A Python Hands-on Tutorial on Network and Topological Neuroscience

Eduarda Gervini Zampieri Centeno[1,2](✉) , Giulia Moreni[1], Chris Vriend[1,3],
Linda Douw[1] , and Fernando Antônio Nóbrega Santos[1,4]

[1] Amsterdam UMC, Vrije Universiteit Amsterdam, Anatomy and Neurosciences,
Amsterdam Neuroscience, De Boelelaan 1117, Amsterdam, Netherlands
e.centeno@amsterdamumc.nl
[2] Bordeaux Neurocampus, Université de Bordeaux, Institut des Maladies
Neurodégénératives, CNRS UMR 5293, 146 Rue Léo Saignat, Bordeaux, France
[3] Amsterdam UMC, Vrije Universiteit Amsterdam, Psychiatry, Amsterdam
Neuroscience, De Boelelaan 1117, Amsterdam, Netherlands
[4] Institute of Advanced Studies, University of Amsterdam, Oude Turfmarkt 147,
1012 Amsterdam, GC, Netherlands

Abstract. Network neuroscience investigates brain functioning through
the prism of connectivity, and graph theory has been the main framework to understand brain networks. Recently, an alternative framework
has gained attention: topological data analysis. It provides a set of metrics that go beyond pairwise connections and offer improved robustness
against noise. Here, our goal is to provide an easy-to-grasp theoretical
and computational tutorial to explore neuroimaging data using these
frameworks, facilitating their accessibility, data visualisation, and comprehension for newcomers to the field. We provide a concise (and by
no means complete) theoretical overview of the two frameworks and a
computational guide on the computation of both well-established and
newer metrics using a publicly available resting-state functional magnetic
resonance imaging dataset. Moreover, we have developed a pipeline for
three-dimensional (3-D) visualisation of high order interactions in brain
networks.

Keywords: Network neuroscience · Data visualisation · Topological
data analysis

1 Introduction

Network neuroscience sees the brain through an integrative lens by mapping
and modelling its elements and interactions [3,21]. The main theoretical framework from complex network science used to model, estimate, and simulate brain
networks is graph theory [9,24]. A graph is comprised of a set of interconnected elements, also known as vertices and edges. Vertices (also known as
nodes) in a network can be, for example, brain areas, while edges (also known
as links) are a representation of the functional connectivity between pairs of

vertices [42]. Several descriptive graph metrics[1] [16] can then be calculated to describe the brain network's characteristic [21,26], and they have consistently allowed researchers to identify non-random features of brain networks. An example is the ground-breaking discovery that the brain (like most other real-world networks) follows a 'small-world network' architecture [2,43], indicating a compromise between wiring cost and optimal efficiency. Using graph theory, many insights have been gathered on the healthy and diseased brain neurobiology [19,26]. Algebraic topological data analysis (TDA) provides another prism on brain connectivity investigation beyond the 'simple' pairwise connections (*i.e.*, higher-order interactions). With TDA, one can identify a network's shape and its invariant properties (*i.e.*, coordinate and deformation invariances [47]). Moreover, TDA often provides more robustness against noise than graph theoretical analysis [41], a significant neuroimaging data issue [30]. Although TDA has only recently been adopted to network neuroscience [15,39], it has already shown exciting results on brain network data [14,18]. However, clinical scientists' comprehension and application can be hindered by TDA's complexity and mathematical abstraction. Here, we want to facilitate the use of network neuroscience and its constituents graph theory and TDA by the general neuroscientific community by providing both computational and theoretical explanation of the primary metrics, strongly inspired by [21]. The work is divided into a longer manuscript [13] containing several resources (see Table in [13]) and more theoretical explanations, and a publicly available Jupyter Notebook online (https://github.com/multinetlab-amsterdam/network_TDA_tutorial). In these notebooks, we use third-party Python packages for part of the computations (*e.g.*, *networkx* [25] for the graph theory metrics and *gudhi* [32] for persistent homology) and provide practical scripts for some TDA metrics and 3-D visualisations of simplicial complexes (a new addition to the field). Our tutorial focuses on resting-state functional magnetic resonance imaging (rsfMRI) data; however, the main concepts and tools discussed in this paper can be extrapolated to other imaging modalities, biological or complex networks. The extended version [13] covers the most commonly used graph metrics in network neuroscience, also in line with reference [21], and TDA. However, due to the size constraint, here we prioritize the latter.

1.1   Starting Point: The Adjacency Matrix

The basic unit on which graph theory and TDA are applied in the context of rsfMRI is the adjacency or functional connectivity matrix [3,21]. Typically, rsfMRI matrices are symmetric and do not specify the direction of connectivity (*i.e.*, activity in area A drives activity in area B), therefore yielding undirected networks. To further analyse the sfMRI connectivity matrix, one has to decide

---

[1] Notice that the notion of metric in mathematics defines distance between two points in a set [16], which is distinct from what we are using in this work. We denote as metric any quantity that can be computed, *i.e.*, "measured", in a brain network or simplicial complex.

whether to keep or not edges' weights (*e.g.*, correlation values in rsfMRI connectivity) or to absolutise negative weights (or anticorrelations) [21,27]. These decisions influence the computation of the different metrics described in the tutorial and matter for the biological interpretation of the results [21]. In this tutorial we use an undirected, weighted, and absolutised connectivity matrix.

1.2   Topological Data Analysis

TDA uses topology and geometry methods to study the shape of the data [11] and can identify a network's different characteristics by addressing a network's high-order structure [4,10,28]. A core success of TDA is the ability to provide robust results when compared with alternative methods, even if the data are noisy [18,41]. One of the benefits of using TDA in network neuroscience is the possibility of finding global properties of a network that are preserved regardless of the way we represent the network [35], as we illustrate below. Those properties are the so-called topological invariants. Here, we cover some fundamental TDA concepts: filtration, simplicial complexes, Euler characteristic, phase-transitions, Betti numbers, curvature, and persistent homology.

Simplicial Complexes. In TDA, we consider that the network as a multidimensional structure called the simplicial complex. Such a network is not only made up of the set of vertices (0-simplex) and edges (1-simplex) but also of triangles (2-simplex), tetrahedrons (3-simplex), and higher k-dimensional structures. In short, a k-simplex is an object in k-dimensions and, in our work, is formed by a subset of $k + 1$ vertices of the network.

Filtration. Consists of a nested sequence of simplicial complexes. Here, a filtration is defined by changing the density d of the network, from $0 \leq d \leq 1$. This yields a nested sequence of networks, in which increasing d leads to a more densely connected network. In neuroscience, It can be used to avoid arbitrary threshold/density choices, which are usually made in the field.

We can encode a network into a simplicial complex in several ways [17,29,31]. Here, we focus on building a simplicial complex only from the brain network's cliques, *i.e.*, we create the so-called clique complex of a brain network. In a network, a k-clique is a subset of the network with k all-to-all connected nodes. 0-clique corresponds to the empty set, 1-cliques correspond to nodes, 2-cliques to links, 3-cliques to triangles, *etc.*. In the clique complex, each $k + 1$ clique is associated with a k-simplex. This choice for creating simplexes from cliques has the advantage that we can still use pairwise signal processing to create a simplicial complex from brain networks, such as in [23]. It is essential to mention that other strategies to build simplicial complexes beyond pairwise signal processing are still under development, such as applications using multivariate information theory together with tools from algebraic topology [1,5–7,22,36]. In our Jupyter Notebook [12], we provide the code to visualise the clique complex developed in [38] (Fig. 1).

The Euler Characteristic. The Euler characteristic is one example of topological invariants: the network properties that do not depend on a specific graph

representation. We first introduce the Euler characteristic for polyhredra. Later, we translate this concept to brain networks. In 3-D convex polyhedra, the Euler characteristic is defined as the numbers of vertices minus edges plus faces. For convex polyhedra without cavities (holes in its shape), which are isomorphous to the sphere, the Euler characteristic is always two. If we take the cube and make a cavity, the Euler drops to zero as it is in the torus. If we make two cavities in a polyhedral (as in the bitorus), the Euler drops to minus two. We can understand that the Euler characteristic tells us something about a polyhedron's topology and its analogous surface. In other words, if we have a surface and we make a discrete representation of it (*e.g.*, a surface triangulation), its Euler characteristic is always the same, regardless of the way we do it. We can now generalise the definition of Euler characteristic to a simplicial complex in any dimension. Thus, the high dimensional version of the Euler characteristic is expressed by the alternate sum of the numbers $Cl_k(d)$ of the k-cliques (which are $(k-1)$-simplexes) present in the network's simplicial complex for a given value of the density threshold d:

$$\chi(d) = Cl_1 - Cl_2 + ... + Cl_n = \sum_{k=1}^{n} (-1)^k Cl_{(k)}(d).$$

Betti Numbers. Another set of topological invariants are the Betti numbers (β). Given that a simplicial complex is a high-dimensional structure, β_k counts the number of k-dimensional holes in the simplicial complex. These are topological invariants that correspond, for each $k \geq 0$, to the number of k-dimensional holes in the simplicial complex [47]. In a simplicial complex, there can be many of these k-holes and counting them provide the Betti number β, e.g., if β_2 is equal to five, there are 5 two-dimensional holes. The Euler characteristics of a simplicial complex can also be computed using the β via the following formula [17]:

$$\chi = \beta_0 - \beta_1 + \beta_2 - \dots (-1)^{k_{\max}} \beta_{k_{\max}} = \sum_{k=0}^{k_{max}} (-1)^k \beta_k ,$$

where $k_{\max}$ the maximum dimension that we are computing the cycles.

Curvature. Curvature is a TDA metric that can link the global network properties described above to local features [20,38,44]. It allows us to compute topological invariants for the whole-brain set of vertices and understand the contribution of specific individual nodal, or subnetwork, geometric proprieties to the brain network's global properties. Several approaches to defining a curvature for networks are available [33,44], including some already used in neuroscientific investigations [38]. We illustrate the curvature approach linked to topological phase transitions, previously introduced for complex systems [20,33,45]. To compute the curvature, filtration is used to calculate the clique participation rank (*i.e.*, the number of k-cliques in which a vertex i participates for density d) [40], which we denote here by $Cl_{ik}(d)$. The curvature of the vertex based on the participation rank is then defined as:
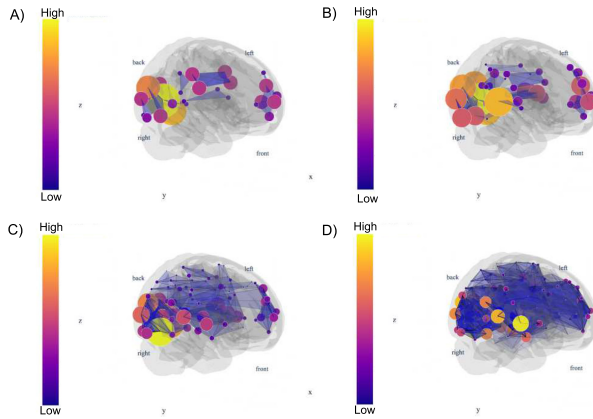
Fig. 1. Simplex 3-D visualisation. Here we visualise the number of 3-cliques (triangles) in a functional brain network as we increase the edge density d (0.01, 0.015, 0.02, and 0.025, from A to D). For higher densities, we have a more significant number of 3-cliques compared to smaller densities. The vertex colour indicates the clique participation rank. We used fMRI data from the 1000 Functional Connectome Project [8].

$$\kappa_i = \sum_{k=1}^{k_{\max}} (-1)^{k+1}, \frac{Cl_{ik}\ (d)}{k}$$

where $Cl_{ik} = 1$ since each vertex i participates in a single 1-clique (the vertex itself), and $k_{\max}$ the maximum number of vertices that are all-to-all connected in the network (see in Fig. 1 the participation in 3-cliques). We use the Gauss-Bonnet theorem for networks to link the local (nodal) curvature to the network's global properties (its Euler characteristic). Conversely, by summing up all the curvatures of the network across different thresholds, one can reach the alternate sum of the numbers Cl_k of k-cliques (a subgraph with k all-to-all connected vertices) present in the network's simplicial complex for a given density threshold $d \in [0,1]$. By doing so, we also write the Euler characteristics as a sum of the curvature of all network vertices, *i.e.*,

$$\chi(d) = \sum_{i=1}^{N} k_i(d).$$

1.3   Discussion

This tutorial explains some of the primary metrics related to two network neuroscience branches - graph theory and TDA -, providing short theoretical backgrounds and code examples accompanied by a publicly available Jupyter Notebook, with a special section on visualisations of simplicial complexes and curvature computation in brain data. Here, we did not aim to provide a use-case

report but rather a hands on computational resource. Finally, we would like to mention some relevant limitations in interpretation when using these metrics in connectivity-based data. Considering that rsfMRI data is often calculated as a temporal correlation between time series using Pearson's correlation coefficient, a bias on the number of triangles can emerge [46]. This affects TDA (where the impact depends on how high-order interactions are defined) and graph-theoretical metrics (such as the clustering coefficient), with networks based on this statistical method being automatically more clustered than random models. The proper way to determine and infer high-order interactions in the brain is an ongoing challenge in network neuroscience [1,5–7,22,36]. Moreover, it is essential to think about the computational cost. The computation of cliques falls in the clique problem, an NP (nonpolynomial time) problem; thus, listing cliques may require exponential time as the size of the cliques or networks grows [34]. What we can do for practical applications is to limit the clique size that can be reached by the algorithm, which determines the dimension of the simplicial complex in which the brain network is represented. This arbitrary constraint implies a theoretical simplification, limiting the space or the dimensionality in which we would analyse brain data. Another issue is that, to finish TDA computations in a realistic time frame, the researcher might need to establish a maximal threshold/density for convergence even after reducing the maximal clique size. Even though TDA approaches lead to substantial improvements in network science, apart from applications using the Mapper algorithm [37], the limitations mentioned above contribute to losing information on the data's shape. In conclusion, graph theory has been widely used in network neuroscience, but newer methods such as TDA are gaining momentum. To further improve the field, especially for users in the domain of clinical network neuroscience, it is imperative to make the computation of the developed metrics accessible and easy to comprehend and visualise. We hope to have facilitated the comprehension of some aspects of network and topological neuroscience, the computation and visualisation of some of its metrics.

References

1. Barbarossa, S., Sardellitti, S.: Topological signal processing over simplicial complexes. IEEE Trans. Sig. Process. **68**, 2992–3007 (2020). https://doi.org/10.1109/TSP.2020.2981920
2. Bassett, D.S., Bullmore, E.T.: Small-world brain networks revisited. Neuroscientist **23**(5), 499–516 (2017). https://doi.org/10.1177/1073858416667720
3. Bassett, D.S., Sporns, O.: Network neuroscience. Nat. Neurosci. **20**(3), 353 (2017). https://doi.org/10.1038/nn.4502
4. Battiston, F., et al.: Networks beyond pairwise interactions: structure and dynamics. Phys. Rep. **874**, 1–92 (2020). https://doi.org/10.1016/j.physrep.2020.05.004
5. Baudot, P.: The poincare-shannon machine: statistical physics and machine learning aspects of information cohomology. Entropy **21**(9), 881 (2019). https://doi.org/10.3390/e21090881
6. Baudot, P., Bennequin, D.: The homological nature of entropy. Entropy **17**(5), 3253–3318 (2015). https://doi.org/10.3390/e17053253

7. Baudot, P., et al.: Topological information data analysis. Entropy **21**(9), 869 (2019). https://doi.org/10.3390/e21090869
8. Biswal, B.B., et al.: Toward discovery science of human brain function. Proc. Nat. Acad. Sci. USA **107**(10), 4734–4739 (2010). https://doi.org/10.1073/pnas.0911855107
9. Bullmore, E., Sporns, O.: Complex brain networks: graph theoretical analysis of structural and functional systems. Nat. Rev. Neurosci. **10**(3), 186–198 (2009). https://doi.org/10.1038/nrn2575
10. Carlsson, G.: Topological methods for data modelling. Nat. Rev. Phys. **2**(12), 697–708 (2019). https://doi.org/10.1038/s42254-020-00249-3
11. Carlsson, G.: Topology and data. Bull. Am. Math. Soc. **46**(2), 255–308 (2009). https://doi.org/10.1090/S0273-0979-09-01249-X
12. Centeno, E., Santos, F.N., MultinetLAB: Notebook for Network and Topological Analysis in Neuroscience (2021). https://doi.org/10.5281/zenodo.4483651
13. Centeno, E.G.Z., et al.: A hands-on tutorial on network and topological neuroscience. bioRxiv. p. 2021.02.15.431255. https://doi.org/10.1101/2021.02.15.431255
14. Curto, C.: What can topology tell us about the neural code? Bull. Am. Math. Soc. **54**(1), 63–78 (2017). https://doi.org/10.1090/bull/1554
15. Curto, C., Itskov, V.: Cell groups reveal structure of stimulus space. PLOS Comput. Biol. **4**(10), e1000205 (2008). https://doi.org/10.1371/journal.pcbi.1000205
16. Do Carmo, M.P.: Differential Geometry of Curves and Surfaces: Revised and Updated Second Edition. Courier Dover Publications, New York (2016)
17. Edelsbrunner, H., Harer, J.: Computational topology: an introduction, vol. 69. 1st edn. American Mathematical Society, Providence, USA (2010). https://doi.org/10.1090/mbk/069
18. Expert, P., et al.: Editorial: topological neuroscience. Network Neurosci. **3**(3), 653–655 (2019). https://doi.org/10.1162/netne00096
19. Farahani, F.V., Karwowski, W., Lighthall, N.R.: Application of graph theory for identifying connectivity patterns in human brain networks: a systematic review. Front. Neurosci. **13**, 585 (2019). https://doi.org/10.3389/fnins.2019.00585
20. Farooq, H., et al.: Network curvature as a hallmark of brain structural connectivity. Nat. Commun. **10**(1), 4937 (2019). https://doi.org/10.1038/s41467-019-12915-x
21. Fornito, A., Zalesky, A., Bullmore, E.: Fundamentals of Brain Network Analysis, 1st edn. Academic Press, San Diego (2016)
22. Gatica, M., et al.: High-order interdependencies in the aging brain. bioRxiv (2020)
23. Giusti, C., Pastalkova, E., Curto, C., Itskov, V.: Clique topology reveals intrinsic geometric structure in neural correlations. Proc. Nat. Acad. Sci. USA **112**(44), 13455–13460 (2015). https://doi.org/10.1073/pnas.1506407112
24. Gross, J.L., Yellen, J.: Handbook of Graph Theory, 1st edn. CRC Press, Boca Raton (2003)
25. Hagberg, A., Swart, P., S Chult, D.: Exploring network structure, dynamics, and function using networkx. In: Varoquaux, G.T., Vaught, J.M. (ed.) Proceedings of the 7th Python in Science Conference (SciPy 2008), pp. 11–15 (2008)
26. Hallquist, M.N., Hillary, F.G.: Graph theory approaches to functional network organization in brain disorders: a critique for a brave new small-world. Network Neurosci. **3**(1), 1–26 (2018). https://doi.org/10.1162/netna00054
27. Jalili, M.: Functional brain networks: does the choice of dependency estimator and binarization method matter? Sci. Rep. **6**, 29780 (2016). https://doi.org/10.1038/srep29780

28. Kartun-Giles, A.P., Bianconi, G.: Beyond the clustering coefficient: a topological analysis of node neighbourhoods in complex networks. Chaos Solitons Fract. X **1**, 100004 (2019). https://doi.org/10.1016/j.csfx.2019.100004

29. Lambiotte, R., Rosvall, M., Scholtes, I.: From networks to optimal higher-order models of complex systems. Nat. Phys. **15**(4), 313–320 (2019). https://doi.org/10.1038/s41567-019-0459-y

30. Liu, T.T.: Noise contributions to the FMRI signal: an overview. Neuroimage **143**, 141–151 (2016). https://doi.org/10.1016/j.neuroimage.2016.09.008

31. Maletić, S., Rajković, M., Vasiljević, D.: Simplicial complexes of networks and their statistical properties. In: Bubak, M., van Albada, G.D., Dongarra, J., Sloot, P.M.A. (eds.) ICCS 2008. LNCS, vol. 5102, pp. 568–575. Springer, Heidelberg (2008). https://doi.org/10.1007/978-3-540-69387-1_65

32. Maria, C., Boissonnat, J.-D., Glisse, M., Yvinec, M.: The Gudhi library: simplicial complexes and persistent homology. In: Hong, H., Yap, C. (eds.) ICMS 2014. LNCS, vol. 8592, pp. 167–174. Springer, Heidelberg (2014). https://doi.org/10.1007/978-3-662-44199-2_28

33. Najman, L., Romon, P. (eds.): Modern Approaches to Discrete Curvature. LNM, vol. 2184. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-58002-9

34. Pardalos, P.M., Xue, J.: The maximum clique problem. J. Glob. Optim. **4**(3), 301–328 (1994)

35. Petri, G., et al.: Homological scaffolds of brain functional networks. J. R. Soc. Interface **11**(101), 20140873 (2014). https://doi.org/10.1098/rsif.2014.0873

36. Rosas, F.E., et al.: Quantifying high-order interdependencies via multivariate extensions of the mutual information. Phys. Rev. E **100**(3), 032305 (2019). https://doi.org/10.1103/PhysRevE.100.032305

37. Saggar, M., et al.: Towards a new approach to reveal dynamical organization of the brain using topological data analysis. Nat. Commun. **9**(1), 1399 (2018). https://doi.org/10.1038/s41467-018-03664-4

38. Santos, F.A.N., et al.: Topological phase transitions in functional brain networks. Phys. Rev. E **100**(3–1), 032414 (2019). https://doi.org/10.1103/PhysRevE.100.032414

39. Singh, G., et al.: Topological analysis of population activity in visual cortex. J. Vis. **8**(8), 11 (2008). https://doi.org/10.1167/8.8.11

40. Sizemore, A.E., Giusti, C., Kahn, A., Vettel, J.M., Betzel, R.F., Bassett, D.S.: Cliques and cavities in the human connectome. J. Comput. Neurosci. **44**(1), 115–145 (2017). https://doi.org/10.1007/s10827-017-0672-6

41. Sizemore Blevins, A., Bassett, D.S.: Reorderability of node-filtered order complexes. Phys. Rev. E **101**(5–1), 052311 (2020). https://doi.org/10.1103/PhysRevE.101.052311

42. Sporns, O.: Graph theory methods: applications in brain networks. Dialogues Clin. Neurosci. **20**(2), 111–121 (2018)

43. Watts, D.J., Strogatz, S.H.: Collective dynamics of 'small-world' networks. Nature **393**(6684), 440 (1998). https://doi.org/10.1038/30918

44. Weber, M., et al.: Curvature-based methods for brain network analysis. arXiv preprint arXiv:1707.00180 (2017)

45. Wu, Z., et al.: Emergent complex network geometry. Sci. Rep. **5**, 10073 (2015). https://doi.org/10.1038/srep10073

46. Zalesky, A., Fornito, A., Bullmore, E.: On the use of correlation as a measure of network connectivity. Neuroimage **60**(4), 2096–2106 (2012). https://doi.org/10.1016/j.neuroimage.2012.02.001
47. Zomorodian, A.J.: Topology for Computing, vol. 16, 1st edn. Cambridge University Press, New York (2005)