Mesh-Enhanced Persistent Homology*

Benoît Hudson
bhudson@tti-c.edu

Gary L. Miller
glmiller@cs.cmu.edu

Steve Y. Oudot
steve.oudot@inria.fr

Donald R. Sheehy
dsheehy@cs.cmu.edu

Abstract

Existing algorithms for computing the persistent homology of a set of n points in $\mathbb{R}^d$ can take up to $n^{\Omega(d)}$ time and space in the worst case. We apply ideas from mesh generation to reduce the time to $2^{O(d)}n + O(n \log \Delta)$ and the space to $2^{O(d)}n^2$, where Δ denotes the spread of the point cloud. This makes the problem of computing the full persistence barcode tractable for data in (moderate) dimensions higher than 3. Our technique is based on a new filtration, the α-mesh filtration, that can be intertwined with the offset filtration and avoids the $n^{O(d)}$ blowup of the α-complex.

1 Introduction

Persistent homology is a powerful tool for understanding the intrinsic structure of a set of data points at different scales. Unfortunately, existing methods for computing the persistent homology of points in $\mathbb{R}^d$ can take up to $n^{\Omega(d)}$ time and space, thus limiting our ability to use it in even moderate dimensions. Here we show how to reduce the dependence of both time and space on the ambient dimension to $2^{O(d)}$. We achieve this by preprocessing the points using Delaunay mesh refinement, then building a filtration on the mesh such that the persistent homology of the mesh closely approximates the persistent homology of the offsets of the input points. This makes it practical to compute persistent homology at all scales in higher dimensions than was previously possible.

The homology groups give an algebraic description of the shape of a topological space. For sets in $\mathbb{R}^3$, the 0th, 1st, and 2nd homology groups describe the connected components, tunnels, and voids respectively. Among topological invariants, homology stands out for its ease of computation: indeed, computing the homology of a simplicial complex reduces to a sequence of simple operations in linear algebra [8].

Persistent homology is useful when the input is only a point sample rather than a full complex. A sequence of complexes are constructed on the points at different scales. The persistent homology captures those topological features that remain over a significant range of scales (see [13, 6] for surveys).
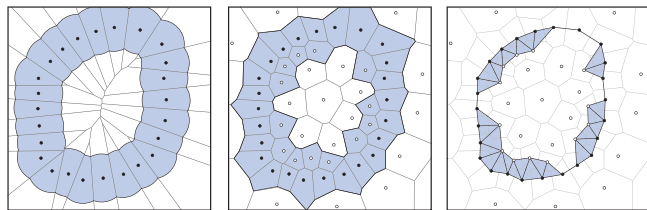

Figure 1: From left to right: The offset filtration, P^α; the α-Voronoi filtration, V^α; and the α-mesh, D^α.

A *filtration* is a nested sequence of topological spaces. The *offset filtration* of a point set P is the filtration whose elements, P^α, are the unions of α-radius balls centered at the points of P (See Figure 1, left). The α-complex filtration is a filtration of the Delaunay triangulation of P that is known to have the same persistent homology as the offset filtration [5]. Unfortunately, the size of the Delaunay triangulation can be up to $\Omega(n^{\lceil d/2 \rceil})$, even when the points of P are distributed along some submanifold of $\mathbb{R}^d$ [1].

Sparse meshing technology can bypass the worst-case blow-up in the size of the Delaunay triangulation while maintaining many of the same properties of the original Delaunay triangulation [9]. It does this by adding extra points to the input in such a way that the Delaunay triangulation complexity is less for the superset than for the input. Under very mild sampling assumptions, the size of the superset is at most $2^{O(d)}n$ [10].

2 The α-mesh filtration

Constructing the α-mesh filtration starts with a quality mesh of the input points. Let $P \subset \mathbb{R}^d$ be the input points. Let $M \supset P$ be a superset of the input with the property that for each v in M, the ratio of the radii of the smallest enclosing ball and largest enclosed ball, both centered at v, of the Voronoi cell of v is bounded by a constant τ. The set M and its Voronoi diagram constitute a τ-quality mesh. The sizes of P and M are

denoted by n and m respectively. The points of $M \setminus P$ are known as *Steiner points*.

The mesh M can be filtered to closely approximate the offset filtration P^α of P. The result is called the α-mesh filtration and is denoted D^α. For each Steiner point s, we assign a value $t(s)$ to be the distance from s to P. For each input point p, we let $t(p)$ be half the distance from p to $M \setminus \{p\}$. The filtration D^α is the set of input vertices as well as all simplices σ of $Del(M)$ such that each $v \in \sigma$ has $t(v) \leq \alpha$ (See Figure 1, right).

Why it works. The α-mesh has a natural dual formed by the union of Voronoi cells of the vertices $v \in M$ with $t(v) \leq \alpha$ (See Fig. 1, middle). For technical reasons, we add to this union the open balls of radius $\min\{\alpha, t(p)\}$ centered at the input points $p \in P$, which does not affect the combinatorial structure of the dual, and we call α-*Voronoi filtration* V^α the obtained filtration. Ignoring some technicalities involving the boundary of the mesh, V^α and P^α are *multiplicatively τ-interleaved*:

$$\forall \alpha \geq 0, \ P^{\alpha/\tau} \subseteq V^\alpha \subseteq P^{\alpha\tau}. \tag{1}$$

Consider now $P^\alpha_{\log}$ and $V^\alpha_{\log}$, the reparametrizations of the filtrations P^α and V^α over a logarithmic scale:

$$\forall \alpha \geq 0, \ P^\alpha_{\log} = P^{2^\alpha} \text{ and } V^\alpha_{\log} = V^{2^\alpha}.$$

Multiplicative τ-interleaving of P^α and V^α in the sense of Eq. 1 implies *additive* $\log(\tau)$-*interleaving* of their reparametrizations $P^\alpha_{\log}$ and $V^\alpha_{\log}$:

$$\forall \alpha \geq 0, \ P^{\alpha - \log \tau}_{\log} \subseteq V^\alpha_{\log} \subseteq P^{\alpha + \log \tau}_{\log}.$$

Additively $\log(\tau)$-interleaved filtrations are known to have $\log(\tau)$-close persistence diagrams in the bottleneck distance [2, 4]. In addition, the Persistent Nerve Lemma of Chazal and Oudot [3, Lem. 3.4] implies that $V^\alpha_{\log}$ and $D^\alpha_{\log}$ (the log-reparametrization of D^α) have identical persistence diagrams. Therefore, the persistence diagrams of $D^\alpha_{\log}$ and $P^\alpha_{\log}$ are $\log(\tau)$-close in the bottleneck distance.

3 Time and Space complexities

Computing the persistence barcode involves a variant of Gaussian elimination of a matrix with dimensions proportional to the number of mesh simplices. Since M is well-spaced, the number of simplices is $2^{O(d)}m$ [12]. Assuming the input point set P is an (ϵ, δ)-sample as is often assumed (or, more generally, a *well-paced point set* [11]), then $m \in 2^{O(d)}n$ [10]. Thus, the persistence computation involves a matrix in $2^{O(d)}n$ dimensions. In theory, this could take as much as $2^{O(d)}n^2$ space and $2^{O(d)}n^3$ time using the persistence algorithm [7, 14]. In practice, near-linear running time and memory usage have been observed.

We know how to compute a small but well-spaced superset and its Delaunay complex in $2^{O(d)}n$ space and $2^{O(d)}n \log(\Delta)$ time [9], where Δ is the spread of the input—the ratio of the largest to smallest interpoint distances. The spread is usually small: on a regular (ϵ, δ)-sample, it is at most $O(n)$ so that meshing takes $2^{O(d)}n \log(n)$ time. These space and time bounds imply our technique is practical in moderate dimensions, certainly up to six.

4 Extensions

The methods presented here have several natural extensions. A tighter approximation to the offset filtration is possible by refining the mesh. The α-mesh and the offset filtration can achieve a multiplicative $(1+\varepsilon)$-interleaving at a cost of $O(\varepsilon^{-d})$ blowup in the complex size. Also, it is possible to compute the persistent homology of the complement of the offset filtration by filtering the Voronoi diagram of V *backwards*. A third extension is the complete elimination of any assumptions on the sampling by applying linear-size Delaunay meshing technology [11].

References

[1] N. Amenta, , D. Attali, and O. Devillers. Complexity of delaunay triangulation for points on lower-dimensional polyhedra. In *SODA*, 2007.

[2] F. Chazal, D. Cohen-Steiner, M. Glisse, L. J. Guibas, and S. Y. Oudot. Proximity of persistence modules and their diagrams. In *SOCG*, 2009.

[3] F. Chazal and S. Y. Oudot. Towards persistence-based reconstruction in Euclidean spaces. In *SOCG*, 2008.

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

[5] H. Edelsbrunner. The union of balls and its dual shape. In *SOCG*, 1993.

[6] H. Edelsbrunner and J. Harer. Persistent homology – a survey. In J. E. Goodman, J. Pach, and R. Pollack, editors, *Twenty Years After*. AMS Press, 2007.

[7] H. Edelsbrunner, D. Letscher, and A. Zomorodian. Topological persistence and simplification. *Discrete Comput. Geom.*, 28:511–533, 2002.

[8] J. Friedman. Computing Betti numbers via combinatorial Laplacians. In *STOC*, 1996.

[9] B. Hudson, G. Miller, and T. Phillips. Sparse Voronoi Refinement. In *Proceeding of the International Meshing Roundtable*, 2006. Long version available as Carnegie Mellon University Technical Report CMU-CS-06-132.

[10] B. Hudson, G. L. Miller, T. Phillips, and D. R. Sheehy. Size complexity of volume meshes vs. surface meshes. In *SODA*, 2009.

[11] G. L. Miller, T. Phillips, and D. R. Sheehy. Linear-size meshes. In *CCCG*, 2008.

[12] G. L. Miller, D. Talmor, S.-H. Teng, and N. Walkington. A Delaunay based numerical method for three dimensions: generation, formulation, and partition. In *STOC*, 1995.

[13] A. Zomorodian. *Topology for Computing*. Cambridge Univ. Press, 2005.

[14] A. Zomorodian and G. Carlsson. Computing persistent homology. *Discrete Comp. Geom.*, 33(2):249–274, 2005.