



ELSEVIER

Contents lists available at ScienceDirect

# Computational Geometry: Theory and Applications

[www.elsevier.com/locate/comgeo](http://www.elsevier.com/locate/comgeo)


## Graph induced complex on point data



Tamal K. Dey\*, Fengtao Fan, Yusu Wang

Department of Computer Science and Engineering, The Ohio State University, Columbus, OH 43210, USA

### ARTICLE INFO

#### Article history:

Available online 14 April 2015

#### Keywords:

Point cloud data  
Homology  
Simplicial complex  
Sparsification  
Topological persistence

### ABSTRACT

The efficiency of extracting topological information from point data depends largely on the complex that is built on top of the data points. From a computational viewpoint, the favored complexes for this purpose have so far been Vietoris–Rips and witness complexes. While the Vietoris–Rips complex is simple to compute and is a good vehicle for extracting topology of sampled spaces, its size becomes prohibitively large for reasonable computations. The witness complex on the other hand enjoys a smaller size because of a subsampling, but fails to capture the topology in high dimensions unless imposed with extra structure. We investigate a complex called the *graph induced complex* that, to some extent, enjoys the advantages of both. It works on a subsample but still retains the power of capturing the topology as the Vietoris–Rips complex. It only needs a graph connecting the original sample points from which it builds a complex on the subsample thus taming the size considerably. We show that, using the graph induced complex one can (i) infer the one dimensional homology of a manifold from a lean subsample, (ii) reconstruct a surface in three dimensions from a sparse subsample without computing Delaunay triangulations, (iii) infer the persistent homology groups of compact sets from a sufficiently dense sample. We provide experimental evidences in support of our theory.

Published by Elsevier B.V.

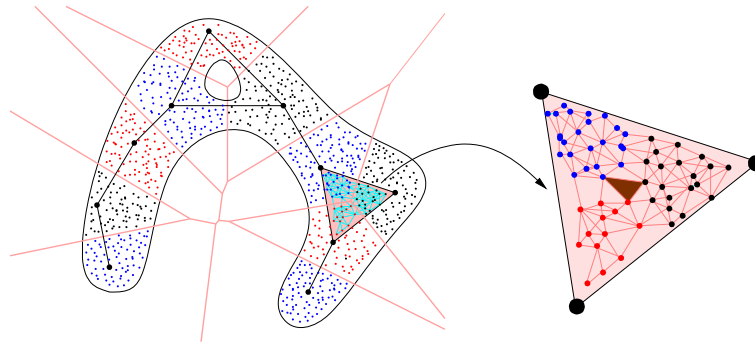
## 1. Introduction

Acquiring knowledge about a sampled space from a set of points has become a key problem in many areas of science and engineering. The sampled space could be a hidden manifold sitting in some high dimensions, or could be a compact subset of some Euclidean space. Topological information such as the rank of the homology groups, or their persistent behavior can divulge important features of the hidden space. Therefore, a considerable effort has ensued to extract topological information from point data in recent years [7,9,16,19]. With the advent of advanced technologies, the data is often generated in abundance. Mixed with the burden of high dimensionality, large data sets pose a challenge to the resource required to process them. As a result, some recent investigations have focused on how to use a lighter data structure or sparsify the input, which aids a faster computation, but still guarantees that the output inference is correct.

Point data by themselves do not have interesting topology. So, a foremost step in inferring topology from data is to impose a structure such as a simplicial complex onto it. The Delaunay, Čech, Vietoris–Rips, and witness complexes are some of the most commonly proposed complexes for this purpose. Among these, Vietoris–Rips (Rips in short) and witness complexes [9] have been favored because they can be constructed with simple computations. Rips complexes are easy to construct as they can be built from a graph by recognizing the cliques in it. However, the presence of simplices correspond-

\* Corresponding author.

E-mail addresses: [tamaldey@cse.ohio-state.edu](mailto:tamaldey@cse.ohio-state.edu) (T.K. Dey), [fanf@cse.ohio-state.edu](mailto:fanf@cse.ohio-state.edu) (F. Fan), [yusu@cse.ohio-state.edu](mailto:yusu@cse.ohio-state.edu) (Y. Wang).



**Fig. 1.** A graph induced complex shown with bold vertices, edges, and a shaded triangle on the left. Input graph within the shaded triangle is shown on the right.

ing to all cliques makes its size quite large. Even in three dimensions with a few thousand points, the size of the Rips complex can be an obstacle, if not a stopper, for further processing. Witness complexes, on the other hand, have too few simplices to capture the topology of the sampled space in dimensions three or more [4]. To tackle this issue, Boissonnat et al. [4] suggested modifications to the original definition of witness complex [12]. This enlarges the witness complexes but makes it more complicated and costly to compute.

We investigate a new complex, a version of which was originally introduced in [18] for the application of sensor network routing. We set up a more general definition and call it the *graph induced complex*. We provide new theoretical understanding of the graph induced complex in terms of topology inference. In particular, we show that, when equipped with an appropriate metric, this complex can help deciphering the topology from data. It retains the simplicity of the Rips complex as well as the sparsity of the witness complex. Its construction resembles the sparsified Rips complex proposed in [24] and also the combinatorial Delaunay triangulation proposed in [6], but it does not build a Rips complex on the subsample and thus is sparser than the Rips complex with the same set of vertices. This fact makes a real difference in practice as our preliminary experiments show. The idea of graph induced complex also bears similarity to the geodesic Delaunay triangulation which was proposed to recover the topology of a bounded planar region (with holes) from point samples [23]. Our work investigates its theoretical properties and generalizes it to settings beyond the planar case.

Given a graph  $G$  on a set of points  $P$  equipped with a metric, one can build a graph induced complex on a subsample  $Q \subseteq P$  as follows. A simplex is in the complex if and only if its vertex set  $V \subseteq Q$  has the property that a set of points in  $P$ , each being closest to exactly one vertex in  $V$ , forms a clique in  $G$ . Fig. 1 shows a graph induced complex for a set of data points in the plane. Subsampled points are the darker vertices. Input points are grouped according to the proximity to the subsampled vertices (indicated with a Voronoi partition). The shaded triangle enlarged on the right-hand side is in the graph induced complex since there is a 3-clique in the input graph whose 3 vertices have 3 different closest point in the subsample. Observe that, in this example, the graph induced complex has the same homology as the sampled space.

Fig. 2 shows experimental results on two data sets, 40,000 sample points from a Klein bottle in  $\mathbb{R}^4$  and 15,000 sample points from the primary circle of natural image data considered in  $\mathbb{R}^{25}$  [1]. The graphs connecting any two points within  $\alpha = 0.05$  unit distance for Klein bottle and  $\alpha = 0.6$  unit distance for the primary circle were taken as input for the graph induced complexes. The 2-skeleton of the Rips complexes for these  $\alpha$  parameters have 608,200 and 1,329,672,867 simplices respectively. These sizes are too large to carry out fast computations.

For comparison, we constructed the graph induced complex, sparsified Rips complex, and the witness complex on the same subsample determined by a parameter  $\delta$ . The parameter  $\delta$  is also used in the graph induced complex (see definitions later) and the witness complex. The edges in the Rips complex built on the same subsample were of lengths at most  $\alpha + 2\delta$  (justified by Proposition 2.8). We varied  $\delta$  and observed the rank of the one dimensional homology group ( $\beta_1$ ). The plots show that the graph induced complex captured  $\beta_1$  correctly for a significantly wider range of  $\delta$  (left plots) while its size remained comparable to that of the witness complex (right plots). In some cases, the graph induced complex could capture the correct  $\beta_1$  with remarkably small number of simplices. For example, it had  $\beta_1 = 2$  for Klein bottle when there were 278 simplices for  $\delta = 0.7$  and 154 simplices for  $\delta = 1.0$ . In both cases  $\beta_1$  for Rips and witness complexes did not match with that of the sampled spaces while the Rips complex had a much larger size ( $\log_e$  scale plot) and the witness complex had comparable size. This illustrates why the graph induced complex can be a better choice than the Rips and witness complexes.

We establish three different results. First, we show that the one-dimensional homology group of surfaces in three dimensions can be determined by graph induced complexes. Even the surface itself can be reconstructed with some post-processing from a sparse subsample of a sample that could be excessively dense. Second, we show that, for higher-dimensional manifolds, one-dimensional homology can still be determined from graph induced complexes with a simple modification of the metric. Finally, we extend our results to other homology groups where we show that the persistent homology groups of a pair of graph induced complexes can determine the homology groups of compact spaces. Experimental results support our theory.

Download English Version:

<https://daneshyari.com/en/article/414213>

Download Persian Version:

<https://daneshyari.com/article/414213>

[Daneshyari.com](https://daneshyari.com)