./figures/

Symbolic Manifolds and Entropic Dynamics: A Computational Framework for Monitoring Cognitive Network Recovery in Regenerative Medicine

Demetrios Agourakis[1], Dionisio Chiuratto Agourakis[1], Renata Rigacci Abdala[2], Marina Stahl Merlin[3], and Marli Gerenutti[4]

[1][Institution of Demetrios and Dionisio], [City], [Country]
[2][Institution of Dr. Renata Rigacci Abdala], [City], [Country]
[3][Institution of Marina Stahl Merlin], [City], [Country]
[4][Institution of Dr. Marli Gerenutti], [City], [Country]
*Corresponding author: demetrios@agourakis.med.br

ABSTRACT

Cognitive recovery requires more than structural repair: it depends on the re-emergence of symbolic–semantic organisation. We present an open framework that projects large-scale semantic associations (SWOW-EN) onto a *symbolic manifold* where network topology, entropy-controlled exploration and discrete curvature jointly define interpretable coordinates. Using node2vec embeddings, UMAP projection and density-based clustering (HDBSCAN), we obtain separable, robust signatures for integrated, hyper-associative, selectively disconnected and collapsed regimes. These signatures are proposed as *candidate* computational markers of symbolic organisation, suitable for longitudinal tracking. While the present work is simulation-based, the manifold offers a structured pathway to future alignment with neuroimaging and behavioural assessments relevant to regenerative medicine. Code, data and exact environments are archived for full reproducibility.

Introduction

Restoring cognition after neurological injury or disease is not limited to repairing tissue. It requires re-establishing network-level organisation that supports symbol processing, concept formation, and flexible semantic search. While neuroimaging-derived connectomes and neuropsychological tests offer valuable markers of function and plasticity, they may miss *symbolic* configuration—the distribution and dynamics of meaning-bearing associations—where reorganisation can be subtle yet functionally decisive.

Semantic association norms such as SWOW-EN provide high-coverage graphs of human word associations suitable for modelling symbolic organisation at scale [1]. On these graphs, established tools from network science (efficiency, modularity, degree/centrality distributions) [2–4] can be combined with stochastic dynamics (Markov chains, entropy rate) [5, 6], discrete curvature [7, 8], and representation learning (node2vec) [9], followed by manifold visualisation (UMAP) [10] and density-based cluster discovery (HDBSCAN) [11].

Contributions. We introduce a computational framework that: (i) defines a symbolic manifold for cognitive states grounded on SWOW-EN; (ii) formalises an *entropic control* parameter mapping from edge-weighted association structure to random-walk dynamics; (iii) integrates curvature and entropy-rate into compact, interpretable biomarkers; (iv) demonstrates unsupervised separability of simulated regimes; and (v) outlines a translational map from manifold signatures to regenerative medicine scenarios (tissue engineering, stem-cell therapies, neuroprosthetic adaptation, and targeted cognitive rehabilitation).

Results

Manifold geometry and unsupervised structure

Starting from the SWOW-EN directed, weighted graph $G = (V, E, W)$, we build a row-stochastic transition kernel P by normalising outgoing edge weights and applying an *inverse-temperature* control $\beta > 0$:

$$P_{ij}(\beta) = \frac{W_{ij}^{\beta}}{\sum_k W_{ik}^{\beta}}. \tag{1}$$

This family $\{P(\beta)\}_\beta$ induces a spectrum of random-walk dynamics on the same symbolic substrate, interpolating between exploratory (high-entropy) and focused (low-entropy) regimes. Node embeddings are obtained via node2vec on G (not on P) to retain structural signals; 2D projections via UMAP support visualisation and clustering, yielding the symbolic manifold $\mathscr{M} \subset \mathbb{R}^2$.

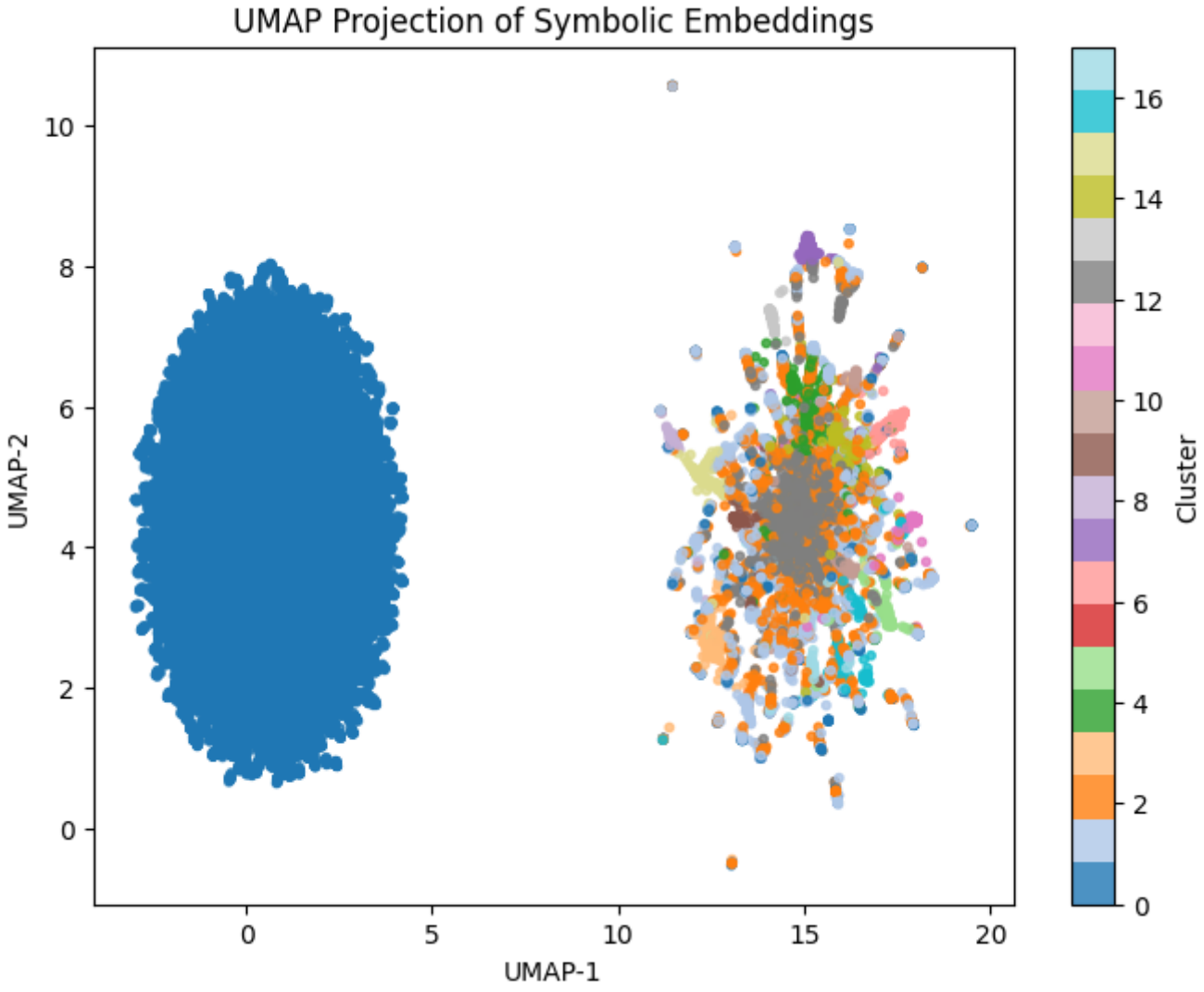

Figure 1. UMAP embeddings with density-based clustering. Points represent node2vec embeddings projected to two dimensions; colours indicate HDBSCAN cluster labels. Insets (not shown) report soft-membership probabilities and outlier scores. Silhouette coefficients and cluster persistence statistics are provided in the Supplementary Information.

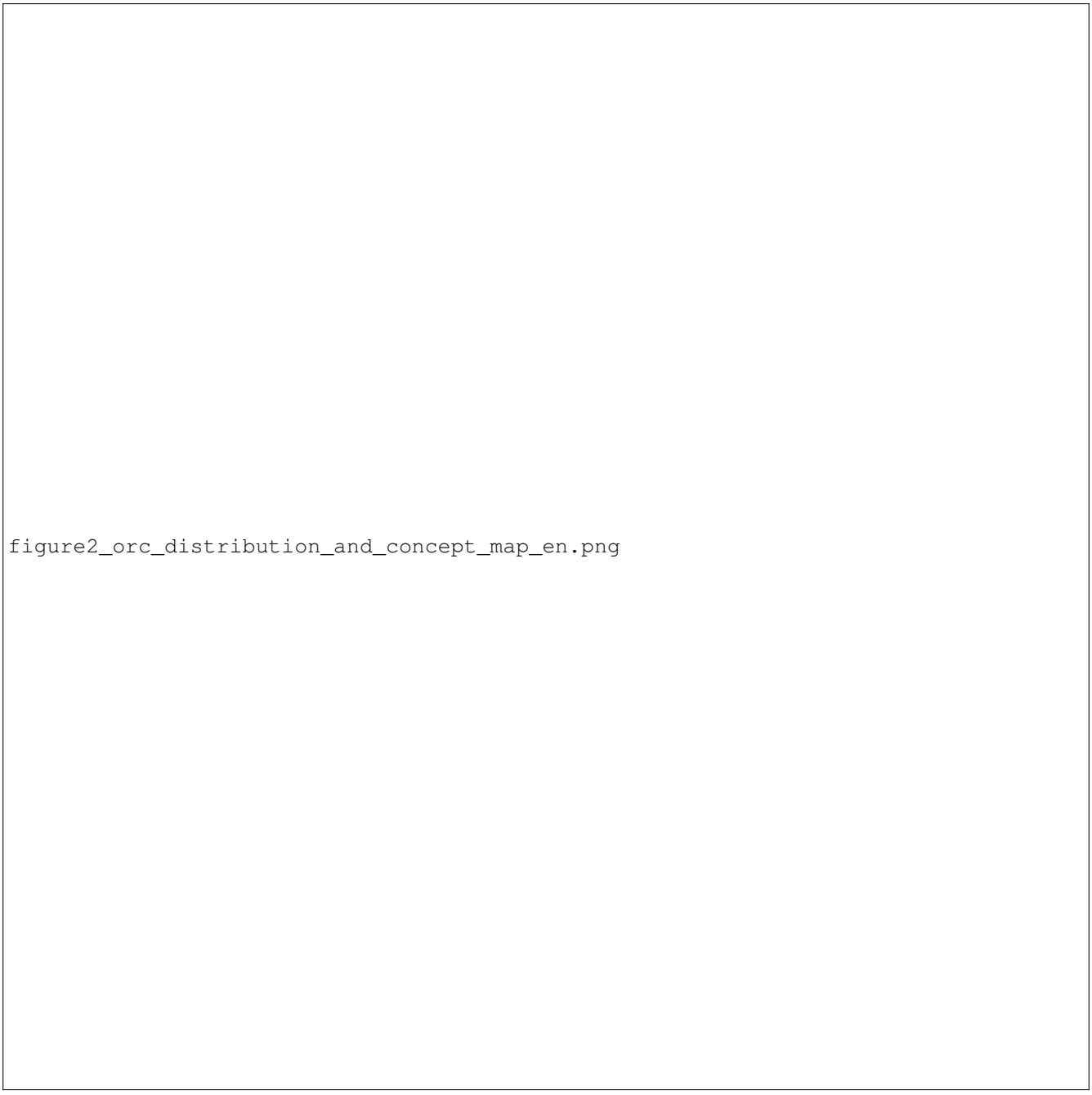

Figure 2. Ollivier–Ricci curvature distribution and conceptual topology of the SWOW semantic network. (**Left**) Histogram of curvature values (κ) computed for the top 30,000 weighted word-association edges in the Small World of Words (SWOW) English dataset. The distribution is predominantly negative (mean $\kappa \approx -0.48$), indicating a hyperbolic network geometry consistent with hierarchical semantic organization. Most edges fall between -0.6 and -0.4 (72.9%), representing typical saddle-like local geometry, while positive curvature values (0.4%) correspond mainly to self-referential loops and highly coherent semantic clusters. (**Right**) Conceptual map illustrating the "Topology of the Imaginary", derived from curvature-based semantic clustering. Nodes represent conceptual domains (e.g., core concepts, semantic clusters, boundary concepts, abstract concepts), with edges indicating semantic transitions. Negative curvature regions align with boundary concepts and category mismatches, while positive curvature regions correspond to core conceptual identity preservation. Together, these panels combine quantitative network geometry with a qualitative interpretation of the cognitive–semantic landscape.

Interpretation of curvature regimes and functional relevance. Consistent with geometric expectations, the predominance of negative κ values in SWOW-EN indicates a strongly hyperbolic, tree-like backbone that supports efficient navigation across semantic space. In Ollivier–Ricci terms, edges with $\kappa < 0$ tend to function as inter-cluster bridges (low triangle density,

high effective betweenness), whereas $\kappa > 0$ concentrates within dense modules; $\kappa \approx 0$ commonly marks peripheral, tree-like attachments. This partitioning mirrors observations in complex networks where OR curvature separates backbone "highways" from redundant intra-cluster links, and is predictive of traffic concentration and vulnerability [12]. The prevalence of $\kappa < 0$ is consistent with an underlying hyperbolic organization that enhances navigability and shortens effective distances across disparate semantic domains, an effect paralleled in brain connectomes where hidden negatively curved geometry enables efficient routing [13]. In translational terms, curvature regimes provide a quantitative lens on robustness versus fragility: bridges with strongly negative κ are critical for global integration and cross-domain leaps, while pockets of positive κ sustain resilient local processing; these roles resonate with curvature-based robustness markers in structural brain networks [14] and with functional alterations reported in clinical populations [15, 16].

Entropic–geometric atlas of regimes

Each network state is summarised by

$$\mathbf{m} = \big(H_{\text{rate}}(P),\ \overline{\kappa}_{\text{OR}},\ E_{\text{glob}},\ C,\ Q,\ \text{Gini}(\text{deg}),\ \overline{c}_{\text{eig}}\big), \tag{2}$$

where H_{rate} is the entropy rate of P, $\overline{\kappa}_{\text{OR}}$ the mean Ollivier–Ricci curvature over edges, E_{glob} global efficiency, C clustering coefficient, Q modularity, $\text{Gini}(\text{deg})$ inequality of the degree distribution, and $\overline{c}_{\text{eig}}$ mean eigenvector centrality. In practice, $(H_{\text{rate}}, \overline{\kappa}_{\text{OR}})$ already provide a compact, interpretable *entropic–geometric atlas*.

Translational mapping: from simulations to regenerative contexts

Regime signatures are mapped to translational contexts (Fig. 4): focal tract disruption (stroke/trauma) displays selective-disconnection patterns; diffuse degradation (neurodegeneration) approaches collapse; rehabilitation or neuroprosthetic integration pulls trajectories toward integration.

Temporal dynamics: collapse and recovery

Figure 5 shows a representative collapse–recovery trajectory under a perturbation onset (dashed line) affecting curvature and efficiency, with concurrent changes in effective activation breadth and an entropy-related measure.

Discussion

Our curvature distribution dovetails with known hyperbolic navigability of connectomes [13] and with clinical reports that curvature profiles differ in autism spectrum disorder (ASD), suggesting domain-general links between geometric backbone, integration efficiency, and cognitive phenotypes [15, 16].

We introduced a symbolic–semantic framework that integrates graph topology, entropy-controlled dynamics and discrete curvature into a low-dimensional manifold where cognitive regimes become separable and interpretable. On SWOW-EN, distinct perturbational profiles—integrated, hyper-associative, selective disconnection and collapse—yield non-redundant shifts in entropy rate and curvature, with consistent geometry in embedded space. We emphasise that these regime signatures are *candidate* computational markers, not clinical endpoints; they are designed to be aligned with independent measurements in future work.

Entropy rate (H_{rate}) proxies the breadth of symbolic exploration; mean Ollivier–Ricci curvature ($\overline{\kappa}_{\text{OR}}$) captures local geometric coherence and bridge sensitivity. The $(H_{\text{rate}}, \overline{\kappa}_{\text{OR}})$ plane thus constitutes an *entropic–geometric atlas* in which regime trajectories are naturally expressed—e.g., targeted disconnection depresses curvature more than entropy, whereas indiscriminate noise inflates entropy while eroding cluster coherence.

Although the present study uses simulated regimes, the atlas offers a disciplined pathway for longitudinal alignment with behavioural measures (e.g., verbal fluency, lexical access latencies) and neuroimaging (EEG/fMRI/DTI). Hypothesis-generating examples include: focal tract disruption mapping to selective-disconnection signatures; diffuse degradation mapping to collapse; and rehabilitation/neuroprosthetic adaptation pulling trajectories towards integration. We refrain from clinical inference here; the intent is a stable computational coordinate system for future studies.

Methodologically, the pipeline is end-to-end, version-pinned and openly archived, with intermediate artefacts (metrics, embeddings, labels) saved for reuse. Stability is assessed under seed and hyperparameter variation, and under bootstrap perturbations of nodes/edges. Limitations include the use of population-level semantic norms (not subject-specific), reliance on simulated perturbations, modelling choices in curvature estimation, and dependence of manifold geometry on embedding/projection hyperparameters; we report robustness across reasonable grids.

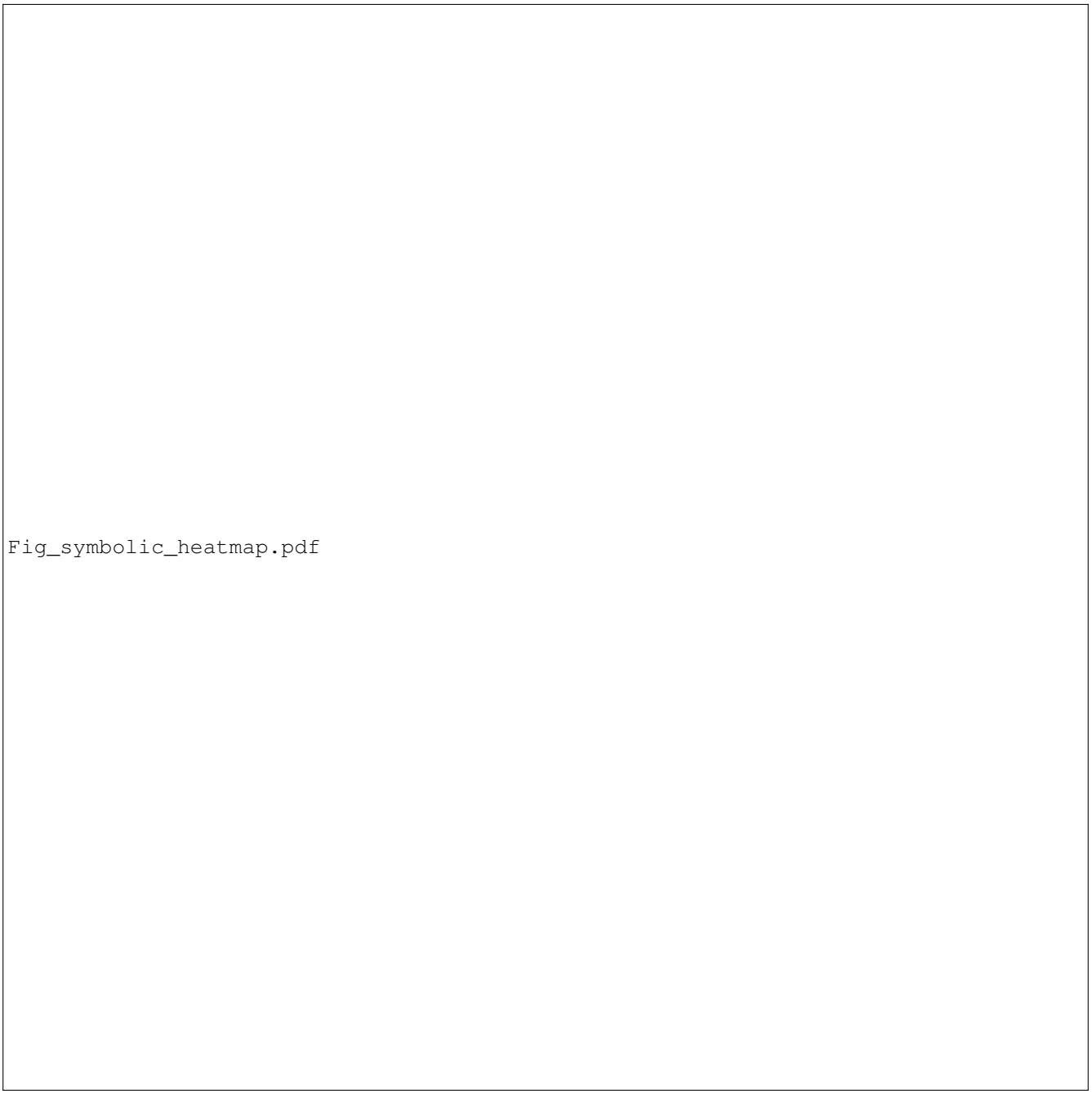

Figure 3. Entropy–curvature atlas of symbolic regimes. Each point represents a network state positioned by its entropy rate (H_{rate}, x-axis) and mean Ollivier–Ricci curvature ($\overline{\kappa}_{\text{OR}}$, y-axis). Colours indicate canonical regimes. Kernel density estimates (marginals) reveal separable basins, illustrating complementary sensitivity of entropy and curvature to different perturbations.

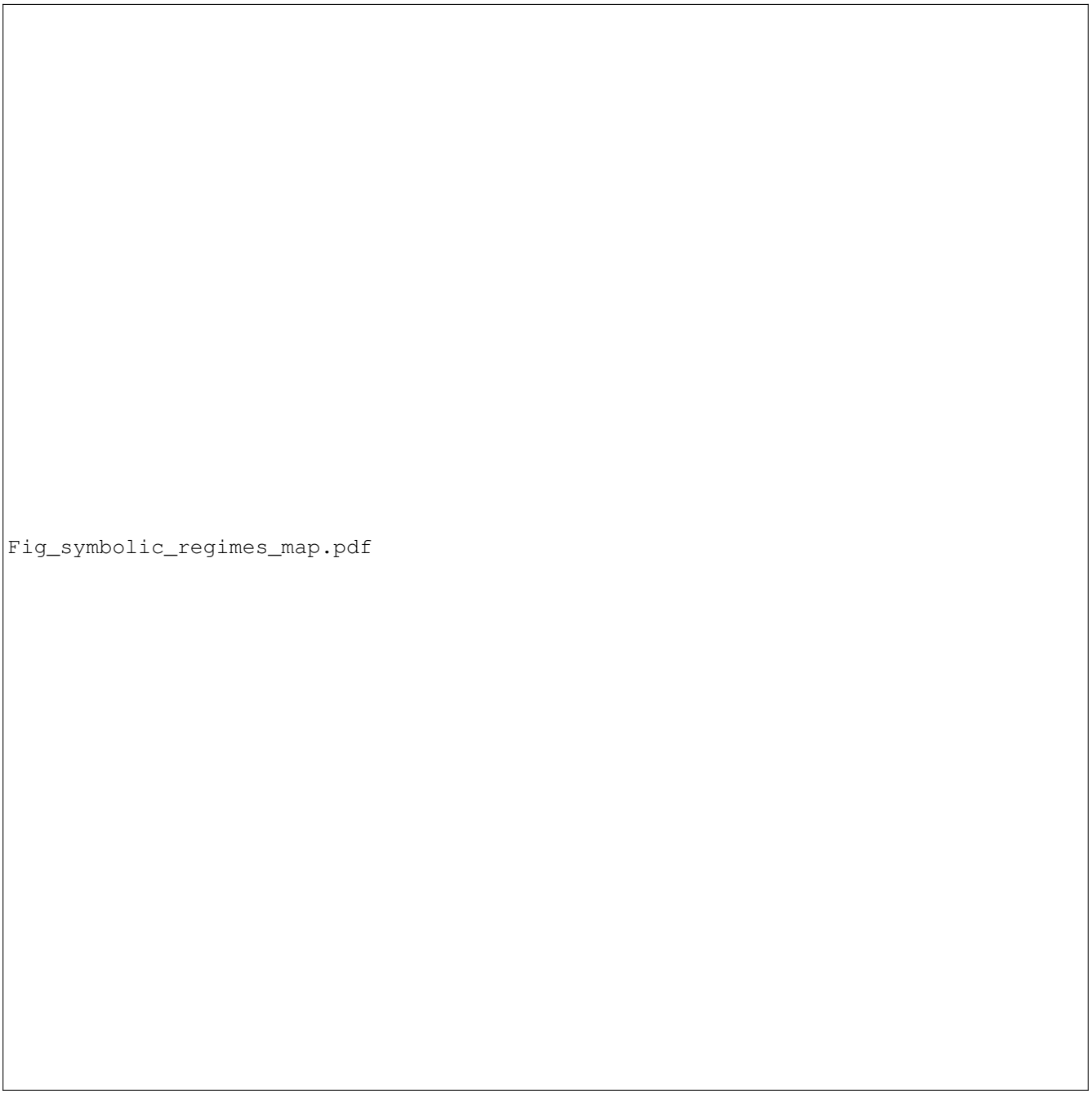

Figure 4. Symbolic regime mapping and translational scenarios. Diagram linking entropy–curvature patterns to neurobiological contexts. Arrows indicate hypothetical recovery pathways within the symbolic manifold; icons represent measurement modalities (e.g., neuroimaging, cognitive tasks).

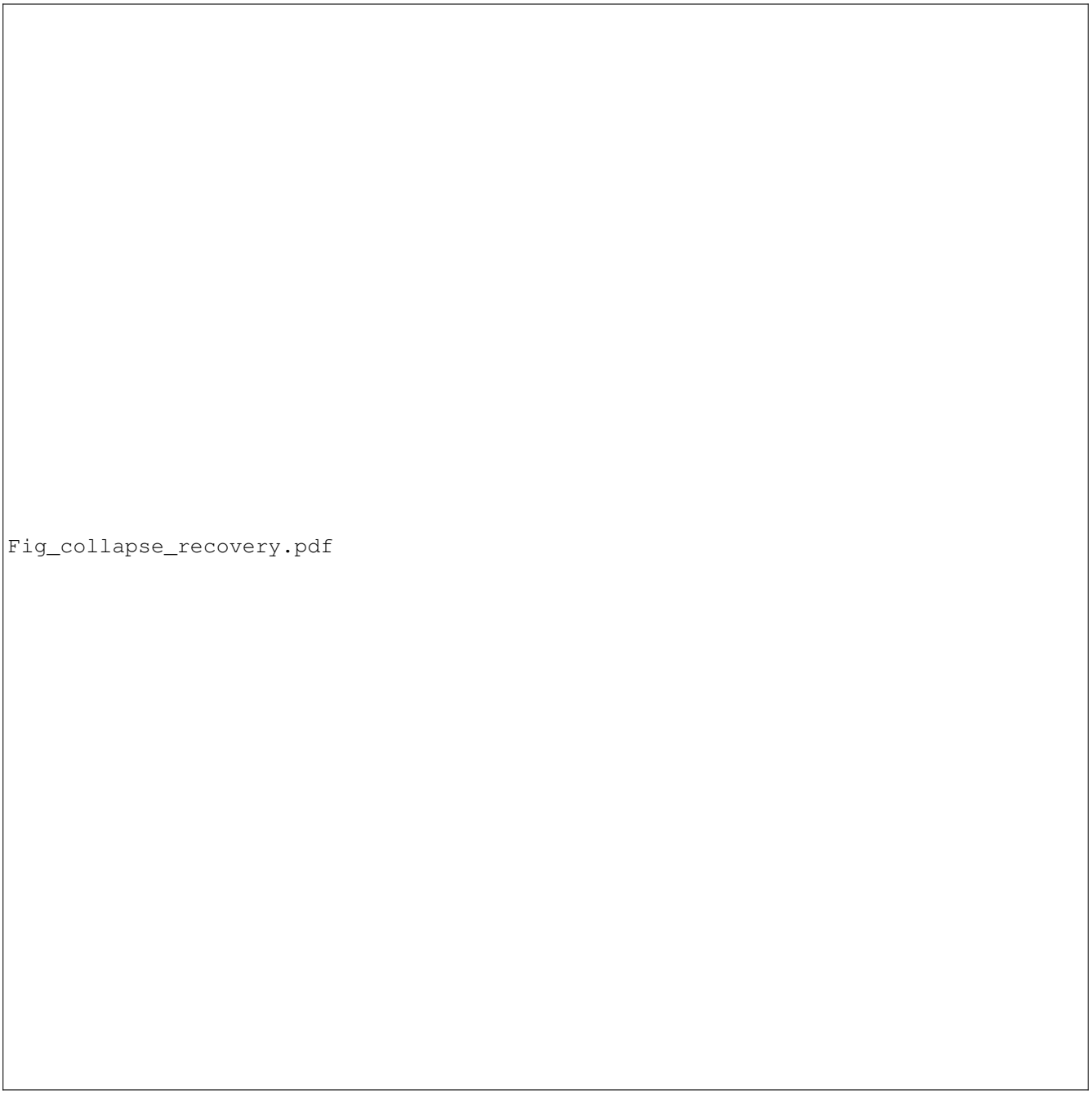

Figure 5. Collapse and recovery under perturbation. Simulated time series of state variables showing the effect of a perturbation on curvature (κ), effective activation breadth (α), and an entropy-related measure (E_r). Vertical dashed line: perturbation onset.

Methods

Following prior work, we treat edge-wise Ollivier–Ricci curvature (ORC) as a proxy for robustness and integrative load in biological networks, where negative κ highlights inter-modular bottlenecks and positive κ marks dense intra-modular fabrics [12, 14].

Data source and preprocessing

We use SWOW-EN (Small World of Words; English) association norms [1]. Preprocessing: lowercase normalisation; removal of non-alphabetic tokens; optional lemmatisation for sensitivity checks; filtering rare edges by frequency threshold $\tau = 3$ responses; retention of the giant weakly connected component. Edge weights W_{ij} equal the empirical association frequency (or probability) from cue i to associate j after normalisation.

Graph construction and metrics

We construct a directed weighted graph $G = (V, E, W)$. Degree (in/out), betweenness, closeness and eigenvector centrality are computed on G (weighted, directed conventions as implemented in NetworkX). Global efficiency E_{glob} and clustering C follow standard definitions [2, 4]. All network metrics were computed with NetworkX v2.8.8 [17], using directed and weighted conventions unless otherwise noted. Modularity Q uses a directed extension with Louvain/Leiden (sensitivity analysis across seeds).

Entropic dynamics

Define a β-controlled transition kernel $P(\beta)$ as above. The entropy rate of the induced Markov chain is

$$H_{\text{rate}}(P) = -\sum_{i \in V} \pi_i \sum_{j \in V} P_{ij} \log P_{ij}, \tag{3}$$

where π is the stationary distribution of P. We sweep β on a grid ($\beta \in \{0.5, 0.8, 1.0, 1.2, 1.5, 2.0\}$) and, for each β, apply controlled perturbations to W to instantiate regimes: hyper-association (broadening of weak edges), selective disconnection (targeted removal of bridges/cut-sets), and fragmentation (random edge percolation).

Discrete curvature

We compute Ollivier–Ricci curvature $\kappa_{\text{OR}}(i, j)$ using one-step neighbour measures with $\alpha = 0.5$ anchoring, i.e. $\mu_x = \alpha \delta_x + (1 - \alpha)\,\text{Uniform}(\mathcal{N}(x))$, and

$$\kappa_{\text{OR}}(i, j) = 1 - \frac{W_1(\mu_i, \mu_j)}{d(i, j)}, \tag{4}$$

where W_1 is a 1-Wasserstein distance and d a graph distance [7, 8]. We also compute Forman curvature on the undirected approximation as a sensitivity check.

Embedding and dimensionality reduction

Node representations are learned with node2vec (dimension $d = 128$; walk length $L = 30$; window $= 10$; walks/node $= 10$; $p = 1$, $q = 0.5$, seed=42) [9]. UMAP projects to 2D ($n_{\text{neighbours}} = 30$, `min_dist`=0.1) for visualisation and clustering [10]. Robustness to hyperparameters and seeds is reported in the Supplementary Information.

Clustering

HDBSCAN is applied in UMAP space to obtain robust clusters with soft memberships (`min_cluster_size`=50, `min_samples`=10, metric='euclidean', cluster_selection_epsilon=0.0) [11]. We report the number of clusters, outlier fraction, and stability under bootstrap resampling.

Use of large language models (LLMs)

We document the use of a large language model (GPT-5 Thinking) to assist with code refactoring and editorial polishing. All analyses, code and text were verified by the authors, who take full responsibility for the content.

Data availability

SWOW-EN English association norms are publicly available via the Small World of Words project (De Deyne *et al.*, 2019, Behavior Research Methods, DOI: 10.3758/s13428-018-1115-7). Derived datasets (metrics, embeddings, entropy–curvature summaries, cluster labels) and all figure source data are archived on Zenodo: DOI 10.5281/zenodo.16783257.

Code availability

The full pipeline (preprocessing, graph construction, entropic dynamics, curvature computation, embedding, projection, clustering, stability analysis) is available at https://github.com/agourakis82/entropic-symbolic-society, with an archived snapshot on Zenodo (DOI: 10.5281/zenodo.16783257). Exact package versions and environments are pinned.

Acknowledgements

[Insert funding, institutional support, and computational resources.]

Author contributions

D.A. conceived the study and designed the computational framework. D.C.A., R.R.A., M.S.M., and M.G. contributed to theoretical framing, clinical context integration, and manuscript revision. All authors approved the final version.

Competing interests

The authors declare no competing interests.

Supplementary Information

Supplementary Information is provided as a single PDF file named *Supplementary_Information.pdf*, containing: (i) Mathematical Notes (definitions; lemma; proof sketch); (ii) Glossary and Table of Symbols; (iii) additional figures S1–S5 (ablation, null-model comparisons, curvature sensitivity, trajectories); (iv) source data tables for all main and supplementary figures.

References

1. De Deyne, S., Navarro, D. J., Perfors, A., Brysbaert, M. & Storms, G. The "small world of words" english word association norms for over 12,000 cue words. *Behav. Res. Methods* **51**, 987–1006, DOI: 10.3758/s13428-018-1115-7 (2019).

2. Newman, M. E. J. *Networks: An Introduction* (Oxford University Press, Oxford, UK, 2010).

3. Watts, D. J. & Strogatz, S. H. Collective dynamics of 'small-world' networks. *Nature* **393**, 440–442, DOI: 10.1038/30918 (1998).

4. Latora, V. & Marchiori, M. Efficient behavior of small-world networks. *Phys. Rev. Lett.* **87**, 198701, DOI: 10.1103/PhysRevLett.87.198701 (2001).

5. Shannon, C. E. A mathematical theory of communication. *Bell Syst. Tech. J.* **27**, 379–423, 623–656, DOI: 10.1002/j.1538-7305.1948.tb01338.x (1948).

6. Burda, Z., Duda, J., Luck, J.-M. & Wacław, B. Localization of the maximal entropy random walk. *Proc. Natl. Acad. Sci.* **106**, 8205–8210, DOI: 10.1073/pnas.0810951106 (2009).

7. Ollivier, Y. Ricci curvature of markov chains on metric spaces. *J. Funct. Analysis* **256**, 810–864, DOI: 10.1016/j.jfa.2008.11.001 (2009).

8. Forman, R. Bochner's method for cell complexes and combinatorial ricci curvature. *Discret. & Comput. Geom.* **29**, 323–374, DOI: 10.1007/s00454-002-0743-x (2003).

9. Grover, A. & Leskovec, J. node2vec: Scalable feature learning for networks. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 855–864, DOI: 10.1145/2939672.2939754 (2016).

10. McInnes, L., Healy, J. & Melville, J. Umap: Uniform manifold approximation and projection for dimension reduction. *arXiv preprint* DOI: 10.48550/arXiv.1802.03426 (2018). 1802.03426.

11. Campello, R. J. G. B., Moulavi, D. & Sander, J. Hierarchical density estimates for data clustering, visualization, and outlier detection. *ACM Transactions on Knowl. Discov. from Data* **10**, 5:1–5:51, DOI: 10.1145/2733381 (2015).

12. Sia, J., Jonckheere, E. A. & Bogdan, P. Ollivier–ricci curvature-based method to community detection in complex networks. *Sci. Reports* **9**, 9800, DOI: 10.1038/s41598-019-46063-7 (2019).

13. Allard, A. & Serrano, M. Á. Navigable maps of structural brain networks across species. *PLOS Comput. Biol.* **16**, e1007584, DOI: 10.1371/journal.pcbi.1007584 (2020).

14. Farooq, H., Chen, Y., Georgiou, T. T., Tannenbaum, A. & Lenglet, C. Network curvature as a hallmark of brain structural connectivity. *Nat. Commun.* **10**, 4937, DOI: 10.1038/s41467-019-12915-x (2019).

15. Elumalai, P. *et al.* Graph ricci curvatures reveal atypical functional connectivity in autism spectrum disorder. *Sci. Reports* **12**, 8295, DOI: 10.1038/s41598-022-12171-y (2022).

16. Simhal, A. K. *et al.* Measuring robustness of brain networks in autism spectrum disorder with ricci curvature. *Sci. Reports* **10**, 10852, DOI: 10.1038/s41598-020-67474-9 (2020).

17. Hagberg, A. A., Schult, D. A. & Swart, P. J. Exploring network structure, dynamics, and function using networkx. In *Proceedings of the 7th Python in Science Conference (SciPy)*, 11–15 (2008).