

Reconstructing knots from point clouds with persistent homology

Maxim Beketov,^{1,2,*} Alexey Gorelov,^{1,3,†} German Magai^{1,2,‡}

¹All authors contributed equally to this work.

²Faculty of Computer Science, HSE University, Moscow, Russia

³Steklov Mathematical Institute of Russian Academy of Sciences

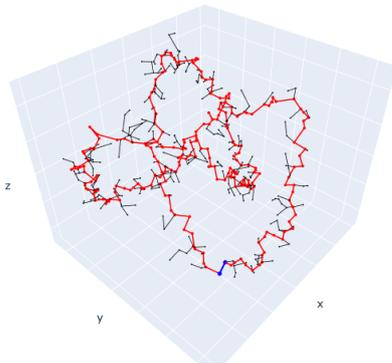


FIG. 1. Noisy trefoil knot reconstructed. The edge shown in blue gives birth to the longest-living one-dimensional persistent homology (red+blue).

The problem of reconstructing a knot from a point cloud followed by recognition of its topological type arises in physics, biology and chemistry. In the present work we propose several new methods to reconstruct a knot from a noisy point cloud with further simplification of the resulting closed polygonal curve. We examine the robustness of the proposed methods with respect to normal noise. We use the data from the KnotProt 2.0 database [1] of proteins that are knotted in certain ways (forming knots of types $0_1, 3_1, 4_1, 5_1$, etc.).

Present solution of this problem comes in 3 steps:

1. Reconstructing a polygonal closed curve from the given point cloud.
2. Simplifying the obtained curve, that is minimising the number of segments forming it.
3. Computing the topological invariants of the curve that distinguish its topological type.

The first step starts with constructing the Euclidean minimum spanning tree (EMST) of the point cloud. Next, there are at least two options for adding an edge to form a cycle that represents the knot (Fig. 1). One is to connect the two points that are most distant with respect to the tree (path) metric. The other option is to connect the two points forming an edge that corresponds to the birth of the most long-living one-dimensional persistent homology (PH). PH is a central concept in topological data analysis (TDA), and applying it in knot reconstruction problem makes one of the contributions of the present paper.

For the second step, we use the KTM algorithm [2] that reduces the number of segments in the polygonal closed curve making the knot by removing certain vertices, in a way that preserves its topological type. There seems to be room for improvement of this method, since it only operates on the original points, while perturbing them in some way may improve the final result.

To distinguish the topological types of knots presented in the KnotProt 2.0 database on the third step, we use calculate knot invariants (e.g. Alexander and Jones polynomials) by means of Topoly software [3].

Another contribution of the present paper is the study this method's performance in presence of noise that models uncertainty in point positions points are considered to be uniformly distributed along the polygonal curve of the knot plus some Gaussian noise in the planes orthogonal to the segments the points are chosen from. By controlling the variance scale of this normal component and the number of points the knot is reconstructed from, we examine the robustness of such invariant-based knot classification.

We thank Anton Ayzenberg and Konstantin Sorokin for fruitful discussions.

- [1] Dabrowski-Tumanski, Pawel, et al. "KnotProt 2.0: a database of proteins with knots and other entangled structures." *Nucleic acids research* 47.D1 (2019): D367-D375.
- [2] Koniaris, Kleantes, and Murugappan Muthukumar. "Selfentanglement in ring polymers." *The Journal of chemical physics* 95.4 (1991): 2873-2881.
- [3] Dabrowski-Tumanski, Pawel, et al. "Topoly: Python package to analyze topology of polymers." *Briefings in Bioinformatics* 22.3 (2021): bbaa196.

* email: mbeketov@hse.ru

† email: algor512@gmail.com

‡ email: gmagaj@hse.ru