BACHELOR'S THESIS

Szabó Zóra

2026

Budapest University of Technology and Economics

Faculty of Natural Sciences

Institute of Mathematics

Analysis of Forward and Backward Pathways in Strongly Recurrent, Directed, and Weighted Cortical Networks

BACHELOR'S THESIS IN MATHEMATICS

Szabó Zóra

External supervisor: László Négyessy, Ph.D.

Internal supervisor: József Pintér

Budapest, 2026

Abstract

Cortical hierarchy is thought to operate through a counterstream of forward and backward projections, but its directed network organisation remains incompletely understood. This study compared anatomical and topology-based forward–backward organisation in a macaque cortical network containing 29 areas and 536 directed projections weighted by fraction of projection neurons. Anatomical subnetworks were defined by the laminar distribution of projection neurons (SLN), whereas topology-based subnetworks were defined by weighted convergence degree (CDw). Subnetworks were examined using extremal arborescences, strongly connected components, Weighted Directed Matrix–Tree Theorem (WDMT) and representational similarity analysis-based permutation tests. Both SLN subnetworks contained global spanning arborescences in both orientations, while each CDw subnetwork lacked a global spanning orientation. Root-dependent reachability varied between subnetworks, which were composed of highly connected components. Significant matched-direction correspondence was detected with the full feature configuration, but not with WDMT-based features alone. Overall, the results suggest that forward and backward counterstreams are based on distinct arborescence properties.

Contents

List of Abbreviations

Abbreviation	**Meaning**
CD	Convergence degree
CDw	Weighted convergence degree
FLN	Fraction of labelled neurons
MDS	Multidimensional scaling
RDM	Representational dissimilarity matrix
RSA	Representational similarity analysis
SCC	Strongly connected component
SLN	Proportion of supragranular labelled neurons
SMACOF	Scaling by MAjorizing a COmplicated Function
WDMT	Weighted Directed Matrix–Tree Theorem

List of Mathematical Notation

Symbol	Description
$G = (V, E, w)$	Full directed weighted anatomical network
V	Set of cortical areas
E	Set of directed anatomical projections
w	Anatomical edge-weight function, with $w(i, j) = \mathrm{FLN}(i, j)$
$i \to j$	Projection from source area i to target area j
$e_{ij} = (i, j)$	Directed edge from area i to area j
$N(i, j)$	Labelled-neuron count for projection $i \to j$
$N_{\mathrm{tot}}(j)$	Total labelled-neuron count following injection into j
$\mathrm{FLN}(i, j)$	Relative anatomical strength of projection $i \to j$
$\mathrm{dist}(i, j)$	Physical distance of projection $i \to j$
$N_{\mathrm{sup}}(i, j)$	Supragranular labelled-neuron count for projection $i \to j$
$N_{\mathrm{infra}}(i, j)$	Infragranular labelled-neuron count for projection $i \to j$
$\mathrm{SLN}(i, j)$	Supragranular labelled-neuron proportion for projection $i \to j$
$c(i, j)$	Weighted path cost of edge (i, j) used for CDw construction
κ	Path-cost exponent, fixed at 0.07
P	Directed path
$L_w(P)$	Weighted length of path P
k	Maximum number of retained weighted shortest paths per reachable ordered pair, fixed at 8
$SP_w(G)$	Selected weighted shortest paths in G
$SP_w(G, e_{ij})$	Selected weighted shortest paths in G traversing edge e_{ij}
$\mathrm{In}_{w,\mathrm{full}}(i, j)$	Input set of edge (i, j) in the full weighted network
$\mathrm{Out}_{w,\mathrm{full}}(i, j)$	Output set of edge (i, j) in the full weighted network
$\mathrm{Int}_{w,\mathrm{full}}(i, j)$	Intersection of the weighted input and output sets of edge (i, j)
$\mathrm{SIn}_{w,\mathrm{full}}(i, j)$	Strict weighted input set of edge (i, j) in the full network
$\mathrm{SOut}_{w,\mathrm{full}}(i, j)$	Strict weighted output set of edge (i, j) in the full network

Symbol	Description
$\mathrm{CDw}_{\mathrm{full}}(i,j)$	Weighted convergence degree of edge (i,j) computed in the full network G
$\mathrm{CDw}_X(i,j)$	Weighted convergence degree of edge (i,j) recomputed within the fixed network G_X for node-level feature construction
$E^{\mathrm{SLN}}_{\mathrm{forward}}$	Anatomical forward edge set, consisting of edges with $\mathrm{SLN}(i,j) \geq 0.5$
$E^{\mathrm{SLN}}_{\mathrm{backward}}$	Anatomical backward edge set, consisting of edges with $\mathrm{SLN}(i,j) < 0.5$
$G^{\mathrm{SLN}}_{\mathrm{forward}}$	Anatomical forward subnetwork
$G^{\mathrm{SLN}}_{\mathrm{backward}}$	Anatomical backward subnetwork
$E^{\mathrm{CDw}}_{\mathrm{forward}}$	Topological forward edge set, consisting of edges with $\mathrm{CDw}_{\mathrm{full}}(i,j) < 0$
$E^{\mathrm{CDw}}_{\mathrm{backward}}$	Topological backward edge set, consisting of edges with $\mathrm{CDw}_{\mathrm{full}}(i,j) > 0$
E^{CDw}_{0}	Neutral edge set, consisting of edges with $\mathrm{CDw}_{\mathrm{full}}(i,j) = 0$
$G^{\mathrm{CDw}}_{\mathrm{forward}}$	Topological forward subnetwork
$G^{\mathrm{CDw}}_{\mathrm{backward}}$	Topological backward subnetwork
$\dot{\cup}$	Disjoint union
$\mathcal{X}$	Set of the five analysed networks: the full network and the four edge-restricted subnetworks
$\mathcal{X}_{\mathrm{sub}}$	Set of the four edge-restricted subnetworks included in the MDS and RSA/permutation analyses
G_X	Network indexed by $X \in \mathcal{X}$
E_X	Edge set of network G_X
G_{full}	Full anatomical network, equal to G
E_{full}	Full anatomical edge set, equal to E
H	Weighted directed subgraph of G
$W(H)$	Total anatomical edge weight of weighted directed subgraph H
$u \rightsquigarrow_X v$	Existence of a directed path from u to v in G_X
r; $r(\mathbf{u},\mathbf{v})$	Root of an arborescence; Pearson correlation coefficient computed across the coordinates of feature vectors $\mathbf{u}$ and $\mathbf{v}$, respectively
$R^{\mathrm{out}}_X(r)$	Vertices reachable from root r in G_X
$R^{\mathrm{in}}_X(r)$	Vertices from which root r is reachable in G_X

Symbol	Description
η	Arborescence orientation, out or in
a	Arborescence optimisation criterion, min or max
$\mathcal{T}_X^\eta$	Set of global spanning η-arborescences in G_X
$\mathcal{T}_{X,r}^\eta$	Set of node-centric η-arborescences rooted at r in G_X
$T_{X,a}^\eta$	Optimised global spanning η-arborescence in G_X under criterion a
$T_{X,r,a}^\eta$	Optimised node-centric η-arborescence rooted at r in G_X under criterion a
$\delta_T(u,v)$	Directed tree-edge distance from vertex u to vertex v in tree T
$\lambda_T(v)$	Level of vertex v in rooted tree T
$L_k(T)$	Set of vertices at tree level k
$D(T)$	Depth of rooted tree T
$\overline{b}(T)$	Mean non-root level width of rooted tree T
$\sigma_b(T)$	Standard deviation of non-root level widths of rooted tree T
$C_T(k)$	Cumulative proportion of vertices in rooted tree T located within at most k tree levels of the root
$\mathcal{R}$	Set of selected roots
m	Number of selected roots
$q_X^\eta(r)$	Order of a node-centric η-arborescence rooted at r in G_X
$\rho_X^\eta(r)$	Relative order of a node-centric η-arborescence rooted at r in G_X
$\overline{q}_X^\eta$	Mean node-centric tree order across selected roots in G_X
$\sigma_{q,X}^\eta$	Standard deviation of node-centric tree order across selected roots in G_X
$\overline{\rho}_X^\eta$	Mean relative node-centric tree order across selected roots in G_X
C	Eligible strongly connected component
$\mathcal{C}_{\text{elig}}(X)$	Set of eligible strongly connected components in G_X
$G_X[C]$	Subgraph of G_X induced by eligible component C
$W_X(C)$	Internal anatomical edge weight of eligible component C in G_X
$\pi_{X,C}$	Normalised internal weight of eligible component C in G_X

Symbol	Description
$A_{X,C}$	Unweighted adjacency matrix of component C in G_X
$A^{w}_{X,C}$	FLN-weighted adjacency matrix of component C in G_X
$D_{\mathrm{in},X,C}$	Unweighted indegree matrix of component C in G_X
$D_{\mathrm{out},X,C}$	Unweighted outdegree matrix of component C in G_X
$D^{w}_{\mathrm{in},X,C}$	FLN-weighted indegree matrix of component C in G_X
$D^{w}_{\mathrm{out},X,C}$	FLN-weighted outdegree matrix of component C in G_X
$L_{1,X,C}$	Unweighted outgoing-arborescence Laplacian of component C in G_X
$L_{2,X,C}$	Unweighted incoming-arborescence Laplacian of component C in G_X
$L^{w}_{1,X,C}$	FLN-weighted outgoing-arborescence Laplacian of component C in G_X
$L^{w}_{2,X,C}$	FLN-weighted incoming-arborescence Laplacian of component C in G_X
$M^{(i)}$	Matrix M with the row and column corresponding to root i removed
$\tau^{\mathrm{out}}_{X,i,C}$	Unweighted outgoing rooted-arborescence count for node i in eligible component C of G_X
$\tau^{\mathrm{in}}_{X,i,C}$	Unweighted incoming rooted-arborescence count for node i in eligible component C of G_X
$\tau^{\mathrm{out},w}_{X,i,C}$	Weighted outgoing rooted-arborescence quantity for node i in eligible component C of G_X
$\tau^{\mathrm{in},w}_{X,i,C}$	Weighted incoming rooted-arborescence quantity for node i in eligible component C of G_X
$\overline{\tau}^{\mathrm{out},w}_{X,i,C}$	Mean weighted product per outgoing rooted arborescence for node i in eligible component C of G_X
$\overline{\tau}^{\mathrm{in},w}_{X,i,C}$	Mean weighted product per incoming rooted arborescence for node i in eligible component C of G_X
$\tau^{\mathrm{out}}_{X,i,\mathrm{sc}}$	SCC-scaled unweighted outgoing rooted-arborescence count for node i in G_X
$\tau^{\mathrm{in}}_{X,i,\mathrm{sc}}$	SCC-scaled unweighted incoming rooted-arborescence count for node i in G_X
$\tau^{\mathrm{out},w}_{X,i,\mathrm{sc}}$	SCC-scaled weighted outgoing rooted-arborescence quantity for node i in G_X

Symbol	Description
$\tau^{\mathrm{in},w}_{X,i,\mathrm{sc}}$	SCC-scaled weighted incoming rooted-arborescence quantity for node i in G_X
$\overline{\tau}^{\mathrm{out},w}_{X,i,\mathrm{sc}}$	SCC-scaled mean weighted product per outgoing rooted arborescence for node i in G_X
$\overline{\tau}^{\mathrm{in},w}_{X,i,\mathrm{sc}}$	SCC-scaled mean weighted product per incoming rooted arborescence for node i in G_X
$N^{\mathrm{out}}_{X,\mathrm{CDw}}(i)$	Outgoing neighbours of node i with defined recomputed CDw_X values in G_X
$N^{\mathrm{in}}_{X,\mathrm{CDw}}(i)$	Incoming neighbours of node i with defined recomputed CDw_X values in G_X
$\Phi_X(i)$	Network-specific normalised CDw-flow feature of node i in G_X
σ^X_{st}	Number of directed unweighted shortest paths from s to t in G_X
$\sigma^X_{st}(i)$	Number of directed unweighted shortest paths from s to t in G_X that pass through node i
$B_X(i)$	Directed unweighted endpoint-excluding node betweenness of node i in G_X
$N^{\mathrm{in}}_X(i)$	Incoming-neighbour set of node i in G_X
$N^{\mathrm{out}}_X(i)$	Outgoing-neighbour set of node i in G_X
$N^{\mathrm{all}}_X(i)$	Union of incoming and outgoing neighbours of node i in G_X
$d^{\mathrm{in}}_X(i)$	Unweighted indegree of node i in G_X
$d^{\mathrm{out}}_X(i)$	Unweighted outdegree of node i in G_X
$d^{\mathrm{tot}}_X(i)$	Total unweighted directed degree of node i in G_X
$\ell_X(N)$	Lobby-index functional evaluated on neighbour set N in G_X
$\ell^{\mathrm{in}}_X(i)$	Incoming lobby-index feature of node i in G_X
$\ell^{\mathrm{out}}_X(i)$	Outgoing lobby-index feature of node i in G_X
$\ell^{\mathrm{all}}_X(i)$	Direction-combined lobby-index feature of node i in G_X
$h^{\mathrm{SLN}}(i)$	Fixed full-network ordinal rank of cortical area i in the SLN-based anatomical hierarchy
$h^{\mathrm{CDw}}(i)$	Fixed full-network ordinal rank of cortical area i in ascending full-network CDw-flow order
$I^{\tau}_X(i)$	Indicator of availability of the six SCC-based WDMT features for node i in G_X

Symbol	Description
f	Numeric feature index used in standardisation and feature-vector construction
$x_{X,i,f}$	Numeric value of feature f for node i in network G_X before standardisation
$\mu_{X,f}$	Within-network mean of numeric feature f in G_X before standardisation
$\sigma_{X,f}$	Within-network population standard deviation of numeric feature f in G_X before standardisation
$z_{X,i,f}$	Standardised value of feature f for node i in network G_X
F	Selected feature configuration, $F \in \{F_{\text{full}}, F_{\text{red}}\}$
F_{red}	Reduced feature configuration containing the six log-transformed SCC-scaled WDMT-based rooted-arborescence features
F_{full}	Full feature configuration containing F_{red} together with the retained network-topological and hierarchy-related features
$\mathbf{z}_{X,i}^{(F)}$	Standardised feature vector of cortical area i in subnetwork G_X under feature configuration F
$\mathcal{O}$	Pooled set of node–subnetwork observations used for MDS, $\mathcal{O} = \{(X,i) : X \in \mathcal{X}_{\text{sub}},\ i \in V\}$
$\delta_{ab}^{\text{corr}}$	Correlation dissimilarity between pooled observations a and b
δ_{ab}^{Euc}	Euclidean dissimilarity between pooled observations a and b
δ_{ab}^{Man}	Manhattan dissimilarity between pooled observations a and b
δ_{ab}^{Mah}	Mahalanobis dissimilarity between pooled observations a and b
Σ_F^+	Moore–Penrose pseudoinverse of the pooled covariance matrix for feature configuration F
Δ	Pooled dissimilarity matrix of node–subnetwork observations for a selected feature configuration and dissimilarity definition
$\overline{\delta}_+$	Mean positive off-diagonal dissimilarity used to rescale Δ before MDS

Symbol	Description
$\widetilde{\Delta}$	Normalised pooled dissimilarity matrix used as input to MDS
Y	Two-dimensional MDS embedding of the pooled node–subnetwork observations
$\mathbf{y}_a$	Two-dimensional MDS coordinate vector of pooled observation a
$\sigma(Y)$	Raw stress objective minimised in the metric MDS embedding
$D_X^{(F)}$	Representational dissimilarity matrix of subnetwork G_X under feature configuration F
$\mathrm{ut}(D)$	Vector of strict upper-triangular entries of a symmetric matrix D in the fixed cortical-area order
$d_F(G_X, G_Y)$	Manhattan distance between the vectorised RDMs of subnetworks G_X and G_Y under feature configuration F
$d_{\mathrm{SLN}}^{(F)}$	RDM distance between the $\mathrm{SLN}_{\mathrm{forward}}$ and $\mathrm{SLN}_{\mathrm{backward}}$ subnetworks under feature configuration F
$d_{\mathrm{CDw}}^{(F)}$	RDM distance between the $\mathrm{CDw}_{\mathrm{forward}}$ and $\mathrm{CDw}_{\mathrm{backward}}$ subnetworks under feature configuration F
$d_{\mathrm{within}}^{(F)}$	Mean within-classification forward–backward RDM distance under feature configuration F
$d_{\mathrm{forward}}^{(F)}$	RDM distance between the $\mathrm{SLN}_{\mathrm{forward}}$ and $\mathrm{CDw}_{\mathrm{forward}}$ subnetworks under feature configuration F
$d_{\mathrm{backward}}^{(F)}$	RDM distance between the $\mathrm{SLN}_{\mathrm{backward}}$ and $\mathrm{CDw}_{\mathrm{backward}}$ subnetworks under feature configuration F
$d_{\mathrm{matched}}^{(F)}$	Mean matched-direction SLN–CDw RDM distance under feature configuration F
$S_{\mathrm{forward}}^{(F)}$	Forward matched-direction correspondence statistic under feature configuration F
$S_{\mathrm{backward}}^{(F)}$	Backward matched-direction correspondence statistic under feature configuration F
$S_{\mathrm{combined}}^{(F)}$	Combined matched-direction correspondence statistic under feature configuration F
P_b	Permutation matrix for the bth sampled cortical-area label permutation

Symbol	Description
$S_{\text{obs}}^{(F)}$	Observed value of a selected correspondence statistic under feature configuration F
$S_b^{(F)}$	Value of a selected correspondence statistic under the bth permutation and feature configuration F
B	Number of sampled non-identity cortical-area permutations in a Monte Carlo permutation test
$\mathbf{1}\{\cdot\}$	Indicator function
$p_{\text{perm}}^{(F,S)}$	One-sided Monte Carlo permutation p-value for statistic S under feature configuration F
α	Significance level used for the permutation tests

1 Introduction

1.1 Cortical hierarchy, forward–backward pathways

1.1.1 Anatomical organisation of cortical signal flow

The present study is situated at the interface of mathematics—specifically graph theory, linear algebra, and statistics—and systems neuroscience. Mathematical methods were applied to the anatomical connectivity of the primate cerebral cortex, represented as a directed and weighted network in which cortical areas formed the vertices and directed interareal projections formed the edges. The investigation was based on an experimentally quantified subnetwork of 29 macaque cortical areas and 536 directed anatomical projections, for which quantitative retrograde tracer data provided projection directions, laminar projection characteristics and relative anatomical projection strengths [7, 4].

In the visual system, cortical areas have been described as hierarchically organised processing regions. Relatively low-level areas are associated with the representation of elementary sensory features, whereas higher-order areas participate in increasingly integrative processing. Felleman and Van Essen formalised this organisation through the laminar patterns of interareal projections: feedforward projections are directed predominantly from lower-order to higher-order areas, whereas feedback projections are directed predominantly in the reverse hierarchical direction [11]. This hierarchical organisation does not imply an acyclic cortical graph. Rather, cortical areas are connected by a dense recurrent system of directed projections in both hierarchical directions [12, 7].

Markov et al. (2014) [7] quantified the feedforward–feedback character of macaque cortical projections by the proportion of supragranular labelled neurons, denoted by SLN. In the present study, projections with at least as many supragranular as infragranular labelled parent neurons were classified as anatomically forward-like, whereas projections with fewer supragranular than infragranular labelled parent neurons were classified as anatomically backward-like. The same tract-tracing data provided the relative anatomical strength of a projection from a source area to a target area, denoted by FLN (fraction of labelled neurons), which is calculated as the number of labelled neurons in the source area projecting to that target area, divided by the total number of labelled neurons projecting to the same target area [7].

1.1.2 Topology-based organisation of cortical signal flow

Whereas SLN describes the forward–backward character of cortical projections from their laminar anatomical origin, convergence degree provides a complementary topology-based description of directed signal-flow organisation. Originally introduced for directed cortical networks by Négyessy et al. (2008) [3] and formalised by Bányai et al. (2011) [1], convergence degree (CD) is an edge-based measure derived from the directed shortest paths

traversing a projection. An edge is characterised as convergent when paths reaching it originate from a larger set of areas than the set reached after traversing it, and as divergent in the opposite case. Thus, positive CD values represent convergent or condensing edges, whereas negative values represent divergent or spreading edges.

Varga et al. (2021) [4] extended the edge-based CD framework to a directed and weighted macaque cortical network by defining weighted convergence degree, denoted by CDw. In this weighted formulation, the selected paths incorporate both anatomical projection strength, quantified by FLN, and physical distance. Positive CDw values identify topology-based convergent edges, whereas negative values identify topology-based divergent edges. In the present study, divergent edges were treated as topology-based forward-like projections and convergent edges as topology-based backward-like projections.

Thus, SLN provided the anatomical forward–backward edge classification, whereas CDw provided the topology-based forward–backward edge classification.

1.1.3 Anatomical and topology-based forward–backward subnetworks

The anatomical forward–backward organisation described by SLN provided one decomposition of the directed cortical network. In this decomposition, projections classified as anatomically forward-like formed the $\mathrm{SLN}_{\mathrm{forward}}$ subnetwork, whereas projections classified as anatomically backward-like formed the $\mathrm{SLN}_{\mathrm{backward}}$ subnetwork.

A second decomposition was defined by CDw. Projections classified as topology-based forward-like formed the $\mathrm{CDw}_{\mathrm{forward}}$ subnetwork, whereas projections classified as topology-based backward-like formed the $\mathrm{CDw}_{\mathrm{backward}}$ subnetwork. Projections with neutral CDw values were excluded from both topology-based subnetworks. In the resulting anatomically and topologically defined forward and backward subnetworks, retained projections preserved their FLN values as edge weights.

In the present study we analyzed four directed and weighted subnetworks derived from the same anatomical network: $\mathrm{SLN}_{\mathrm{forward}}$, $\mathrm{SLN}_{\mathrm{backward}}$, $\mathrm{CDw}_{\mathrm{forward}}$, and $\mathrm{CDw}_{\mathrm{backward}}$. The two classifications were not assumed to identify identical sets of projections. Instead, the present study investigated whether the anatomically defined and topology-based subnetworks exhibited corresponding directed network organisations in matched forward and backward classes.

1.2 Directed arborescences as representations of directed network organisation

The four forward–backward subnetworks were analysed using rooted directed arborescences. An out-arborescence rooted at a cortical area r was a directed acyclic tree in which each included area was reachable from r by a unique directed path. Conversely, an in-arborescence rooted at r was a directed acyclic tree in which r was reachable from

each included area by a unique directed path. Thus, the root was considered a directed source in an out-arborescence and a directed sink in an in-arborescence.

Arborescences were suitable for the present analysis because they provided cycle-free directed structures with a specified root and orientation within subnetworks derived from a recurrent cortical network. They were not interpreted as actual routes followed by neuronal activity; rather, they characterised directed organisations permitted by the edge sets of the anatomically or topologically defined subnetworks.

Two forms of arborescence analysis were used. Global spanning arborescences tested whether each subnetwork permitted directed reachability across all 29 cortical areas in either orientation and characterised the structure of such network-spanning organisations. In the node-centric analysis, the root area was prescribed, and the arborescence vertex set was determined by directed reachability before optimisation: an out-arborescence included the areas reachable from the root, whereas an in-arborescence included the areas from which the root was reachable. These rooted arborescences were used to characterise how the directed organisation associated with individual cortical areas varied across the four subnetworks.

For each admissible global or node-centric optimisation problem, the minimum- and maximum-weight arborescences were those that minimised and maximised, respectively, the sum of the FLN weights of their constituent edges. Since FLN quantified relative anatomical projection strength [7], these extremal trees identified admissible directed organisations with low or high total FLN weight. Because the same anatomical weighting scheme was used in all four subnetworks, differences between the resulting trees reflected the effects of the two forward–backward edge-selection criteria. These trees did not imply that cortical signal transmission was restricted to any single selected arborescence.

The Weighted Directed Matrix–Tree Theorem provided a complementary root-specific characterisation of arborescence organisation [2]. In the present analysis, WDMT-based quantities were computed within eligible strongly connected components, in which rooted spanning in- and out-arborescences existed for every included cortical area. The unweighted τ-quantities counted rooted spanning in- and out-arborescences for each cortical area. The weighted τ-quantities instead summed the products of FLN edge weights over all rooted spanning arborescences of the corresponding orientation. Mean weighted values per rooted spanning arborescence were obtained by dividing each weighted τ-quantity by its corresponding unweighted count. Thus, the extremal-tree analysis optimised sums of FLN edge weights for individual arborescences, whereas the WDMT-based quantities characterised each eligible root area by its rooted spanning arborescence count, aggregate FLN-weighted value, and mean FLN-weighted value in each orientation.

The WDMT-based quantities provided rooted-arborescence variables for the node-level feature-vector comparison of the SLN- and CDw-defined forward and backward subnetworks, while the extremal-tree and reachability analyses were interpreted separately as

structural summaries.

1.3 Comparison of subnetwork-specific node organisations

Beyond the arborescence-based characterisation of the four subnetworks, node-level feature vectors were constructed to compare the organisation of individual cortical areas more generally.

Because convergence degree was defined for individual edges, a node-centric interpretation was also introduced to characterise the signal-flow organisation associated with cortical areas. Négyessy et al. (2008) [3] represented each area according to the convergence–divergence properties of its incoming and outgoing projections, while Bányai et al. (2011) [1] subsequently formalised this approach within a general flow-representation framework for directed networks. In this interpretation, areas with convergent incoming and divergent outgoing projections exhibited a source-like organisation, spreading directed paths towards a larger set of target areas. Conversely, areas with divergent incoming and convergent outgoing projections exhibited a more allocating organisation, collecting directed paths from larger sets of source areas and directing them towards fewer targets. Varga et al. (2021) [4] extended this node-centric framework to weighted convergence degree in the macaque cortical network. In the present study, the resulting weighted node-centric CD-flow measure was included as a topology-based descriptor of the directed signal-flow position of each cortical area within each analysed network.

For each of the four edge-restricted subnetworks, node-level feature vectors were constructed from the WDMT-based rooted-arborescence quantities described above and complementary network-topological node features. The reduced feature configuration contained only the WDMT-based rooted-arborescence quantities. The full feature configuration additionally included node-centric CD-flow, directed node betweenness, directed in- and out-degree, incoming, outgoing and direction-combined lobby-index [6] measures, and fixed anatomical and topology-based hierarchy ranks. These vectors enabled comparison of how the same cortical areas were organised under the anatomically defined SLN forward–backward classification and the topology-based CDw forward–backward classification.

Multidimensional scaling (MDS) was used for exploratory visualisation of the dissimilarity relationships among node–subnetwork feature vectors. Representational similarity analysis (RSA) was then adapted to compare the patterns of pairwise dissimilarities among cortical areas induced by the four subnetworks. Following the RDM-based comparison principle of Kriegeskorte et al. (2008) [8], each subnetwork was represented by a symmetric representational dissimilarity matrix (RDM) containing pairwise correlation dissimilarities between cortical-area feature vectors.

Statistical evidence for matched-direction correspondence between the SLN-based and

CDw-based subnetwork organisations was assessed using one-sided Monte Carlo permutation tests. Two primary feature configurations were analysed: the full node-level feature set and the reduced WDMT-based rooted-arborescence feature set. For each feature configuration, distances between subnetwork-specific RDMs were measured by the Manhattan distance between their vectorised strict upper-triangular entries. Three matched-direction statistics were evaluated: a forward-specific statistic, a backward-specific statistic, and a combined statistic based on the mean matched-direction distance.

1.4 Aim and research questions

The aim of the present study was to compare anatomically defined and topology-based forward–backward organisations in a directed and weighted subnetwork of the macaque cerebral cortex. Specifically, the following questions were addressed:

1. Do the four forward–backward subnetworks permit global spanning out- and in-arborescences across all 29 cortical areas, and how do their minimum- and maximum-FLN global arborescences differ in directed organisation?

2. How does the root-specific arborescence organisation of individual cortical areas differ across the four subnetworks, as characterised by node-centric extremal arborescences and WDMT-based rooted spanning arborescence quantities?

3. In exploratory analyses, what Pearson correlations are observed between WDMT-based rooted-arborescence quantities and selected node-level topological and hierarchy-related features within the analysed networks, and how are node–subnetwork representations arranged under the full and reduced feature configurations in multidimensional scaling representations?

4. Do the matched-direction subnetwork pairs, $\mathrm{SLN}_{\mathrm{forward}}$ with $\mathrm{CDw}_{\mathrm{forward}}$ and $\mathrm{SLN}_{\mathrm{backward}}$ with $\mathrm{CDw}_{\mathrm{backward}}$, exhibit smaller node-level RDM distances than the within-classification forward–backward reference, and are the corresponding matched-direction statistics larger than expected under random cortical-area alignment, for each of the two primary feature configurations?

2 Methods

2.1 Definition of anatomical and topological forward–backward subnetworks

2.1.1 Anatomical measures

The analysed anatomical network was a directed and weighted subnetwork of 29 macaque cortical areas containing 536 measured interareal projections [4]. Let $G = (V, E, w)$ denote this network, where V is the set of cortical areas, $E \subseteq V \times V$ is the set of directed anatomical projections, and $w : E \to \mathbb{R}_{>0}$ assigns an anatomical weight to each existing projection. For an edge $i \to j$, area i is the source and area j is the target.

Following retrograde tracer injection into target area j, let $N(i, j)$ denote the number of labelled projection neurons observed in source area i, and let $N_{\mathrm{tot}}(j)$ denote the total number of retrogradely labelled cortical neurons counted in the cortical hemisphere following that injection. The fraction of labelled neurons (FLN) was defined as [7]

$$\mathrm{FLN}(i, j) = \frac{N(i, j)}{N_{\mathrm{tot}}(j)},$$

and was used as the anatomical edge weight in the analysed network [4]:

$$w(i, j) = \mathrm{FLN}(i, j), \qquad (i, j) \in E.$$

Thus, $\mathrm{FLN}(i, j)$ measures the relative anatomical strength of the projection $i \to j$ among the cortical projections labelled by the injection into j.

Each projection $i \to j$ was additionally associated with its physical distance $\mathrm{dist}(i, j)$, measured in millimetres through the white matter between the geometric centres of the corresponding cortical areas, and with the proportion of supragranular labelled neurons (SLN), an anatomical index of cortical hierarchy [7]. Let $N_{\mathrm{sup}}(i, j)$ and $N_{\mathrm{infra}}(i, j)$ denote the numbers of labelled projection neurons in the supragranular and infragranular layers, respectively. Then

$$\mathrm{SLN}(i, j) = \frac{N_{\mathrm{sup}}(i, j)}{N_{\mathrm{sup}}(i, j) + N_{\mathrm{infra}}(i, j)}.$$

Here, $\mathrm{SLN}(i, j) \in [0, 1]$ is a graded measure of laminar origin: values above 0.5 indicate a predominance of supragranular labelled neurons, values below 0.5 indicate a predominance of infragranular labelled neurons, and $\mathrm{SLN}(i, j) = 0.5$ indicates equal contributions from the two compartments. For the present analysis, this measure was dichotomised into anatomical forward–backward subnetworks by classifying projections with $\mathrm{SLN}(i, j) \geq 0.5$ as forward and projections with $\mathrm{SLN}(i, j) < 0.5$ as backward.

2.1.2 Weighted convergence degree

In addition to SLN, weighted convergence degree (CDw) was used as a topological measure of directed signal-flow organisation. Following Négyessy et al. (2008) [3] and Bányai et al. (2011) [1], with the weighted extension introduced by Varga et al. (2021) [4], CDw was computed from selected weighted shortest paths in the full anatomical network G.

For each edge $(i, j) \in E$, the path cost was defined as

$$c(i, j) = \left(\frac{\mathrm{dist}(i, j)}{\mathrm{FLN}(i, j)} \right)^{\kappa}, \qquad \kappa = 0.07.$$

Thus, stronger projections had lower cost and longer projections had higher cost. The exponent κ controlled the influence of this weighted cost relative to edge count: $\kappa = 1$ gives the standard weighted-cost formulation, whereas $\kappa = 0$ recovers binary shortest-path structure.

For a directed path $P = (v_0, v_1, \ldots, v_m)$, its weighted length was defined as

$$L_w(P) = \sum_{r=0}^{m-1} c(v_r, v_{r+1}).$$

For every ordered pair of distinct cortical areas, up to $k = 8$ shortest simple directed paths with respect to L_w were retained. Let $SP_w(G)$ denote the collection of these selected paths, and for an edge $e_{ij} = (i, j) \in E$, let $SP_w(G, e_{ij}) = \{P \in SP_w(G) : e_{ij} \in P\}$ denote the selected paths traversing that edge.

For each edge $(i, j) \in E$, let $\mathrm{In}_{w,\mathrm{full}}(i, j)$ be the set of initial vertices and $\mathrm{Out}_{w,\mathrm{full}}(i, j)$ the set of terminal vertices of paths in $SP_w(G, e_{ij})$. Their overlap was defined as $\mathrm{Int}_{w,\mathrm{full}}(i, j) = \mathrm{In}_{w,\mathrm{full}}(i, j) \cap \mathrm{Out}_{w,\mathrm{full}}(i, j)$, and the strict input and output sets were $\mathrm{SIn}_{w,\mathrm{full}}(i, j) = \mathrm{In}_{w,\mathrm{full}}(i, j) \setminus \mathrm{Int}_{w,\mathrm{full}}(i, j)$ and $\mathrm{SOut}_{w,\mathrm{full}}(i, j) = \mathrm{Out}_{w,\mathrm{full}}(i, j) \setminus \mathrm{Int}_{w,\mathrm{full}}(i, j)$.

The full-network weighted convergence degree was defined as

$$\mathrm{CDw}_{\mathrm{full}}(i, j) = \frac{|\mathrm{SIn}_{w,\mathrm{full}}(i, j)| - |\mathrm{SOut}_{w,\mathrm{full}}(i, j)|}{|\mathrm{In}_{w,\mathrm{full}}(i, j) \cup \mathrm{Out}_{w,\mathrm{full}}(i, j)|} = \frac{|\mathrm{In}_{w,\mathrm{full}}(i, j)| - |\mathrm{Out}_{w,\mathrm{full}}(i, j)|}{|\mathrm{In}_{w,\mathrm{full}}(i, j) \cup \mathrm{Out}_{w,\mathrm{full}}(i, j)|}.$$

The value of $\mathrm{CDw}_{\mathrm{full}}(i, j)$ lies in $[-1, 1]$: positive values indicate convergent or condensing edges, negative values indicate divergent or spreading edges, and zero indicates equal cardinalities of the input and output sets induced by the selected weighted paths. For the present analysis, this measure was dichotomised into topological forward–backward subnetworks by classifying projections with $\mathrm{CDw}_{\mathrm{full}}(i, j) < 0$ as forward and projections with $\mathrm{CDw}_{\mathrm{full}}(i, j) > 0$ as backward. Projections with $\mathrm{CDw}_{\mathrm{full}}(i, j) = 0$ were treated as neutral and excluded from both CDw-based subnetworks.

2.1.3 Forward–backward subnetwork definitions

The four forward–backward subnetworks were obtained from the full anatomical network $G = (V, E, w)$ by restricting its edge set according to either SLN or $\mathrm{CDw}_{\mathrm{full}}$. In all subnetworks, retained edges preserved their anatomical weights, $w(i, j) = \mathrm{FLN}(i, j)$.

The anatomical forward and backward edge sets were defined as $E^{\mathrm{SLN}}_{\mathrm{forward}} = \{(i, j) \in E : \mathrm{SLN}(i, j) \geq 0.5\}$ and $E^{\mathrm{SLN}}_{\mathrm{backward}} = \{(i, j) \in E : \mathrm{SLN}(i, j) < 0.5\}$, respectively. These sets form a disjoint partition of the full edge set:

$$E = E^{\mathrm{SLN}}_{\mathrm{forward}} \,\dot\cup\, E^{\mathrm{SLN}}_{\mathrm{backward}}.$$

The corresponding directed weighted subnetworks were $G^{\mathrm{SLN}}_{\mathrm{forward}} = (V, E^{\mathrm{SLN}}_{\mathrm{forward}}, w|_{E^{\mathrm{SLN}}_{\mathrm{forward}}})$ and $G^{\mathrm{SLN}}_{\mathrm{backward}} = (V, E^{\mathrm{SLN}}_{\mathrm{backward}}, w|_{E^{\mathrm{SLN}}_{\mathrm{backward}}})$.

The topological forward and backward edge sets were defined as $E^{\mathrm{CDw}}_{\mathrm{forward}} = \{(i, j) \in E : \mathrm{CDw}_{\mathrm{full}}(i, j) < 0\}$ and $E^{\mathrm{CDw}}_{\mathrm{backward}} = \{(i, j) \in E : \mathrm{CDw}_{\mathrm{full}}(i, j) > 0\}$, respectively. The neutral edge set was $E^{\mathrm{CDw}}_{0} = \{(i, j) \in E : \mathrm{CDw}_{\mathrm{full}}(i, j) = 0\}$; its edges were excluded from both CDw-based subnetworks. Hence,

$$E = E^{\mathrm{CDw}}_{\mathrm{forward}} \,\dot\cup\, E^{\mathrm{CDw}}_{\mathrm{backward}} \,\dot\cup\, E^{\mathrm{CDw}}_{0},$$

where $\dot\cup$ denotes disjoint union. The corresponding directed weighted subnetworks were $G^{\mathrm{CDw}}_{\mathrm{forward}} = (V, E^{\mathrm{CDw}}_{\mathrm{forward}}, w|_{E^{\mathrm{CDw}}_{\mathrm{forward}}})$ and $G^{\mathrm{CDw}}_{\mathrm{backward}} = (V, E^{\mathrm{CDw}}_{\mathrm{backward}}, w|_{E^{\mathrm{CDw}}_{\mathrm{backward}}})$.

For concise reference in prose, tables and figure labels, the four edge-restricted subnetworks are denoted by $\mathrm{SLN}_{\mathrm{forward}}$, $\mathrm{SLN}_{\mathrm{backward}}$, $\mathrm{CDw}_{\mathrm{forward}}$ and $\mathrm{CDw}_{\mathrm{backward}}$. Let $\mathcal{X}_{\mathrm{sub}} = \{\mathrm{SLN}_{\mathrm{forward}}, \mathrm{SLN}_{\mathrm{backward}}, \mathrm{CDw}_{\mathrm{forward}}, \mathrm{CDw}_{\mathrm{backward}}\}$. denote the subnetworks used in the arborescence analysis, and let $\mathcal{X} = \{\mathrm{full}\} \cup \mathcal{X}_{\mathrm{sub}}$ denote the five networks used in the node-level feature construction. For each $X \in \mathcal{X}$, write $G_X = (V, E_X, w|_{E_X})$, where $E_{\mathrm{full}} = E$ and $G_{\mathrm{full}} = G$.

For any weighted directed subgraph $H = (V_H, E_H, w|_{E_H})$ of G, its total anatomical edge weight was defined as

$$W(H) = \sum_{(i,j)\in E_H} w(i, j) = \sum_{(i,j)\in E_H} \mathrm{FLN}(i, j).$$

This notation was used for the subnetworks and arborescences considered in the subsequent analyses.

2.2 Directed rooted arborescences and reachability analysis

Each edge-restricted subnetwork G_X, $X \in \mathcal{X}_{\mathrm{sub}}$, was analysed using the retained anatomical edge weights $w(i, j) = \mathrm{FLN}(i, j)$. The path cost $c(i, j)$ used to construct $\mathrm{CDw}_{\mathrm{full}}$ was

not used for arborescence optimisation.

In G_X, write $u \rightsquigarrow_X v$ if a directed path existed from u to v. For each root $r \in V$, the outward and inward reachable vertex sets were

$$R_X^{\mathrm{out}}(r) = \{v \in V : \ r \rightsquigarrow_X v\}, \qquad R_X^{\mathrm{in}}(r) = \{v \in V : \ v \rightsquigarrow_X r\}.$$

Reachability included the length-zero path, so $r \in R_X^{\mathrm{out}}(r) \cap R_X^{\mathrm{in}}(r)$.

Let $\eta \in \{\mathrm{out}, \mathrm{in}\}$ denote arborescence orientation. A global spanning out-arborescence in G_X was a rooted directed tree on V in which every vertex was uniquely reachable from its root. A global spanning in-arborescence was a rooted directed tree on V in which the root was uniquely reachable from every vertex. Let $\mathcal{T}_X^{\eta}$ denote the set of admissible global spanning η-arborescences in G_X, allowing the root to vary. Such an out-arborescence rooted at r could exist only if $R_X^{\mathrm{out}}(r) = V$, and such an in-arborescence only if $R_X^{\mathrm{in}}(r) = V$. If no admissible root existed, then $\mathcal{T}_X^{\eta} = \varnothing$, which was treated as a structural property of the subnetwork.

Let $a \in \{\min, \max\}$ denote the optimisation criterion. Whenever $\mathcal{T}_X^{\eta} \neq \varnothing$, the global minimum- and maximum-weight arborescences were defined by

$$T_{X,\min}^{\eta} \in \operatorname*{arg\,min}_{T \in \mathcal{T}_X^{\eta}} W(T), \qquad T_{X,\max}^{\eta} \in \operatorname*{arg\,max}_{T \in \mathcal{T}_X^{\eta}} W(T).$$

For node-centric analysis, the root $r \in V$ was fixed before optimisation. Let $\mathcal{T}_{X,r}^{\mathrm{out}}$ denote the out-arborescences rooted at r and spanning $R_X^{\mathrm{out}}(r)$, and let $\mathcal{T}_{X,r}^{\mathrm{in}}$ denote the in-arborescences rooted at r and spanning $R_X^{\mathrm{in}}(r)$. Thus, every vertex in a node-centric out-arborescence had a unique directed path from r, whereas every vertex in a node-centric in-arborescence had a unique directed path to r. The node-centric minimum- and maximum-weight arborescences were

$$T_{X,r,\min}^{\eta} \in \operatorname*{arg\,min}_{T \in \mathcal{T}_{X,r}^{\eta}} W(T), \qquad T_{X,r,\max}^{\eta} \in \operatorname*{arg\,max}_{T \in \mathcal{T}_{X,r}^{\eta}} W(T).$$

Because their vertex sets were fixed by directed reachability, minimum and maximum node-centric arborescences with the same root and orientation had the same order.

Computationally, an out-arborescence rooted at r was obtained from the subgraph induced by $R_X^{\mathrm{out}}(r)$ after removing all edges entering r. An in-arborescence rooted at r was obtained by restricting the graph to the subgraph induced by $R_X^{\mathrm{in}}(r)$, reversing its edges, solving the corresponding out-arborescence problem, and reversing the resulting tree back. The resulting trees were verified to have indegree zero at the root in the out-orientation and outdegree zero at the root in the in-orientation.

For any rooted arborescence T with root r, let $\delta_T(u, v)$ denote the number of directed tree edges on the unique directed path from u to v. This tree distance was distinct from

the physical projection distance $\mathrm{dist}(i,j)$. The level of a vertex $v \in V(T)$ was defined as

$$\lambda_T(v) = \begin{cases} \delta_T(r,v), & \text{if } T \text{ was an out-arborescence,} \\ \delta_T(v,r), & \text{if } T \text{ was an in-arborescence.} \end{cases}$$

Let $L_k(T) = \{v \in V(T) : \lambda_T(v) = k\}$ denote the set of vertices at level k, with $L_0(T) = \{r\}$, and let $D(T) = \max_{v \in V(T)} \lambda_T(v)$ denote tree depth. For $D(T) > 0$, the mean and standard deviation of non-root level width were defined by

$$\overline{b}(T) = \frac{1}{D(T)} \sum_{k=1}^{D(T)} |L_k(T)|, \qquad \sigma_b(T) = \sqrt{\frac{1}{D(T)} \sum_{k=1}^{D(T)} \big(|L_k(T)| - \overline{b}(T)\big)^2}.$$

For visual comparison of vertex distributions over tree levels, the cumulative tree-distance profile of a rooted arborescence T was defined as

$$C_T(k) = \frac{|\{v \in V(T) : \lambda_T(v) \le k\}|}{|V(T)|}, \qquad k = 0, 1, \ldots, D(T).$$

Thus, $C_T(k)$ denoted the proportion of vertices lying within at most k tree levels of the root, including the root itself at $k = 0$. For node-centric arborescences, the denominator was the order of the corresponding reachable-set tree rather than the order of the full network.

For each global spanning arborescence $T^{\eta}_{X,a}$, tree order was fixed at $|V|$ and was therefore not used as a distinguishing feature. The global summary reported its root, total anatomical weight $W(T^{\eta}_{X,a})$, depth $D(T^{\eta}_{X,a})$, mean non-root level width $\overline{b}(T^{\eta}_{X,a})$, and level-width standard deviation $\sigma_b(T^{\eta}_{X,a})$.

Node-centric quantities were computed for selected roots representing low-, intermediate- and high-level cortical areas. Let

$$\mathcal{R} = \{\mathrm{V1}, \mathrm{V2}, \mathrm{MT}, \mathrm{V4}, \mathrm{7A}, \mathrm{TEpd}, \mathrm{9/46v}, \mathrm{9/46d}\}, \qquad m = |\mathcal{R}|.$$

For root $r \in \mathcal{R}$, the tree order and relative tree order were defined as

$$q^{\eta}_X(r) = |R^{\eta}_X(r)| = |V(T^{\eta}_{X,r,a})|, \qquad \rho^{\eta}_X(r) = \frac{q^{\eta}_X(r)}{|V|}.$$

Both quantities were independent of a, since minimum and maximum trees with the same root and orientation spanned the same reachable set. Across selected roots, the mean and standard deviation of tree order and the mean relative tree order were

$$\overline{q}^{\eta}_X = \frac{1}{m} \sum_{r \in \mathcal{R}} q^{\eta}_X(r), \qquad \sigma^{\eta}_{q,X} = \sqrt{\frac{1}{m} \sum_{r \in \mathcal{R}} (q^{\eta}_X(r) - \overline{q}^{\eta}_X)^2}, \qquad \overline{\rho}^{\eta}_X = \frac{1}{m} \sum_{r \in \mathcal{R}} \rho^{\eta}_X(r).$$

Node-centric total anatomical weight, depth, and mean non-root level width were defined by applying W, D and $\bar{b}$ to $T^{\eta}_{X,r,a}$ and were averaged across selected roots separately for each orientation η and optimisation criterion a.

2.3 Construction and scaling of node-level feature vectors

For each cortical area in each fixed network G_X, $X \in \mathcal{X}$, a node-level feature vector was constructed for exploratory visualisations, representational dissimilarity matrix (RDM) comparisons and permutation tests. The feature vector combined Weighted Directed Matrix–Tree Theorem-based rooted-arborescence quantities [2], network-topological node measures derived from directed signal-flow and graph structure [1, 4, 6, 5], and whole-network anatomical and topological hierarchy ranks [7, 4].

2.3.1 WDMT-based features in eligible strongly connected components

Features based on the Weighted Directed Matrix–Tree Theorem (WDMT) [2] were computed within eligible strongly connected components (SCCs) of each network G_X. An SCC C was eligible if $|C| \geq 7$. Under this criterion, singleton SCCs and one two-node SCC were excluded from the WDMT-based feature construction.

Let $\mathcal{C}_{\mathrm{elig}}(X)$ denote the set of eligible SCCs in G_X, and let $G_X[C]$ be the subgraph induced by C. Using the total anatomical edge-weight function W defined above, the internal weight of C was $W_X(C) = W(G_X[C])$, and its normalised eligible-SCC weight was

$$\pi_{X,C} = \frac{W_X(C)}{\sum_{C' \in \mathcal{C}_{\mathrm{elig}}(X)} W_X(C')}.$$

Thus, $\pi_{X,C} = 1$ when G_X contained only one eligible SCC.

For an eligible SCC C, let $A_{X,C}$ and $A^{w}_{X,C}$ denote its unweighted and FLN-weighted adjacency matrices, respectively, and let $D_{\mathrm{in},X,C}$, $D_{\mathrm{out},X,C}$, $D^{w}_{\mathrm{in},X,C}$ and $D^{w}_{\mathrm{out},X,C}$ denote the corresponding diagonal in- and out-degree matrices. The unweighted and weighted directed Laplacians were

$$\begin{aligned} L_{1,X,C} &= D_{\mathrm{in},X,C} - A_{X,C}, \quad L_{2,X,C} = D_{\mathrm{out},X,C} - A^{T}_{X,C}, \\ L^{w}_{1,X,C} &= D^{w}_{\mathrm{in},X,C} - A^{w}_{X,C}, \quad L^{w}_{2,X,C} = D^{w}_{\mathrm{out},X,C} - (A^{w}_{X,C})^{T}. \end{aligned}$$

For a root $i \in C$, let the superscript (i) denote deletion of row i and column i. Following the WDMT [2], the unweighted rooted-arborescence counts were

$$\tau^{\mathrm{out}}_{X,i,C} = \det\left(L^{(i)}_{1,X,C}\right), \qquad \tau^{\mathrm{in}}_{X,i,C} = \det\left(L^{(i)}_{2,X,C}\right),$$

and the corresponding weighted quantities were

$$\tau^{\mathrm{out},w}_{X,i,C} = \det\big((L^{w}_{1,X,C})^{(i)}\big)\,, \qquad \tau^{\mathrm{in},w}_{X,i,C} = \det\big((L^{w}_{2,X,C})^{(i)}\big)\,.$$

Here, $\tau^{\mathrm{out}}_{X,i,C}$ and $\tau^{\mathrm{in}}_{X,i,C}$ counted outgoing and incoming directed spanning arborescences rooted at i, while $\tau^{\mathrm{out},w}_{X,i,C}$ and $\tau^{\mathrm{in},w}_{X,i,C}$ summed the products of FLN edge weights over the corresponding rooted arborescences.

The mean weighted product per rooted arborescence was defined as

$$\overline{\tau}^{\mathrm{out},w}_{X,i,C} = \frac{\tau^{\mathrm{out},w}_{X,i,C}}{\tau^{\mathrm{out}}_{X,i,C}}, \qquad \overline{\tau}^{\mathrm{in},w}_{X,i,C} = \frac{\tau^{\mathrm{in},w}_{X,i,C}}{\tau^{\mathrm{in}}_{X,i,C}}.$$

Because each eligible component was strongly connected, both denominators were positive for every node in that component.

For node-level feature construction, the six WDMT quantities were scaled by the normalised internal weight of the node's eligible SCC. Specifically, for

$$z \in \left\{\tau^{\mathrm{out}}, \tau^{\mathrm{in}}, \tau^{\mathrm{out},w}, \tau^{\mathrm{in},w}, \overline{\tau}^{\mathrm{out},w}, \overline{\tau}^{\mathrm{in},w}\right\},$$

the SCC-scaled counterpart was defined as

$$z_{X,i,\mathrm{sc}} = \pi_{X,C}\, z_{X,i,C}.$$

The six SCC-scaled quantities were included in the node-level feature vector, whereas their unscaled counterparts were retained only for reporting. WDMT-based features were undefined for nodes outside eligible SCCs and were handled as described in Subsection 2.3.4.

A complete dictionary of the WDMT-based quantities is provided in Appendix A.

2.3.2 Network-topological node features

For each $X \in \mathcal{X}$, weighted convergence degree was recomputed within the fixed network G_X using the path cost c, weighted path length L_w, and the first $k = 8$ weighted shortest simple directed paths per reachable ordered pair, as defined in Subsection 2.1.2. The resulting edge-level quantity was denoted by CDw_X. This recomputation was used only for node-level feature construction: membership in $G^{\mathrm{CDw}}_{\mathrm{forward}}$ and $G^{\mathrm{CDw}}_{\mathrm{backward}}$ remained determined by $\mathrm{CDw}_{\mathrm{full}}$.

For a node $i \in V$, let

$$N^{\mathrm{out}}_{X,\mathrm{CDw}}(i) = \{j \in V : (i,j) \in E_X \text{ and } \mathrm{CDw}_X(i,j) \text{ was defined}\},$$

and

$$N^{\mathrm{in}}_{X,\mathrm{CDw}}(i) = \{j \in V : (j,i) \in E_X \text{ and } \mathrm{CDw}_X(j,i) \text{ was defined}\}.$$

Following the normalised CD-flow construction of Varga et al. (2021) [4], the network-specific CD-flow feature was computed as

$$\Phi_X(i) = \frac{1}{|N^{\mathrm{out}}_{X,\mathrm{CDw}}(i)|} \sum_{j \in N^{\mathrm{out}}_{X,\mathrm{CDw}}(i)} \mathrm{CDw}_X(i,j) - \frac{1}{|N^{\mathrm{in}}_{X,\mathrm{CDw}}(i)|} \sum_{j \in N^{\mathrm{in}}_{X,\mathrm{CDw}}(i)} \mathrm{CDw}_X(j,i),$$

whenever both neighbour sets were non-empty. Thus, $\Phi_X(i)$ contrasted the mean convergence–divergence character of outgoing and incoming projections within G_X.

Directed node betweenness was computed from unweighted directed shortest paths in G_X. Let σ^X_{st} denote the number of shortest directed paths from s to t, and let $\sigma^X_{st}(i)$ denote the number of such paths passing through i. The unnormalised endpoint-excluding node betweenness of i was

$$B_X(i) = \sum_{\substack{s,t \in V \\ s \neq t,\ s \neq i,\ t \neq i \\ \sigma^X_{st} > 0}} \frac{\sigma^X_{st}(i)}{\sigma^X_{st}},$$

where unreachable ordered pairs were omitted by the condition $\sigma^X_{st} > 0$.

The lobby index [6] was adapted here to the directed networks by distinguishing incoming and outgoing neighbour sets; this direction-specific adaptation used neighbours' unweighted total degrees rather than edge weights. For $u \in V$, define

$$N^{\mathrm{in}}_X(u) = \{v \in V : (v,u) \in E_X\}, \qquad N^{\mathrm{out}}_X(u) = \{v \in V : (u,v) \in E_X\},$$

and $N^{\mathrm{all}}_X(u) = N^{\mathrm{in}}_X(u) \cup N^{\mathrm{out}}_X(u)$. The unweighted degrees were

$$d^{\mathrm{in}}_X(u) = |N^{\mathrm{in}}_X(u)|, \qquad d^{\mathrm{out}}_X(u) = |N^{\mathrm{out}}_X(u)|, \qquad d^{\mathrm{tot}}_X(u) = d^{\mathrm{in}}_X(u) + d^{\mathrm{out}}_X(u).$$

Thus, reciprocal projections contributed twice to $d^{\mathrm{tot}}_X(u)$, although a reciprocally connected neighbour occurred only once in $N^{\mathrm{all}}_X(u)$.

For a neighbour set $N \subseteq V$, define

$$\ell_X(N) = \max\left\{h \in \mathbb{N}_0 : \left|\left\{v \in N : d^{\mathrm{tot}}_X(v) \geq h\right\}\right| \geq h\right\}.$$

The retained lobby-index features were

$$\ell^{\mathrm{in}}_X(i) = \ell_X\left(N^{\mathrm{in}}_X(i)\right), \qquad \ell^{\mathrm{out}}_X(i) = \ell_X\left(N^{\mathrm{out}}_X(i)\right), \qquad \ell^{\mathrm{all}}_X(i) = \ell_X\left(N^{\mathrm{all}}_X(i)\right).$$

Finally, $d^{\mathrm{in}}_X(i)$ and $d^{\mathrm{out}}_X(i)$ were retained as direct node-level connectivity descriptors in G_X.

2.3.3 Whole-network hierarchy ranks

Two full-network ordinal reference ranks were assigned to each cortical area and retained unchanged across the analysed networks. The anatomical rank $h^{\mathrm{SLN}}(i) \in \{1, \ldots, 29\}$ was assigned according to the ordering of areas in the SLN-based hierarchy of the complete 29-area network shown in Figure 10B of Varga et al. (2021) [4], with increasing rank corresponding to increasing anatomical hierarchical position. The topological rank $h^{\mathrm{CDw}}(i) \in \{1, \ldots, 29\}$ was obtained by ordering areas in ascending full-network CD-flow, $\Phi_{\mathrm{full}}(i)$, from more source-like to more allocating-like positions. Thus, $h^{\mathrm{SLN}}(i)$ and $h^{\mathrm{CDw}}(i)$ were fixed ordinal reference features, whereas $\Phi_X(i)$ was recomputed within each network G_X.

2.3.4 Feature transformation, availability indicators and standardisation

The unscaled node-level feature table retained the computed values and undefined entries. For distance-based analyses, the six SCC-scaled WDMT-based features

$$\tau^{\mathrm{out}}_{X,i,\mathrm{sc}}, \quad \tau^{\mathrm{in}}_{X,i,\mathrm{sc}}, \quad \tau^{\mathrm{out},w}_{X,i,\mathrm{sc}}, \quad \tau^{\mathrm{in},w}_{X,i,\mathrm{sc}}, \quad \overline{\tau}^{\mathrm{out},w}_{X,i,\mathrm{sc}}, \quad \overline{\tau}^{\mathrm{in},w}_{X,i,\mathrm{sc}}$$

were transformed by $x \mapsto \log_{10}(1 + x)$.

WDMT-feature availability was recorded in the exported scaled feature table by

$$I^{\tau}_X(i) = \begin{cases} 1, & \text{if } i \in C \text{ for some } C \in \mathcal{C}_{\mathrm{elig}}(X), \\ 0, & \text{otherwise.} \end{cases}$$

Because each eligible component was strongly connected, every node in an eligible SCC admitted incoming and outgoing rooted arborescences. Consequently, all six WDMT-based quantities were jointly available within eligible SCCs, and a single availability indicator was sufficient.

Before standardisation, undefined log-transformed WDMT-based values were coded as zero, while $I^{\tau}_X(i)$ retained their availability information in the exported feature table. Undefined values of the network-specific CD-flow feature $\Phi_X(i)$ were also coded as zero in the implemented full feature configuration.

Numeric feature columns, excluding the binary availability indicator, were standardised separately within each network. For each non-constant standardised feature f, node i, and network X,

$$z_{X,i,f} = \frac{x_{X,i,f} - \mu_{X,f}}{\sigma_{X,f}},$$

where $\mu_{X,f}$ and $\sigma_{X,f}$ were the within-network mean and population standard deviation of feature f, respectively. Constant standardised feature columns were mapped to zero. The binary indicator $I^{\tau}_X(i)$ remained unscaled and was retained only for verification; it

was excluded from the interpreted multidimensional scaling, representational similarity analysis and permutation-test feature sets.

2.4 Exploratory analysis of node-level feature vectors

2.4.1 Feature-pair correlation heatmaps

Pairwise linear correlations were examined separately in each network G_X. Here, $i \in V$ denotes a cortical area. Network-specific CD-flow $\Phi_X(i)$ and node betweenness $B_X(i)$ were compared with the log-transformed SCC-scaled unweighted and average weighted WDMT-based tau features of the same network. The hierarchy ranks $h^{\mathrm{SLN}}(i)$ and $h^{\mathrm{CDw}}(i)$ were retained as the previously defined fixed full-network ordinal reference ranks.

For each orientation $\eta \in \{\mathrm{out}, \mathrm{in}\}$, Pearson correlation coefficients were computed across cortical areas for which both quantities in the corresponding pair were defined. The selected feature pairs were

$$
\begin{aligned}
&\big(\Phi_X(i),\, B_X(i)\big),\\
&\big(\Phi_X(i),\, \log_{10}\big(1+\tau^{\eta}_{X,i,\mathrm{sc}}\big)\big)\,, \qquad \big(\Phi_X(i),\, \log_{10}\big(1+\overline{\tau}^{\eta,w}_{X,i,\mathrm{sc}}\big)\big)\,,\\
&\big(B_X(i),\, \log_{10}\big(1+\tau^{\eta}_{X,i,\mathrm{sc}}\big)\big)\,, \qquad \big(B_X(i),\, \log_{10}\big(1+\overline{\tau}^{\eta,w}_{X,i,\mathrm{sc}}\big)\big)\,,\\
&\big(h^{\mathrm{SLN}}(i),\, \log_{10}\big(1+\tau^{\eta}_{X,i,\mathrm{sc}}\big)\big)\,, \qquad \big(h^{\mathrm{SLN}}(i),\, \log_{10}\big(1+\overline{\tau}^{\eta,w}_{X,i,\mathrm{sc}}\big)\big)\,,\\
&\big(h^{\mathrm{CDw}}(i),\, \log_{10}\big(1+\tau^{\eta}_{X,i,\mathrm{sc}}\big)\big)\,, \qquad \big(h^{\mathrm{CDw}}(i),\, \log_{10}\big(1+\overline{\tau}^{\eta,w}_{X,i,\mathrm{sc}}\big)\big)\,.
\end{aligned}
$$

Because $h^{\mathrm{SLN}}(i)$ and $h^{\mathrm{CDw}}(i)$ were ordinal reference ranks, correlations involving these features were interpreted descriptively as linear summaries of their numerical rank encodings.

Correlations involving WDMT-based quantities were computed from valid paired observations before undefined WDMT-based values were replaced for the subsequent distance-based analyses. For each network, the coefficients were displayed in separate outgoing- and incoming-tau heatmaps; the orientation-independent coefficient for $\big(\Phi_X(i), B_X(i)\big)$ was shown in both panels for comparison. Diagonal entries were set to 1, while grey off-diagonal cells labelled `NaN` denoted feature pairs that were not included among the selected comparisons. The heatmaps were used descriptively and were not treated as inferential tests.

2.4.2 Multidimensional scaling of subnetwork feature spaces

Multidimensional scaling (MDS) [9] was used to visualise the dissimilarity structure of node–subnetwork feature vectors in the four edge-restricted networks

$$\mathcal{X}_{\mathrm{sub}} = \{\mathrm{SLN}_{\mathrm{forward}}, \mathrm{SLN}_{\mathrm{backward}}, \mathrm{CDw}_{\mathrm{forward}}, \mathrm{CDw}_{\mathrm{backward}}\}.$$

The full network was excluded from this comparison.

Two feature configurations were considered. The reduced WDMT-based rooted-arborescence feature set was

$$F_{\mathrm{red}} = \{\log_{10}(1+\tau^{\mathrm{out}}_{X,i,\mathrm{sc}}), \log_{10}(1+\tau^{\mathrm{in}}_{X,i,\mathrm{sc}}), \log_{10}(1+\tau^{\mathrm{out},w}_{X,i,\mathrm{sc}}), \log_{10}(1+\tau^{\mathrm{in},w}_{X,i,\mathrm{sc}}),$$
$$\log_{10}(1+\overline{\tau}^{\mathrm{out},w}_{X,i,\mathrm{sc}}), \log_{10}(1+\overline{\tau}^{\mathrm{in},w}_{X,i,\mathrm{sc}})\}.$$

The full node-level feature set additionally included the network-topological and hierarchy-related features:

$$F_{\mathrm{full}} = F_{\mathrm{red}} \cup \{\Phi_X,\ B_X,\ \ell^{\mathrm{in}}_X,\ \ell^{\mathrm{out}}_X,\ \ell^{\mathrm{all}}_X,\ d^{\mathrm{in}}_X,\ d^{\mathrm{out}}_X,\ h^{\mathrm{SLN}},\ h^{\mathrm{CDw}}\}.$$

Thus, $|F_{\mathrm{red}}| = 6$ and $|F_{\mathrm{full}}| = 15$. The binary WDMT-availability indicator I^{τ}_X was excluded from both configurations.

For $F \in \{F_{\mathrm{full}}, F_{\mathrm{red}}\}$, let

$$\mathbf{z}^{(F)}_{X,i} = \big(z_{X,i,f}\big)_{f\in F} \in \mathbb{R}^{|F|}$$

denote the standardised feature vector of cortical area i in subnetwork G_X. The pooled observation set was

$$\mathcal{O} = \{(X,i) : X \in \mathcal{X}_{\mathrm{sub}},\ i \in V\},$$

with $|\mathcal{O}| = 4 \cdot 29 = 116$. For observations $a, b \in \mathcal{O}$, exploratory dissimilarities were computed using correlation, Euclidean, Manhattan and Mahalanobis dissimilarities:

$$\delta^{\mathrm{corr}}_{ab} = 1 - r(\mathbf{z}_a, \mathbf{z}_b), \qquad \delta^{\mathrm{Euc}}_{ab} = \|\mathbf{z}_a - \mathbf{z}_b\|_2, \qquad \delta^{\mathrm{Man}}_{ab} = \|\mathbf{z}_a - \mathbf{z}_b\|_1,$$

and

$$\delta^{\mathrm{Mah}}_{ab} = \big[(\mathbf{z}_a - \mathbf{z}_b)^{\mathsf{T}} \Sigma^{+}_{F} (\mathbf{z}_a - \mathbf{z}_b)\big]^{1/2},$$

where r denotes the Pearson correlation coefficient across feature coordinates and Σ^{+}_{F} is the Moore–Penrose pseudoinverse of the pooled covariance matrix for feature set F.

For each feature set and dissimilarity definition, the pooled dissimilarity matrix

$$\Delta = (\delta_{ab})_{a,b\in\mathcal{O}}$$

was rescaled by its mean positive off-diagonal dissimilarity. Specifically, with

$$\overline{\delta}_{+} = \frac{\sum_{\substack{a<b \\ \delta_{ab}>0}} \delta_{ab}}{|\{(a,b) : a < b,\ \delta_{ab} > 0\}|}, \qquad \widetilde{\delta}_{ab} = \frac{\delta_{ab}}{\overline{\delta}_{+}},$$

the normalised matrix $\widetilde{\Delta} = (\widetilde{\delta}_{ab})$ had mean positive off-diagonal dissimilarity equal to 1.

Two-dimensional metric MDS was applied to each precomputed normalised dissimilar-

ity matrix using the Scaling by MAjorizing a COmplicated Function (SMACOF) stress-majorisation algorithm. For an embedded configuration

$$Y = (\mathbf{y}_1, \ldots, \mathbf{y}_{116})^{\mathsf{T}} \in \mathbb{R}^{116\times 2},$$

the fitted coordinates minimised the raw stress

$$\sigma(Y) = \sum_{a<b} \left(\widetilde{\delta}_{ab} - \|\mathbf{y}_a - \mathbf{y}_b\|_2\right)^2.$$

The optimisation used eight initialisations, at most 1000 iterations and a fixed random seed of 42. Correlation- and Manhattan-based embeddings were retained for presentation for both F_{full} and F_{red}.

Following Kriegeskorte et al. (2008) [8], MDS was used only as an exploratory visualisation of dissimilarity structure. Because a two-dimensional embedding necessarily distorted part of the original higher-dimensional structure, visible separation in the MDS plots was not interpreted as confirmatory statistical evidence.

2.5 Representational dissimilarity analysis and Monte Carlo permutation testing

Representational similarity analysis (RSA) [8] was used to compare the node-level feature-space organisation of the four edge-restricted subnetworks in $\mathcal{X}_{\text{sub}}$. Analyses were performed for the full node-level feature set F_{full} and the reduced WDMT-based rooted-arborescence feature set F_{red} defined in Subsection 2.4.2. The reduced configuration was retained as a second primary feature configuration because the preceding exploratory MDS analysis indicated that exclusion of the network-topological and hierarchy-related features materially altered the arrangement of node–subnetwork observations.

For each $F \in \{F_{\text{full}}, F_{\text{red}}\}$ and each $X \in \mathcal{X}_{\text{sub}}$, a representational dissimilarity matrix (RDM) was constructed from the standardised node-level feature vectors $\mathbf{z}_{X,i}^{(F)}$ using correlation distance:

$$D_X^{(F)}(i,j) = \begin{cases} 1 - r\left(\mathbf{z}_{X,i}^{(F)}, \mathbf{z}_{X,j}^{(F)}\right), & i \neq j, \\ 0, & i = j, \end{cases} \qquad i,j \in V,$$

where r denotes the Pearson correlation coefficient across feature coordinates. Thus, each $D_X^{(F)}$ was a symmetric 29×29 matrix.

Let $\mathrm{ut}(D)$ denote the vector of strict upper-triangular entries of an RDM in the fixed cortical-area order. Since $|V| = 29$, $\mathrm{ut}(D) \in \mathbb{R}^{406}$. For subnetworks G_X and G_Y, the

RDM-to-RDM dissimilarity under feature set F was defined as the Manhattan distance

$$d_F(G_X, G_Y) = \left\| \mathrm{ut}\left(D_X^{(F)}\right) - \mathrm{ut}\left(D_Y^{(F)}\right) \right\|_1 .$$

The within-classification forward–backward distances were

$$d_{\mathrm{SLN}}^{(F)} = d_F\left(G_{\mathrm{forward}}^{\mathrm{SLN}}, G_{\mathrm{backward}}^{\mathrm{SLN}}\right), \qquad d_{\mathrm{CDw}}^{(F)} = d_F\left(G_{\mathrm{forward}}^{\mathrm{CDw}}, G_{\mathrm{backward}}^{\mathrm{CDw}}\right),$$

and their mean was

$$d_{\mathrm{within}}^{(F)} = \frac{d_{\mathrm{SLN}}^{(F)} + d_{\mathrm{CDw}}^{(F)}}{2}.$$

The matched-direction distances were

$$d_{\mathrm{forward}}^{(F)} = d_F\left(G_{\mathrm{forward}}^{\mathrm{SLN}}, G_{\mathrm{forward}}^{\mathrm{CDw}}\right), \qquad d_{\mathrm{backward}}^{(F)} = d_F\left(G_{\mathrm{backward}}^{\mathrm{SLN}}, G_{\mathrm{backward}}^{\mathrm{CDw}}\right),$$

with mean

$$d_{\mathrm{matched}}^{(F)} = \frac{d_{\mathrm{forward}}^{(F)} + d_{\mathrm{backward}}^{(F)}}{2}.$$

Three matched-direction correspondence statistics were computed:

$$S_{\mathrm{forward}}^{(F)} = d_{\mathrm{within}}^{(F)} - d_{\mathrm{forward}}^{(F)}, \qquad S_{\mathrm{backward}}^{(F)} = d_{\mathrm{within}}^{(F)} - d_{\mathrm{backward}}^{(F)},$$

and

$$S_{\mathrm{combined}}^{(F)} = d_{\mathrm{within}}^{(F)} - d_{\mathrm{matched}}^{(F)}.$$

For each statistic, positive values indicated that the corresponding matched-direction SLN–CDw RDM dissimilarity was smaller than the mean within-classification forward–backward reference dissimilarity.

Statistical evidence for matched-direction organisation was assessed by one-sided Monte Carlo permutation tests. Under the null hypothesis, cortical-area labels in the two CDw-based RDMs were exchangeable relative to the fixed SLN-based RDMs. For each sampled permutation matrix P_b, the same cortical-area permutation was applied jointly to both CDw-based RDMs:

$$D_{\mathrm{CDw}_{\mathrm{forward}},b}^{(F)} = P_b D_{\mathrm{CDw}_{\mathrm{forward}}}^{(F)} P_b^{\mathsf{T}}, \qquad D_{\mathrm{CDw}_{\mathrm{backward}},b}^{(F)} = P_b D_{\mathrm{CDw}_{\mathrm{backward}}}^{(F)} P_b^{\mathsf{T}}.$$

This joint permutation preserved $d_{\mathrm{CDw}}^{(F)}$, and hence $d_{\mathrm{within}}^{(F)}$, while randomising the cortical-area alignment between the SLN-based and CDw-based RDMs.

For a generic statistic

$$S^{(F)} \in \left\{ S_{\mathrm{forward}}^{(F)}, S_{\mathrm{backward}}^{(F)}, S_{\mathrm{combined}}^{(F)} \right\},$$

let $S_{\text{obs}}^{(F)}$ denote its observed value and let $S_b^{(F)}$ denote its value under permutation P_b. The right-tailed Monte Carlo permutation p-value was

$$p_{\text{perm}}^{(F,S)} = \frac{1 + \sum_{b=1}^{B} \mathbf{1}\left\{S_b^{(F)} \geq S_{\text{obs}}^{(F)}\right\}}{B+1},$$

where $\mathbf{1}\{\cdot\}$ denotes the indicator function. The plus-one calculation followed Phipson and Smyth [10].

In the implementation, $B = 10{,}000$ unique non-identity cortical-area permutations were sampled without replacement using random seed 42. The same sampled permutations were used for both feature sets. For each of F_{full} and F_{red}, the tests based on $S_{\text{forward}}^{(F)}$, $S_{\text{backward}}^{(F)}$ and $S_{\text{combined}}^{(F)}$ were treated as primary inferential tests. Thus, six one-sided permutation tests were performed in total, and each test was evaluated individually at significance level $\alpha = 0.05$.

3 Results

3.1 Forward–backward subnetwork construction and arborescence analysis

3.1.1 Construction of the four forward–backward subnetworks

The full directed anatomical network contained 29 cortical areas and 536 directed projections. Applying the edge classifications defined in Subsection 2.1.3 produced four directed weighted subnetworks, in which retained edges preserved their FLN weights. The SLN-based decomposition assigned 265 projections to $\text{SLN}_{\text{forward}}$ and 271 projections to $\text{SLN}_{\text{backward}}$, thereby partitioning the complete edge set. The CDw-based decomposition assigned 264 projections to $\text{CDw}_{\text{forward}}$ and 246 projections to $\text{CDw}_{\text{backward}}$; the remaining 26 projections satisfied $\text{CDw}_{\text{full}}(i,j) = 0$ and were not included in either CDw-based subnetwork.

3.1.2 Relationship between the two hierarchy ranks

To compare the two hierarchy definitions at the node level, the SLN-based and CDw-based hierarchy ranks were plotted against each other (Figure 1). Across the full set of 29 cortical areas, the correlation was weak, with Pearson $r = 0.060$ and Spearman $\rho = 0.060$. This provided evidence against treating the anatomical SLN-based hierarchy and the topological CDw-based hierarchy as equivalent global orderings of the cortical areas.

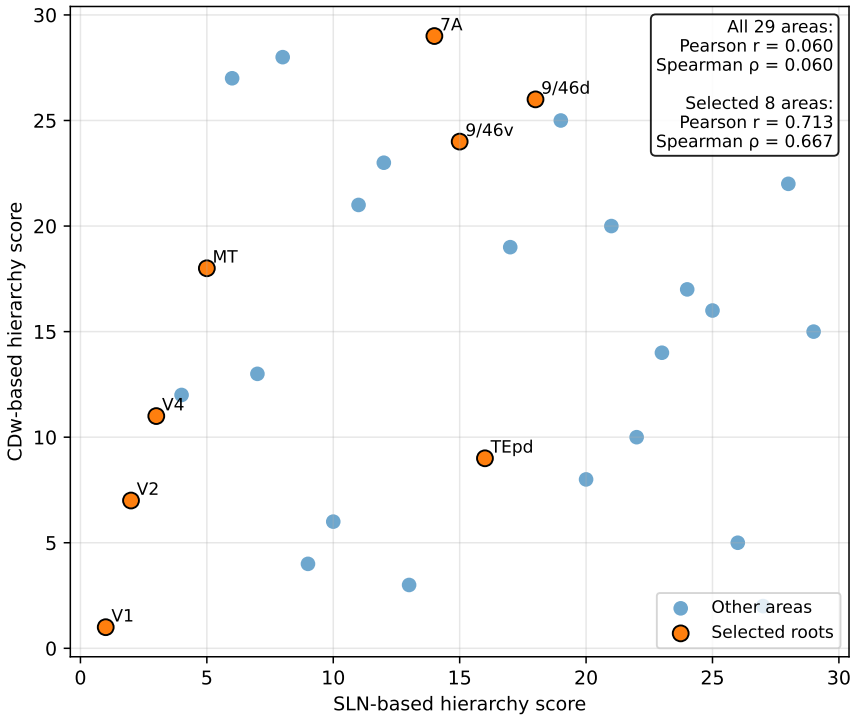

Figure 1: Comparison of SLN-based and CDw-based hierarchy ranks across the 29 cortical areas. The labelled points represent the eight preselected roots used in the node-centric arborescence analysis. The correlation of the total set of areas is weak ($r = 0.060$, $\rho = 0.060$), whereas the selected roots show a stronger positive correlation ($r = 0.713$, $\rho = 0.667$).

The weak correlation between the two hierarchy ranks provided a basis for examining whether forward–backward structural organisation was nevertheless similar when subnetworks were defined using the anatomical SLN criterion or the topology-based CDw criterion. To this end, the edge-restricted subnetworks $\mathrm{SLN}_{\mathrm{forward}}$, $\mathrm{SLN}_{\mathrm{backward}}$, $\mathrm{CDw}_{\mathrm{forward}}$, and $\mathrm{CDw}_{\mathrm{backward}}$ were analysed. Their node-level structural correspondence was subsequently examined in the RSA analysis.

The eight preselected roots provided a useful set of examples in which the two hierarchy definitions were not identical, but still assigned comparable hierarchical positions to the chosen areas across the SLN- and CDw-based ranks. In addition to the comparison of whole subnetworks, this correlation enabled a node-centric comparison of rooted arborescences with respect to the hierarchical position of their roots.

3.1.3 Global spanning arborescences in the four subnetworks

After the four edge-restricted subnetworks had been constructed, it was first examined whether each subnetwork supported global spanning arborescences. The purpose of this analysis was to determine whether the SLN- and CDw-based forward–backward decompositions yielded different directed tree organisations.

Both SLN-based subnetworks supported global spanning arborescences in both orientations. In contrast, $\mathrm{CDw}_{\mathrm{forward}}$ did not support a global spanning out-arborescence, and $\mathrm{CDw}_{\mathrm{backward}}$ did not support a global spanning in-arborescence. Hence, the corresponding admissible sets were empty, so neither minimum- nor maximum-weight global spanning arborescences existed in these orientations; this was a structural property of the directed subnetworks.

Table 1: Summary statistics of existing global minimum and maximum spanning arborescences in the four forward–backward subnetworks, computed using FLN as edge weight. Total FLN, depth, mean non-root level width, and standard deviation of non-root level widths are defined in Subsection 2.2.

Network	Extremum type	Direction	Root	Total FLN	Depth	Mean non-root level width	SD
CDw_backward	Minimum	Out	8m	0.023	6	4.667	1.795
CDw_backward	Maximum	Out	8m	8.942	6	4.667	3.197
CDw_forward	Minimum	In	7A	0.004	9	3.111	1.197
CDw_forward	Maximum	In	8m	8.239	7	4.000	2.000
SLN_forward	Minimum	Out	V1	0.804	13	2.154	0.769
SLN_forward	Maximum	Out	V1	11.931	9	3.111	1.663
SLN_forward	Minimum	In	24c	0.002	11	2.545	1.373
SLN_forward	Maximum	In	ProM	9.975	10	2.800	1.720
SLN_backward	Minimum	Out	2	0.002	9	3.111	1.286
SLN_backward	Maximum	Out	9/46d	6.688	14	2.000	1.309
SLN_backward	Minimum	In	10	0.002	15	1.867	1.024
SLN_backward	Maximum	In	V1	6.934	8	3.500	1.936

Of the 16 nominal combinations of subnetwork, orientation, and extremum type, 12 optimal global spanning arborescences existed and were reported in Table 1; the remaining

four corresponded to the two inadmissible CDw-based orientations, for which neither a minimum- nor a maximum-weight global spanning arborescence existed.

To place the total FLN edge weights of the optimal arborescences in the context of the corresponding subnetworks, the full network had total edge weight $W(G_{\text{full}}) = 29.000$. The SLN-based decomposition partitioned this total into $W(G^{\text{SLN}}_{\text{forward}}) = 18.664$ (64.4%) and $W(G^{\text{SLN}}_{\text{backward}}) = 10.336$ (35.6%). For the CDw-based decomposition, $W(G^{\text{CDw}}_{\text{forward}}) = 15.368$ (53.0%) and $W(G^{\text{CDw}}_{\text{backward}}) = 12.114$ (41.8%), while the excluded neutral subgraph had total edge weight $W(G^{\text{CDw}}_{0}) = 1.518$ (5.2%).

For the CDw-based subnetworks, the existing maximum-weight global spanning arborescences had opposite orientations but the same root: the $\text{CDw}_{\text{backward}}$ maximum-weight out-arborescence and the $\text{CDw}_{\text{forward}}$ maximum-weight in-arborescence were both rooted at $8m$, with depths 6 and 7, respectively. Among the SLN-based maximum-weight global spanning arborescences, the $\text{SLN}_{\text{forward}}$ maximum-weight out-arborescence was rooted at $V1$, had $W(T) = 11.931$, and had depth 9, whereas the corresponding maximum-weight in-arborescence was rooted at $ProM$, had $W(T) = 9.975$, and had depth 10. In $\text{SLN}_{\text{backward}}$, the maximum-weight out-arborescence was rooted at $9/46d$ and had depth 14, while the maximum-weight in-arborescence was rooted at $V1$ and had depth 8. The structural descriptors of the global arborescences reported in Table 1 varied across the four subnetworks.

Figure 2 displays one representative maximum-weight global spanning arborescence from each forward–backward subnetwork. The displayed arborescences showed that the edge-restricted subnetworks yielded different global spanning tree organisations. In particular, the CDw-based examples selected $8m$ as their root but had opposite orientations, whereas the displayed SLN-based examples were rooted at $V1$. The $\text{SLN}_{\text{forward}}$ maximum-weight out-arborescence was rooted at the low-level area $V1$ and directed paths outward from it, which was consistent with a low-level source-like role in the anatomical hierarchy. Conversely, the $\text{SLN}_{\text{backward}}$ maximum-weight in-arborescence was rooted at the same area and represented convergence towards this low-level root. These observations supported the separate treatment of the SLN- and CDw-based subnetworks: their differences arose from the retained edge sets.

3.1.4 Rooted node-centric arborescences in the four subnetworks

The node-centric analysis examined how the four edge-restricted subnetworks behaved when the root was prescribed. Unlike global spanning arborescences, node-centric arborescences were computed on the maximal reachable set associated with the chosen root and orientation. Thus, tree order was used as a reachability descriptor, while the main structural comparison concerned the organisation of the reachable vertices.

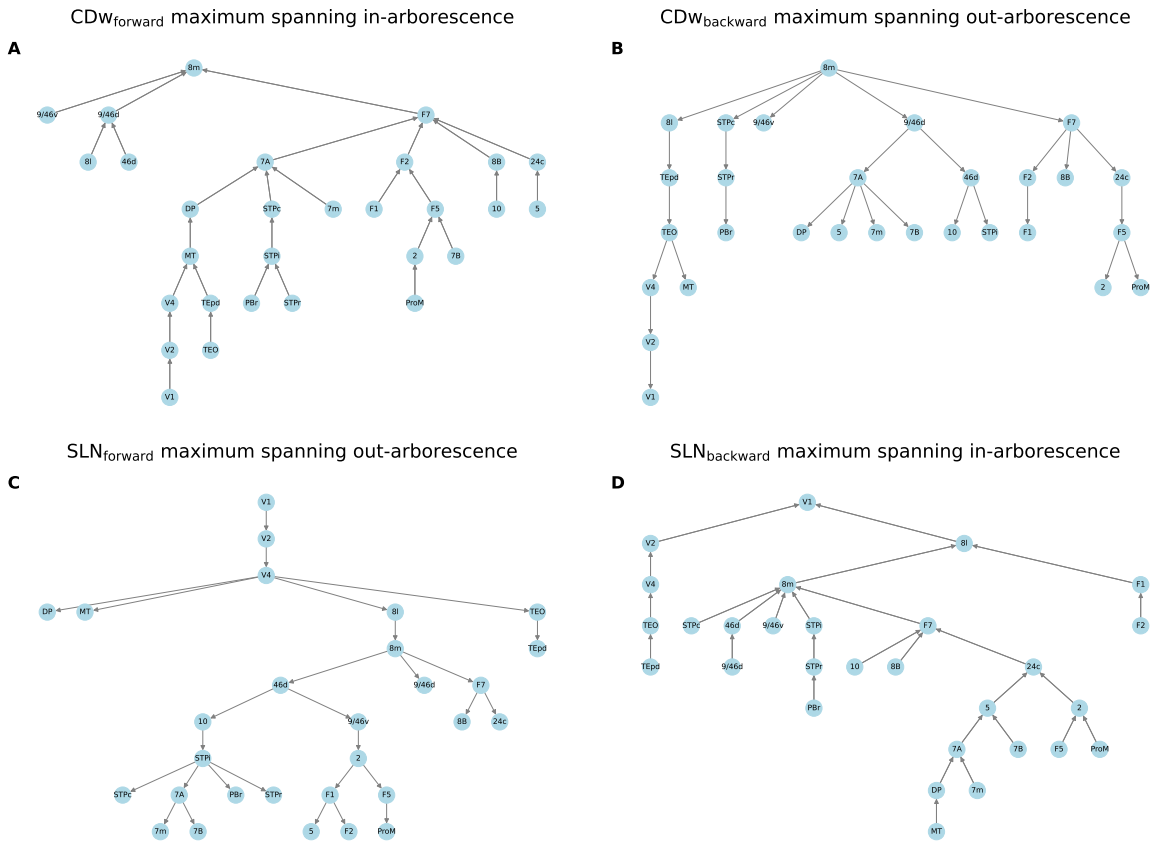

Figure 2: Representative maximum-weight global spanning arborescences computed using FLN as edge weight. (A) $\mathrm{CDw}_{\mathrm{forward}}$ maximum spanning in-arborescence. (B) $\mathrm{CDw}_{\mathrm{backward}}$ maximum spanning out-arborescence. (C) $\mathrm{SLN}_{\mathrm{forward}}$ maximum spanning out-arborescence. (D) $\mathrm{SLN}_{\mathrm{backward}}$ maximum spanning in-arborescence.

Table 2: Summary statistics of node-centric rooted arborescences in the four forward–backward subnetworks, computed over the eight preselected roots using FLN as edge weight. Tree order, relative tree order, total FLN, depth, and mean non-root level width were defined in Subsection 2.2.

Network	Extremum	Direction	Mean order	SD order	Min. order	Max. order	Mean relative order	Mean total FLN	Mean depth	Mean non-root level width
$SLN_{forward}$	Minimum	Out	27.375	0.696	27	29	0.944	0.105	11.375	2.397
$SLN_{forward}$	Minimum	In	22.125	11.910	1	29	0.763	0.102	6.375	2.752
$SLN_{forward}$	Maximum	Out	27.375	0.696	27	29	0.944	10.550	8.875	3.061
$SLN_{forward}$	Maximum	In	22.125	11.910	1	29	0.763	7.351	7.000	2.521
$SLN_{backward}$	Minimum	Out	29.000	0.000	29	29	1.000	0.003	12.375	2.374
$SLN_{backward}$	Minimum	In	29.000	0.000	29	29	1.000	0.002	14.875	1.902
$SLN_{backward}$	Maximum	Out	29.000	0.000	29	29	1.000	6.475	12.875	2.228
$SLN_{backward}$	Maximum	In	29.000	0.000	29	29	1.000	6.593	10.500	2.876
$CDw_{forward}$	Minimum	Out	22.750	1.090	22	25	0.784	1.291	9.000	2.604
$CDw_{forward}$	Minimum	In	18.875	13.081	1	29	0.651	0.266	6.125	2.311
$CDw_{forward}$	Maximum	Out	22.750	1.090	22	25	0.784	5.611	6.625	3.499
$CDw_{forward}$	Maximum	In	18.875	13.081	1	29	0.651	5.211	5.375	2.595
$CDw_{backward}$	Minimum	Out	21.125	8.268	1	28	0.728	0.023	5.750	3.264
$CDw_{backward}$	Minimum	In	19.375	8.817	8	27	0.668	0.162	7.375	2.376
$CDw_{backward}$	Maximum	Out	21.125	8.268	1	28	0.728	5.462	6.875	2.696
$CDw_{backward}$	Maximum	In	19.375	8.817	8	27	0.668	3.800	9.250	2.006

Table 2 showed that the four subnetworks differed in node-centric arborescence order and relative order, total FLN weight, depth, and non-root level width.

For a fixed subnetwork and orientation, the minimum-weight and maximum-weight rows had identical order statistics, as expected: the reachable vertex set was fixed before optimisation, and the minimum–maximum distinction affected only the selected edge set.

The most consistent node-centric reachability was observed in $\mathrm{SLN}_{\mathrm{backward}}$. Both its out- and in-arborescences had mean order 29.000, standard deviation 0.000, and mean relative tree order 1.000. Thus, for the eight preselected roots, $\mathrm{SLN}_{\mathrm{backward}}$ supported full node-centric reachability in both orientations. In $\mathrm{SLN}_{\mathrm{forward}}$, the out-arborescences also covered most of the graph, with mean order 27.375, whereas the in-arborescences were more variable, with mean order 22.125, standard deviation 11.910, and minimum order 1.

The CDw-based subnetworks showed more restricted and more variable node-centric reachability. In $\mathrm{CDw}_{\mathrm{forward}}$, out-arborescences had mean order 22.750, while in-arborescences had mean order 18.875 and standard deviation 13.081. In $\mathrm{CDw}_{\mathrm{backward}}$, both orientations were variable: the mean orders were 21.125 for out-arborescences and 19.375 for in-arborescences. These differences indicated that the four edge-restricted subnetworks imposed different root-dependent reachability constraints.

Figure 3 displays one selected node-centric maximum-weight arborescence from each forward–backward subnetwork, rooted at either the low-level area $V1$ or the high-level area $9/46d$. The CDw-based examples are out-arborescences, whereas the SLN-based examples are in-arborescences, illustrating subnetwork-dependent differences in reachable sets and branching structure.

The selected node-centric examples were consistent with the summary table. The $\mathrm{CDw}_{\mathrm{backward}}$ out-arborescence rooted at $9/46d$ reached 27 non-root vertices, and the $\mathrm{CDw}_{\mathrm{forward}}$ out-arborescence rooted at $V1$ reached 24 non-root vertices. Thus, these CDw-based node-centric examples covered large, but not complete, reachable sets. By contrast, the two selected SLN-based in-arborescences each reached 28 non-root vertices, so their node-centric reachable sets coincided with the full 29-node vertex set.

The branching patterns also differed across the four examples. The $\mathrm{CDw}_{\mathrm{backward}}$ out-arborescence rooted at $9/46d$ branched broadly near the root. The $\mathrm{CDw}_{\mathrm{forward}}$ out-arborescence rooted at $V1$ contained longer directed paths with limited early branching before several terminal vertices were reached. The two SLN-based in-arborescences both spanned the full 29-area vertex set, but differed in their convergent branching organisation towards the root. These examples showed that node-centric tree organisation reflected the combined effects of subnetwork definition, orientation, root-dependent reachability, and maximum-FLN-weight optimisation.

Figure 4 compares selected global and node-centric maximum-weight arborescences by their cumulative tree-level profiles. For each displayed arborescence, the threshold k denotes the maximum tree level included from the root.

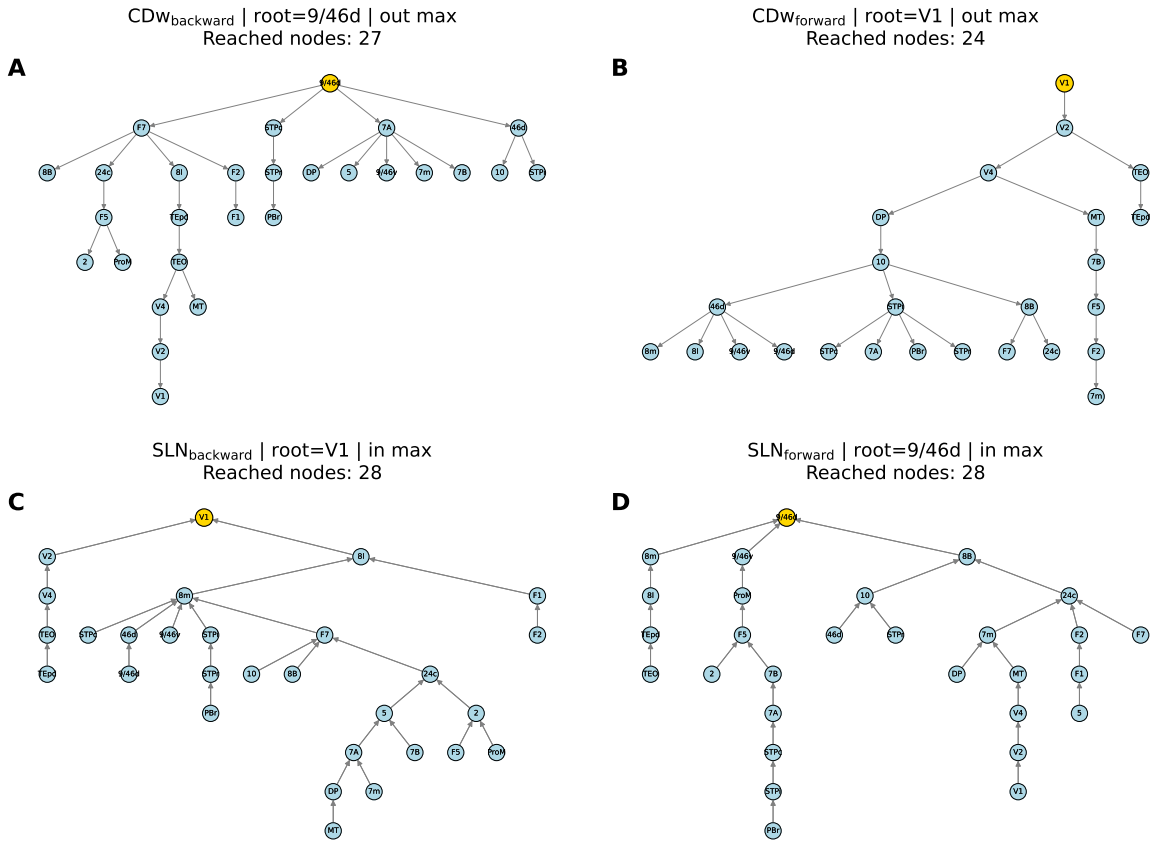

Figure 3: Selected node-centric maximum-weight arborescences computed using FLN as edge weight after fixing the root and reachable vertex set. (A) $\mathrm{CDw}_{\mathrm{backward}}$ out-arborescence rooted at $9/46d$. (B) $\mathrm{CDw}_{\mathrm{forward}}$ out-arborescence rooted at $V1$. (C) $\mathrm{SLN}_{\mathrm{backward}}$ in-arborescence rooted at $V1$. (D) $\mathrm{SLN}_{\mathrm{forward}}$ in-arborescence rooted at $9/46d$. Reported reached-node counts exclude the root, whereas tree order includes the root. In all four displayed trees, the non-root member of the $V1$–$9/46d$ pair occurs as a terminal vertex, reflecting the large hierarchical separation between the selected nodes.

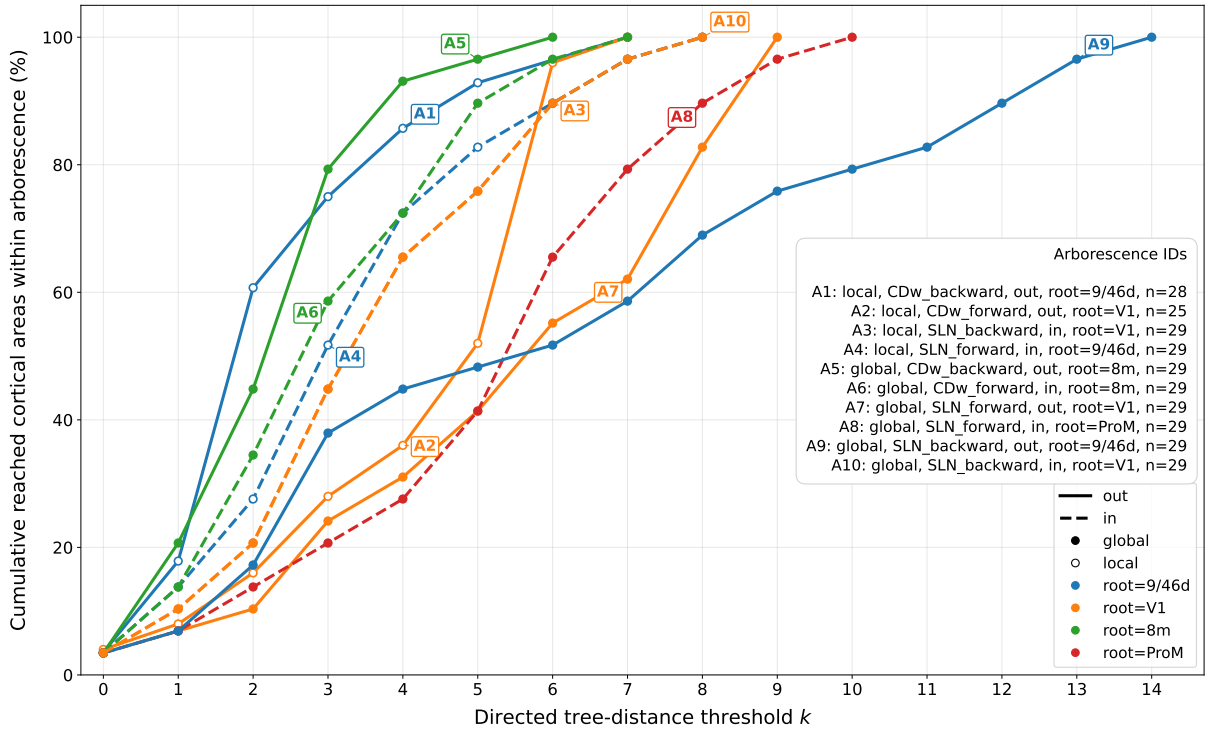

Figure 4: Joint cumulative reachability profiles of selected global and node-centric maximum-weight arborescences. For each curve, k denotes the tree-distance threshold within the corresponding arborescence, and the plotted percentage is the fraction of that tree's vertices located within at most k tree levels of the root. The root is counted at $k = 0$. Solid and dashed curves denote out- and in-arborescences, respectively; filled and open markers denote global and node-centric arborescences, respectively. Labels A1–A10 identify the trees listed in the figure legend.

Because each cumulative profile was normalised by the order of the corresponding tree, reaching 100% did not necessarily indicate coverage of the complete 29-area network for node-centric arborescences. In particular, A1 and A2 contained $|V(T)| = 28$ and $|V(T)| = 25$ vertices, respectively. For A1, the node-centric $\mathrm{CDw}_{\mathrm{backward}}$ out-arborescence rooted at $9/46d$, 60.7% of the vertices in $V(T)$ were at directed tree distance at most $k = 2$ from the root. For A2, the node-centric $\mathrm{CDw}_{\mathrm{forward}}$ out-arborescence rooted at $V1$, the corresponding proportion was 16.0%. By contrast, the global arborescences A5, A6 and A9 all satisfied $|V(T)| = 29$ and had depths 6, 7, and 14, respectively. Taken together, for a fixed 60% cumulative-coverage threshold, the required value of k varied across the displayed arborescences, ranging from 2 to 7 tree levels. Thus, the cumulative profiles complement the summary tables by resolving the distribution of arborescence vertices over directed distances from the root.

Overall, the global and node-centric arborescence analyses supported the separate investigation of the SLN- and CDw-based forward–backward subnetworks, since the four retained edge sets induced distinct directed tree organisations under the same FLN-weighted optimisation procedure.

3.2 Results of feature-vector construction

3.2.1 Strong connectivity of the full network and the four subnetworks

Before computing WDMT-based features, the strongly connected components of the full network and of each edge-restricted subnetwork were examined. As specified in Subsection 2.3.1, these WDMT-based features were computed only within eligible SCCs containing at least seven cortical areas. The numbers and sizes of eligible SCCs are summarised in Table 3.

Table 3: Strong-connectivity summary of the full network and the four edge-restricted subnetworks. Eligible SCCs are defined as strongly connected components containing at least seven cortical areas.

Network	Nodes	Directed edges	SCCs	Largest SCC size	Eligible SCCs	Nodes in eligible SCCs	Fraction of nodes retained	Total eligible SCC weight
full	29	536	1	29	1	29	1.000000	29.000000
$\mathrm{SLN}_{\mathrm{forward}}$	29	265	3	27	1	27	0.931000	16.953070
$\mathrm{SLN}_{\mathrm{backward}}$	29	271	1	29	1	29	1.000000	10.335648
$\mathrm{CDw}_{\mathrm{forward}}$	29	264	7	22	1	22	0.759000	9.577803
$\mathrm{CDw}_{\mathrm{backward}}$	29	246	6	18	2	25	0.862000	4.472487

Notably, the $\mathrm{CDw}_{\mathrm{backward}}$ subnetwork was the only analysed network that contained two eligible SCCs, of sizes 18 and 7. In each of the four edge-restricted subnetworks, more than 75% of the 29 cortical areas belonged to eligible SCCs.

Table 4: Eligible SCCs used for WDMT-based feature computation. The normalised SCC weight is one in networks containing a single eligible SCC; in $CDw_{backward}$, the two eligible SCCs contribute according to their internal FLN-weight shares.

Network	SCC ID	SCC size	Internal SCC weight	Total eligible SCC weight	Normalised SCC weight	Nodes in SCC
full	`full_scc_1`	29	29.000000	29.000000	1.000000	10, 2, 24c, 46d, 5, 7A, 7B, 7m, 8B, 8l, 8m, 9/46d, 9/46v, DP, F1, F2, F5, F7, MT, PBr, ProM, STPc, STPi, STPr, TEO, TEpd, V1, V2, V4
$SLN_{forward}$	`SLN_forward_scc_1`	27	16.953070	16.953070	1.000000	10, 2, 24c, 46d, 5, 7A, 7B, 7m, 8B, 8l, 8m, 9/46d, 9/46v, DP, F1, F2, F5, F7, MT, PBr, ProM, STPc, STPi, STPr, TEO, TEpd, V4
$SLN_{backward}$	`SLN_backward_scc_1`	29	10.335648	10.335648	1.000000	10, 2, 24c, 46d, 5, 7A, 7B, 7m, 8B, 8l, 8m, 9/46d, 9/46v, DP, F1, F2, F5, F7, MT, PBr, ProM, STPc, STPi, STPr, TEO, TEpd, V1, V2, V4
$CDw_{forward}$	`CDw_forward_scc_1`	22	9.577803	9.577803	1.000000	10, 24c, 46d, 7A, 7B, 7m, 8B, 8l, 8m, 9/46d, 9/46v, DP, F2, F5, F7, MT, PBr, STPc, STPi, STPr, TEO, TEpd
$CDw_{backward}$	`CDw_backward_scc_1`	18	3.622531	4.472487	0.809959	10, 46d, 5, 7B, 7m, 8B, DP, F1, F2, F5, MT, PBr, STPi, STPr, TEO, TEpd, V2, V4
$CDw_{backward}$	`CDw_backward_scc_2`	7	0.849955	4.472487	0.190041	24c, 7A, 8l, 9/46d, 9/46v, F7, STPc

Across the five networks, 132 of the $145 = 5 \cdot 29$ node–network rows belonged to eligible SCCs. Consequently, WDMT-based quantities were available for all 29 nodes in the full network and in $\mathrm{SLN}_{\mathrm{backward}}$, for 27 nodes in $\mathrm{SLN}_{\mathrm{forward}}$, for 22 nodes in $\mathrm{CDw}_{\mathrm{forward}}$, and for 25 nodes in $\mathrm{CDw}_{\mathrm{backward}}$.

3.2.2 WDMT-based features within eligible SCCs

The six SCC-scaled WDMT-based features defined in Subsection 2.3.1 were computed for each node belonging to an eligible SCC. In the four networks containing a single eligible SCC, the normalised eligible-SCC weight $\pi_{X,C}$ was 1. In $\mathrm{CDw}_{\mathrm{backward}}$, the 18-node and 7-node eligible SCCs had normalised internal FLN edge-weight shares of 0.810 and 0.190, respectively; this was therefore the only network with nontrivial SCC scaling.

Table 5: Summary of the six SCC-scaled WDMT-based features before logarithmic transformation and within-network standardisation. Values are reported only for nodes belonging to eligible SCCs in the corresponding network.

Network	Feature	Valid nodes	Minimum	Median	Maximum
$\mathrm{CDw}_{\mathrm{backward}}$	Scaled average weighted in-tau	25	2.809×10^{-40}	2.968×10^{-36}	1.469×10^{-11}
$\mathrm{CDw}_{\mathrm{backward}}$	Scaled average weighted out-tau	25	1.020×10^{-37}	1.076×10^{-33}	1.010×10^{-10}
$\mathrm{CDw}_{\mathrm{backward}}$	Scaled unweighted in-tau	25	1.140	1.323×10^{8}	5.286×10^{9}
$\mathrm{CDw}_{\mathrm{backward}}$	Scaled unweighted out-tau	25	3.041	1.353×10^{10}	1.877×10^{11}
$\mathrm{CDw}_{\mathrm{backward}}$	Scaled weighted in-tau	25	1.009×10^{-32}	1.436×10^{-27}	1.467×10^{-8}
$\mathrm{CDw}_{\mathrm{backward}}$	Scaled weighted out-tau	25	3.549×10^{-29}	6.510×10^{-23}	2.502×10^{-8}
$\mathrm{CDw}_{\mathrm{forward}}$	Scaled average weighted in-tau	22	3.208×10^{-39}	2.435×10^{-36}	1.869×10^{-31}
$\mathrm{CDw}_{\mathrm{forward}}$	Scaled average weighted out-tau	22	1.336×10^{-36}	2.976×10^{-34}	5.149×10^{-31}
$\mathrm{CDw}_{\mathrm{forward}}$	Scaled unweighted in-tau	22	1.476×10^{13}	4.732×10^{15}	3.941×10^{17}
$\mathrm{CDw}_{\mathrm{forward}}$	Scaled unweighted out-tau	22	1.108×10^{12}	2.350×10^{14}	1.470×10^{16}
$\mathrm{CDw}_{\mathrm{forward}}$	Scaled weighted in-tau	22	2.205×10^{-24}	1.982×10^{-21}	7.366×10^{-14}
$\mathrm{CDw}_{\mathrm{forward}}$	Scaled weighted out-tau	22	9.691×10^{-24}	9.175×10^{-20}	1.411×10^{-15}
$\mathrm{SLN}_{\mathrm{backward}}$	Scaled average weighted in-tau	29	2.972×10^{-52}	1.409×10^{-50}	2.151×10^{-48}
$\mathrm{SLN}_{\mathrm{backward}}$	Scaled average weighted out-tau	29	2.624×10^{-51}	1.793×10^{-48}	6.602×10^{-47}
$\mathrm{SLN}_{\mathrm{backward}}$	Scaled unweighted in-tau	29	4.979×10^{23}	3.422×10^{24}	4.391×10^{25}
$\mathrm{SLN}_{\mathrm{backward}}$	Scaled unweighted out-tau	29	5.123×10^{23}	5.038×10^{24}	3.102×10^{25}
$\mathrm{SLN}_{\mathrm{backward}}$	Scaled weighted in-tau	29	1.279×10^{-27}	3.251×10^{-26}	9.446×10^{-23}
$\mathrm{SLN}_{\mathrm{backward}}$	Scaled weighted out-tau	29	1.344×10^{-27}	5.928×10^{-24}	6.079×10^{-22}
$\mathrm{SLN}_{\mathrm{forward}}$	Scaled average weighted in-tau	27	1.168×10^{-41}	1.307×10^{-39}	2.795×10^{-37}
$\mathrm{SLN}_{\mathrm{forward}}$	Scaled average weighted out-tau	27	3.027×10^{-39}	2.531×10^{-37}	1.758×10^{-33}
$\mathrm{SLN}_{\mathrm{forward}}$	Scaled unweighted in-tau	27	2.233×10^{21}	7.244×10^{22}	3.984×10^{23}
$\mathrm{SLN}_{\mathrm{forward}}$	Scaled unweighted out-tau	27	7.283×10^{20}	6.818×10^{21}	1.355×10^{23}
$\mathrm{SLN}_{\mathrm{forward}}$	Scaled weighted in-tau	27	2.609×10^{-20}	1.477×10^{-16}	2.337×10^{-14}
$\mathrm{SLN}_{\mathrm{forward}}$	Scaled weighted out-tau	27	2.333×10^{-17}	1.634×10^{-15}	2.382×10^{-10}
full	Scaled average weighted in-tau	29	3.862×10^{-41}	2.757×10^{-40}	1.398×10^{-39}
full	Scaled average weighted out-tau	29	7.945×10^{-40}	4.971×10^{-39}	3.392×10^{-38}
full	Scaled unweighted in-tau	29	5.141×10^{33}	1.158×10^{34}	1.922×10^{34}
full	Scaled unweighted out-tau	29	2.508×10^{33}	4.348×10^{33}	9.060×10^{33}
full	Scaled weighted in-tau	29	3.544×10^{-7}	3.166×10^{-6}	1.078×10^{-5}
full	Scaled weighted out-tau	29	2.972×10^{-6}	2.174×10^{-5}	1.434×10^{-4}

The summary values spanned several orders of magnitude. The unweighted τ features counted rooted spanning arborescences within eligible SCCs and attained very large values. The FLN-weighted and average weighted quantities were substantially smaller because they were based on products of anatomical edge weights. This numerical range

motivated the predefined transformation $x \mapsto \log_{10}(1 + x)$ before within-network standardisation.

3.2.3 Resulting node-level feature table

The unscaled feature table retained all 145 node–network rows. The six SCC-scaled WDMT-based features were undefined in the 13 rows lying outside eligible SCCs; these rows were retained rather than removed. In addition, the network-specific CD-flow feature $\Phi_X(i)$ was undefined in 6 node–network rows.

Following the preprocessing procedure defined in Subsection 2.3.4, a single indicator $I_X^{\tau}(i)$ recorded whether the six SCC-scaled WDMT-based features were available for a row. Undefined entries in the six log-transformed SCC-scaled WDMT feature columns and undefined $\Phi_X(i)$ values were replaced by zero before continuous features were standardised separately within each network. The resulting scaled feature table therefore retained all 145 rows and contained no missing values. The availability indicator was retained as a verification variable and was excluded from the interpreted feature subsets used in the subsequent distance-based analyses. Although the exported scaled feature table contained all five networks, the subsequent MDS and RSA analyses included only the $116 = 4 \cdot 29$ node–subnetwork rows from the four edge-restricted subnetworks.

3.3 Exploratory analysis of node-level feature vectors

3.3.1 Feature-pair correlation heatmaps

An exploratory correlation analysis of feature associations was conducted for the full network and all four edge-restricted subnetworks. However, due to page limitations, only the results obtained for the CDw-defined subnetworks are presented here, while the full-network and SLN-defined heatmaps are provided in the Appendix. The SLN-defined subnetworks also showed prominent feature correlations, although their detailed correlation patterns differed from those observed in the CDw-defined subnetworks.

In $\mathrm{CDw}_{\mathrm{forward}}$, the strongest displayed correlations involving WDMT-based features were observed between unweighted τ and the fixed full-network CDw hierarchy rank h^{CDw} (Figure 5). Unweighted outgoing τ was strongly negatively correlated with h^{CDw} ($r = -0.92$), whereas unweighted incoming τ was strongly positively correlated with the same rank ($r = 0.87$). Thus, the two orientations showed opposite relations with the fixed full-network topology-based hierarchy reference. Their relations with network-specific CD-flow were weaker: $r = 0.36$ for unweighted outgoing τ and $r = 0.01$ for unweighted incoming τ.

In $\mathrm{CDw}_{\mathrm{backward}}$, the largest absolute correlation involving WDMT-based features was again obtained for unweighted τ and h^{CDw} (Figure 6). In this subnetwork, both orientations were negatively correlated with the fixed topological hierarchy rank: $r = -0.67$ for

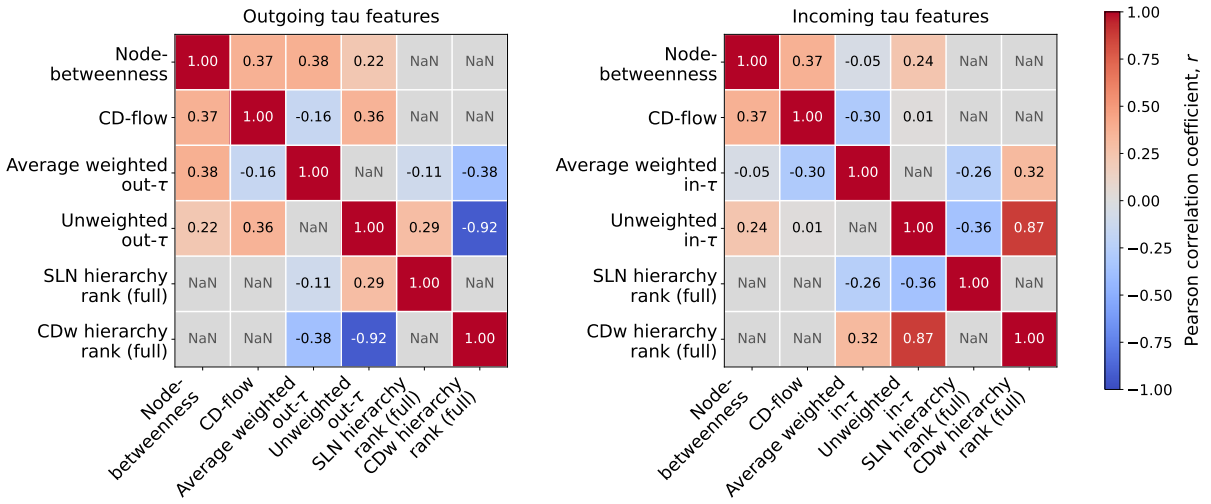

Figure 5: Pearson correlation heatmaps for selected feature pairs in $\mathrm{CDw}_{\mathrm{forward}}$. Outgoing- and incoming-tau relations are displayed separately. Tau quantities are the log-transformed SCC-scaled features defined in Subsection 2.3.1. Grey cells denote feature pairs not included among the selected comparisons.

unweighted outgoing tau and $r = -0.87$ for unweighted incoming tau. Network-specific node betweenness was moderately positively correlated with unweighted outgoing and incoming tau ($r = 0.53$ and $r = 0.49$, respectively), whereas the corresponding correlations with network-specific CD-flow were $r = -0.23$ and $r = -0.48$.

Descriptively, several displayed WDMT-based correlations were more similar in sign and magnitude between orientations in $G^{\mathrm{CDw}}_{\mathrm{backward}}$ than in $G^{\mathrm{CDw}}_{\mathrm{forward}}$.

In $\mathrm{CDw}_{\mathrm{backward}}$, unweighted τ was negatively correlated with the fixed full-network CDw hierarchy rank in both orientations, with $r = -0.67$ for outgoing tau and $r = -0.87$ for incoming tau. The correlations between network-specific node betweenness and average weighted τ were also nearly identical ($r = -0.30$ for outgoing tau and $r = -0.29$ for incoming tau). Similarly, the correlations between network-specific CD-flow and unweighted τ were negative in both orientations, although stronger for incoming tau ($r = -0.23$ for outgoing tau and $r = -0.48$ for incoming tau).

By contrast, in $\mathrm{CDw}_{\mathrm{forward}}$, the correlation between unweighted τ and the fixed full-network CDw hierarchy rank h^{CDw} had opposite signs in the two orientations: unweighted outgoing τ was strongly negatively correlated with the rank ($r = -0.92$), whereas unweighted incoming τ was strongly positively correlated with it ($r = 0.87$). The correlation between network-specific node betweenness and average weighted τ also differed in sign between orientations, from $r = 0.38$ for outgoing τ to $r = -0.05$ for incoming τ. Thus, the displayed WDMT-based relations were descriptively more similar between orientations in $\mathrm{CDw}_{\mathrm{backward}}$, whereas $\mathrm{CDw}_{\mathrm{forward}}$ exhibited a pronounced orientation-dependent sign reversal in the correlation between unweighted τ and the fixed full-network CDw hierarchy rank.

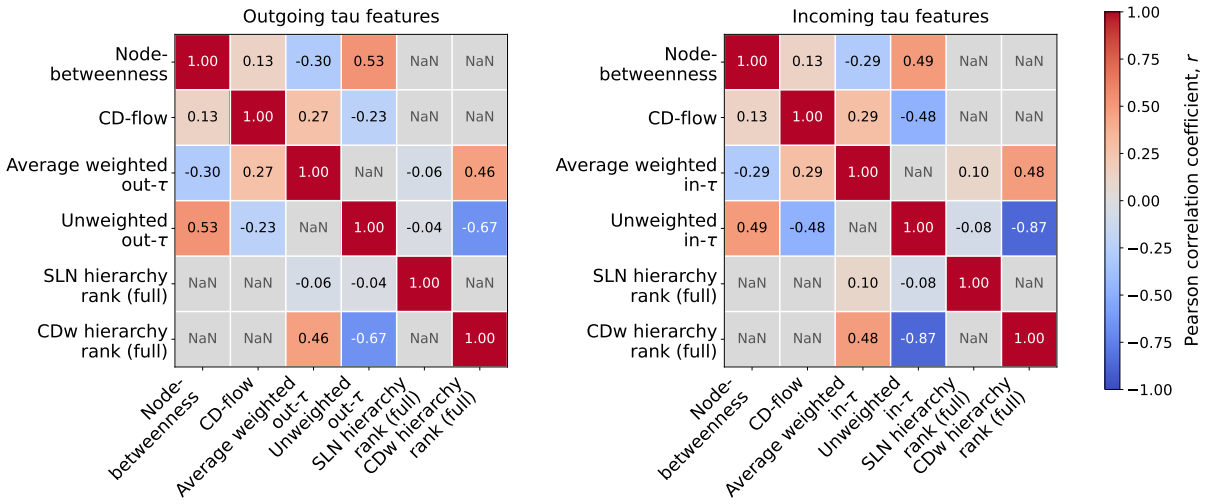

Figure 6: Pearson correlation heatmaps for selected feature pairs in $\mathrm{CDw}_{\mathrm{backward}}$. Outgoing- and incoming-tau relations are displayed separately. Tau quantities are the log-transformed SCC-scaled features defined in Subsection 2.4.1. Grey cells denote feature pairs not included among the selected comparisons.

In the SLN-defined subnetworks, the strongest displayed correlations involving WDMT-based features were observed between unweighted τ and the fixed full-network SLN hierarchy rank. In $\mathrm{SLN}_{\mathrm{backward}}$, these correlations were $r = 0.84$ for outgoing τ and $r = -0.79$ for incoming τ, whereas in $\mathrm{SLN}_{\mathrm{forward}}$ they were $r = -0.77$ and $r = 0.73$, respectively (Figures 13 and 14). By contrast, unweighted τ did not show comparably strong correlations with network-specific node betweenness, network-specific CD-flow, or the fixed full-network CDw hierarchy rank in either SLN-defined subnetwork. An interesting observation was that the network-specific CD-flow and node betweenness were positively correlated in the SLN-defined subnetworks, with $r = 0.76$ in $\mathrm{SLN}_{\mathrm{backward}}$ and $r = 0.72$ in $\mathrm{SLN}_{\mathrm{forward}}$. These correlations were still positive, but weaker in the CDw-defined subnetworks, particularly in $\mathrm{CDw}_{\mathrm{backward}}$ ($r = 0.13$).

3.3.2 Multidimensional scaling of subnetwork feature spaces

The MDS embeddings visualised the pooled node–subnetwork observations from the four edge-restricted subnetworks under correlation and Manhattan dissimilarities, using both the full node-level feature set F_{full} and the reduced WDMT-based rooted-arborescence feature set F_{red} (Figure 7).

With F_{full}, observations from the four subnetworks overlapped substantially under both displayed dissimilarities. Neither the correlation-based nor the Manhattan-based embedding showed complete separation of the four subnetworks. After the network-topological and hierarchy-related features had been excluded in F_{red}, the visual arrangement and dispersion of the observations changed under both displayed dissimilarities. This indicated that the excluded features contributed to the pairwise dissimilarity structure represented

in the MDS embeddings.

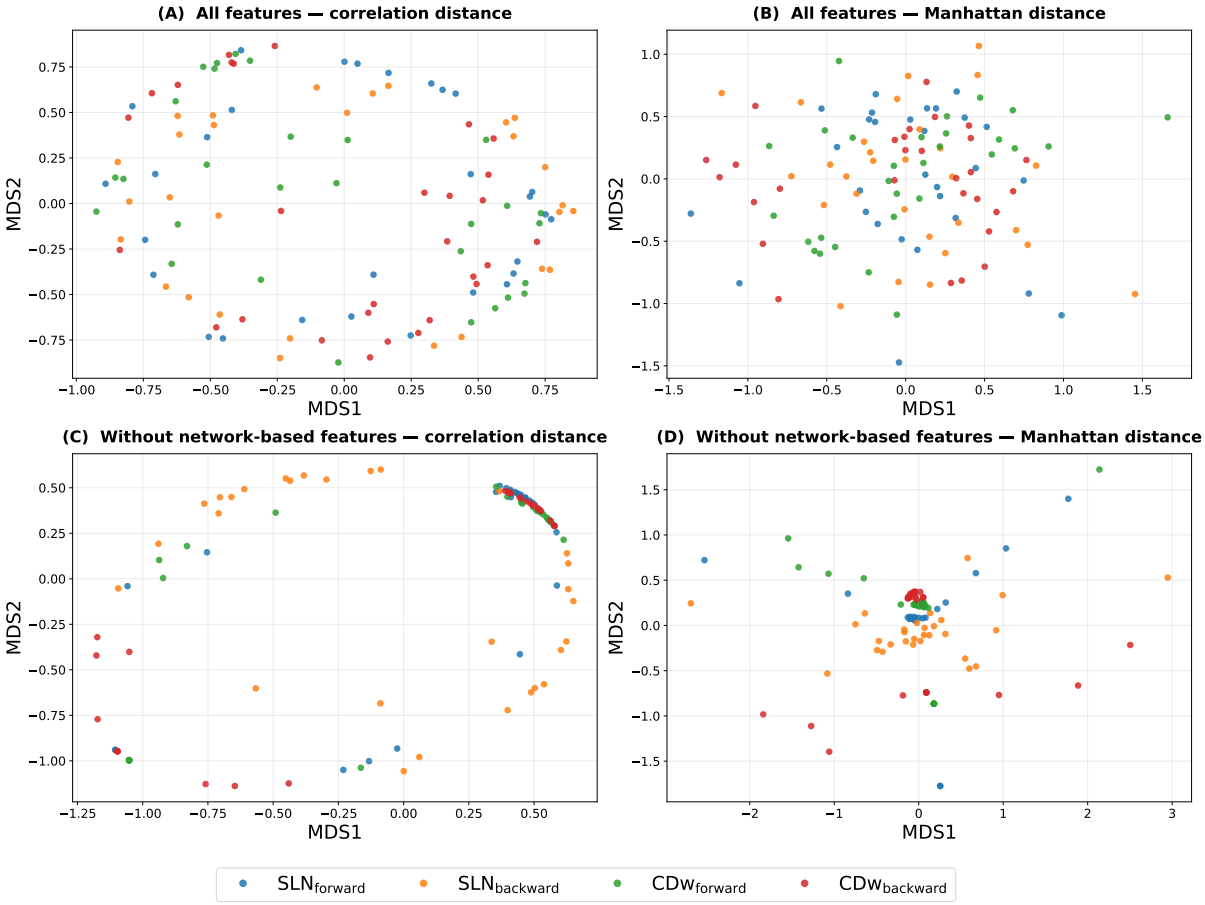

Figure 7: Two-dimensional metric MDS representations of node–subnetwork feature vectors in the four edge-restricted subnetworks. Panels (A) and (B) display the full node-level feature set F_{full} under correlation and Manhattan dissimilarities, respectively. Panels (C) and (D) display the reduced WDMT-based rooted-arborescence feature set F_{red}, obtained after exclusion of the network-topological and hierarchy-related features specified in Subsection 2.4.2. The MDS coordinates are interpreted only through their pairwise arrangement; the axes have no separate substantive meaning.

Exploratory MDS embeddings were computed using correlation, Euclidean, Mahalanobis and Manhattan dissimilarities. For F_{red}, the Euclidean- and Manhattan-based embeddings showed the clearest visual separation of node–subnetwork observations according to subnetwork identity among the examined dissimilarities, whereas the Mahalanobis-based embedding showed a broadly similar but less distinct arrangement. Manhattan dissimilarity was therefore carried forward as the distance measure for the subsequent comparison of vectorised RDMs. The Euclidean- and Mahalanobis-based embeddings are not displayed here.

Correlation and Manhattan dissimilarities were retained for presentation because they paralleled the two dissimilarity constructions subsequently used in the RDM-based analysis: correlation distance was used within each subnetwork to construct the RDMs, whereas Manhattan distance was used between subnetworks to compare their vectorised strict

upper-triangular RDM entries. The visible change in arrangement after exclusion of the network-topological and hierarchy-related features motivated the separate examination of F_{red} in the subsequent RDM comparison and permutation testing.

3.4 RDM comparison and permutation tests

3.4.1 Representational dissimilarity matrices

The RDMs obtained from the full node-level feature set F_{full} are shown in Figure 8. Each matrix contained the pairwise correlation dissimilarities between the standardised node-level feature vectors of the 29 cortical areas within one of the four edge-restricted subnetworks. Each RDM was symmetric, with zeros on its main diagonal, because the pairwise dissimilarity between two cortical areas was independent of their ordering and each area had zero dissimilarity from itself.

The full-feature RDMs formed the basis of three of the six primary inferential tests. The corresponding RDMs for the reduced WDMT-based rooted-arborescence feature set F_{red}, which formed the basis of the remaining three primary tests, are shown in Figure 9. This feature set contained the six log-transformed SCC-scaled WDMT-based rooted-arborescence features after exclusion of the network-topological and hierarchy-related features, as motivated by the earlier exploratory MDS comparison.

Comparison of Figures 8 and 9 shows marked differences in the dissimilarity patterns. The full node-level feature set F_{full} yielded a richer dissimilarity pattern than the reduced feature set F_{red}.

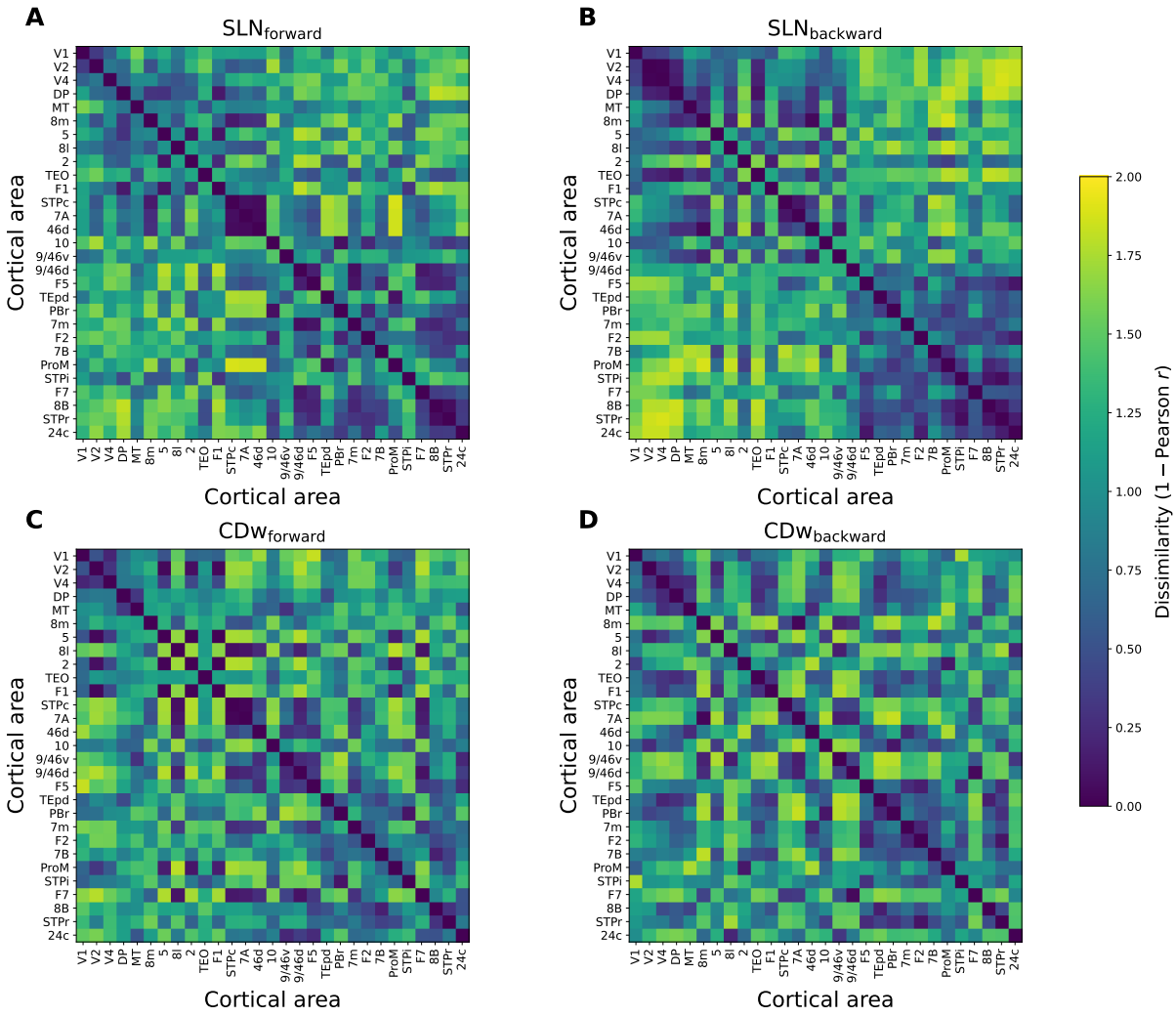

Figure 8: Representational dissimilarity matrices for the four edge-restricted subnetworks based on the full node-level feature set F_{full}. Entries represent $1 - r$ dissimilarity, where r is the Pearson correlation coefficient between the standardised feature vectors of two cortical areas. Each RDM is symmetric with a zero main diagonal. Panels display (A) SLN$_{\text{forward}}$, (B) SLN$_{\text{backward}}$, (C) CDw$_{\text{forward}}$, and (D) CDw$_{\text{backward}}$. All panels use the same dissimilarity scale.

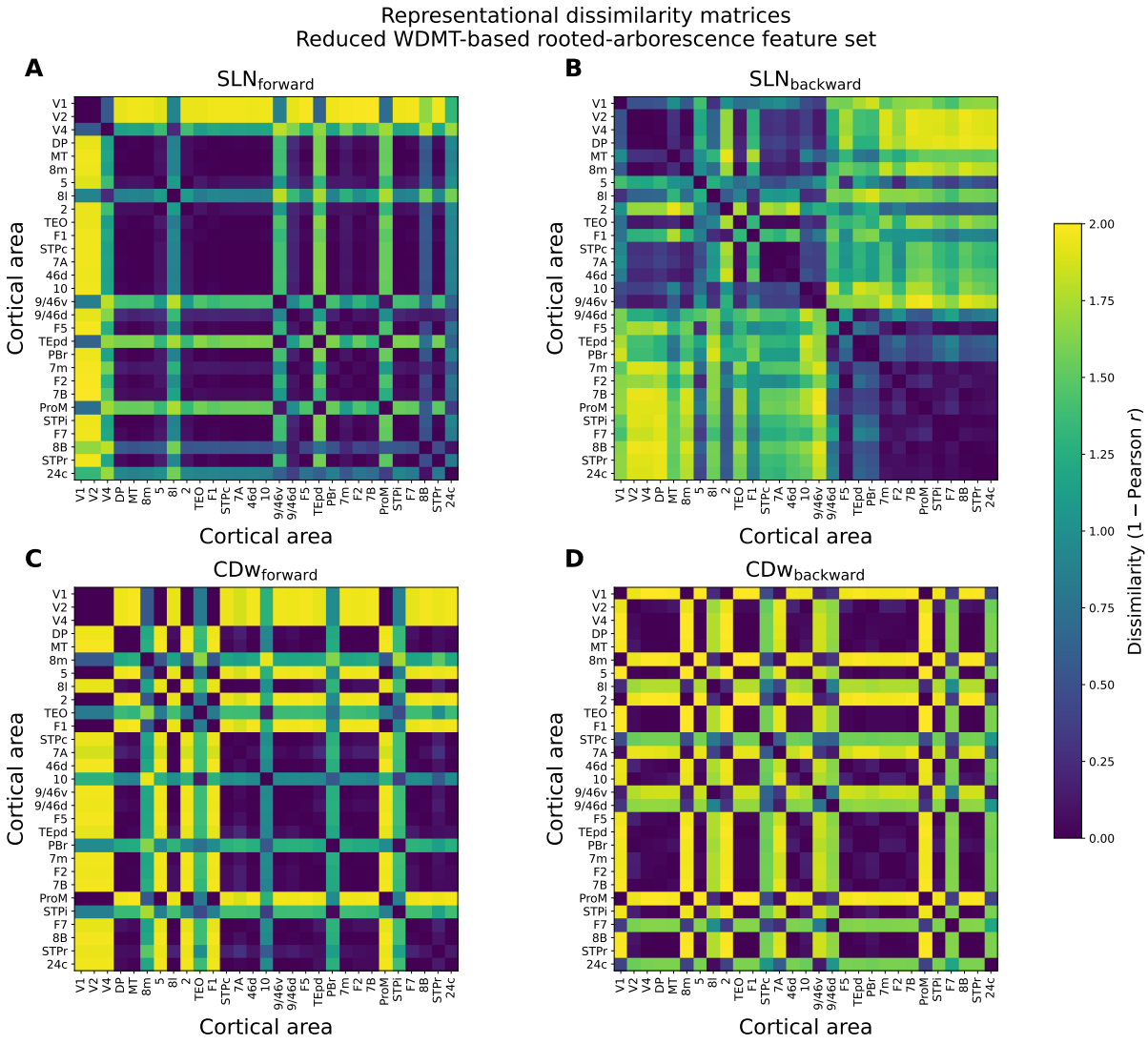

Figure 9: Representational dissimilarity matrices for the four edge-restricted subnetworks based on the reduced WDMT-based rooted-arborescence feature set F_{red}. Entries represent $1 - r$ dissimilarity, where r is the Pearson correlation coefficient between the standardised feature vectors of two cortical areas. Each RDM is symmetric with a zero main diagonal. Panels display (A) $\text{SLN}_{\text{forward}}$, (B) $\text{SLN}_{\text{backward}}$, (C) $\text{CDw}_{\text{forward}}$, and (D) $\text{CDw}_{\text{backward}}$. The dissimilarity scale is identical to that used for the full-feature RDMs.

3.4.2 Monte Carlo permutation tests

For each feature set $F \in \{F_{\text{full}}, F_{\text{red}}\}$, each subnetwork-specific RDM was represented by its 406 unique strict upper-triangular dissimilarity entries, and Manhattan distances were computed between the vectorised RDMs. The matched-direction distances compared $\text{SLN}_{\text{forward}}$ with $\text{CDw}_{\text{forward}}$ and $\text{SLN}_{\text{backward}}$ with $\text{CDw}_{\text{backward}}$, whereas the within-classification reference distance was the mean of the forward–backward distances for the SLN- and CDw-defined pairs. The forward-specific and backward-specific correspondence statistics contrasted this reference distance with the corresponding matched-direction distance, while the combined statistic contrasted it with the mean matched-direction distance. Larger values therefore represented stronger matched-direction correspondence relative to the within-classification reference.

For the full node-level feature set F_{full}, the observed correspondence statistics and corresponding one-sided Monte Carlo permutation p-values were

$$S_{\text{forward}}^{(F_{\text{full}})} = 3.576, \qquad p_{\text{perm}} = 0.0001,$$

$$S_{\text{backward}}^{(F_{\text{full}})} = -25.836, \qquad p_{\text{perm}} = 0.0003,$$

and

$$S_{\text{combined}}^{(F_{\text{full}})} = -11.130, \qquad p_{\text{perm}} = 0.0001.$$

All three observed statistics lay in the extreme right tails of their respective Monte Carlo permutation distributions (Figure 10). Thus, at significance level $\alpha = 0.05$, the corresponding null hypotheses of arbitrary cortical-area alignment between the SLN-based and CDw-based RDMs were rejected for all three full-feature primary tests.

The positive forward statistic indicated that the observed $\text{SLN}_{\text{forward}}$–$\text{CDw}_{\text{forward}}$ RDM distance was smaller than the within-classification forward–backward reference distance and was also smaller than expected under random cortical-area alignment. By contrast, the negative backward statistic indicated that the observed $\text{SLN}_{\text{backward}}$–$\text{CDw}_{\text{backward}}$ RDM distance was larger than the within-classification reference distance. Nevertheless, its significant right-tailed permutation result indicated that the backward matched-direction distance was still smaller than expected under random cortical-area alignment. The combined statistic was also negative because the negative backward contribution outweighed the positive forward contribution. Therefore, the full-feature analysis identified non-random matched-direction correspondence for all three full-feature statistics, although only the forward matched-direction distance was smaller than the within-classification forward–backward reference distance.

The reduced WDMT-based rooted-arborescence feature set F_{red} was examined separately. Its observed statistics and corresponding one-sided Monte Carlo permutation

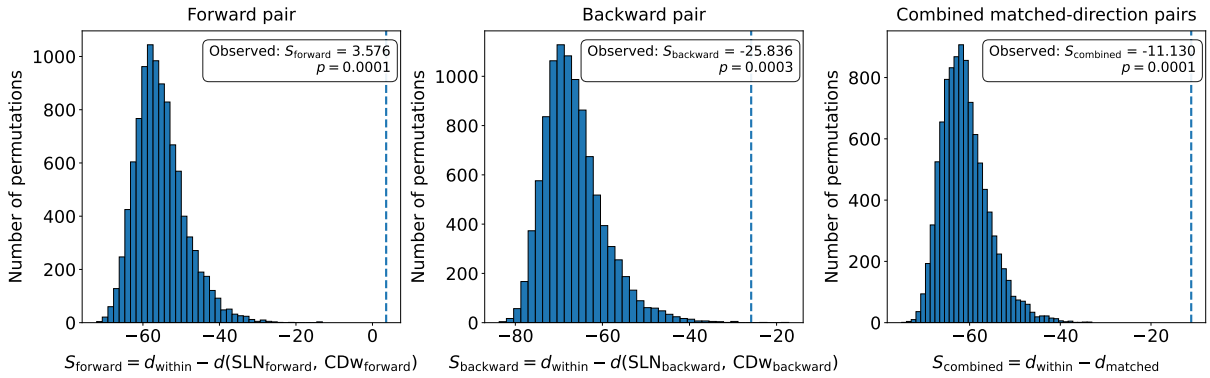

Figure 10: One-sided Monte Carlo permutation distributions for the full node-level feature set F_{full}. Dashed vertical lines indicate the observed statistics. The forward-specific, backward-specific, and combined matched-direction statistics constitute three of the six primary inferential tests. For each statistic, larger values indicate stronger matched-direction correspondence relative to randomly permuted cortical-area alignments. Numerical permutation p-values are reported within the panels.

p-values were

$$S^{(F_{\text{red}})}_{\text{forward}} = 15.815, \qquad p_{\text{perm}} = 0.0968,$$

$$S^{(F_{\text{red}})}_{\text{backward}} = -9.805, \qquad p_{\text{perm}} = 0.2652,$$

and

$$S^{(F_{\text{red}})}_{\text{combined}} = 3.005, \qquad p_{\text{perm}} = 0.0927.$$

The positive forward and combined statistics indicated that the corresponding matched-direction RDM distances were smaller than the within-classification forward–backward reference distance, whereas the negative backward statistic indicated the opposite pattern for the backward pair. However, none of the three reduced-feature primary tests reached significance at $\alpha = 0.05$. Thus, unlike the full-feature analysis, the reduced-feature analysis did not provide statistical evidence for matched-direction correspondence beyond that expected under random cortical-area alignment.

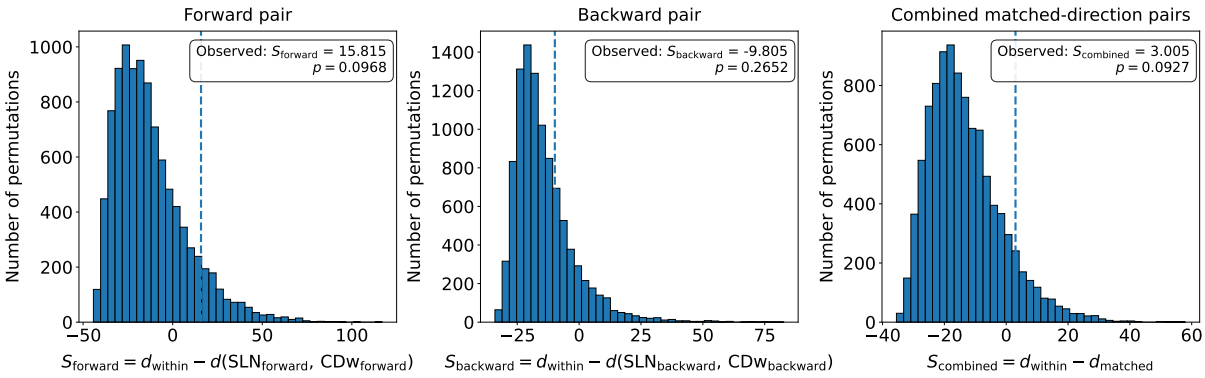

Figure 11: One-sided Monte Carlo permutation distributions for the reduced WDMT-based rooted-arborescence feature set F_{red}. Dashed vertical lines indicate the observed statistics. The forward-specific, backward-specific, and combined matched-direction statistics constitute three of the six primary inferential tests. Numerical permutation p-values are reported within the panels.

Table 6: Observed matched-direction correspondence statistics and one-sided Monte Carlo permutation p-values. For each of the full node-level feature set F_{full} and the reduced WDMT-based rooted-arborescence feature set F_{red}, the forward-specific, backward-specific, and combined statistics are treated as primary inferential statistics, giving six primary tests in total.

Feature set	Statistic	Matched comparison	Observed S	p_{perm}
Full node-level feature set	Forward	$\mathrm{SLN}_{\mathrm{forward}}$ – $\mathrm{CDw}_{\mathrm{forward}}$	3.576	0.0001
Full node-level feature set	Backward	$\mathrm{SLN}_{\mathrm{backward}}$ – $\mathrm{CDw}_{\mathrm{backward}}$	-25.836	0.0003
Full node-level feature set	Combined	Matched forward and backward pairs	-11.130	0.0001
Reduced WDMT-based rooted-arborescence feature set	Forward	$\mathrm{SLN}_{\mathrm{forward}}$ – $\mathrm{CDw}_{\mathrm{forward}}$	15.815	0.0968
Reduced WDMT-based rooted-arborescence feature set	Backward	$\mathrm{SLN}_{\mathrm{backward}}$ – $\mathrm{CDw}_{\mathrm{backward}}$	-9.805	0.2652
Reduced WDMT-based rooted-arborescence feature set	Combined	Matched forward and backward pairs	3.005	0.0927

4 Discussion

4.1 Principal findings

Using network pathfinding and directed weighted graph analysis, the present study provides a novel systems-level view of the structural foundations of distributed—that is, parallel and segregated—as well as integrative cortical processing. More specifically, it examined elementary structural features of forward and backward subnetworks, whose organisation is crucial for understanding hierarchical processing in the cerebral cortex. Two complementary definitions of forward–backward organisation were applied to the same anatomical network: an anatomical definition based on the laminar origin of projections, quantified by SLN, and a topology-based definition based on weighted convergence degree, quantified by CDw. Accordingly, the cortical network was decomposed into anatomically defined and topology-based forward and backward subnetworks, with the SLN-based decomposition assigning all projections to an anatomical forward or backward class and the CDw-based decomposition excluding neutral edges from both topology-based subnetworks. The weak correlations between the fixed full-network SLN and CDw hierarchy ranks further indicated that the anatomical SLN-based definition and the topological CDw-based definition captured distinct aspects of cortical hierarchical organisation and should therefore be regarded as complementary, rather than interchangeable, definitions of forward–backward organisation.

Interestingly, despite restriction of the cortical network to forward or backward edge classes, all four subnetworks retained substantial strongly connected structure, with their eligible strongly connected components collectively containing more than three quarters of the cortical areas in every case. Thus, anatomically and topologically defined forward and backward subnetworks were not fragmented collections of isolated directed pathways, but preserved recurrent directed accessibility across large parts of the network. Since strong connectivity forms a structural basis for integrated functioning [14], its preservation within both forward and backward subnetworks is biologically relevant. To gain deeper insight into directed hierarchical cortical organisation, global spanning arborescences were compared across the four subnetworks in both outward and inward orientations. Of the sixteen nominal global minimum- and maximum-weight arborescences, twelve existed, while the four absent cases all occurred in the topology-based subnetworks: $\mathrm{CDw}_{\mathrm{forward}}$ did not support a global spanning out-arborescence, whereas $\mathrm{CDw}_{\mathrm{backward}}$ did not support a global spanning in-arborescence. These orientation-specific absences indicate that the topology-based decomposition imposed stronger directional constraints on whole-network spanning organisation than the anatomical decomposition.

Beyond the existence of global spanning arborescences, their architecture and the node-centric trees revealed differences in how the subnetworks organised directed reachability. Despite the constraints imposed by edge restriction and FLN-based optimisation, the se-

lected example trees followed cortical hierarchy approximately, and hierarchical ordering appeared more prevalent in the node-centric trees constructed for the preselected roots. One may also note tendencies in the differences between arborescences of the different subnetworks and orientations: among the reported global arborescences, the admissible CDw arborescences had larger mean width than the SLN arborescences, and width varied more along depth in the maximum-weight than in the corresponding minimum-weight arborescence set. Although exceptions occurred, forward arborescences and inward orientations tended to have larger depth than backward arborescences and outward orientations, while the inward node-centric arborescences of the two forward subnetworks showed very large variance in tree order. These features may be considered structural bases of computational length, represented by tree depth, and of integration, represented by tree width through convergence and divergence. In this structural sense, the $\mathrm{CDw}_{\mathrm{backward}}$ and outward arborescence organisations were interpreted as more efficient than the other trees, since information can reach the root or spread from it through fewer levels when the tree is broader.

Regarding node-level network organisation, the hierarchy measure corresponding to the definition of each subnetwork family was highly correlated with arborescence-related τ features (especially with unweighted τ), either positively or negatively depending on the orientation. Comparing the similarity of the subnetworks, the full feature set provided a richer pattern at the level of cortical areas than the reduced feature set. Statistical comparisons supported this observation by detecting significant matched-direction correspondence between the anatomically and topologically defined organisations only when the full feature set was used, that is, when network-topological and hierarchy-related indices were included in addition to the WDMT-based arborescence features. It is important to note that the subnetworks differed according to both means of selecting the edges, anatomical versus topological and forward versus backward. Notably, in contrast to the permutation-test results, visual clustering of node–subnetwork feature vectors emerged in the MDS plots using the reduced feature set. Decreased dissimilarity in the reduced feature space could be responsible for this unexpected difference between the MDS visualisation and the permutation analysis.

In conclusion, the structural heterogeneity shown at the subnetwork level allows a highly flexible functional organisation of cortical processing. More specifically, from the perspective of the aim of the present study, it is important to stress that the structural basis of forward and backward processing was markedly different irrespective of whether forward and backward connections were defined anatomically or topologically. Thus, although the SLN-based and CDw-based definitions captured distinct and complementary aspects of cortical hierarchical organisation, both revealed substantial structural differences between forward and backward subnetworks.

4.2 Mathematical applicability and limitations of the analytical framework

The directed and weighted graph representation was appropriate for the present study because cortical projections have direction and experimentally quantified anatomical strength. The SLN-based decomposition selected edges according to laminar projection characteristics and the CDw-based decomposition selected edges according to weighted path convergence–divergence organisation [3, 1, 7, 4]. In all subsequent arborescence analyses, retained edges continued to carry their anatomical FLN weights. Consequently, differences between the resulting trees reflected the effect of the two different forward–backward edge-selection under a common anatomical weighting scheme.

The analytical methods characterised complementary aspects of subnetwork organisation. Strongly connected components identified recurrently accessible domains, whereas global and node-centric arborescences described cycle-free outward and inward directed organisations permitted by each subnetwork. Minimum- and maximum-weight trees represented extremal FLN-weighted organisations within the retained edge sets, not actual routes of cortical activity. WDMT-based quantities complemented these selected trees by summarising the multiplicity and FLN weighting of rooted spanning arborescences within eligible strongly connected components [2].

The feature-vector analyses extended this structural comparison to node-level organisation. Correlation heatmaps and MDS were exploratory: correlations involving fixed hierarchy ranks were descriptive relations with ordinal reference variables, while the two-dimensional MDS embeddings necessarily approximated and may have distorted the original higher-dimensional dissimilarity structure and therefore could not be treated as inferential evidence [8]. Following the RDM-based comparison principle of Kriegeskorte et al. (2008) [8], RDM comparison with Monte Carlo permutation testing assessed whether matched SLN- and CDw-defined subnetworks showed non-random correspondence in their cortical-area dissimilarity structures. Significant results therefore indicated correspondence in the analysed structural feature representations, not functional equivalence of the subnetworks. The six permutation tests addressed separately specified comparison questions and were evaluated individually at significance level $\alpha = 0.05$. Since each p-value was interpreted only with respect to its corresponding hypothesis, and no single joint inferential decision was made based on the complete set of tests, no multiple-testing correction was applied.

Several limitations should be considered. The analysis concerned a quantified 29-area macaque cortical subnetwork rather than the complete cortical network. In addition, WDMT-based quantities were computed only within eligible strongly connected components, whose sizes and compositions differed between subnetworks. For distance-based analyses, unavailable WDMT-based values and undefined network-specific CD-flow values

were coded as zero before standardisation; therefore, parts of the MDS and RDM dissimilarity structure may reflect both structural differences and the adopted preprocessing rule.

4.3 Biological relevance and potential future applications

The findings of the present study, showing that forward and backward subnetworks are characterised by different arborescence organisations, offer a new perspective on the large-scale organisation of hierarchical cortical processing. Based on previous findings, cortical hierarchy has been considered as a counterstream architecture formed by feedforward and feedback pathways through cortical areas [12, 7]. However, it remains unknown how this counterstream is organised at the level of directed network structure. The present findings suggest that forward–backward interactions are supported by forward and backward subnetworks with different topological organisations. Based on this observation, we propose that forward and backward pathways, represented by arborescences in the present study, differently influence large-scale cortical activity.

This interpretation is further supported by the qualitative observations concerning the architecture of the selected example arborescences. In particular, the broader, more divergent or convergent $\mathrm{CDw}_{\mathrm{backward}}$ trees tended to support a more efficient structural organisation of signal flow than forward trees containing longer directed chains. This observation is consistent with the suggestion that, during the evolution of the cerebral cortex, selection pressure was greater on high-level "sink" areas coordinating signal flow in the network [3, 1]. Accordingly, feedback connections predominantly originate from higher-level areas in the cortical hierarchy [12, 7].

Insufficient systems-level knowledge of the brain, and particularly of the cerebral cortex, remains a major obstacle to understanding neuropsychiatric illness. Friston (2023) [14] not only states that psychosis is caused by cortical disconnection, but he proposes that the illness originates from the inappropriate interactions between forward and backward connections. The prefrontal cortex is a key region in psychiatric disorders and is also a major source of feedback projections, with an important role in cognitive control [13, 14]. The approach outlined in the present study may therefore provide a fruitful starting point for future investigations aimed at a better systems-level understanding of hierarchical brain functions and their clinical relevance.

4.4 Conclusions

The four research questions formulated in the Introduction can be answered as follows.

1. The four subnetworks differed in their capacity for global directed spanning organisation. Both SLN-defined subnetworks supported global spanning out- and in-arborescences, whereas $\mathrm{CDw}_{\mathrm{forward}}$ lacked a global spanning out-arborescence and

$\mathrm{CDw}_{\mathrm{backward}}$ lacked a global spanning in-arborescence. The admissible minimum- and maximum-weight trees also differed in their roots and architecture. Moreover, the absence of one global spanning orientation in each CDw-defined subnetwork indicated stronger orientation-specific constraints than in the SLN-defined subnetworks.

2. Root-specific organisation also differed across subnetworks. $\mathrm{SLN}_{\mathrm{backward}}$ supported complete node-centric reachability in both orientations for all selected roots, whereas the other subnetworks showed more restricted or variable reachability, particularly the CDw-defined subnetworks. WDMT-based quantities could nevertheless be computed for large eligible strongly connected components in all four subnetworks, confirming substantial retained recurrent organisation.

3. Exploratory feature analyses showed that WDMT-based rooted-arborescence quantities were most strongly related to the hierarchy rank corresponding to the defining criterion of the given subnetwork family: SLN ranks in the SLN-defined subnetworks and CDw ranks in the CDw-defined subnetworks. The MDS representations further indicated that the network-topological and hierarchy-related features materially altered the node–subnetwork dissimilarity structure relative to the reduced WDMT-only configuration.

4. RDM-based permutation testing supported matched-direction correspondence only under the full feature configuration. For F_{full}, the forward matched pair was closer than the within-classification forward–backward reference and showed significant non-random alignment. The backward and combined statistics also showed significant non-random alignment, although their matched-direction distances were not smaller than the within-classification reference. For F_{red}, none of the three tests reached significance.

Overall, the results showed that SLN-based and CDw-based classifications captured complementary aspects of forward–backward cortical organisation. Both separated structurally different forward and backward subnetworks. In the present RDM-based analysis, significant matched-direction correspondence between the two classifications was detected under the full feature configuration, but not when the comparison was restricted to the WDMT-based rooted-arborescence features.

Acknowledgements

I would like to thank my external supervisor, László Négyessy, for his help throughout my BSc thesis project. I am especially grateful for the time he spent explaining the biological background of the research, for always answering my emails and questions quickly, sometimes even late at night, and for the weekly meetings, during which his enthusiasm for the project was truly motivating. I would also like to thank Bálint Varga, who always participated in these meetings, contributed valuable biological and mathematical insights, and shared the code from his previous research project with me. Finally, I would like to thank my internal supervisor, József Pintér, for helping with the mathematical revision of my thesis.

References

[1] M. Bányai, L. Négyessy, and F. Bazsó. Organization of signal flow in directed networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2011(06):P06001, 2011. doi:10.1088/1742-5468/2011/06/P06001.

[2] P. De Leenheer. An elementary proof of a matrix tree theorem for directed graphs. *SIAM Review*, 62(3):716–726, 2020. doi:10.1137/19M1265193.

[3] L. Négyessy, T. Nepusz, L. Zalányi, and F. Bazsó. Convergence and divergence are mostly reciprocated properties of the connections in the network of cortical areas. *Proceedings of the Royal Society B: Biological Sciences*, 275:2403–2410, 2008. doi:10.1098/rspb.2008.0629.

[4] B. Varga, B. Soós, B. Jákli, E. Bálint, Z. Somogyvári, and L. Négyessy. Network path convergence shapes low-level processing in the visual cortex. *Frontiers in Systems Neuroscience*, 15:645709, 2021. doi:10.3389/fnsys.2021.645709.

[5] S. X. Zhao and F. Y. Ye. Exploring the directed h-degree in directed weighted networks. *Journal of Informetrics*, 6:619–630, 2012. doi:10.1016/j.joi.2012.06.007.

[6] A. Korn, A. Schubert, and A. Telcs. Lobby index in networks. *Physica A: Statistical Mechanics and its Applications*, 388(11):2221–2226, 2009. doi:10.1016/j.physa.2009.02.013.

[7] N. T. Markov, J. Vezoli, P. Chameau, A. Falchier, R. Quilodran, C. Huissoud, C. Lamy, P. Misery, P. Giroud, S. Ullman, P. Barone, C. Dehay, K. Knoblauch, and H. Kennedy. Anatomy of hierarchy: Feedforward and feedback pathways in macaque visual cortex. *The Journal of Comparative Neurology*, 522:225–259, 2014. doi:10.1002/cne.23458.

[8] N. Kriegeskorte, M. Mur, and P. A. Bandettini. Representational similarity analysis – connecting the branches of systems neuroscience. *Frontiers in Systems Neuroscience*, 2:4, 2008. doi:10.3389/neuro.06.004.2008.

[9] I. Borg and P. J. F. Groenen. *Modern Multidimensional Scaling: Theory and Applications*. Springer, New York, 2nd edition, 2005.

[10] B. Phipson and G. K. Smyth. Permutation p-values should never be zero: Calculating exact p-values when permutations are randomly drawn. *Statistical Applications in Genetics and Molecular Biology*, 9(1):Article 39, 2010. doi:10.2202/1544-6115.1585.

[11] D. J. Felleman and D. C. Van Essen. Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex*, 1:1–47, 1991. doi:10.1093/cercor/1.1.1-a.

[12] N. T. Markov, M. Ercsey-Ravasz, D. C. Van Essen, K. Knoblauch, Z. Toroczkai, and H. Kennedy. Cortical high-density counterstream architectures. *Science*, 342:1238406, 2013. doi:10.1126/science.1238406.

[13] L. Négyessy, M. Bányai, T. Nepusz, and F. Bazsó. What makes the prefrontal cortex so appealing in the era of brain imaging? A network analytical perspective. *Acta Biologica Hungarica*, 63(Suppl. 1):38–53, 2012. doi:10.1556/ABiol.63.2012.Suppl.1.5.

[14] K. J. Friston. Computational psychiatry: from synapses to sentience. *Molecular Psychiatry*, 28:256–268, 2023. doi:10.1038/s41380-022-01743-z.

A Dictionary of WDMT-based rooted-arborescence features

Table 7: WDMT tau quantities computed from rooted directed spanning arborescences. Unscaled quantities are component-level intermediate measures retained for interpretation, while the six SCC-scaled quantities were included in the node-level feature vector before transformation and standardisation.

Feature	Notation	Orient.	Weighting	SCC-scaled	Short description
Unweighted out-tau	$\tau^{\mathrm{out}}_{X,i,C}$	Out	Unweighted	No	Number of outgoing directed spanning arborescences rooted at node i within eligible SCC C of G_X.
Unweighted in-tau	$\tau^{\mathrm{in}}_{X,i,C}$	In	Unweighted	No	Number of incoming directed spanning arborescences rooted at node i within eligible SCC C of G_X.
Weighted out-tau	$\tau^{\mathrm{out},w}_{X,i,C}$	Out	FLN	No	Sum of FLN-weight products over outgoing directed spanning arborescences rooted at i within C.
Weighted in-tau	$\tau^{\mathrm{in},w}_{X,i,C}$	In	FLN	No	Sum of FLN-weight products over incoming directed spanning arborescences rooted at i within C.
Average weighted out-tau	$\overline{\tau}^{\mathrm{out},w}_{X,i,C}$	Out	FLN average	No	Weighted out-tau divided by unweighted out-tau within eligible SCC C.
Average weighted in-tau	$\overline{\tau}^{\mathrm{in},w}_{X,i,C}$	In	FLN average	No	Weighted in-tau divided by unweighted in-tau within eligible SCC C.

Feature	Notation	Orient.	Weighting	SCC-scaled	Short description
Scaled unweighted out-tau	$\tau^{\mathrm{out}}_{X,i,\mathrm{sc}}$	Out	Unweighted	Yes	Unweighted out-tau multiplied by the normalised internal SCC weight $\pi_{X,C}$.
Scaled unweighted in-tau	$\tau^{\mathrm{in}}_{X,i,\mathrm{sc}}$	In	Unweighted	Yes	Unweighted in-tau multiplied by the normalised internal SCC weight $\pi_{X,C}$.
Scaled weighted out-tau	$\tau^{\mathrm{out},w}_{X,i,\mathrm{sc}}$	Out	FLN	Yes	Weighted out-tau multiplied by the normalised internal SCC weight $\pi_{X,C}$.
Scaled weighted in-tau	$\tau^{\mathrm{in},w}_{X,i,\mathrm{sc}}$	In	FLN	Yes	Weighted in-tau multiplied by the normalised internal SCC weight $\pi_{X,C}$.
Scaled average weighted out-tau	$\overline{\tau}^{\mathrm{out},w}_{X,i,\mathrm{sc}}$	Out	FLN average	Yes	Average weighted out-tau multiplied by the normalised internal SCC weight $\pi_{X,C}$.
Scaled average weighted in-tau	$\overline{\tau}^{\mathrm{in},w}_{X,i,\mathrm{sc}}$	In	FLN average	Yes	Average weighted in-tau multiplied by the normalised internal SCC weight $\pi_{X,C}$.

B Feature-pair correlation heatmaps

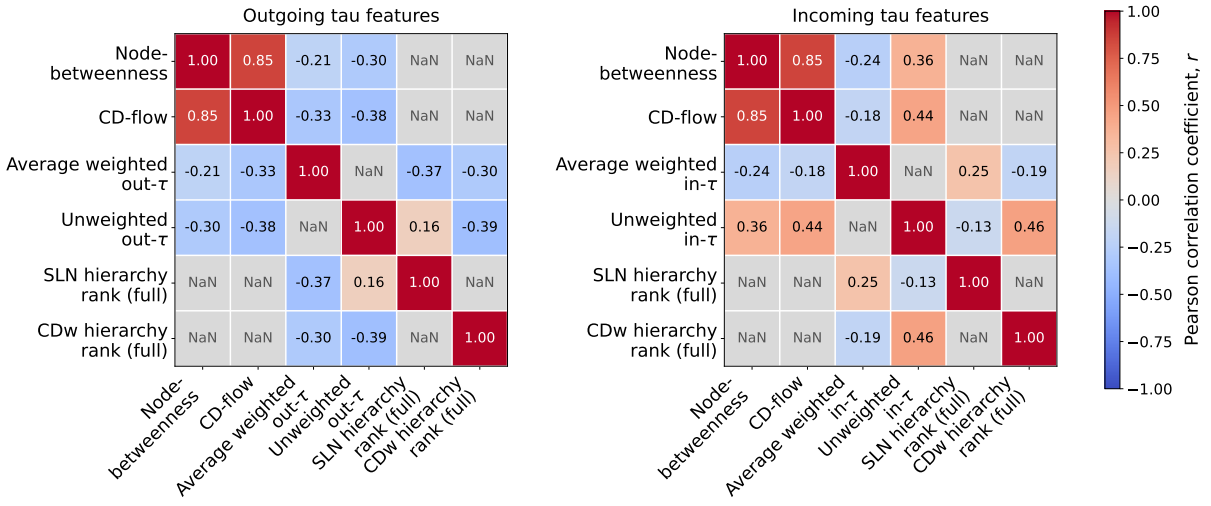

Figure 12: Pearson-correlation heatmaps for the full anatomical network. The left panel shows correlations involving outgoing τ-based features, and the right panel shows correlations involving incoming τ-based features.

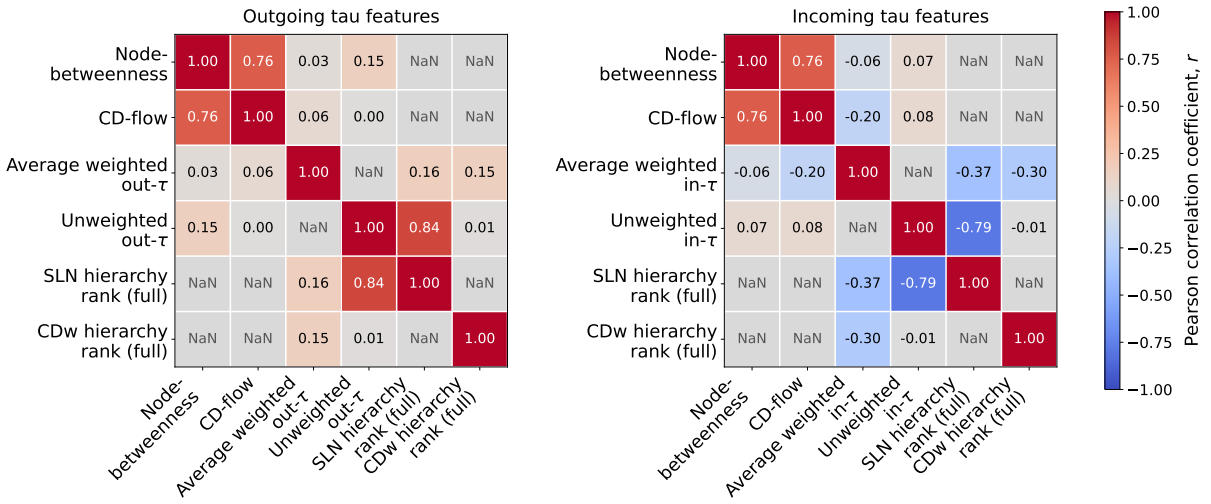

Figure 13: Pearson-correlation heatmaps for the $\mathrm{SLN}_{\mathrm{backward}}$ subnetwork. The left panel shows correlations involving outgoing τ-based features, and the right panel shows correlations involving incoming τ-based features.

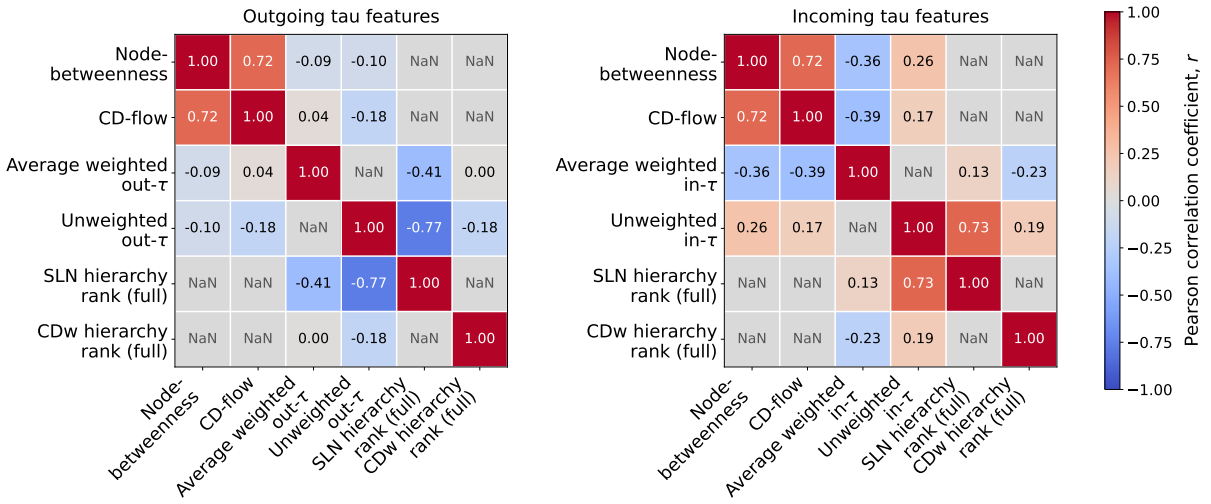

Figure 14: Pearson-correlation heatmaps for the $\text{SLN}_{\text{forward}}$ subnetwork. The left panel shows correlations involving outgoing τ-based features, and the right panel shows correlations involving incoming τ-based features.