0-dimensional Homology Preserving Dimensionality Reduction with TopoMap

Harish Doraiswamy
New York University
harishd@nyu.edu

Julien Tierny
CNRS / Sorbonne Universite
julien.tierny@sorbonne-universite.fr

Paulo J. S. Silva
University of Campinas
pjssilva@ime.unicamp.br

Luis Gustavo Nonato
University of Sao Paulo
gnonato@icmc.usp.br

Claudio Silva
New York University
csilva@nyu.edu

Abstract

This note presents *TopoMap* [20], a novel dimensionality reduction technique which provides topological guarantees during the mapping process. In particular, TopoMap performs the mapping from a high-dimensional space to a visual space, while preserving the 0-dimensional persistence diagram of the Rips filtration of the high-dimensional data, ensuring that the filtrations generate the same connected components when applied to the original as well as projected data. The presented case studies show that the topological guarantee provided by TopoMap not only brings confidence to the visual analytic process but also can be used to assist in the assessment of other projection methods.

1 Introduction

Dimensionality reduction, mapping data points from a d-dimensional to a d'-dimensional Cartesian space (with $d' << d$), is an important tool for visual data analysis. Over the last decades, multidimensional projection (MDP) methods [51, 66] have focused on preserving *geometric* properties (such as the Euclidean distance between data points). This note presents *TopoMap* [20], a novel MDP technique that is guaranteed to preserve *topological* structures during the dimensionality reduction process (specifically, preserving the 0-dimensional persistence diagram of the Rips filtration of the input data points). The topological guarantee provided by TopoMap allows analysts to confidently explore high-dimensional data by visualizing which groups of objects are more tightly connected in the high-dimensional space. The main contributions of this work are:

- A dimensionality reduction technique called TopoMap, which is guaranteed to preserve the 0-dimensional persistence of the Rips filtration of the input high-dimensional point cloud data.
- An optimization procedure that minimizes pairwise distances in the output, while ensuring the correct mapping of the connected components resulting from the filtration.
- An exhaustive evaluation using both labeled and unlabeled data, showing the potential of TopoMap to support the analysis of high-dimensional data.

Topological Data Analysis and Beyond Workshop at the 34th Conference on Neural Information Processing Systems (NeurIPS 2020), Vancouver, Canada.

To the best of our knowledge, TopoMap is the first dimensionality reduction technique to provide constructive guarantees as to the preservation of the topological properties of the Rips filtration of the data under analysis. This note briefly presents an overview of TopoMap. Complete technical details, arguments and experiments are available in the original publication introducing TopoMap [20].

2 Related Work

Multi-dimensional projection (MDP) has been a fundamental analytical tool for a long time, mainly in the context of data visualization [3, 28, 37, 38, 40, 46, 51]. The extensive literature about MDP techniques has been organized over several books [10, 43] and surveys [18, 51, 78]. In the following, we focus on MDP methods which include topological considerations. We refer interested readers to the above books and surveys for a broader discussion about MDP methods.

In general, topology-based methods [22] have been very popular in the last two decades to support advanced data analysis and visualization tasks [36]. By providing a concise, structural representation of the data, these techniques greatly help in the visualization and analysis of the data. They have been applied successfully to a variety of domains, such as astrophysics [69,73], biological imaging [2,9,14], chemistry [8, 29, 56], fluid dynamics [39], material sciences [33, 34, 72], or turbulent combustion [11,31,42]. The Rips filtration [6,22] is often used to analyze the topology of high dimensional point clouds. In particular, it has been shown that it can reliably capture the homology of the manifold sampled by the point cloud [15]. Given the increasing popularity of topological methods, several open source software packages have been made available over the years [1,7,12,24,45,49,50,74,76].

Several MDP approaches aimed at preserving some topological information during the projection [4]. Topological concepts have also been used to evaluate the quality of dimensionality reduction techniques [58,63,64]. Sharing motivations with previous work on skeletonization [41], Yan et al. [80] introduce a variant of Landmark Isomap [70, 75], which exploits the *Mapper* [71] (an approximation of the Reeb graph [17, 30, 57, 62]), to identify potential handles in the data and to preserve them as much as possible by enforcing landmark constraints in 2D. Gerber et al. [26, 27] introduce projection methods driven by the cells of maximum dimension (called crystals) of the Morse-Smale complex [25, 32, 65, 68] of high dimensional scalar functions. In a similar context, Weber et al. [79] introduce a terrain metaphor to provide an intuitive visualization of the topological features present in a volume scalar field. The concept has been extended to high-dimensional point clouds in a series of papers [35, 52–55], which present techniques generating 2D terrains admitting the same contour tree (loop-free Reeb graph) as the input high dimensional point cloud (with regard to a kernel density estimation function for instance). More recently, McInnes et al. introduce UMAP [47], an MDP approach based on category theory, which is now often considered as a more modern and scalable alternative to t-SNE [77]. Moor et al. [48] introduce a numerical framework aiming at minimizing a distance between the persistence diagrams of the input and projected point clouds. In contrast, TopoMap specifically focuses on the preservation of the 0-dimensional persistence diagram, which is strictly enforced constructively (and not approximated via numerical optimization). Shieh et al. [67] introduce a method (with cubic time complexity) aiming at preserving the single-linkage dendrogram produced by hierarchical clustering, which, as discussed in section 3, has some connection to our work. However, their approach is also based on numerical optimization (extended from multi-dimensional scaling by the introduction of topological constraints) and the documented greedy approximation scheme tries to satisfy the topological constraints as much as possible (as discussed by the authors). In contrast, TopoMap provides strong guarantees as the preservation of the 0-dimensional persistence diagram is strictly enforced constructively (i.e. from a combinatorial point of view).

3 Topology Preserving Projection

Let $P = \{p_1, p_2, \dots, p_n\}$ be a set of points in $\mathbb{R}^d$. Given a distance threshold δ, the Vietoris–Rips complex [22] is the collection of all k-simplexes K obtained from P ($k \geq 0$), such that $d(p_i, p_j) \leq \delta$, $\forall p_i, p_j \in K$. Here, $d(\cdot,\cdot)$ is the Euclidean distance (Figure 1). Consider the growth as defined by the Rips filtration, wherein the simplexes are added one at a time. That is, the i^{th} iteration in this growth yields a complex $\mathbb{K}_P^i = \{K_0, K_1, \dots, K_{i-1}\}$. The addition of each new simplex can change the topology of $\mathbb{K}_P^i$, where the topology is captured by the set of k-cycles (i.e. formal sums with empty boundary of k simplices [22]) in $\mathbb{K}_P^i$. Given a k-cycle ($k \geq 0$), let δ_c be the threshold at which this

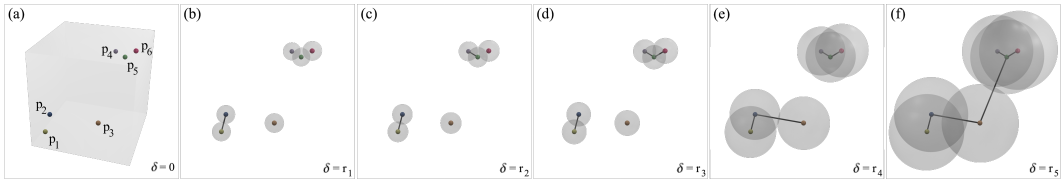

Figure 1: Different stages of the Rips filtration of a 3D point cloud P **(a)**, with increasing diameter δ **(b)–(f)**. These stages correspond to the instants in the filtration when two components (0-cycles) merge into one. The edge from the Rips filtration responsible for this merge is also shown. Note that this collection of edges correspond to the minimum spanning tree of the input point cloud.

cycle is created, and δ_d the threshold at which it is destroyed. Then the *topological persistence* [23] of this k-cycle is defined as $\delta_d - \delta_c$, and intuitively captures the lifetime of this cycle in the given filtration. The k-dimensional *persistence diagram* [16], noted PD_P^k, plots each persistent k-cycle in 2D as a point at coordinates (δ_c, δ_d). Our goal is to construct a set of points $P' = \{p'_1, p'_2, \ldots, p'_n\}$ in $\mathbb{R}^2$ such that $PD_P^0 = PD_{P'}^0$.

Given a Rips filtration defined over a set of n points, there is exactly $(n-1)$ *topology changing* edges that result in reducing the number of 0-cycles. Specifically, the set of topology changing edges ordered by increasing length $E_P^{n-1} = \{e_1, e_2, \ldots, e_{n-1}\}$ is precisely the minimum spanning tree (MST) of P [20] (see Figure 1). Let C_P^i be the set of connected components (i.e. set of connected vertex sets) of $\mathbb{K}_P^i$ and $PD_P^0(i)$ the 0-dimensional persistence diagram of $\mathbb{K}_P^i$. It follows that there exists at least one mapping $\mathscr{M} : \mathbb{R}^d \rightarrow \mathbb{R}^{d'}$ with $\mathscr{M}(P) = P'$, such that (see [20] for more details):

(a) $|e'_i| = |e_i|, \quad \forall i \in [0, n-1], \quad e_i \in \mathbb{K}_P^i, \quad e'_i \in \mathbb{K}_{P'}^i$

(b) $C_{P'}^i = C_P^i, \quad \forall i \in [0, n-1]$

(c) and thus $PD_{P'}^0(i) = PD_P^0(i), \quad \forall i \in [0, n-1]$

The existence of $\mathscr{M}$ can be shown constructively [20]. In particular, TopoMap projects P to $\mathbb{R}^2$ (i.e., $d' = 2$), while enforcing the above conditions (a) and (b) by iteratively placing the points of P' while guaranteeing that the lengths of the edges of both minimum spanning trees (E_P^i and $E_{P'}^i$) exactly coincide for each i. As illustrated in Figure 2, TopoMap iteratively introduces the edges of $E_{P'}^i$. In particular let C_1 and C_2 be two connected components of $\mathbb{K}_P^i$ merged by the edge $e_i \in E_P^i$ and C'_1 and C'_2 their images by $\mathscr{M}$. TopoMap starts by rotating C'_1 and C'_2 such that an edge of their convex hull (for instance the longest) gets aligned with the x-axis. Then C'_2 is translated along the y-axis such that the minimum distance between the vertices of C'_1 and C'_2 equals $|e_i|$ (therefore enforcing $|e'_i| = |e_i|$ and maintaining the components C_1 and C_2 of the Rips filtration). In general, the rotations of C'_1 and C'_2 can be chosen arbitrarily and the vertices v'_1 and v'_2 (respectively belonging to the convex hulls of C'_1 and C'_2) linked by the edge e'_i can also be chosen arbitrarily. In particular, these variables can be optimized to minimize the overall sum of pairwise distances in P', as further described in [20].

Given the equivalence between the 0-dimensional homology of the Rips filtration and hierarchical clustering using single-linkage criterion [13], it follows that TopoMap also provides 2D layouts for which the hierarchical clustering will be identical to that computed with the input high-dimensional data. In other words, TopoMap preserves, constructively, hierarchical clustering with single-linkage.

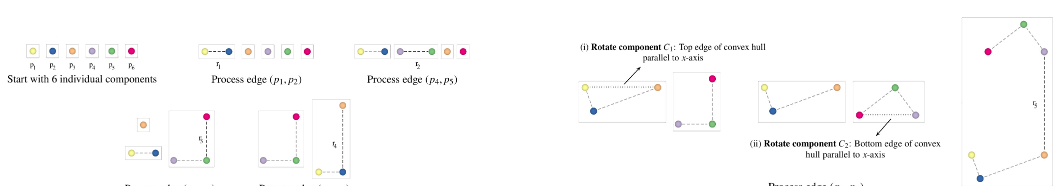

Figure 2: Iterative placement of the points of P' in 2D (left to right) given the input P from Figure 1.

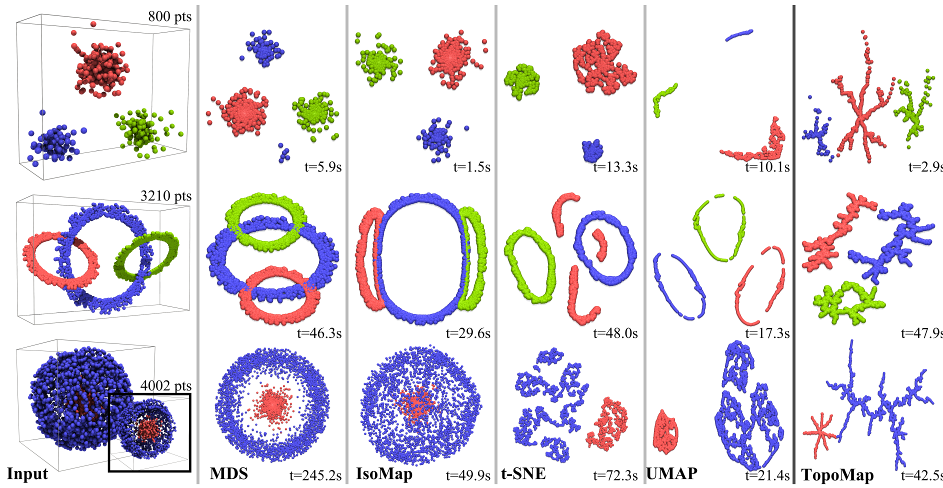

Figure 3: Projecting three-dimensional data to a 2D space using geometry preserving projections: Classical MDS, Isomap, tSNE, UMAP; and the proposed topology preserving TopoMap method.

4 Experimental Results

TopoMap was implemented in C++ using the *mlpack* library [19] (for the efficient computation of the Euclidean distance minimum spanning tree [44], in $\mathcal{O}\big(nlogn\alpha(n)\big) \approx \mathcal{O}(nlogn)$ steps, where n is the number of input points) and the *qhull* library [5] (for the computation of the convex hulls, in $\mathcal{O}(n'logn')$, where n' is the number of points in the considered component, with $n' << n$). For the optimization procedure (further described in [20]), the Algencan [60, 61] library was used. All experiments were run on a machine with an Intel Xeon CPU E5-2630 v2 running at 2.60GHz and 64GB of memory. Figure 3 presents some visual comparisons on synthetic three-dimensional examples between the projections computed by TopoMap (1 color per connected component) and these computed by state-of-the-art methods (using the scikit-learn [59] and the UMAP authors' implementations). This figure shows that geometry preserving methods tend either to split connected components or mix them up, while TopoMap is guaranteed to preserve them, leveraging more reliable analysis. Figure 4 presents similar comparisons on public labeled data sets [21] (1 color per label). As shown in Figure 4, TopoMap produces characteristic star-shaped projections, where the centers of the stars correspond to the clusters which can be identified independently by hierarchical clustering with single-linkage. In contrast, points located at the tips of long branches correspond to outliers or boundaries between clusters. To further stress this, a transparent grey color map can be applied to the output (right side of the TopoMap column) by considering, in the planar projection, a kernel density estimate with a Gaussian kernel. Overall, TopoMap then produces 2D layouts where important clusters can be readily identified visually, and where clustering uncertainties can be conveyed via the grey color map in the less dense regions of the projections. We refer the reader to the original TopoMap publication [20], for further evaluations, including detailed use cases on unlabeled data, where TopoMap is shown to nicely preserve meaningful clusters at the centers of its stars.

5 Conclusions

This note presented TopoMap [20], to our knowledge the first planar projection technique that is guaranteed to preserve the homology of 0-cycles of the Rips filtration. Experiments using a variety of data sets demonstrated several key properties that are desirable in visual analysis: the layout is easy to understand while its theoretical guarantees provide confidence to the users. In the future, we would like to explore ways in which 1-cycles can be preserved as well in the projection. Analyzing the effectiveness of TopoMap to assist clustering mechanisms is another direction we will pursue.

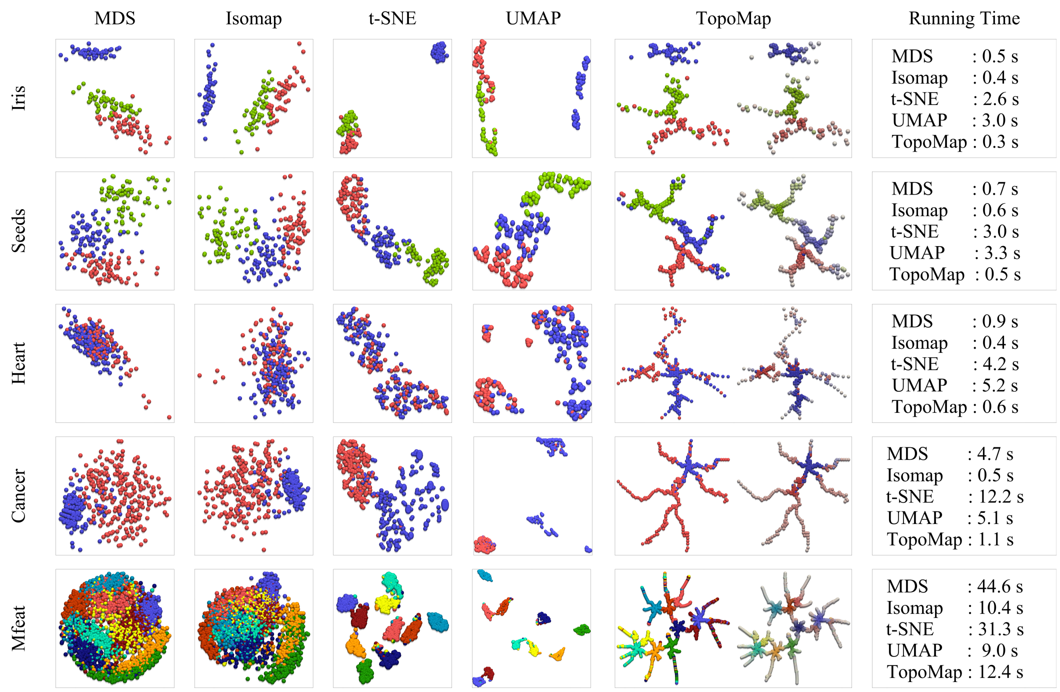

Figure 4: Planar projections produced by MDS, Isomap, t-SNE, UMAP and TopoMap when applied to five labeled data sets (Iris: 150 points in $\mathbb{R}^5$, Seeds: 210 points in $\mathbb{R}^8$, Heart: 261 points in $\mathbb{R}^{11}$, Cancer: 699 points in $\mathbb{R}^{11}$, Mfeat: 2,000 points in $\mathbb{R}^{64}$). The right images in the TopoMap column highlight with a transparent grey color map the denser areas in the left images.

Acknowledgments and Disclosure of Funding

The authors would like to thank the reviewers for their thoughtful remarks and suggestions. This work was partially supported by the DARPA D3M program; Moore Sloan Data Science Environment at NYU; NSF awards CNS-1229185, CCF-1533564, CNS-1544753, CNS-1730396, CNS-1828576; European Commission grant ERC-2019-COG "TORI" (ref. 863464), CNPq-Brazil (303552/2017-4, 304301/2019-1); and the Sao Paulo Research Foundation (FAPESP) - Brazil (2013/07375-0, 2016/04190-7, 2018/07551-6, 2018/24293-0). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of NSF and DARPA.

References

[1] H. Adams, A. Tausz, and M. Vejdemo-Johansson. Javaplex: A research software package for persistent (co)homology. In *ICMS*, 2014. https://github.com/appliedtopology/javaplex.

[2] K. Anderson, J. Anderson, S. Palande, and B. Wang. Topological data analysis of functional MRI connectivity in time and space domains. In *MICCAI Workshop on Connectomics in NeuroImaging*, 2018.

[3] M. Aupetit. Visualizing distortions and recovering topology in continuous projection techniques. *Neurocomputing*, 70(7):1304–1330, 2007.

[4] M. Aupetit and T. Catz. High-dimensional labeled data analysis with topology representing graphs. *Neurocomputing*, 63:139–169, 2005.

[5] C. B. Barber, D. P. Dobkin, and H. Huhdanpaa. The quickhull algorithm for convex hulls. *ACM Trans. Math. Softw.*, 22(4):469–483, Dec. 1996.

[6] U. Bauer. Ripser: efficient computation of vietoris-rips persistence barcodes, Aug. 2019. Preprint.

[7] U. Bauer, M. Kerber, J. Reininghaus, and H. Wagner. PHAT - persistent homology algorithms toolbox. In *ICMS*, 2014. https://github.com/blazs/phat.

[8] H. Bhatia, A. G. Gyulassy, V. Lordi, J. E. Pask, V. Pascucci, and P.-T. Bremer. Topoms: Comprehensive topological exploration for molecular and condensed-matter systems. *J. Comput. Chem.*, 39(16):936–952, 2018.

[9] A. Bock, H. Doraiswamy, A. Summers, and C. T. Silva. Topoangler: Interactive topology-based extraction of fishes. *IEEE Trans. Comp. Graph.*, 24(1):812–821, 2018.

[10] I. Borg and G. P. *Modern Multidimensional Scaling - Theory and Applications*. Springer Series in Statistics, 1997.

[11] P. Bremer, G. Weber, J. Tierny, V. Pascucci, M. Day, and J. Bell. Interactive exploration and analysis of large scale simulations using topology-based data segmentation. *IEEE Trans. Comp. Graph.*, 17(9):1307–1324, 2011.

[12] P. Bubenik and P. Dłotko. A persistence landscapes toolbox for topological statistics. *Symb. Comp.*, 78:91 – 114, 2017. https://www.math.upenn.edu/~dlotko/persistenceLandscape.html.

[13] G. Carlsson. Topology and Data. *Bulletin of the American Mathematical Society*, 46(2):255–308, 2009.

[14] H. A. Carr, J. Snoeyink, and M. van de Panne. Simplifying Flexible Isosurfaces Using Local Geometric Measures. In *IEEE VIS*, pages 497–504, 2004.

[15] F. Chazal and S. Oudot. Towards persistence-based reconstruction in euclidean spaces. In *Symp. on Comp. Geom.*, pages 232–241, 2008.

[16] D. Cohen-Steiner, H. Edelsbrunner, and J. Harer. Stability of persistence diagrams. *Disc. Comput. Geom.*, 37(1):103–120, 2007.

[17] K. Cole-McLaughlin, H. Edelsbrunner, J. Harer, V. Natarajan, and V. Pascucci. Loops in Reeb graphs of 2-manifolds. In *Symp. on Comp. Geom.*, 2003.

[18] Z. Cunninghamn, J. P. Ghahramani. Linear dimensionality reduction: Survey, insights, and generalizations. *J. Mach. Learn. Res.*, 16(89):2859–2900, 2015.

[19] R. R. Curtin, M. Edel, M. Lozhnikov, Y. Mentekidis, S. Ghaisas, and S. Zhang. mlpack 3: a fast, flexible machine learning library. *Journal of Open Source Software*, 3:726, 2018.

[20] H. Doraiswamy, J. Tierny, P. J. S. Silva, L. G. Nonato, and C. Silva. TopoMap: A 0-dimensional Homology Preserving Projection of High-Dimensional Data. *IEEE Transactions on Visualization and Computer Graphics (Proc. of IEEE VIS)*, 2020. https://arxiv.org/abs/2009.01512.

[21] D. Dua and C. Graff. UCI machine learning repository. https://archive.ics.uci.edu/ml/machine-learning-databases/, 2017.

[22] H. Edelsbrunner and J. Harer. *Computational Topology. An Introduction*. Amer. Math. Society, Jan. 2010.

[23] H. Edelsbrunner, D. Letscher, and A. Zomorodian. Topological Persistence and Simplification. *Disc. Compu. Geom.*, 28(4):511–533, 2002.

[24] B. T. Fasy, J. Kim, F. Lecci, and C. Maria. Introduction to the R package TDA. *CoRR*, abs/1411.1830, 2014. https://cran.r-project.org/web/packages/TDA/index.html.

[25] R. Forman. A User's Guide to Discrete Morse Theory. *Advances in Mathematics*, 1998.

[26] S. Gerber, P. Bremer, V. Pascucci, and R. Whitaker. Visual Exploration of High Dimensional Scalar Functions. *IEEE Trans. Comp. Graph.*, 16(6):1271–1280, 2010.

[27] S. Gerber, O. Rübel, P.-T. Bremer, V. Pascucci, and R. T. Whitaker. Morse–smale regression. *J. Comput. Graph. Stat.*, 22(1):193–214, 2013.

[28] E. Gomez-Nieto, W. Casaca, D. Motta, I. Hartmann, G. Taubin, and L. G. Nonato. Dealing with multiple requirements in geometric arrangements. *IEEE Trans. Comp. Graph.*, 22(3):1223–1235, 2016.

[29] D. Guenther, R. Alvarez-Boto, J. Contreras-Garcia, J.-P. Piquemal, and J. Tierny. Characterizing molecular interactions in chemical systems. *IEEE Transactions on Visualization and Computer Graphics (Proc. of IEEE VIS)*, 20(12):2476–2485, 2014.

[30] C. Gueunet, P. Fortin, J. Jomier, and J. Tierny. Task-based Augmented Reeb Graphs with Dynamic ST-Trees. In *Eurographics Symposium on Parallel Graphics and Visualization*, 2019.

[31] A. Gyulassy, P. Bremer, R. Grout, H. Kolla, J. Chen, and V. Pascucci. Stability of dissipation elements: A case study in combustion. *Comput. Graph. Forum*, 33(3):51–60, 2014.

[32] A. Gyulassy, P. T. Bremer, B. Hamann, and V. Pascucci. A practical approach to Morse-Smale complex computation: Scalability and generality. *IEEE Transactions on Visualization and Computer Graphics (Proc. of IEEE VIS)*, 2008.

[33] A. Gyulassy, M. A. Duchaineau, V. Natarajan, V. Pascucci, E. Bringa, A. Higginbotham, and B. Hamann. Topologically clean distance fields. *IEEE Transactions on Visualization and Computer Graphics (Proc. of IEEE VIS)*, 13(6):1432–1439, 2007.

[34] A. Gyulassy, A. Knoll, K. Lau, B. Wang, P. Bremer, M. Papka, L. A. Curtiss, and V. Pascucci. Interstitial and interlayer ion diffusion geometry extraction in graphitic nanosphere battery materials. *IEEE Transactions on Visualization and Computer Graphics (Proc. of IEEE VIS)*, 22(1):916–925, 2016.

[35] W. Harvey and Y. Wang. Topological Landscape Ensembles for Visualization of Scalar-Valued Functions. *Comput. Graph. Forum*, 29:993–1002, 2010.

[36] C. Heine, H. Leitte, M. Hlawitschka, F. Iuricich, L. De Floriani, G. Scheuermann, H. Hagen, and C. Garth. A survey of topology-based methods in visualization. *Comput. Graph. Forum*, 35(3):643–667, 2016.

[37] P. Joia, D. Coimbra, J. Cuminato, F. Paulovich, and L. Nonato. Local affine multidimensional projection. *IEEE Trans. Comp. Graph.*, 17:2563–2571, 2011.

[38] P. Joia, F. Petronetto, and L. G. Nonato. Uncovering representative groups in multidimensional projections. *Comput. Graph. Forum*, 34(3):281–290, 2015.

[39] J. Kasten, J. Reininghaus, I. Hotz, and H. Hege. Two-dimensional time-dependent vortex regions based on the acceleration magnitude. *IEEE Trans. Comp. Graph.*, 17(12):2080–2087, 2011.

[40] J. Krause, A. Dasgupta, J. Fekete, and E. Bertini. Seekaview: An intelligent dimensionality reduction strategy for navigating high-dimensional data spaces. In M. Hadwiger, R. Maciejewski, and K. Moreland, editors, *IEEE LDAV.*, pages 11–19, 2016.

[41] V. Kurlin. A one-dimensional homologically persistent skeleton of an unstructured point cloud in any metric space. *Comput. Graph. Forum*, 34(5):253–262, 2015.

[42] D. E. Laney, P. Bremer, A. Mascarenhas, P. Miller, and V. Pascucci. Understanding the structure of the turbulent mixing layer in hydrodynamic instabilities. *IEEE Transactions on Visualization and Computer Graphics (Proc. of IEEE VIS)*, 12(5):1053–1060, 2006.

[43] J. A. Lee and M. Verleysen. *Nonlinear Dimensionality Reduction*. Springer, 2007.

[44] W. B. March, P. Ram, and A. G. Gray. Fast euclidean minimum spanning tree: Algorithm, analysis, and applications. In *Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '10, pages 603–612. ACM, 2010.

[45] C. Maria, J. Boissonnat, M. Glisse, and M. Yvinec. The gudhi library: Simplicial complexes and persistent homology. In *ICMS*, 2014. http://gudhi.gforge.inria.fr/.

[46] R. Martins, D. Coimbra, R. Minghim, and A. Telea. Visual analysis of dimensionality reduction quality for parameterized projections. *Comp. & Graph.*, 41:26–42, 2014.

[47] L. McInnes, J. Healy, and J. Melville. UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. *ArXiv e-prints*, Feb. 2018.

[48] M. Moor, M. Horn, B. Rieck, and K. M. Borgwardt. Topological autoencoders. *CoRR*, abs/1906.00722, 2019.

[49] D. Morozov. Dionysus. http://www.mrzv.org/software/dionysus, 2010. Accessed: 2016-09-15.

[50] V. Nanda. Perseus, the persistent homology software. http://www.sas.upenn.edu/~vnanda/perseus, 2013. Accessed: 2016-09-15.

[51] L. G. Nonato and M. Aupetit. Multidimensional projection for visual analytics: Linking techniques with distortions, tasks, and layout enrichment. *IEEE Trans. Comp. Graph.*, 25(8):2650–2673, 2019.

[52] P. Oesterling, C. Heine, H. Jänicke, and G. Scheuermann. Visual analysis of high dimensional point clouds using topological landscapes. In *Proc. PacificVis*, pages 113–120, 2010.

[53] P. Oesterling, C. Heine, H. Jänicke, G. Scheuermann, and G. Heyer. Visualization of High Dimensional Point Clouds Using their Density Distribution's Topology. *IEEE Trans. Comp. Graph.*, 17(11):1547–1559, 2011.

[54] P. Oesterling, C. Heine, G. H. Weber, and G. Scheuermann. Visualizing nd point clouds as topological landscape profiles to guide local data analysis. *IEEE Trans. Vis. Comput. Graph.*, 19(3):514–526, 2013.

[55] P. Oesterling, G. Scheuermann, S. Teresniak, G. Heyer, S. Koch, T. Ertl, and G. H. Weber. Two-stage framework for a topology-based projection and visualization of classified document collections. In *Proc. IEEE VAST*, pages 91–98, 2010.

[56] M. Olejniczak, A. S. P. Gomes, and J. Tierny. A Topological Data Analysis Perspective on Non-Covalent Interactions in Relativistic Calculations. *Int. J. Quantum Chem.*, 120(8):e26133, 2020.

[57] S. Parsa. A deterministic O (m log m) time algorithm for the reeb graph. *Discret. Comput. Geom.*, 49(4):864–878, 2013.

[58] R. Paul and S. K. Chalup. A study on validating non-linear dimensionality reduction using persistent homology. *Pattern Recognition Letters*, 100:160–166, 2017.

[59] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. VanderPlas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in python. *JMLR*, 2011.

[60] J. M. M. R. Andreani, E. G. Birgin and M. L. Schuverdt. On augmented lagrangian methods with general lower-level constraints. *SIAM Opt.*, 18:1286–1309, 2007.

[61] J. M. M. R. Andreani, E. G. Birgin and M. L. Schuverdt. Augmented lagrangian methods under the constant positive linear dependence constraint qualification. *Math. Prog.*, 111:5–32, 2008.

[62] G. Reeb. Sur les points singuliers dune forme de Pfaff complètement intégrable ou d'une fonction numérique. *Comptes Rendus des séances de l'Académie des Sciences*, 222(847-849):76, 1946.

[63] B. Rieck and H. Leitte. Agreement analysis of quality measures for dimensionality reduction. In *Topological Methods in Data Analysis and Visualization*, pages 103–117. Springer, 2015.

[64] B. Rieck and H. Leitte. Persistent homology for the evaluation of dimensionality reduction schemes. *Comput. Graph. Forum*, 34(3):431–440, 2015.

[65] V. Robins, P. J. Wood, and A. P. Sheppard. Theory and Algorithms for Constructing Discrete Morse Complexes from Grayscale Digital Images. *IEEE Trans. Pattern Anal. Mach. Intell.*, 2011.

[66] D. Sacha, L. Zhang, M. Sedlmair, J. A. Lee, J. Peltonen, D. Weiskopf, S. C. North, and D. A. Keim. Visual interaction with dimensionality reduction: A structured literature analysis. *IEEE Trans. Comp. Graph.*, 23(1):241–250, 2017.

[67] Shieh, Albert D. and Hashimoto, Tatsunori B. and Airoldi, Edoardo M. Tree Preserving Embedding. In *ICML*, 2011.

[68] N. Shivashankar and V. Natarajan. Parallel Computation of 3D Morse-Smale Complexes. 2012.

[69] N. Shivashankar, P. Pranav, V. Natarajan, R. van de Weygaert, E. P. Bos, and S. Rieder. Felix: A topology based framework for visual exploration of cosmic filaments. *IEEE Trans. Comp. Graph.*, 22(6):1745–1759, 2016. http://vgl.serc.iisc.ernet.in/felix/index.html.

[70] V. D. Silva and J. B. Tenenbaum. Global versus local methods in nonlinear dimensionality reduction. In *NIPS*, pages 721–728, 2003.

[71] G. Singh, F. Memoli, and G. Carlsson. Topological Methods for the Analysis of High Dimensional Data Sets and 3D Object Recognition. In *Eurographics Symposium on Point-Based Graphics*, 2007.

[72] M. Soler, M. Petitfrere, G. Darche, M. Plainchault, B. Conche, and J. Tierny. Ranking Viscous Finger Simulations to an Acquired Ground Truth with Topology-Aware Matchings. In *IEEE LDAV.*, pages 62–72, 2019.

[73] T. Sousbie. The persistent cosmic web and its filamentary structure: Theory and implementations. *Royal Astronomical Society*, 414:350 – 383, 06 2011. http://www2.iap.fr/users/sousbie/web/html/indexd41d.html.

[74] G. Tauzin, U. Lupo, L. Tunstall, J. B. Pérez, M. Caorsi, A. Medina-Mardones, A. Dassatti, and K. Hess. giotto-tda: A topological data analysis toolkit for machine learning and data exploration, 2020.

[75] J. B. Tenenbaum, V. De Silva, and J. C. Langford. A global geometric framework for nonlinear dimensionality reduction. *science*, 290(5500):2319–2323, 2000.

[76] J. Tierny, G. Favelier, J. A. Levine, C. Gueunet, and M. Michaux. The Topology ToolKit. *IEEE Transactions on Visualization and Computer Graphics (Proc. of IEEE VIS)*, 24(1):832–842, 2018. https://topology-tool-kit.github.io/.

[77] L. van der Maaten and G. Hinton. Visualizing high-dimensional data using t-sne. *J. Mach. Learn. Res.*, 9:2579–2605, 2008.

[78] L. van der Maaten, E. Postma, and J. van den Herik. Dimensionality reduction: A comparative review. Technical report, Tilburg University, 2007.

[79] G. Weber, P.-T. Bremer, and V. Pascucci. Topological Landscapes: A Terrain Metaphor for Scientific Data. *IEEE Trans. Comp. Graph.*, 13(6):1416–1423, Nov. 2007.

[80] L. Yan, Y. Zhao, P. Rosen, C. Scheidegger, and B. Wang. Homology-preserving dimensionality reduction via manifold landmarking and tearing. *CoRR*, abs/1806.08460, 2018.