Open Problems Column
Edited by William Gasarch

1 This Issues Column!

This issue's Open Problem Column is by Brittany Terese Fasy and Bei Wang and is on *Open Problems in Computational Topology*; however, they also use acknowledged contributions from others. They also include references for someone who wants to start learning about the area.

2 Request for Columns!

I invite any reader who has knowledge of some area to contact me and arrange to write a column about open problems in that area. That area can be (1) broad or narrow or anywhere inbetween, and (2) really important or really unimportant or anywhere inbetween.

Open Problems in Computational Topology

By Brittany Terese Fasy and Bei Wang
with contributions by members of the WinCompTop community

3 Computational Topology in a Nutshell

Computational topology lays the theoretical and algorithmic foundations for topological data analysis, an emerging area in exploratory data analysis and data mining. In a nutshell, computational topology combines mathematical concepts from algebraic topology with computational techniques in computer science to address the following question: *What is the shape of data?*[1] In general, we assume that the data is a *point cloud*, that is a set of objects, or *data points*, with pairwise distances defined between them. Sometimes, this point cloud is geometric (embedded in $\mathbb{R}^n$), but not necessarily. Given a point cloud, we ask: can we describe the structure of this data? How many clusters are there? Do any (large) loops or holes exist? Can we reconstruct the underlying space or infer its geometric and topological structures? With dynamic point clouds, we can ask, how the structure changes with time.

For instance, the algebraic notion of *homology* captures the topological features of a space in terms of its connectivity, treating its connected components, tunnels, voids and k-dimensional notions of holes as zero-, one-, two-, and k-dimensional features. The point cloud extension in topological data analysis, *persistent homology* [16], gives practical algorithms that infer such information of the space from its point samples. The structure of the point cloud becomes apparent if we replace the points with balls, and allow the radii of the balls to increase from zero to infinity; we call this sequence a point cloud. We often think about the radii increasing linearly with time, and note that the topology of the union of balls will only change a finite number of times. In

[1] In fact, one of the earliest conversations about persistent homology began when Herbert Edelsbrunner asked David Letscher that exact question. While a ludicrous thing to ask at the time (of course, a point set has the topology of a set of points: uninteresting), this question is why so many researchers today–both applied and theoretical–are interested in computational topology. To emphasize this, Gunnar Carlsson often begins his public lectures with *Data has shape, and that shape matters.*

particular, we focus on important events when the balls merge with one another to form features such as components (i.e., clusters) or holes. We track the appearance (birth), the disappearance (death), and the lifetime (i.e., *persistence*) of topological features in the filtration via *persistence barcodes* [18], or equivalently, *persistence diagrams* [14], which serve as topological descriptors of the data.

In this paper, we present a sampling of the open problems in computational topology. For the interested reader, we suggest [6, 18, 21] for a gentle introduction to computational topology, [15] for algorithmic foundations, and [19] for a (free online) textbook with many examples.

4 Understanding Homotopy Types of Vietoris–Rips Complexes

The first collection of open problems was submitted by Henry Adams from Colorado State University.

Given a metric space X and a scale parameter $r > 0$, the *Vietoris–Rips simplicial complex* $\mathrm{VR}(X;r)$ has X as its vertex set, and as its simplices the finite subsets of X of diameter less than r (or alternatively, less than or equal to r). The *intrinsic Čech simplicial complex* $\check{\mathrm{C}}(X;r)$ has X as its vertex set, and as its simplices the finite subsets σ of X such that $\cap_{x\in\sigma}B(x,r) \neq \emptyset$, where $B(x,r)$ is the open (or alternatively, the closed) ball in X centered about x and of radius r.

Let M be a compact Riemannian manifold. If the scale parameter r is sufficiently small, then a theorem of Jean-Claude Hausmann states that $\mathrm{VR}(M;r)$ is homotopy equivalent to M [20, Theorem 3.5]. Similarly, if r is sufficiently small so that the balls $\{B(x,r)\}_{x\in M}$ form a good cover of M, then the Nerve Theorem [4] implies that $\check{\mathrm{C}}(M;r)$ is homotopy equivalent to M. Little is known about the theory of Vietoris–Rips or intrinsic Čech complexes when scale r is large enough so that Hausmann's result or the Nerve Theorem don't apply, even though these complexes arise naturally in applications of persistent homology. Below are some open questions along these lines.

1. Let $\mathbb{S}^n$ be the n-dimensional sphere equipped with either the Euclidean or the geodesic metric. What are the homotopy types of $\mathrm{VR}(\mathbb{S}^n;r)$ and $\check{\mathrm{C}}(\mathbb{S}^n;r)$ for all r? When $n = 1$, the simplicial complexes $\mathrm{VR}(\mathbb{S}^1;r)$ and $\check{\mathrm{C}}(\mathbb{S}^1;r)$ obtain the homotopy types of $\mathbb{S}^1, \mathbb{S}^3, \mathbb{S}^5, \mathbb{S}^7, \ldots$ as r increases [1]. However, for $n \geq 2$, essentially nothing is known about the homotopy types of $\mathrm{VR}(\mathbb{S}^n;r)$ or $\check{\mathrm{C}}(\mathbb{S}^n;r)$, except that they obtain the homotopy type of $\mathbb{S}^n$ when r is sufficiently small.

2. Let $T^n = (\mathbb{S}^1)^n$ be the n-dimensional torus equipped with the flat metric. What are the homotopy types of $\mathrm{VR}(T^n;r)$ and $\check{\mathrm{C}}(T^n;r)$ when r increases beyond sufficiently small scales? To our knowledge there are no well-accepted conjectures, though the answer is known when instead the L^∞ metric is used ([1, Proposition 10.2]).

3. Questions (1) and (2) above can be asked not only for $M = \mathbb{S}^n$ or $M = T^n$, but for any Riemannian manifold M.

4. If M is a compact connected Riemannian manifold, then is the homotopy connectivity of $\mathrm{VR}(M;r)$ (resp., $\check{\mathrm{C}}(M;r)$) a non-decreasing function of r? Here, we define the homotopy connectivity of $\mathrm{VR}(M;r)$ to be the largest n such that homotopy group $\pi_k(\mathrm{VR}(M;r))$ (resp., $\pi_k(\check{\mathrm{C}}(M;r))$) is trivial for all $k \leq n$. See [20, Problem 3.12].

5. If $X \subseteq \mathbb{R}^2$ is a subset of the plane equipped with the Euclidean metric, then is $\mathrm{VR}(X;r)$ homotopy equivalent to a wedge of spheres? It is known that the fundamental group of

VR$(X;r)$ is a free group [10], and [3, Theorem 6.3] provides evidence towards an affirmative answer to this question.

6. At WinCompTop 2016, a working group led by Yusu Wang characterized the one-dimensional persistent homology of the intrinsic Čech complexes of metric graphs [17]. What can be said about higher-dimensional persistent homology of Čech or Vietoris-Rips complexes of metric graphs? The case of the circle follows from [1, 2].

5 Homotopic Reconstruction

This open problem was submitted by Kate Turner from EPFL.

In computational topology, we often assume that we have a point cloud sampled from a probability distribution supported on a manifold (often Euclidean); see [5, 13] for example.

Let X be a manifold embedded in $\mathbb{R}^n$, for some $n \geq 2$. Then, we define the *local feature size* of X at $x \in X$, denoted $\textsc{lfs}(x)$, as the distance from x to the medial axis of X. We call the minimum local feature size over the whole manifold the *reach* of X, denoted $\textsc{reach}(M)$. If $\varepsilon < \textsc{reach}(M)$, then the tubular neighborhood of radius ε around M deformation retracts onto M. A similar concept is the *weak feature size*, which is defined to be the smallest value ε for which $\check{\mathrm{C}}(X;\varepsilon)$ is not homologous with X; we denote this by $\textsc{wfs}(X)$. Let $\tau = \frac{1}{4}\textsc{wfs}(X)$ and suppose S is a finite τ-sampling from X. Then, the homology of X is isomorphic to the image of the following map induced by inclusion:

$$H(\check{\mathrm{C}}(S;\tau)) \to H(\check{\mathrm{C}}(S;3\tau)).$$

An open question of interest here is *does there exist a subcompex of* $\check{\mathrm{C}}(S;3\tau)$ *that realizes this homology*? And, if so, *can such a subcomplex be found in polynomial time?*

6 Homotopy Height

This open problem was submitted by Erin Wolf Chambers from Saint Louis University.

Data analysis is often about measuring the distance (and the similarity) between data. Topological data analysis is no exception to this. Trajectory data is one type of data that has a rich potential for geometric and topological data analysis.

Given two curves, f and g, embedded in a (metricized) topological space, a 'nice' distance would be one that is realized via a *minimal* homotopy H such that $H(\cdot,0) = f$ and $H(\cdot,1) = g$. However, how do we define (or compute) the cost of a homotopy? One way would be to integrate the area *swept out* by the homotopy. The so-called minimal area homotopy is related to the *winding area* and can be computed efficiently when f and g are embedded in an orientable surface [7] (here, efficiently is with respect to the complexity of the curves and the number of intersections between them). Another option is to measure the homotopy height, which is the length of the longest intermediate curve [11]; a dual, in a way, to the Fréchet distance. While [8] proved various properties of minimum homotopies (including the fact that the minimal homotopy will not 'backtrack'), a longstanding problem is to find a polynomial-time algorithm to compute the homotopy height between two curves embedded on a triangulated manifold.

Progress. In February 2017, a working group at Dagstuhl Seminar 17072 (comprised of Erin Wolf Chambers, Arnaud de Mesmay, and Tim Ophelders) made progress on this open problem, showing that it is - in fact, in NP [9]. It still remains open whether or not this problem is NP-complete.

7 Computational Topology in Practice

This open problem was inspired by a problem submitted by Mimi Tsuruga from UC Davis.

One of the benefits of the tools of computational topology is that *shape* and *structure* can be found in all types of data. For this reason, the popularity of Computational Topology has risen in application domains over the past few years; see [12, 22], for example. However, what does it mean for your data to have *interesting five-dimensional homology*? The answer, thus far, is unclear. In fact, a recent panel at a workshop organized by one of the authors (Bei Wang) at the ACM-BCB 2016 annual conference discussed just this: how can we make computational topology a feasible toolset for those wanting to analyze data in various domains? In fact, this is the most pressing open problem in our community today: how do we narrow the gap between theory and practice? To properly solve this problem will take the partnership between domain experts and researchers from computational topology. In sum, we ask: How do we make sense of the topological features that we can extract from data? In other words, *now what?*

Acknowledgements This collection of open problems was generated from an open problems session at the Women in Computational Topology (WinCompTop) 2016 workshop, as well as from the WinCompTop listserv. WinCompTop 2016 was funded through NSF grant DMS 1619908. Special thanks to all who contributed problems, including: Henry Adams, Erin Wolf Chambers, Maia Grudzien, Emilie Purvine, Mimi Tsuruga, Altaa Tumurbaatar, and Katherine Felicity Turner. While not all problems proposed are included in this article, they did help in understanding the scope of problems that our colleagues are interested in.

The authors also gratefully acknowledge the support of grants: NSF CCF 1618605 and NSF IIS 1513616. We also thank Dagstuhl seminar 17072, for providing an opportunity to work on one of these problems.

References

[1] M. Adamaszek and H. Adams. The Vietoris–Rips complexes of a circle. *Pacific Journal of Mathematics*, 290:1–40, 2017.

[2] M. Adamaszek, H. Adams, F. Frick, C. Peterson, and C. Previte-Johnson. Nerve complexes of circular arcs. *Discrete & Computational Geometry*, 56(2):251–273, 2016.

[3] M. Adamaszek, F. Frick, and A. Vakili. On homotopy types of Euclidean Rips complexes. Accepted by Discrete & Computational Geometry, `arxiv/1602.04131`, 2016.

[4] K. Borsuk. On the imbedding of systems of compacta in simplicial complexes. *Fundamenta Mathematicae*, 35(1):217–234, 1948.

[5] P. Bubenik, P. T. Kim, et al. A statistical approach to persistent homology. *Homology, Homotopy and Applications*, 9(2):337–362, 2007.

[6] G. Carlsson. Topology and data. *Bull. AMS*, 46(2):255–308, 2009.

[7] E. Chambers and Y. Wang. Measuring similarity between curves on 2-manifolds via homotopy area. In *Proc. 29th Annual ACM Symposium on Computational Geometry*, pages 425–434, 2013.

[8] E. W. Chambers, G. R. Chambers, A. de Mesmay, T. Ophelders, and R. Rotman. Monotone contractions of the boundary of the disc. arXiv preprint arXiv:1704.06175, 2017.

[9] E. W. Chambers, A. de Mesmay, and T. Ophelders. Dagstuhl working group result, in prepration.

[10] E. W. Chambers, V. De Silva, J. Erickson, and R. Ghrist. Vietoris–Rips complexes of planar point sets. *Discrete & Computational Geometry*, 44:75–90, 2010.

[11] E. W. Chambers and D. Letscher. On the height of a homotopy. In *CCCG*, volume 9, pages 103–106, 2009.

[12] J. M. Chan, G. Carlsson, and R. Rabadan. Topology of viral evolution. *Proceedings of the National Academy of Sciences*, 110(46):18566–18571, 2013.

[13] T. K. Dey, J. Sun, and Y. Wang. Approximating loops in a shortest homology basis from point data. In *Proceedings of the twenty-sixth annual symposium on Computational geometry*, pages 166–175. ACM, 2010.

[14] H. Edelsbrunner and J. Harer. Persistent homology - a survey. *Contemporary Mathematics*, 453:257, 2008.

[15] H. Edelsbrunner and J. Harer. *Computational Topology: An Introduction.* AMS, 2010.

[16] H. Edelsbrunner, D. Letscher, and A. J. Zomorodian. Topological persistence and simplification. *Discrete and Computational Geometry*, 28:511–533, 2002.

[17] E. Gasparovic, M. Gommel, E. Purvine, R. Sazdanovic, B. Wang, Y. Wang, and L. Ziegelmeier. Complete characterization of the 1-dimensional intrinsic Čech persistence diagrams for metric graphs. Preprint, `arxiv/1702.07379`, 2016.

[18] R. Ghrist. Barcodes: The persistent topology of data. *Bull. AMS*, 45:61–75, 2008.

[19] R. W. Ghrist. *Elementary Applied Topology.* Createspace, 2014.

[20] J.-C. Hausmann. On the Vietoris–Rips complexes and a cohomology theory for metric spaces. *Annals of Mathematics Studies*, 138:175–188, 1995.

[21] E. Munch. A user's guide to topological data analysis. *Journal of Learning Analytics*, 4(2):47–61, 2017.

[22] X. Zhu. Persistent homology: An introduction and a new text representation for natural language processing. In *IJCAI*, pages 1953–1959, 2013.