# Consistency-based thresholding of the human connectome

James A. Roberts[a,b,*], Alistair Perry[a,c,d], Gloria Roberts[d,e], Philip B. Mitchell[d,e], Michael Breakspear[a,b,f]

[a] Systems Neuroscience Group, QIMR Berghofer Medical Research Institute, Herston, QLD 4006, Australia
[b] Centre for Integrative Brain Function, QIMR Berghofer Medical Research Institute, Herston, QLD 4006, Australia
[c] Centre for Healthy Brain Ageing (CHeBA), School of Psychiatry, University of New South Wales, Sydney, NSW, Australia
[d] School of Psychiatry, University of New South Wales, Sydney, NSW 2052, Australia
[e] Black Dog Institute, Prince of Wales Hospital, Hospital Road, Randwick, NSW 2031, Australia
[f] Metro North Mental Health Service, Royal Brisbane and Women's Hospital, Brisbane, QLD, Australia

## ABSTRACT

Densely seeded probabilistic tractography yields weighted networks that are nearly fully connected, hence containing many spurious fibers. It is thus necessary to prune spurious connections from probabilistically-derived networks to obtain a more reliable overall estimate of the connectivity. A standard method is to threshold by weight, keeping only the strongest edges. Here, by measuring the consistency of edge weights across subjects, we propose a new thresholding method that aims to reduce the rate of false-positives in group-averaged connectivity matrices. Close inspection of the relationship between consistency, weight, and distance suggests that the most consistent edges are in fact those that are strong for their length, rather than simply strong overall. Hence retaining the most consistent edges preserves more long-distance connections than traditional weight-based thresholding, which penalizes long connections for being weak regardless of anatomy. By comparing our thresholded networks to mouse and macaque tracer data, we also show that consistency-based thresholding exhibits the species-invariant exponential decay of connection weights with distance, while weight-based thresholding does not. We also show that consistency-based thresholding can be used to identify highly consistent and highly inconsistent subnetworks across subjects, enabling more nuanced analyses of group-level connectivity than just the mean connectivity.

## Introduction

Tractography is a widely-used method for inferring white matter connectivity from diffusion imaging data, and is central to the field of connectomics (Fornito et al., 2015). Various algorithms use voxel-based estimates of water diffusion to infer the likely paths of white matter bundles, either deterministically or probabilistically (Behrens et al., 2003; Descoteaux et al., 2009). Regardless of the method, there is uncertainty over which connections are "true" connections and which are spurious (de Reus and van den Heuvel, 2013; Girard et al., 2014; Smith et al., 2012). In particular, while deterministic tractography has high specificity at the cost of sensitivity to crossing fibers and hence has a high rate of false negatives, probabilistic algorithms yield inherently noisy connection matrices, at least at the single subject level, and hence likely contain numerous false positives (Thomas et al., 2014).

Pooling data over subjects is a common way to reduce the signal to noise ratio, such as by averaging connectivity matrices across subjects (Hagmann et al., 2008; Perry et al., 2015), or determining a consensus connectivity by selecting edges that appear in at least some fraction of the subjects (de Reus and van den Heuvel, 2013; van den Heuvel and Sporns, 2011). Seeking a consensus in this way is problematic for networks derived from densely-seeded probabilistic tractography, where all individual subject-wise networks are densely connected. The most common method in this setting is to "threshold" networks to some desired density by keeping only the strongest links (Rubinov and Sporns, 2010). This method is applicable to dense networks, but it is not at all clear that the strongest links are always the most accurate for inferring white matter connectivity (Gigandet et al., 2008).

Besides reducing spurious connections, thresholding connectivity matrices also plays an important role in graph-based characterization of connection topology (Bullmore and Bassett, 2011; Van Wijk et al., 2010). Thresholding is used to determine binary adjacency matrices associated with weighted networks, enabling use of the full armory of graph-theoretic tools for unweighted networks (Rubinov and Sporns, 2010). Thresholding can also be used to identify subnetworks composed of the strongest (or weakest) edges, whether for analysis of these

subnetworks or simply for ease of visualization.

In this technical note, we propose a hybrid thresholding method that seeks a group consensus connectivity by thresholding the averaged network to retain only those connections whose weights are the most consistent across the group. Unlike the popular weight-based thresholding, our consistency-based thresholding pays heed to the within-group intersubject variability when deriving a group-averaged matrix. We show that our consistency-based approach avoids the "hard threshold edge" imposed by traditional thresholding and preserves the role of long-range connections. We also study the influence of thresholding strategy on the network topology of the ensuing structural connectome.

## Methods

We derived estimates of whole brain structural connectivity from diffusion images of 75 healthy subjects (aged 17–30 years, 47 females). The structural connectivity matrices were derived in a recent study; we briefly present the methods here but see Roberts et al. (2016) for full details on the acquisition and tractography details.

Diffusion MRI data were acquired from all participants on a Philips 3 T Achieva Quasar Dual MRI scanner (Philips Medical System, Best, The Netherlands) using a single-shot echo-planar imaging (EPI) sequence (TR=7767 ms, TE=68 ms). For each diffusion scan, 32 gradient directions ($b$=1000 s/mm²) and a non-diffusion-weighted acquisition ($b$=0 s/mm²) were acquired over a 96×96 image matrix (field of view 240 mm×240 mm×137.5 mm), with a slice thickness of 2.5 mm and no gap, reconstructed to yield 1 mm×1 mm×2.5 mm voxels (where the longer dimension is along the dorsoventral axis). Two sets of diffusion scans were acquired for each subject.

We employed a probabilistic streamline algorithm (Tournier et al., 2012) to generate high-resolution whole-brain fiber tracks. The fiber orientation distribution (FOD) within each voxel was estimated using MRtrix software (Tournier et al., 2012) by performing constrained spherical deconvolution (Tournier et al., 2007) with a maximum spherical harmonic order (*lmax*) of 6. As an intermediate step to constrain the spherical deconvolution, a single-fiber response kernel was estimated from all white matter voxels with fractional anisotropy FA > 0.7. Streamlines were seeded using the skull-stripped brain mask together with a restriction to voxels with FOD amplitude > 0.1. Streamlines will not start outside this region and terminate if they reach the boundary. Tractograms were generated using a probabilistic streamlines algorithm (Tournier et al., 2012), which produces a set of connection trajectories by randomly sampling from the orientation uncertainty inherent in each FOD along the streamline paths. Although non-isotropic voxels were used within the analysis, we subsequently checked the resultant fiber orientations and tractograms, finding no issue with potential biases on quality. To confirm this we repeated our analysis in an independent dataset acquired with isotropic voxels (Supplementary information 1).

Our connectivity matrices were reconstructed from densely seeded tractography ($10^8$ seeds) and parcellated into a relatively fine representation of 513 uniformly sized cortical and sub-cortical regions (Zalesky et al., 2010). The resulting weighted, undirected matrices were nearly fully connected in each subject. The weights are the number of streamlines linking each pair of regions. The spatial connection distance between all nodes was obtained using a streamline-based quantification of distance, in addition to the more traditionally-used Euclidean distance. In six subjects, for each pathway in the connectome, the shortest streamline between the node pair was found, and its length determined; these minimum lengths were then averaged across the six subjects. The minimum streamline length per subject was used here as streamlines are more likely to erroneously continue beyond the length of the connecting pathway (and hence overestimate the actual connection length) rather than provide an erroneously short pathway. In previous work (Roberts et al., 2016) it was

observed that the length distribution converged after averaging over a small number of subject streamline lengths. Indeed, the precise details of the length distribution do not influence our results, as demonstrated in an independent dataset where we used every individual's own set of streamlines (Supplementary information 1).

This combination of streamline generation and anatomical parcellation yields a weighted structural connectivity graph within each subject, which we denote $W$, and a corresponding matrix of streamline lengths, $F$. Within each $W$, a weighted connection $w_{ij}$ represents the number of streamlines from region $i$ terminating within a 2 mm radius of region $j$, with corresponding streamline length $f_{ij}$. The 2 mm radius ensures that fiber terminations near the gray-matter boundary, where the diffusion signal becomes noisier and weaker, are adequately captured. While this could theoretically lead to streamlines being counted twice, our 513 node parcellation is sufficiently coarse that this occurs very infrequently. To improve signal to noise ratio, the corresponding $w_{ij}$ were summed across each subject's two diffusion scans. The non-directional nature of tractography implies that $W$ is symmetric – that is $w_{ij}=w_{ji}$. The larger number of likely random seeds located along longer fiber bundles is well known to result in over-defined fiber densities (Smith et al., 2013). To reduce this confounding effect, $w_{ij}$ were adjusted by dividing the raw count by the streamline length $f_{ij}$ between nodes $i$ and $j$, $w_{ij} \rightarrow w_{ij}/f_{ij}$ (cf. Hagmann et al., 2008). That is, because we seed the white matter uniformly, a tract that is twice as long will have received twice as many seeds. Its weight (streamline count) will thus have been biased relative to a shorter tract with the same true fiber density.

We note that the very dense seeding of our probabilistic tractography yields connectivity matrices that are fully connected (or very nearly fully connected) in all subjects. Whilst the biological connectome at this level of resolution is likely not fully connected, estimates of connection density in the field range very broadly, from < 5% (Hagmann et al., 2008) to 13–36% for the entire brain and 32–52% for cortico-cortical connections in the mouse (Oh et al., 2014). Seeding probabilistic tractography densely, then setting a post-hoc threshold, allows investigation of topology over a range of connection densities, and not that dictated by the acquisition and reconstruction technique.

We estimated the consistency of every edge weight by measuring the coefficient of variation across subjects. We then compared networks thresholded by weight to networks thresholded by consistency. To quantify differences in network topology, we calculated graph metrics (clustering, rich club, modularity) using the Brain Connectivity Toolbox (Rubinov and Sporns, 2010).

## Results

We begin by characterizing the consistency of edge weights across subjects. We quantify consistency by calculating the coefficient of variation (CV) of the weights (SD/mean) across subjects. A low CV corresponds to a high consistency. Edge CV broadly decreases with increasing edge weight (Fig. 1), thus showing that consistency increases with edge weight. This is the widely-assumed justification for weight-based thresholding; here we verify that the very strongest edges are indeed the most consistent. However, this inverse relationship between CV and weight does not follow a simple linear trend. Rather, edges with weights two orders of magnitude below the strongest are almost as consistent.

This slow fall-off of consistency with weight is particularly clear when taking the spatial dimension into account. Weights decrease roughly exponentially with streamline length as shown previously (Roberts et al., 2016). However, CV increases more slowly with distance, such that some of the longest connections are in fact the most consistent (Fig. 2A). In a sense this is unsurprising: true long connections would be expected to have consistent weights between subjects. Grouping the edges by quartiles in consistency (Fig. 2A–D) shows that each quartile's cloud of points spans almost the full range of