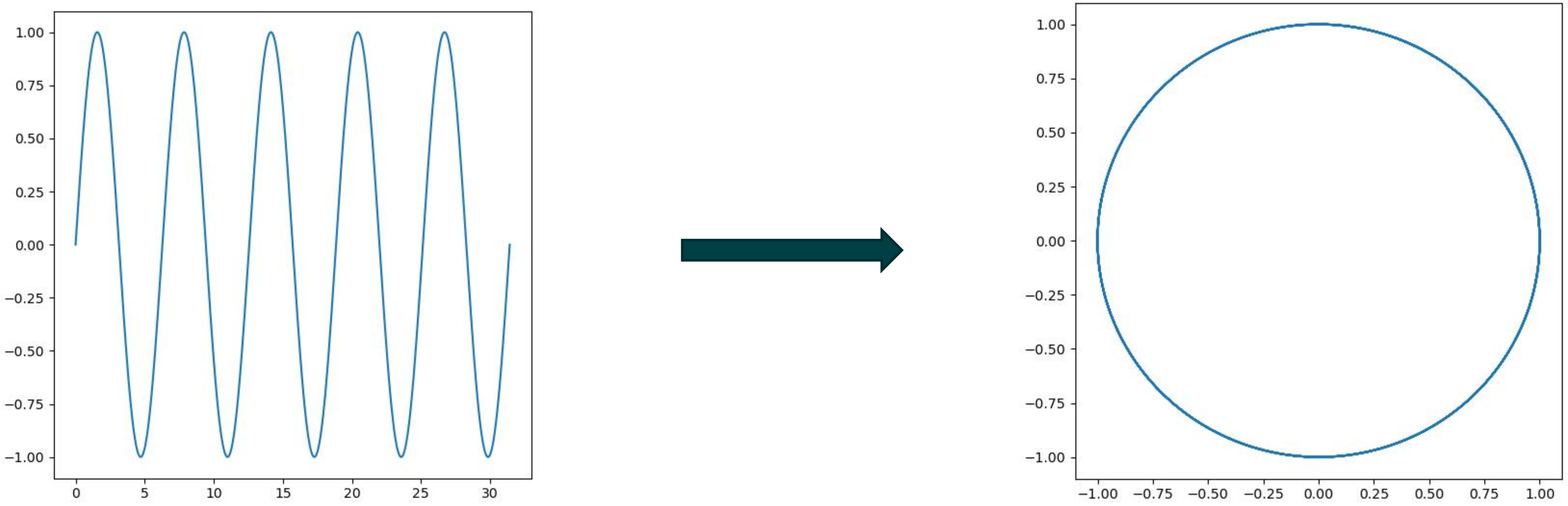
Detecting Periodic Phenomena via Topological Data Analysis

August 25, 2022

Ideas of TDA: from time series to topological shapes

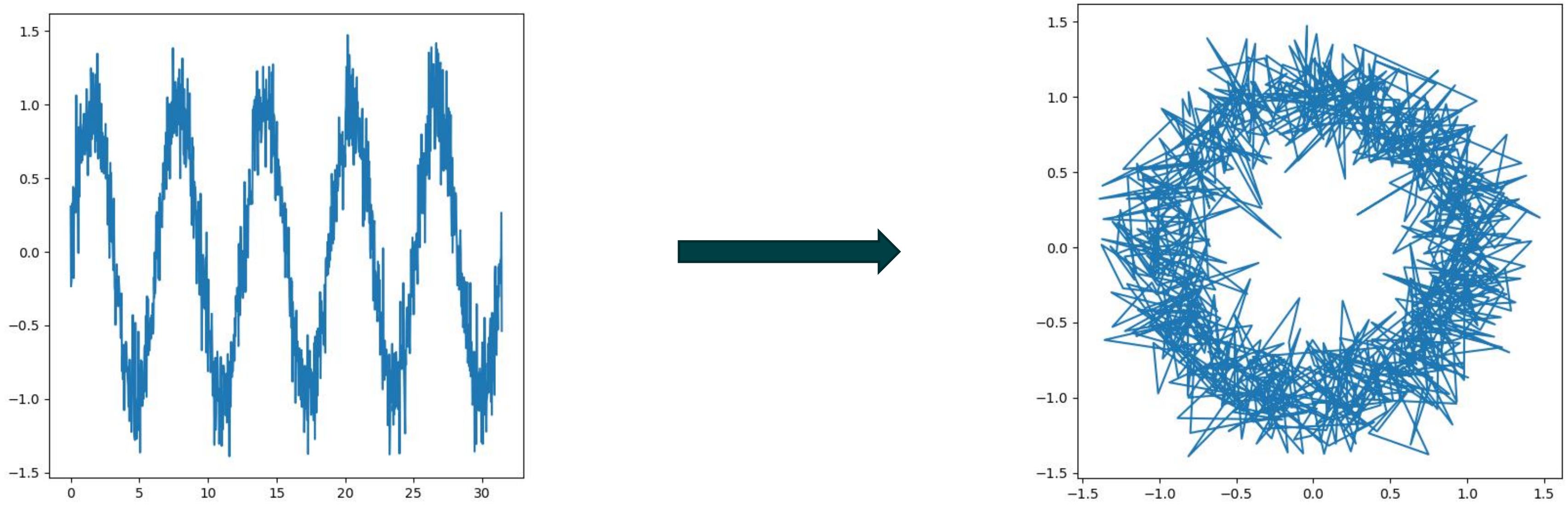
- Most periodic time series can be realized by a topological circle embedded in a Euclidean space of higher dimension.
- The topological type is **robust** against perturbations.



$y = \sin x$

Ideas of TDA: from time series to topological shapes

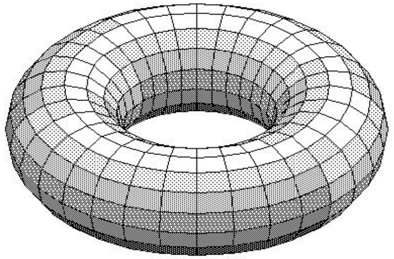
- Most periodic time series can be realized by a topological circle embedded in a Euclidean space of higher dimension.
- The topological type is **robust** against perturbations.



$y = \sin x$

Ideas of TDA: algebraic invariants attached to topological spaces

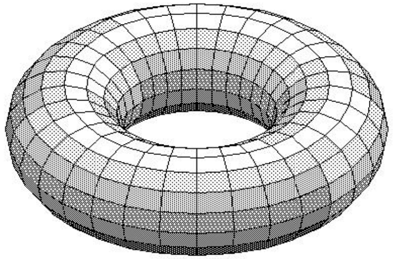
- Features of topological shapes, such as the number of holes, can be captured by algebraic invariants that are **computable**.
- Comparing the invariants effiectively distinguishes the topological types of shapes.



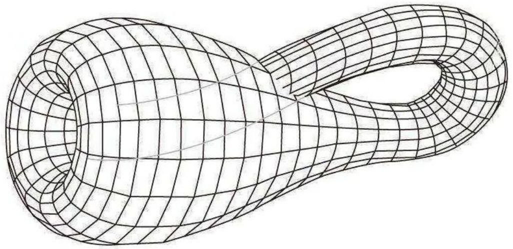
$$H_k(\text{donut surface}) = \begin{cases} \mathbb{Z} \ , & k=0 \\ \mathbb{Z}\oplus\mathbb{Z} \ , & k=1 \\ \mathbb{Z} \ , & k=2 \\ 0 \ , & k>2 \end{cases}$$

Ideas of TDA: algebraic invariants attached to topological spaces

- Features of topological shapes, such as the number of holes, can be captured by algebraic invariants that are **computable**.
- Comparing the invariants effiectively distinguishes the topological types of shapes.

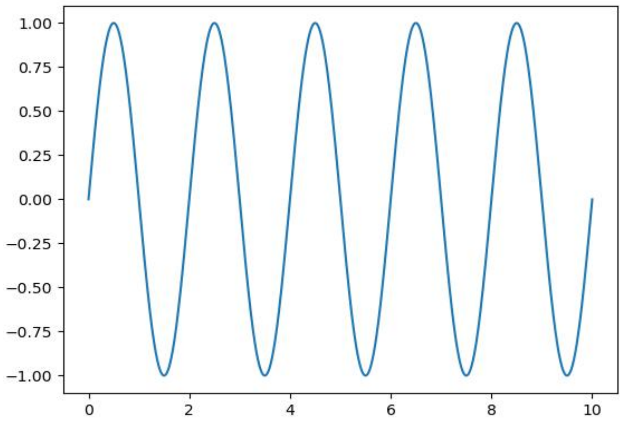


$$H_k(\text{donut surface}) = \begin{cases} \mathbb{Z}, & k=0 \\ \mathbb{Z}\oplus\mathbb{Z}, & k=1 \\ \mathbb{Z}, & k=2 \\ 0, & k>2 \end{cases}$$

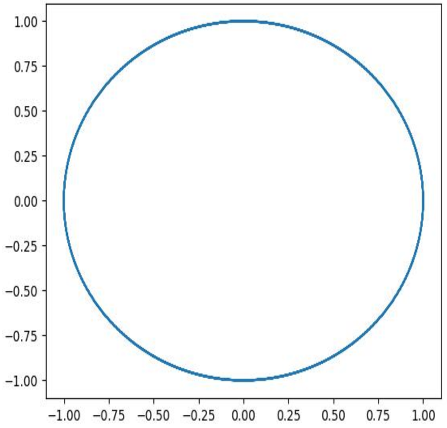


$$H_k(\text{Klein bottle}) = \begin{cases} \mathbb{Z}, & k=0 \\ \mathbb{Z}\oplus\mathbb{Z}/2, & k=1 \\ 0, & k=2 \\ 0, & k>2 \end{cases}$$

Ideas of TDA: algebraic invariants and periodicity



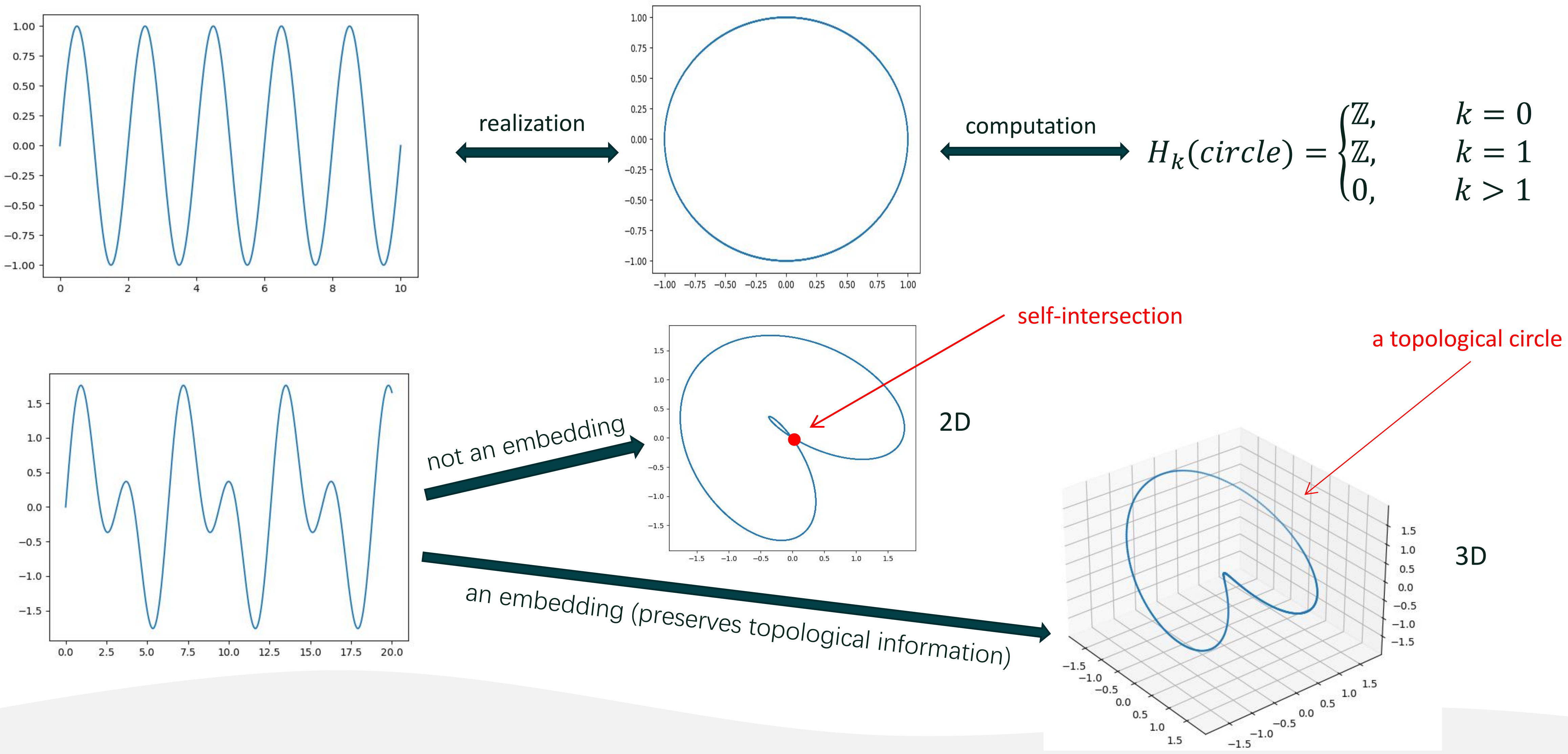
realization



computation

$$H_k(circle) = \begin{cases} \mathbb{Z}, & k = 0 \\ \mathbb{Z}, & k = 1 \\ 0, & k > 1 \end{cases}$$

Ideas of TDA: algebraic invariants and periodicity



$$H_k(circle) = \begin{cases} \mathbb{Z}, & k = 0 \\ \mathbb{Z}, & k = 1 \\ 0, & k > 1 \end{cases}$$

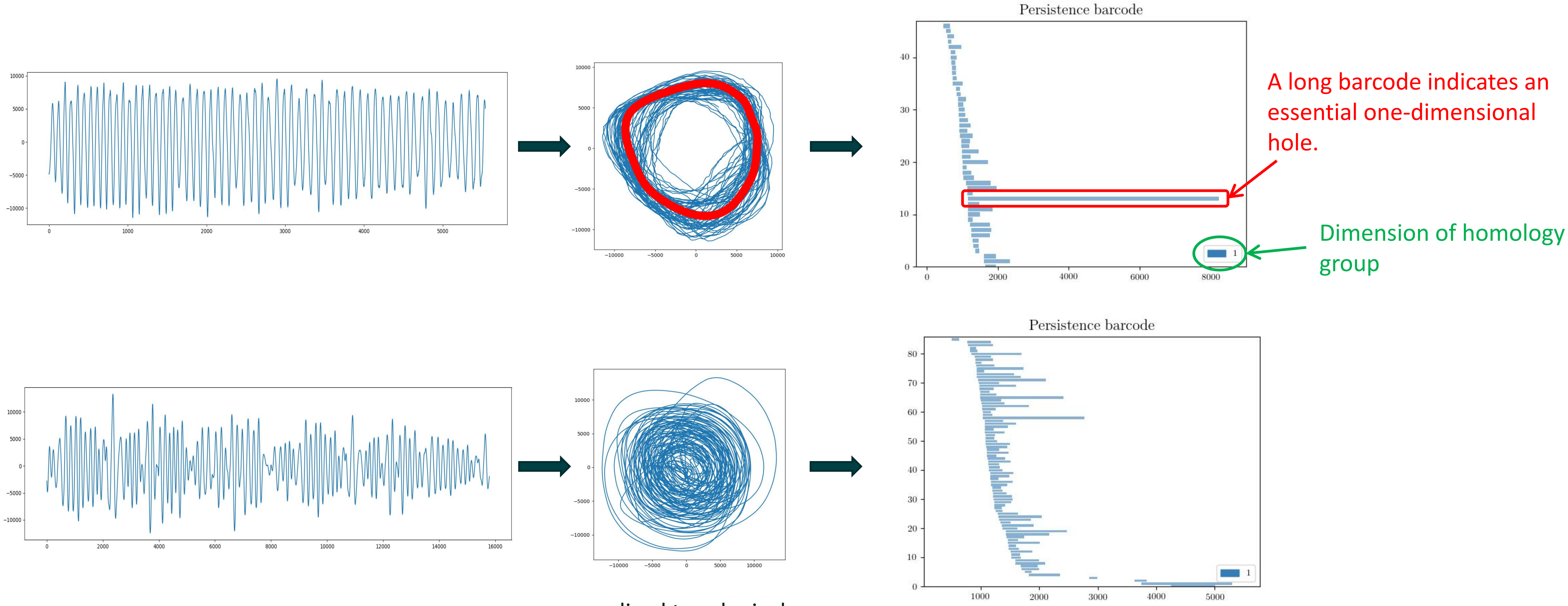
Applications: wheeze detection (Emrani *et al*, IEEE Signal Processing Letters, 2014)

- Wheezes are abnormal lung sounds and usually imply obstructive airway diseases.
- The most important characteristic of wheeze signals is their periodic behavior.
- The accuracy of topological periodicity detection is 98.39%, while in two earlier papers with different techniques it is 86.2% (Homs-Corbera *et al*, IEEE Trans. Biomed. Eng, 2004) and 95.5% (Taplidou *et al*, Comput. Biol. Med., 2007).
- Our research group has reproduced their results using the original data and open-source TDA programing package.

Applications: wheeze detection (Emrani *et al*, IEEE Signal Processing Letters, 2014)

wheeze

normal



original sound signals

realized topological shapes embedded in 2D Euclidean space

"Persistence barcode" as a representations of the algebraic invariant (1D homology group)

Summary: a flow chart for topological time series analysis

Time series data: $x_1, x_2, x_3, x_4, \ldots$

Preprocessing → Sliding window embedding

2D Euclidean shape: point cloud $(x_1, x_{1+\tau}), (x_2, x_{2+\tau}), (x_3, x_{3+\tau}), \ldots$

Topological feature extraction → Persistent homology

Algebraic invariants: homology groups, persistence barcodes, ...

Statistical inference → Machine learning, persistence scoring, etc.

Characteristics conclusions

Current work and proposed collaboration

- Currently, we aim to apply TDA to speech signal and distinguish vowels and consonants based on their topological features.
- Because a video encodes a multidimensional time serie data, one can use analogous methods to distinguish periodic from non-periodic segments in the video (Perea *et al*, SIAM J. Imaging Sci., 2018).
- Topological time series analysis has the advantages of **stability/robustness** against noises and relative computation **time efficiency**. And, there are also some challenges in topological time series analysis, that is, the choice of embedding mode.

More details: persistent homology

Sample

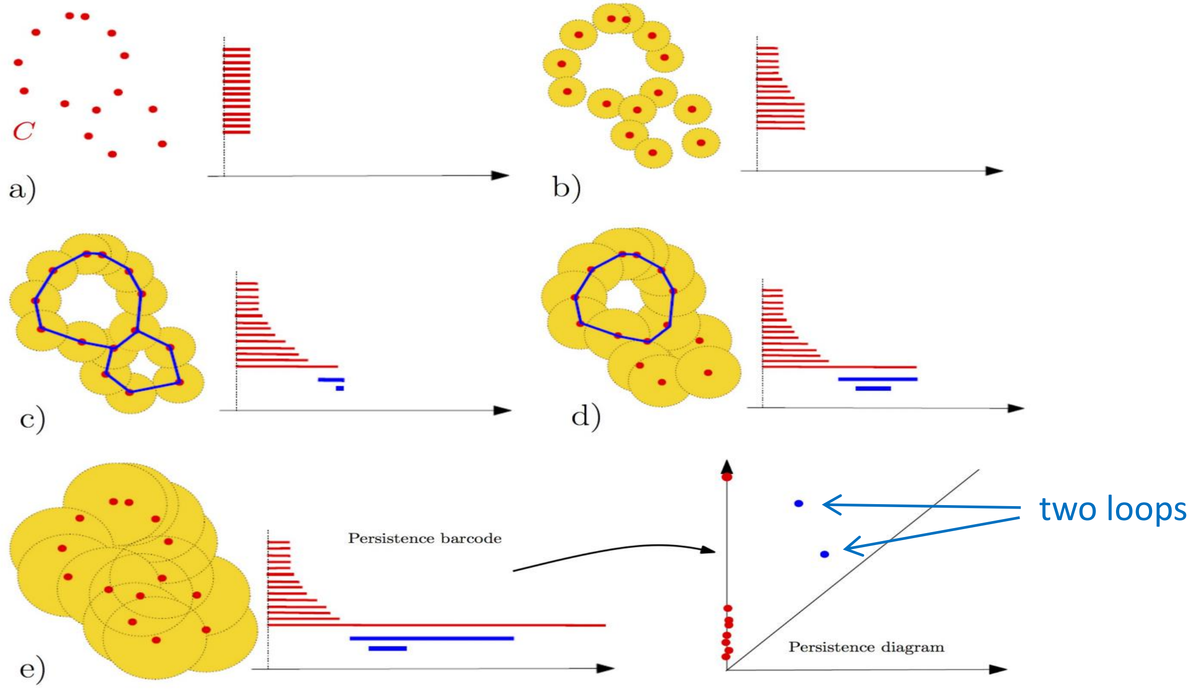
↓

Construct filtration of simplicial complexes and compute homology groups of complexes

↓

Produce barcodes/diagrams

How filtration through varying distance measure reveals essential topological features



More details: sliding window embedding

- Euclidean embedding of time seires data dates back to Takens's work on fluid turbulence in the 1980s.
- Theorem (Takens, 1981): Let M be a compact manifold of dimension n. For pairs (φ, y), $\varphi: M \to M$ a smooth diffeomorphism and $y: M \to \mathbb{R}$ a smooth function, it is a generic property that the map $\Phi_{(\varphi,y)}: M \to \mathbb{R}^{2n+1}$ defined by

$$\Phi_{(\varphi,y)}(x) = (y(x), y(\varphi(x)), \cdots, y(\varphi^{2n}(x)))$$

is an embedding; by "smooth" we mean at least C^2.