Persistent Homology and Nested Dissection

Michael Kerber* Donald R. Sheehy† Primoz Skraba‡

1 Abstract

Nested dissection exploits underlying topology to do matrix reductions while persistent homology exploits matrix reductions to reveal underlying topology. It seems natural that one should be able to combine these techniques to beat the currently best bound of matrix multiplication time for computing persistent homology. However, nested dissection works by fixing a reduction order, whereas persistent homology generally constrains the ordering. Despite this obstruction, we show that it is possible to combine these two theories. This shows that one can improve the computation of persistent homology of a filtration by exploiting information about the underlying space. It gives reasonable geometric conditions under which one can beat the matrix multiplication bound for persistent homology.

2 Overview

Given geometric data, persistent homology gives a way to extract multiscale shape information with the goal of understanding the underlying shape of the distribution from which the data was drawn. The data induces a function on the underlying space, usually a distance-like function to the data. The persistent homology measures the changes in topology of the sublevel sets of the function. To do this, one firsts constructs a filtered simplicial complex, that is, a simplicial complex with an ordering on the simplices.

The persistence algorithm is a restricted form of Gaussian elimination on the boundary matrix of the simplicial complex. As with standard Gaussian elimination, persistent homology can be computed in matrix multiplication time. We let ω denote the smallest exponent such that matrix multiplication can be computed in $O(n^\omega)$ time. It is likely that this is also the best possible running time for computing persistent homology of general filtered simplicial complexes. However, if the simplicial complex is coming from low-dimensional geometric data, one might hope to exploit that structure to improve the running time. We show that this is indeed possible.

Our approach combines several different ideas. We use the output-sensitive algorithm of Chen and Kerber to reduce the persistence computation to rank computations [1]. These ranks can be computed by a divide and conquer method called *nested dissection*, which dates back to the 1970s [3] but has recently been extended to the relevant case of singular matrices over finite fields by Yuster [6]. However, nested dissection requires that the matrices have a certain separability property which is related to cuts in a graph associated with the matrix. We limit our attention to filtrations built on quality meshes, which have been shown to give good approximations to persistent homology for the Euclidean distance to a point set by Hudson et al. [2] and also for more general distance-like functions by Sheehy [5]. As shown by Miller et al. [4], graphs coming from quality meshes satisfy the desired separability property required for nested dissection. We show that this separability can be extended to the case of boundary matrices of these meshes. Putting all of these pieces together, we get an improvement over matrix multiplication

*Max-Planck-Center for Visual Computing and Communication, Saarbrüken, Germany, mkerber@mpi-inf.mpg.de

†UConn, don.r.sheehy@gmail.com

‡Jožef Stefan Institute, Ljubljana, Slovenia, primoz.skraba@ijs.si

time for computing the persistent homology of filtrations on quality meshes as long as the output size is sufficiently bounded.

3 The Main Result

We need a couple definitions in order to state the main result.

The output of the persistent homology algorithm is a set of pairs of real numbers indicating the birth and death of a topological feature. The difference between the birth and death time is the persistence of the pair. These pairs are often represented as points in the plane or as intervals. The former representation is called a *persistence diagram* and the latter is called a *persistence barcode*. For a constant $\Gamma \geq 0$, let D_Γ be the subset of the pairs with persistence at least Γ.

A point set P is *τ-well-spaced* for a constant τ if in the Voronoi diagram of P, the Voronoi cell V_p of each $p \in P$ satisfies

$$\frac{\max\limits_{v \in V(V_p)} \|p - v\|}{\min\limits_{q \in P \setminus \{p\}} \frac{1}{2}\|p - q\|} \leq \tau,$$

where $V(V_p)$ is the set of vertices (0-cells) of the polyhedron V_p. The ratio above is called the aspect ratio of V_p. The *quality* meshes used for mesh-based persistence [2, 5] are Delaunay triangulations of well-spaced point sets.

We are now ready to state the main result.

Theorem 3.1. *Let F be a filtration on the Delaunay triangulation of a set of τ-well-spaced points. For a constant $\Gamma \geq 0$, the subset D_Γ of the persistence diagram of F consisting of those pairs with persistence at least Γ can be computed in $O(|D_{(1-\delta)\Gamma}| n^{\omega(1-\frac{1}{d})})$ time, where $D_{(1-\delta)\Gamma}$ is the set of pairs in the persistence diagram with persistence at least $(1-\delta)\Gamma$ for any constant $\delta > 0$.*

4 Going forward

Beyond putting together the many different pervious results to obtain this theorem, we also needed to extend the existing geometric separator theory for τ-quality point sets from graphs to complexes. Separators provide the dissection in nested dissection. It may be possible to extend this further to find separators of other intermediate matrices that appear in the direct computation of persistent homology using matrix multiplication. This might allow us to remove the output-sensitive term in the running time of Theorem 3.1.

We also believe that the language of algebraic topology may be useful in understanding and generalizing other algorithmic results related to nested dissection. Specifically, we are interested in exploring the relationship between nested dissection and discrete Morse theory.

The separators we use depend on the geometric separator theory. Thus, we exploit the geometry to improve the running time of the persistence computation. Although previous work showed how to use the geometry of the input to decide *what* to compute, this work shows how to use the geometry to decide *how* to compute it. This breaks the usual separation between a geometric phase and a topological phase in the persistent homology pipeline.

References

[1] C. Chen and M. Kerber. An output-sensitive algorithm for persistent homology. *Comput. Geom*, 46(4):435–447, 2013.

[2] B. Hudson, G. L. Miller, S. Y. Oudot, and D. R. Sheehy. Topological inference via meshing. In *Proceedings of the 26th ACM Symposium on Computational Geometry*, pages 277–286, 2010.

[3] R. J. Lipton, D. J. Rose, and R. E. Tarjan. Generalized nested dissection. *SIAM Journal on Numerical Analysis*, 16(2):346–358, 1979.

[4] G. L. Miller, S.-H. Teng, W. Thurston, and S. A. Vavasis. Geometric separators for finite-element meshes. *SIAM J. Sci. Comput.*, 19(2):364–386, 1998.

[5] D. R. Sheehy. *Mesh Generation and Geometric Persistent Homology*. PhD thesis, Carnegie Mellon University, 2011.

[6] R. Yuster. Matrix sparsification for rank and determinant computations via nested dissection. In *FOCS: IEEE Symposium on Foundations of Computer Science*, pages 137–145, 2008.