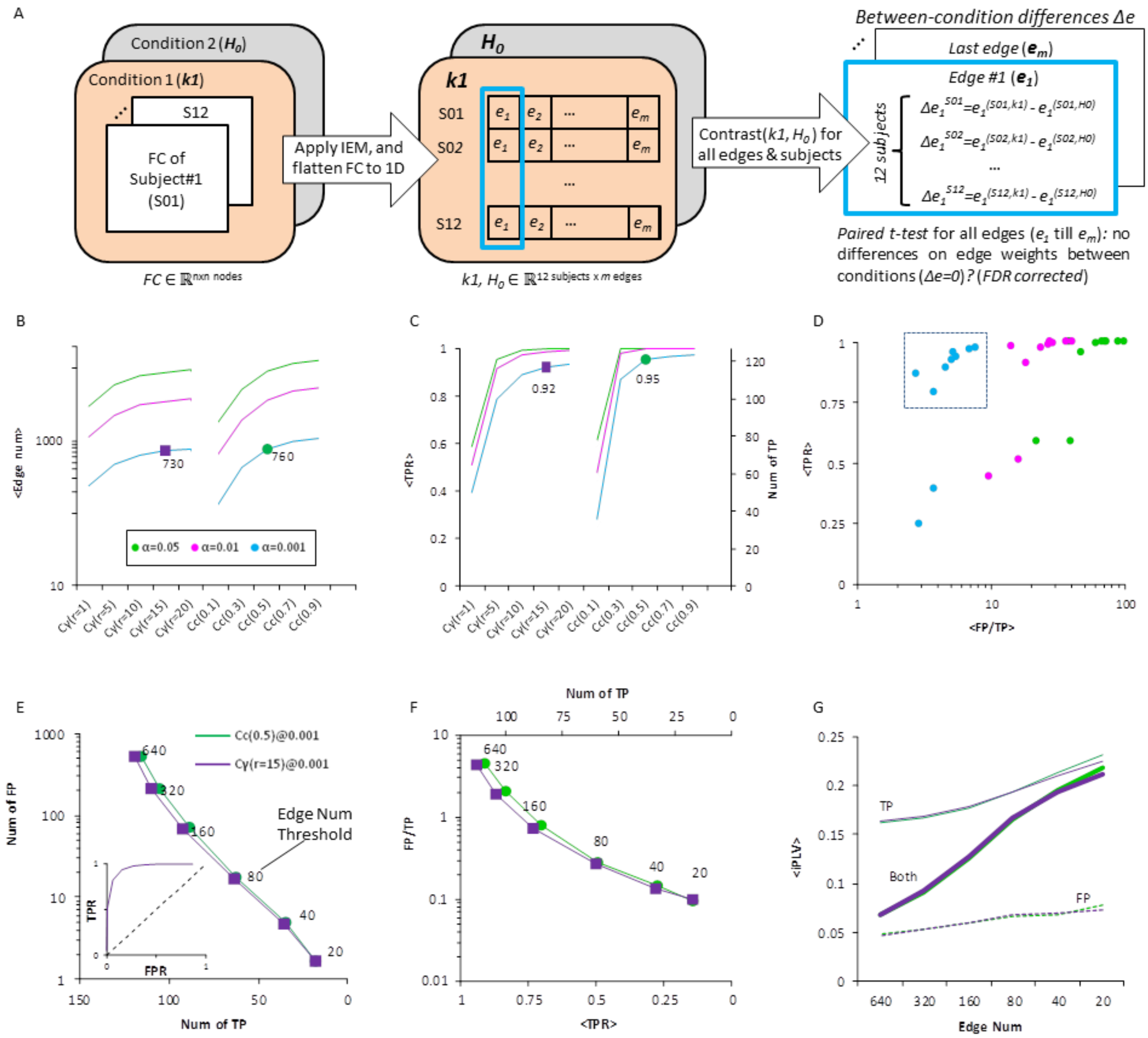


Fig2

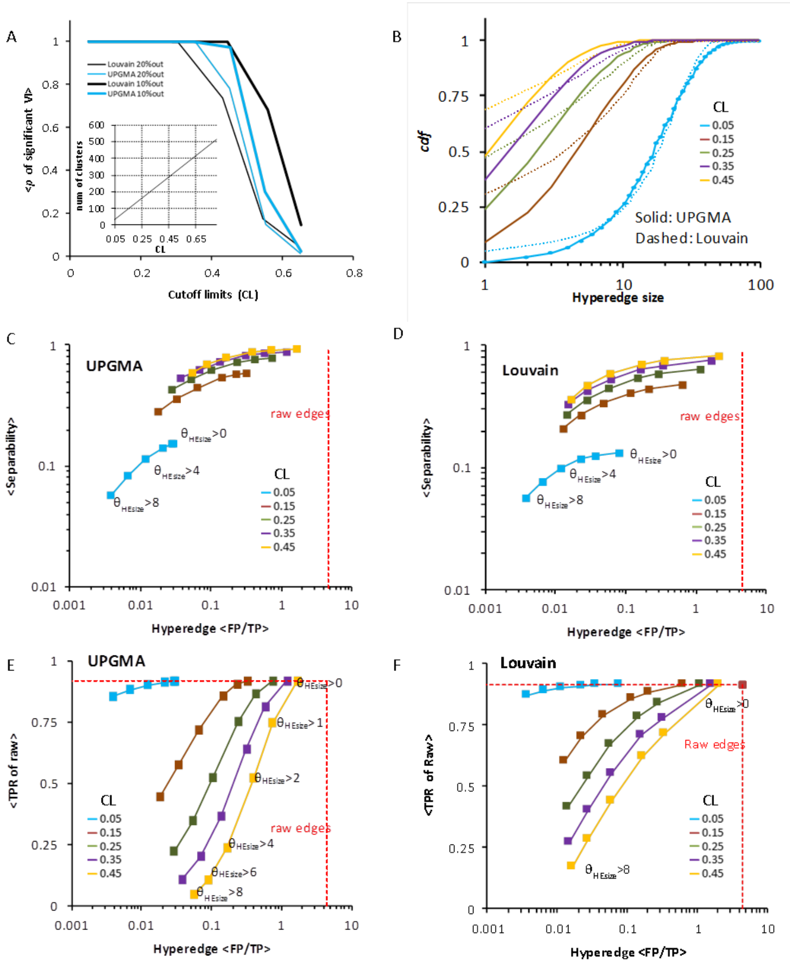

A
B
C
D
E
F
G
H
I
Edge1 (E1)
Edge2 (E2)
Node
Parse
All edges in AFC
Nsrc
Ntgt
E640
Lookup
Edge adjacency
All-to-all adjacency
Edge adjacency (AE)
Edge
Mixing Profiles
Each row of AE is one edge's mixing profile
Similar
Not Similar
Edges
Correlation
r(AE(i), AE(j)) = 0.97
r(AE(i), AE(k)) = -0.04
All-to-all similarity
Mixing similarity (SE)
Clustering
Low resolution
High Resolution

627 Fig3

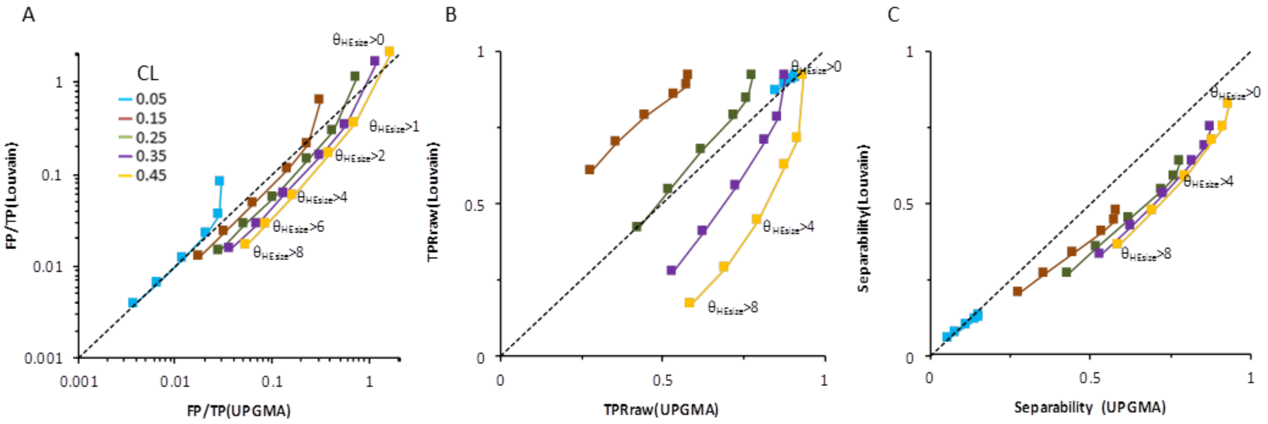

A
Condition 2 (H0)
Condition 1 (k1)
S12
FC of Subject#1 (S01)
Apply IEM, and flatten FC to 1D
FC ∈ ℝ n×n nodes
H0
k1
S01
S02
S12
k1, H0 ∈ ℝ 12 subjects x m edges
Contrast(k1, H0) for all edges & subjects
Between-condition differences Δe
Last edge (em)
Edge #1 (e1)
12 subjects
Paired t-test for all edges (e1 till em): no differences on edge weights between conditions (Δe=0)? (FDR corrected)
B
<Edge num>
730
760
α=0.05
α=0.01
α=0.001
C
<TPR>
0.92
0.95
Num of TP
D
<TPR>
<FP/TP>
E
Num of FP
Num of TP
Cc(0.5)@0.001
Cγ(r=15)@0.001
Edge Num Threshold
TPR
FPR
F
FP/TP
<TPR>
G
<iPLV>
TP
Both
FP
Edge Num

628

Fig4


A
B
C
D
E
F
Louvain 20%out
UPGMA 20%out
Louvain 10%out
UPGMA 10%out
<p of significant V>
num of clusters
Cutoff limits (CL)
cdf
Hyperedge size
Solid: UPGMA
Dashed: Louvain
CL
0.05
0.15
0.25
0.35
0.45
UPGMA
Louvain
raw edges
Raw edges
<Separability>
<TPR of raw>
Hyperedge <FP/TP>

631 Fig 5


A
B
C
CL
0.05
0.15
0.25
0.35
0.45
FP/TP(Louvain)
FP/TP(UPGMA)
TPRraw(Louvain)
TPRraw(UPGMA)
Separability(Louvain)
Separability (UPGMA)
θHEsize>0
θHEsize>1
θHEsize>2
θHEsize>4
θHEsize>6
θHEsize>8

632

References

Banerjee, O., El Ghaoui, L., d'Aspremont, A. (2008). Model selection through sparse maximum likelihood estimation for multivariate gaussian or binary data. *J of Machine Learning Research, 9*, 485-516.

Bastos, A.M. and Schoffelen, J. (2016). A tutorial review of functional connectivity analysis methods and their interpretational pitfalls. *Frontiers in Systems Neuroscience*, *9*, 1662-5137.

Ben-Hur, A., Elisseeff, A., Guyon, I. (2002). A stability based method for discovering structure in clustered data. *Pac Symp Biocomput., 6-17*.

Blondel, V.D., Guillaume, J.L., Lambiotte, R., Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment, 2008*, P10008.

Brookes, M.J., Woolrich, M.W., Barnes, G.R. (2012). Measuring functional connectivity in MEG: A multivariate approach insensitive to linear source leakage. *NeuroImage, 63,* 910-920.

Brookes, M.J., Tewarie, P.K., Hunt, B.A.E., Robson, S.E., Gascoyne, L.E., Liddle, E.B., Liddle, P.F., Morris, P.G. (2016). A multi-layer network approach to MEG connectivity analysis. *NeuroImage, 132*, 425-438.

Brookes, M.J., O'Neill, G.C., Hall, E.L., Woolrich, M.W., Baker, A., Palazzo Corner, S., Robson, S.E., Morris, P.G., Barnes, G.R. (2014). Measuring temporal, spectral and spatial changes in electrophysiological brain network connectivity. *NeuroImage, 91*, 282-299.

Bullmore, E. and Sporns, O. (2012). The economy of brain network organization. *Nature Reviews.Neuroscience, 13*, 336-349.

Bullmore, E. and Sporns, O. (2009). Complex brain networks: Graph theoretical analysis of structural and functional systems. *Nature Reviews.Neuroscience, 10,* 186-198.

Buzsaki, G., Anastassiou, C.A., Koch, C. (2012). The origin of extracellular fields and currents â€” EEG, ECoG, LFP and spikes. *Nature Reviews Neuroscience, 13*, 407-420.

Colclough, G.L., Brookes, M.J., Smith, S.M., Woolrich, M.W. (2015). A symmetric multivariate leakage correction for MEG connectomes. *NeuroImage, 117*, 439-448.

Colclough, G.L., Woolrich, M.W., Tewarie, P.K., Brookes, M.J., Quinn, A.J., Smith, S.M. (2016). How reliable are MEG resting-state connectivity metrics? *NeuroImage, 138*, 284-293.

Drakesmith, M., El-Deredy, W., Welbourne, S. (2013). Reconstructing coherent networks from electroencephalography and magnetoencephalography with reduced contamination from volume conduction or magnetic field spread. *PLoS ONE, 8*, e81553.

Drakesmith, M., Caeyenberghs, K., Dutt, A., Lewis, G., David, A.S., Jones, D.K. (2015). Overcoming the effects of false positives and threshold bias in graph theoretical analyses of neuroimaging data. *NeuroImage, 118*, 313-333.

Everson, R. (1999). Orthogonal, but not orthonormal, procrustes problems. *Imperial College Technical Report, TR99-6*.

Farahibozorg, S.R., Henson, R.N., Hauk, O. Adaptive cortical parcellations for source reconstructed EEG/MEG connectomes. *BioRxiv Doi: 10.1101/123456, .*

Fornito, A., Zalesky, A., Breakspear, M. (2015). The connectomics of brain disorders. *Nature Reviews.Neuroscience, 16*, 159-172.

Friedman, J., Hastie, T., Tibshirani, R. (2008). Sparse inverse covariance estimation with the graphical lasso. *Biostatistics, 9(3)*, 432–441.

Gross, J., Baillet, S., Barnes, G.R., Henson, R.N., Hillebrand, A., Jensen, O., Jerbi, K., Litvak, V., Maess, B., Oostenveld, R., Parkkonen, L., Taylor, J.R., van Wassenhove, V., Wibral, M., Schoffelen, J.M. (2013). Good practice for conducting and reporting MEG research. *NeuroImage, 65*, 349-363.

Hamalainen, M., Hari, R., Ilmoniemi, R.J., Knuutila, J., Lounasmaa, O.V. (1993). Magnetoencephalographytheory, instrumentation, and applications to noninvasive studies of the working human brain. *Reviews of Modern Physics*, *65*, 413-497.

Hamalainen, M.S. and Ilmoniemi, R.J. (1994). Interpreting magnetic fields of the brain: Minimum norm estimates. *Medical & Biological Engineering & Computing, 32*, 35-42.

Hamalainen, M.S. and Sarvas, J. (1989). Realistic conductivity geometry model of the human head for interpretation of neuromagnetic data. *IEEE Transactions on Bio-Medical Engineering, 36*, 165-171.

Hauk, O. and Stenroos, M. (2014). A framework for the design of flexible cross-talk functions for spatial filtering of EEG/MEG data: DeFleCT. *Human Brain Mapping, 35,* 1642-1653.

Hauk, O., Wakeman, D.G., Henson, R. (2011). Comparison of noise-normalized minimum norm estimates for MEG analysis using multiple resolution metrics. *NeuroImage, 54,* 1966-1974.

Haumann, N.T., Parkkonen, L., Kliuchko, M., Vuust, P., Brattico, E. (2016). Comparing the performance of popular MEG/EEG artifact correction methods in an evoked-response study. *Computational Intelligence and Neuroscience*, .

Hillebrand, A., Barnes, G.R., Bosboom, J.L., Berendse, H.W., Stam, C.J. (2012). Frequency-dependent functional connectivity within resting-state networks: An atlas-based MEG beamformer solution. *NeuroImage*, *59*, 3909-3921.

Hincapié, A., Kujala, J., Mattout, J., Pascarella, A., Daligault, S., Delpuech, C., Mery, D., Cosmelli, D., Jerbi, K. (2017). The impact of MEG source reconstruction method on source-space connectivity estimation: A comparison between minimum-norm solution and beamforming. *NeuroImage*, *156*, 29-42.

Hipp, J.F., Hawellek, D.J., Corbetta, M., Siegel, M., Engel, A.K. (2012). Large-scale cortical correlation structure of spontaneous oscillatory activity. *Nature Neuroscience*, *15*, 884-890.

Holten, D. and Wijk, J.J.V. (2009). Force-directed edge bundling for graph visualization. .

Jain, A.K., Murty, M.N., Flynn, P.J. (1999). Data clustering: A review. *ACM Comput.Surv.*, *31*, 264-323.

Karl J., F. (2011). Functional and effective connectivity: A review. *Brain Connectivity*, *1 1*, 13 -36.

Kreuz, T. (2011). Measures of neuronal signal synchrony. *Scholarpedia*, *6*, 11922.

Kujala, J., Gross, J., Salmelin, R. (2008). Localization of correlated network activity at the cortical level with MEG. *NeuroImage*, *39*, 1706-1720.

Lin, F.H., Witzel, T., Ahlfors, S.P., Stufflebeam, S.M., Belliveau, J.W., Hamalainen, M.S. (2006). Assessing and improving the spatial accuracy in MEG source localization by depth-weighted minimum-norm estimates. *NeuroImage*, *31*, 160-171.

Liu, A.K., Dale, A.M., Belliveau, J.W. (2002). Monte carlo simulation studies of EEG and MEG localization accuracy. *Human Brain Mapping*, *16*, 47-62.

Lobier, M., Palva, J.M., Palva, S. (2017). High-alpha band synchronization across frontal, parietal and visual cortex mediates behavioral and neuronal effects of visuospatial attention. *NeuroImage*, *165*, 222-237.

Löwdin, P.O. (1950). On the nonorthogonality problem connected with the use of atomic wave functions in the theory of molecules and crystals. *J. Chem. Phys.*, *18*, 365–375.

Mannino, M. and Bressler, S.L. (2015). Foundational perspectives on causality in large-scale brain networks. *Physics of Life Reviews, 15*, 107-123.

Marrelec, G., Krainik, A., Duffau, H., Pélégrini-Issac, M., Lehéricy, S., Doyon, J., Benali, H. (2006). Partial correlation for functional brain interactivity investigation in functional MRI. *NeuroImage*, *32*, 228-237.

Meilă, M. (2007). Comparing clusterings—an information based distance. *Journal of Multivariate Analysis, 98*, 873-895.

Nolte, G., Bai, O., Wheaton, L., Mari, Z., Vorbach, S., Hallett, M. (2004). Identifying true brain interaction from EEG data using the imaginary part of coherency. *Clinical Neurophysiology : Official Journal of the International Federation of Clinical Neurophysiology, 115*, 2292-2307.

O'Neill, G.C., Barratt, E.L., Hunt, B.A., Tewarie, P.K., Brookes, M.J. (2015). Measuring electrophysiological connectivity by power envelope correlation: A technical review on MEG methods. *Phys Med Biol., 60(21)*, 271-95.

Palva, S. and Palva, J.M. (2012). Discovering oscillatory interaction networks with M/EEG: Challenges and breakthroughs. *Trends in Cognitive Sciences, 16*, 219-230.

Papo, D., Buld\'u Javier M., Boccaletti, S., Bullmore, E.T. (2014). Complex network theory and the brain. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *369*.

Petersen, S. . and Sporns, O. (2015). Brain networks and cognitive architectures. *Neuron, 88*, 207-219.

Preti, M.G., Bolton, T.A., Van De Ville, D. (2016). The dynamic functional connectome: State-of-the-art and perspectives. *NeuroImage, .*

Rubinov, M. and Sporns, O. (2010). Complex network measures of brain connectivity: Uses and interpretations. *NeuroImage, 52*, 1059-1069.

Rubinov, M. (2015). Neural networks in the future of neuroscience research. *Nature Reviews. Neuroscience, advance online publication*, 1471-0048.

S. Baillet, J. C. Mosher, R. M. Leahy. (2001). Electromagnetic brain mapping. *IEEE Signal Processing Magazine*, *18*, 14-30.

Salmelin, R. and Baillet, S. (2009). Electromagnetic brain imaging. *Human Brain Mapping*, *30*, 1753-1757.

Santarnecchi, E., Galli, G., Polizzotto, N.R., Rossi, A., Rossi, S. (2014). Efficiency of weak brain connections support general cognitive functioning. *Human Brain Mapping*, *35*, 4566–4582.

Schoffelen, J.M. and Gross, J. (2009). Source connectivity analysis with MEG and EEG. *Human Brain Mapping*, *30*, 1857-1865.

Siebenhuhner, F., Wang, S.H., Palva, J.M., Palva, S. (2016). Cross-frequency synchronization connects networks of fast and slow oscillations during visual working memory maintenance. *ELife*, *5*, 10.7554/eLife.13451.

Sporns, O., Tononi, G., Kotter, R. (2005). The human connectome: A structural description of the human brain. *PLoS Computational Biology*, *1*, e42.

Sporns, O. (2014). Contributions and challenges for network models in cognitive neuroscience. *Nature Neuroscience*, *17*, 652-660.

Uhlhaas, P.J. and Singer, W. (2010). Abnormal neural oscillations and synchrony in schizophrenia. *Nat Rev Neurosci*, *11*, 100-113.

Uhlhaas, P.J. and Singer, W. (2006). Neural synchrony in brain disorders: Relevance for cognitive dysfunctions and pathophysiology. *Neuron*, *52*, 155-168.

van Wijk, Bernadette C. M., Stam, C.J., Daffertshofer, A. (2010). Comparing brain networks of different size and connectivity density using graph theory. *PLoS ONE*, *5*, e13701.

Wens, V. (2015). Investigating complex networks with inverse models: Analytical aspects of spatial leakage and connectivity estimation. *91*, 17.

Williams, N.J., Nasuto, S.J., Saddy, J.D. (2015). Method for exploratory cluster analysis and visualisation of single-trial ERP ensembles. *Journal of Neuroscience Methods*, *250*, 22-33.

Wollstadt, P., Meyer, U., Wibral, M. (2015). **A graph algorithmic approach to separate direct from indirect neural interactions**. *PLOS One, arXiv:1504.00156*.

Zalesky, A., Fornito, A., Cocchi, L., Gollo, L.L., Breakspear, M. (2014). Time-resolved resting-state brain networks. *Proceedings of the National Academy of Sciences of the United States of America*, *111*, 10341-10346.

Zhigalov, A., Arnulfo, G., Nobili, L., Palva, S., Palva, J.M. (2017). Modular co-organization of functional connectivity and scale-free dynamics in the human brain. *Network Neuroscience, 1*.