Response Letter

Manuscript ID: NIMG-23-337
Unified Topological Inference for Brain Networks
in Temporal Lobe Epilepsy Using the Wasserstein Distance

We thank editors and reviewers for constructive comments and corrections that improved the quality of revision substantially. We addressed every comment and question very carefully in the resubmission. The edited parts as well as new materials are colored in green in the main manuscript. Any major changes are all listed in this response letter.

Reviewer 1

R-1: *However, even though the authors claim to use TDA- or PH-derived features, what they in fact do is to consider graph filtrations (as opposite to filtrations of simplicial complexes). This came long before TDA methods appeared.*

Response. Following the reviewers suggestion, we added more explanations. Graph filtration is a special case of Rips filtration. Also an additional schematic contrasting between the two filtrations is provided. The response below are incorporated into the revision.

Rips filtration does not scale well with increased data size. The computational complexity of Rips filtration grows exponentially with the number of nodes making it impractical for large datasets [Topaz et al., 2015]. The Rips filtration on p number of nodes indexed as $\{1, 2, \cdots, p\}$ can have up to $(p-1)$-simplex that connect all the nodes in a single connected component. In practice, the Rips filtration is restricted to k-skeletons, simplicial complexes containing simplices of dimension at most k. If we denote $\mathcal{R}_{k,\epsilon}$ to be the Rips filtration restricted to k-skeletons, we obtain the sequence of nested simplices

$$\mathcal{R}_{k,\epsilon_0} \subset \mathcal{R}_{k,\epsilon_1} \subset \mathcal{R}_{k,\epsilon_2} \subset \cdots \subset \mathcal{R}_{k,e_{\max}}$$

for filtration values $0 = \epsilon_0 < \epsilon_1 < \epsilon_2 < \cdots < e_{\max}$, where $\epsilon_{\max}$ is the largest pairwise distance d_{ij} between nodes i and j, i.e., $\epsilon_{max} = \max_{ij} d_{ij}$. $\mathcal{R}_{k,\epsilon_{\max}}$ is a single connected $(p-1)$-simplex that contains all the nodes.

The graph filtration is a special case of Rips filtration restricted to 1-simplices given by $\mathcal{X}_\epsilon = \mathcal{R}_{1,\epsilon_{\max}-\epsilon}$. Then the graph filtration yields the sequence of nested graphs

$$\mathcal{X}_{\epsilon_0} \supset \mathcal{X}_{\epsilon_1} \supset \mathcal{X}_{\epsilon_2} \supset \cdots \supset \mathcal{X}_{e_{\max}}.$$

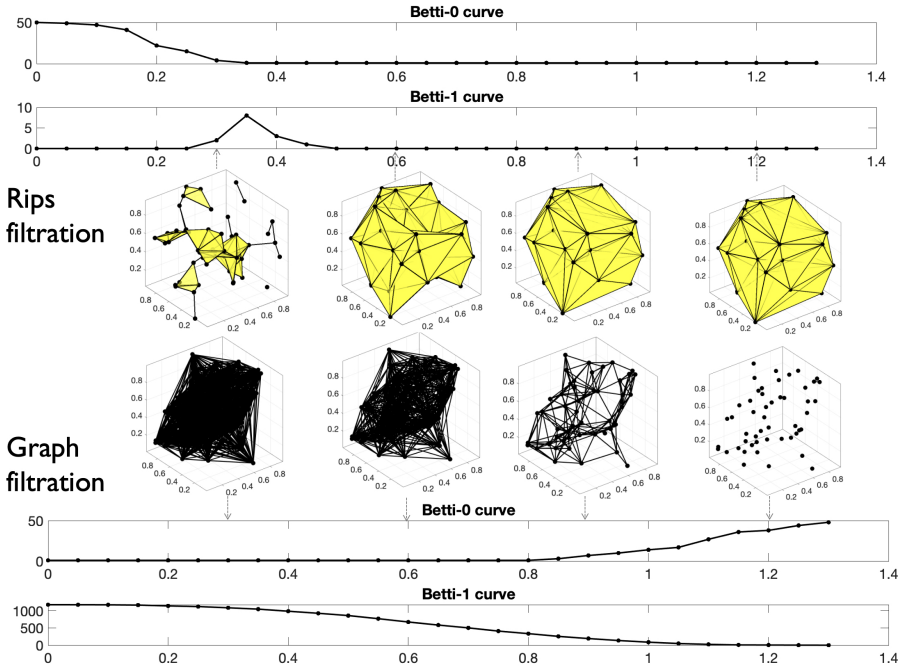

Figure 1: The comparison between the Rips and graph filtrations performed on scatter points. The Euclidean distance between points are used as edge weights. Unlike Rips filtrations, β_0 and β_1 curves for the graph filtration are monotone.

Figure 1 displays the comparison of the two filtrations for randomly generated 50 nodes in a cube. The Euclidean distance between points are used as edge weights. In the both filtrations, β_0-curves are monotone. β_1-curve for the Rips filtration is not monotone while the β_1-curve for the graph filtration is monotone and gradually changing over the whole range of filtration values [Chung et al., 2019]. Thus, the 0D persistent diagrams for the both Rips and graph filtrations are 1D scatter points. While the 1D persistent diagram for Rips filtration is 2D scatter points, the 1D persistent diagram for graph filtration is 1D scatter points. This allows the proposed scalable computation of the Wasserstein distance for graph filtrations.

R-2: *From a statistical perspective, it is not clear to me what the authors mean with "A key obstacle of applying persistent homology to brain network studies has always been the lack of coherent statistical inference framework". The authors should explain better what they mean, considering that PH-diagrams have plenty of representations in Hilbert spaces, where, as far as I know, it is possible (and actually is the right framework) to perform statistics.*

Response. Following reviewer's suggestion, we have rephrased the abstract and additional literature review is added in the introduction as follows.

Persistent homology can extract hidden topological signals present in brain networks. Persistent homology summarizes the changes of topological structures over multiple different scales called filtrations. Doing so detect hidden topological signals that persist over multiple scales. The topological changes are then summarized into topological features such as the persistent diagrams. The topological dissimilarity is then measured through the Wasserstein distance between the persistent diagrams. Since the Wasserstein distance does not follow a known distribution, it is challenging to apply existing parametric methods for statistically quantifying the distance in brain network studies. To address this problem, we present a unified topological inference framework on the Wasserstein distance. Our approach has no explicit models and distributional assumptions. The inference is performed in a completely data driven fashion.

The Wasserstein distance is commonly utilized to measure the dissimilarity between persistent diagrams. Despite its widespread use in comparing persistent diagrams, the Wasserstein distance is rarely employed in statistical inference due to its computational complexity, lack of scalability, and the absence of well-known statistical distributions associated with it. Instead, researchers have turned to the vectorization of persistent diagrams as a more practical and efficient alternative for statistical inference. Vectorization involves transforming a persistent diagram into a vector representation, making it amenable to standard machine learning and statistical techniques. [Chung et al., 2009] vectorized the persistent diagram into images by counting the number of scatter points in the unit squares. [Bubenik, 2015] vectorized the persistent diagram into a sequence of tent functions, known as the persistence landscape. [Adams et al., 2017] converted the persistent diagram into a discrete, grid-based representation referred to as the persistence image. In this paper, we demonstrate the feasibility of developing a coherent statistical inference framework based on the Wasserstein distance. Our method simply bypasses the need for vectorization of persistent diagrams.

R-3: *Coming to the neuroimaging side, this is not the first work involving PH-derived methods in applications to neuroimaging. Such studies have shown that, when TDA is involved, epileptic cases are "easier", if compared to other brain networks coming from other neurological problems. On this regard, I do not see such considerations and connections to recent literature in the manuscript. How would the methods compare to those? And how would it perform in, say, schizophrenia patients? Another point not clear to me, but related, concerns the topological localizations in TDA analysis. I invite the authors to give a look at the paper "Stolz, B. J., Emerson, T., Nahkuri, S., Porter, M. A., Harrington, H. A., 2021. Topological data analysis of task-based fMRI data from experiments on schizophrenia. J. Phys. Complex. 2 035006". In the work, in fact, a localization of "best cylces" is provided and related to the brain areas.*

Response. Following reviewer's suggestion, we provided additional recent literature review on the use of TDA in neuroimaging.

Starting with the first TDA application in brain imaging in [Chung et al., 2009], where the Morse filtration is used in characterizing the cortical thickness of autistic children, there have been numerious application of TDA in brain imaging. [Lee et al., 2011a,b] demonstrated the use persistent homology in modeling functional brain network for the first time. In recent years, TDA has gained increasing interest in the neuroimaging community, with various applications in different modalities and disorders. For instance, TDA has also been employed in the analysis of resting-state fMRI data [Petri et al., 2014], where TDA was used to identify topological changes in brain networks under the influence of the psychedelic compound psilocybin. [Lord et al., 2016] applied TDA to investigate the topological properties of resting-state networks in comparison against graph theory approaches. [Giusti et al., 2016] proposed to use the simplical homology in modeling higher order brain connectivity. In [Wang et al., 2017, 2018], persistent homology was shown to outperform topographic power maps, power spectral density and local variance methods in an EEG study. In [Yoo et al., 2017], center persistency was shown to outperform the network-based statistic and element-wise multiple corrections. [Saggar et al., 2018] applied TDA to task fMRI for tracking within- and between-task transitions. [Stolz et al., 2021] demonstrated TDA in characterizing the task-based fMRI of schizophrenia patients.

TDA has been used to analyze structural connectivity as well. Starting with the first structural connectivity study with TDA in [Chung et al., 2011], where the Rips filtration is used to model white matter fiber tracts. [Reimann et al., 2017] used Betti numbers in modeling synaptic neural networks. [Sizemore et al., 2018] used TDA to explore the topological cavities that exists across subjects in structural connectivity using diffusion spectrum imaging.

R-4: *Finally, as a personal point, a comparison to other existing methods should be considered in the study.*

Since the real dataset does not have the ground truth, we compared the our method against existing methods in a simulation study with the ground truth. We compared the the proposed Wasserstein distance against three established topological distances: bottleneck, Gromov-Hausdorff (GH) and Kolmogorov-Smirnov (KS). The bottleneck distance perhaps the most often used distance in persistent homology [Cohen-Steiner et al., 2007, Edelsbrunner and Harer, 2008]. The Gromov-Hausdorff (GH) distance is possibly the most popular distance that is originally used to measure distance between two metric spaces [Tuzhilin, 2016]. It was later adapted to measure distances in persistent homology, dendrograms [Carlsson and Memoli, 2008, Carlsson and Mémoli, 2010, Chazal et al., 2009] and brain networks [Lee et al., 2011b, 2012]. The Kolmogorov-Smirnov (KS) distance was introduced in [Chung, 2012, Chung et al., 2013, 2017b, Lee et al.,

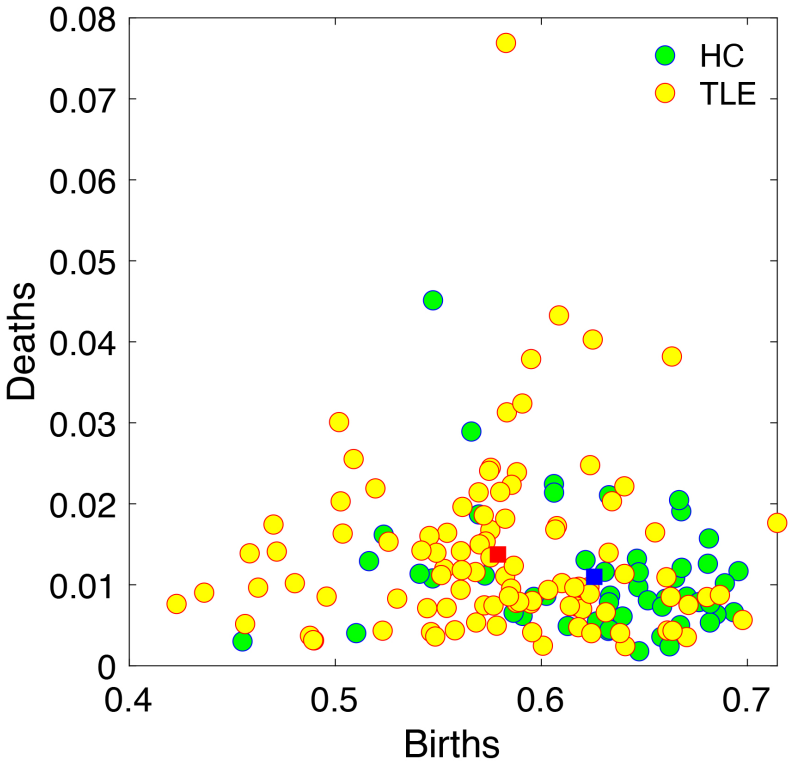

Figure 2: Topological embedding of 151 subjects. Green circles are HC and yellow circles are TLE. The blue square is the topological center of HC while the red square is the topological center of TLE. The horizontal axis represents 0D topology (connected components) through birth values while the vertical axis represents 1D topology (circles) through death values.

2017] to measure the distance between Betti curves. The probability distributions of all the distances are unknown. Thus, the statistical inference has been done through the permutation test.

Reviewer 3

R-5: *The within-group distance for TLE is 114.62 +-147.67 and for HC is 110.65+-124.78. The p-value is quite significant. The authors should give some insights on why the data are so similar and both variances are so large.*

Sorry for confusion. The p-value is not obtained testing for the within-group distance difference between TLE and HC. The p-value is obtained testing for the significance of the between-group distance over the within-group difference.

To answer this in a more intuitive fashion, we developed a new topological embedding technique (Figure 2). We responded as follows in the mainbody.

Figure 2 displays the spread of each subject with respect to the group topological mean (blue square for HC and red square for TLE), where the x-axis shows the spread with respect to the 0D topology and the y-axis shows the spread with respect to the 1D topology. Given sorted birth values $b^k_{(i)}$ for k-th subject, the x-coordinates of the group topological mean is given by

$$\mu_b = \frac{1}{n}\sum_{k=1}^{n} b^k_{(i)}.$$

The y-coordinates of the group topological mean is obtained similarly using death values. The embedding x-coordinate of k-th subject is then

$$\frac{1}{n}\sum_{k=1}^{n}(b^k_{(i)} - \mu_b) = \frac{1}{n}\sum_{k=1}^{n}(b^k_i - \mu_b).$$

The embedding y-coordinate of k-th subject is similarly given using the death values. The embedding shows the relative topological distance with respect to the topological center of each group. We can clearly see more spread for TLE compared to HC. We also can see that 0D topology is the main topological discriminator separating the two groups while 1D topology may not be able to discriminate the groups. Similar observation was made in the huge β_0-curve shape difference while there is almost no shape difference in β_1-curve.

The pairwise distance within TLE is 114.62±147.67 while the pairwise distance within HC is 110.65±124.78. The average pairwise distance within a group should not be too different between TLE and HC. Since the pairwise distance varies from the smallest pairwise distance to the large pairwise distance, variability can be fairly large as illustrated in Figure 2. What separates TLE and HC is the between-group distance which measures the sum of all possible pairwise distance between a TLE subject and a HC subject.

R-6: It would be more interesting if the authors could compare their work with some other topology analysis tools and demonstrate their superiority.

See our response in **R-4:**, where we compared our method against existing topological methods.

References

H. Adams, T. Emerson, M. Kirby, R. Neville, C. Peterson, P. Shipman, S. Chepushtanova, E. Hanson, F. Motta, and L. Ziegelmeier. Persistence images: A stable vector representation of persistent homology. *The Journal of Machine Learning Research*, 18:218–252, 2017.

W. Böhm and K. Hornik. A Kolmogorov-Smirnov test for r samples. *Institute for Statistics and Mathematics*, Research Report Series:Report 105, 2010.

P. Bubenik. Statistical topological data analysis using persistence landscapes. *Journal of Machine Learning Research*, 16:77–102, 2015.

G. Carlsson and F. Memoli. Persistent clustering and a theorem of J. Kleinberg. *arXiv preprint arXiv:0808.2241*, 2008.

G. Carlsson and F. Mémoli. Characterization, stability and convergence of hierarchical clustering methods. *Journal of Machine Learning Research*, 11: 1425–1470, 2010.

F. Chazal, D. Cohen-Steiner, L. Guibas, F. Mémoli, and S. Oudot. Gromov-Hausdorff stable signatures for shapes using persistence. In *Computer Graphics Forum*, volume 28, pages 1393–1403, 2009.

M. Chung. *Computational Neuroanatomy: The Methods*. World Scientific, Singapore, 2012.

M. Chung, P. Bubenik, and P. Kim. Persistence diagrams of cortical surface data. *Proceedings of the 21st International Conference on Information Processing in Medical Imaging (IPMI), Lecture Notes in Computer Science (LNCS)*, 5636:386–397, 2009.

M. Chung, N. Adluru, K. Dalton, A. Alexander, and R. Davidson. Scalable brain network construction on white matter fibers. In *Proc. of SPIE*, volume 7962, page 79624G, 2011.

M. Chung, J. Hanson, H. Lee, N. Adluru, A. L. Alexander, R. Davidson, and S. Pollak. Persistent homological sparse network approach to detecting white matter abnormality in maltreated children: MRI and DTI multimodal study. *MICCAI, Lecture Notes in Computer Science (LNCS)*, 8149:300–307, 2013.

M. Chung, J. Hanson, J. Ye, R. Davidson, and S. Pollak. Persistent homology in sparse regression and its application to brain morphometry. *IEEE Transactions on Medical Imaging*, 34:1928–1939, 2015.

M. Chung, H. Lee, V. Solo, R. Davidson, and S. Pollak. Topological distances between brain networks. *International Workshop on Connectomics in Neuroimaging*, 10511:161–170, 2017a.

M. Chung, V. Vilalta-Gil, H. Lee, P. Rathouz, B. Lahey, and D. Zald. Exact topological inference for paired brain networks via persistent homology. *Information Processing in Medical Imaging (IPMI), Lecture Notes in Computer Science (LNCS)*, 10265:299–310, 2017b.

M. Chung, H. Lee, A. DiChristofano, H. Ombao, and V. Solo. Exact topological inference of the resting-state brain networks in twins. *Network Neuroscience*, 3:674–694, 2019.

D. Cohen-Steiner, H. Edelsbrunner, and J. Harer. Stability of persistence diagrams. *Discrete and Computational Geometry*, 37:103–120, 2007.

A. Cole. *Identifying and exploiting structure in cosmological and string theoretic data*. Ph.D. Thesis, University of Wisconsin-Madison, 2020.

H. Edelsbrunner and J. Harer. Persistent homology - a survey. *Contemporary Mathematics*, 453:257–282, 2008.

J. D. Gibbons and S. Chakraborti. *Nonparametric Statistical Inference*. Chapman & Hall/CRC Press, 2011.

C. Giusti, R. Ghrist, and D. Bassett. Two's company, three (or more) is a simplex. *Journal of computational neuroscience*, 41:1–14, 2016.

H. Lee, M. Chung, H. Kang, B.-N. Kim, and D. Lee. Discriminative persistent homology of brain networks. In *IEEE International Symposium on Biomedical Imaging (ISBI)*, pages 841–844, 2011a.

H. Lee, M. Chung, H. Kang, B.-N. Kim, and D. Lee. Computing the shape of brain networks using graph filtration and Gromov-Hausdorff metric. *MICCAI, Lecture Notes in Computer Science*, 6892:302–309, 2011b.

H. Lee, H. Kang, M. Chung, B.-N. Kim, and D. Lee. Persistent brain network homology from the perspective of dendrogram. *IEEE Transactions on Medical Imaging*, 31:2267–2277, 2012.

H. Lee, H. Kang, M. Chung, S. Lim, B.-N. Kim, and D. Lee. Integrated multimodal network approach to PET and MRI based on multidimensional persistent homology. *Human Brain Mapping*, 38:1387–1402, 2017.

L.-D. Lord, P. Expert, H. Fernandes, G. Petri, T. Van H., F. Vaccarino, G. Deco, F. Turkheimer, and M. Kringelbach. Insights into brain architectures from the homological scaffolds of functional connectivity networks. *Frontiers in systems neuroscience*, 10:85, 2016.

G. Petri, P. Expert, F. Turkheimer, R. Carhart-Harris, D. Nutt, P. Hellyer, and F. Vaccarino. Homological scaffolds of brain functional networks. *Journal of The Royal Society Interface*, 11:20140873, 2014.

M. Reimann, M. Nolte, M. Scolamiero, K. Turner, R. Perin, G. Chindemi, P. Dłotko, R. Levi, K. Hess, and H. Markram. Cliques of neurons bound into cavities provide a missing link between structure and function. *Frontiers in Computational Neuroscience*, 11:48, 2017.

M. Saggar, O. Sporns, J. Gonzalez-Castillo, P. Bandettini, G. Carlsson, G. Glover, and A. Reiss. Towards a new approach to reveal dynamical organization of the brain using topological data analysis. *Nature Communications*, 9:1399, 2018.

A. Sizemore, C. Giusti, A. Kahn, J. Vettel, R. Betzel, and D. Bassett. Cliques and cavities in the human connectome. *Journal of computational neuroscience*, 44:115–145, 2018.

B. Stolz, T. Emerson, S. Nahkuri, M. Porter, and H. Harrington. Topological data analysis of task-based fMRI data from experiments on Schizophrenia. *Journal of Physics: Complexity*, 2:035006, 2021.

C. Topaz, L. Ziegelmeier, and T. Halverson. Topological data analysis of biological aggregation models. *PLoS One*, page e0126383, 2015.

A. Tuzhilin. Who invented the Gromov-Hausdorff distance? *arXiv preprint arXiv:1612.00728*, 2016.

Y. Wang, M. Chung, D. Dentico, A. Lutz, and R. Davidson. Topological network analysis of electroencephalographic power maps. In *International Workshop on Connectomics in NeuroImaging, Lecture Notes in Computer Science (LNCS)*, volume 10511, pages 134–142, 2017.

Y. Wang, H. Ombao, and M. Chung. Topological data analysis of single-trial electroencephalographic signals. *Annals of Applied Statistics*, 12:1506–1534, 2018.

K. Yoo, P. Lee, M. Chung, W. Sohn, S. Chung, D. Na, D. Ju, and Y. Jeong. Degree-based statistic and center persistency for brain connectivity analysis. *Human Brain Mapping*, 38:165–181, 2017.