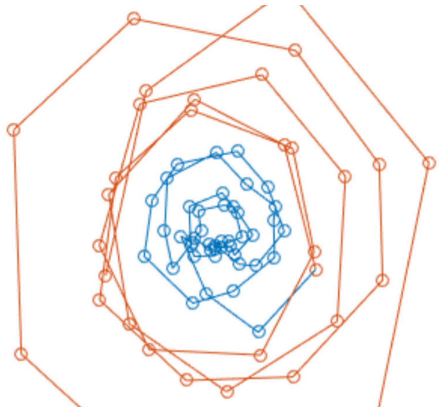
Topological time series analysis: with applications to biomedical and speech signal processing



朱一飞 (SUSTech)

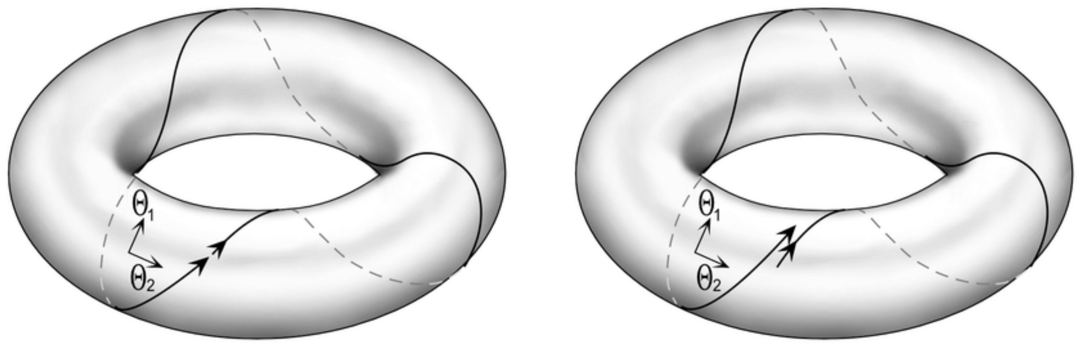
杭州师范大学数学学院陈建功大讲堂, 2022.11

Periodic phenomena: a motivating example

Let $\mathbb{T}^2 = (\mathbb{R}/\mathbb{Z})^2$ be the 2D torus. Consider the dynamical system given by

$$\begin{aligned} \Phi_\sigma \colon \mathbb{T}^2 \times \mathbb{R} &\to \mathbb{T}^2 \\ \big((a,b),t\big) &\mapsto (a+t, b+\sigma t) \end{aligned}$$

If σ is rational, then every orbit is periodic. Otherwise every orbit is dense in $\mathbb{T}^2$.



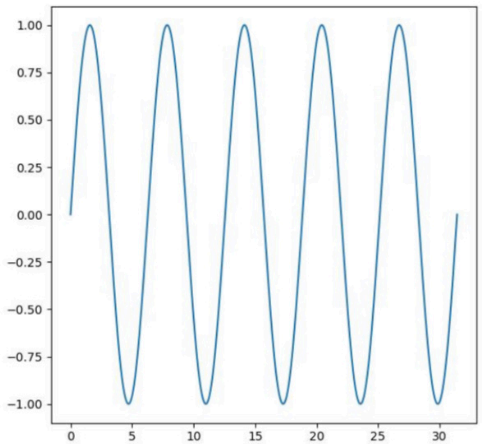
(a) (b)

From time series to topological shapes

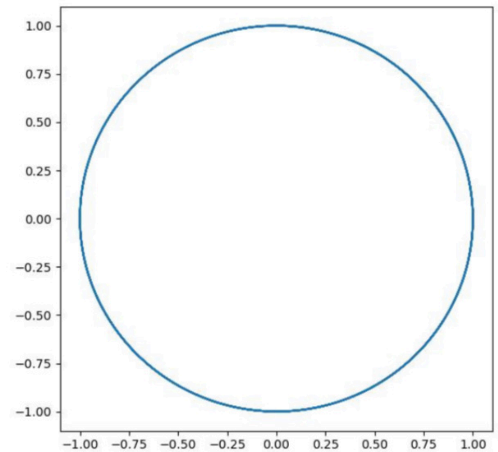
Most periodic time series can be realized by a topological circle S^1 embedded in a Euclidean space of higher dimension.

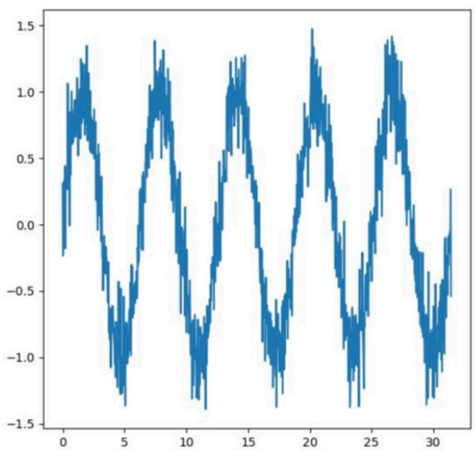
Ideas of topological data analysis (TDA)

The topological type (more precisely, homotopy type) is robust against perturbations.

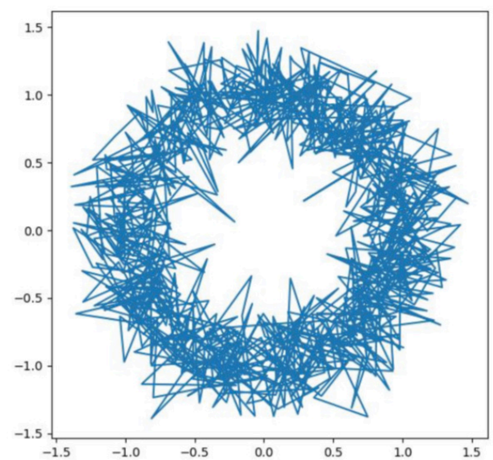


$y = \sin x$





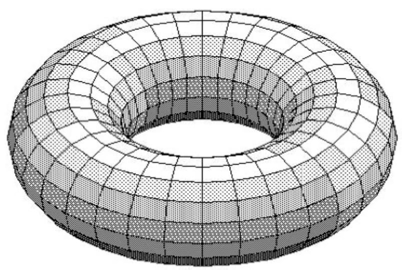
$y = \sin x + \epsilon(x)$

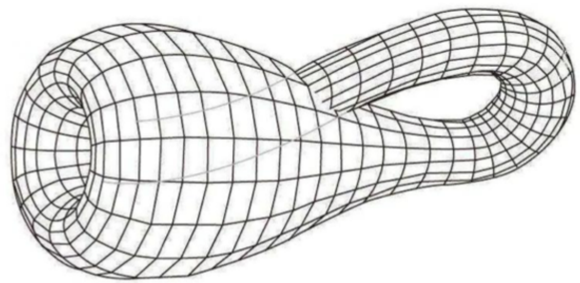


Ideas of topological data analysis (TDA)

Features of topological shapes, such as the number of holes, can be captured by algebraic invariants that are computable.

Comparing these invariants effectively distinguishes the topological types of shapes.





$$H_k(\mathbb{T}^2) = \begin{cases} \mathbb{Z} & k = 0 \\ \mathbb{Z} \oplus \mathbb{Z} & k = 1 \\ \mathbb{Z} & k = 2 \\ 0 & k > 2 \end{cases}$$

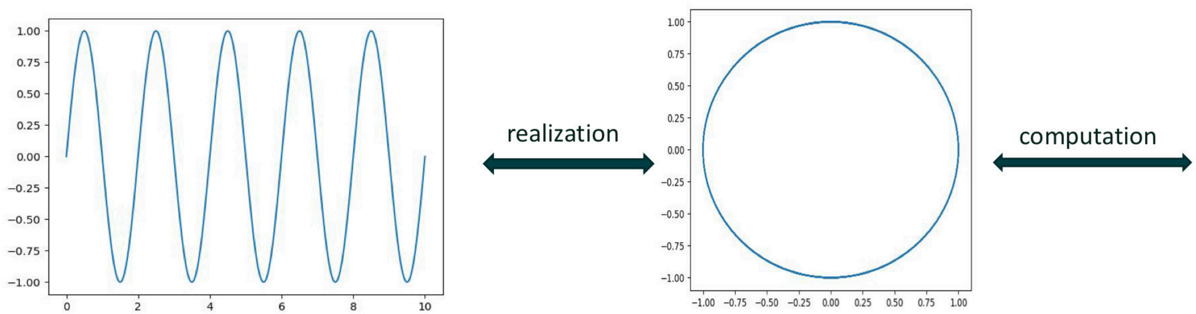
$$H_k(\text{Klein bottle}) = \begin{cases} \mathbb{Z} & k = 0 \\ \mathbb{Z} \oplus \mathbb{Z}/2 & k = 1 \\ 0 & k = 2 \\ 0 & k > 2 \end{cases}$$

Topological time series analysis

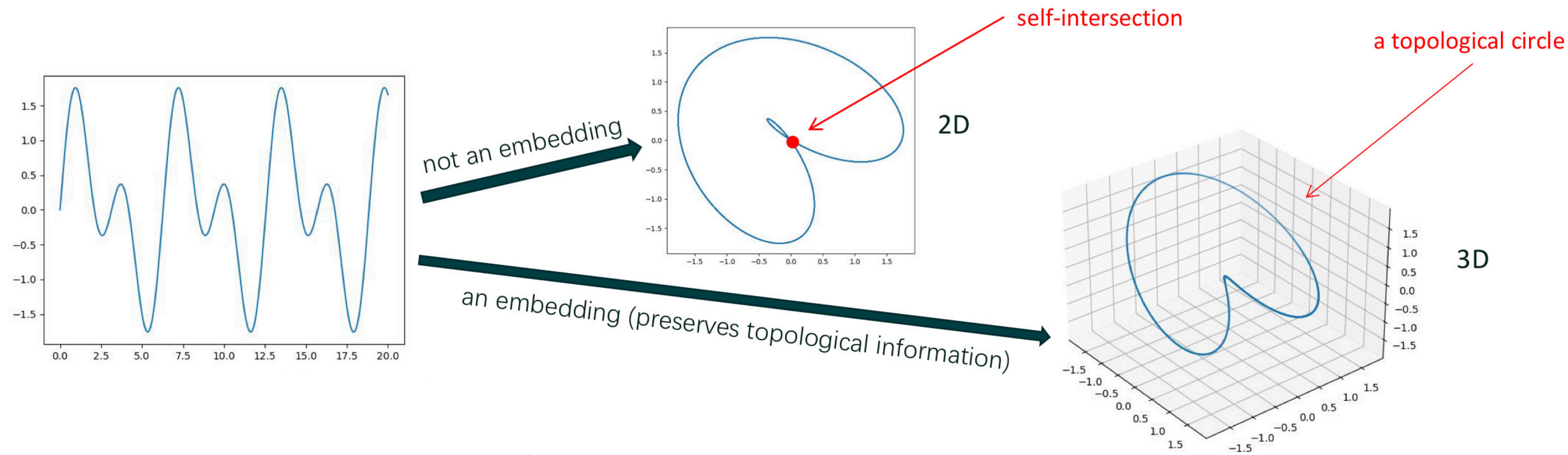
Let us make the assumption that sampled signals are distributed over a manifold. To topologically analyze time series, we then proceed as follows:

<u>Step1</u> Embed the data into a Euclidean space of suitable dimension;

<u>Step 2</u> Compute the algebraic invariants for statistical inference.



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



An application: detection of wheeze in medical science (pulmonology)

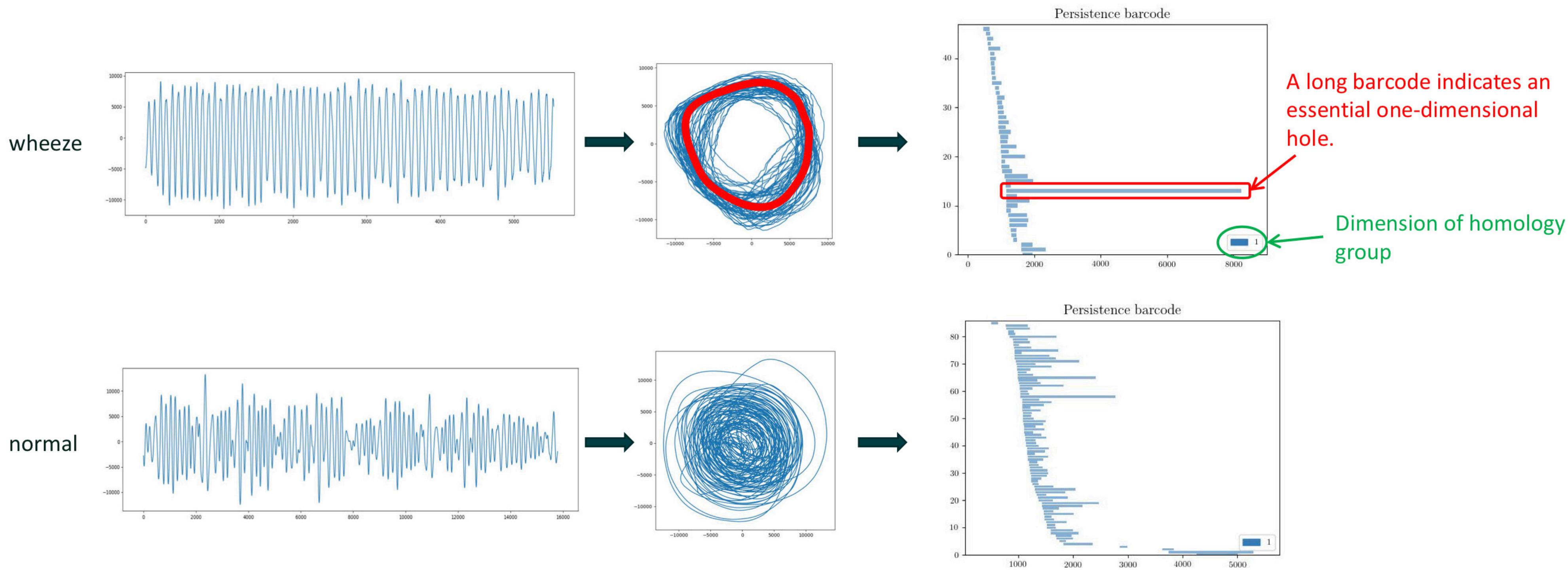
Wheezes are abnormal lung sounds and usually imply obstructive airway diseases.

The most important characteristic of wheeze signals is their periodic patterns.

The accuracy of topological periodicity detection is 98.39% (Emrani et al., IEEE Signal Processing Letters, 2014), while in two earlier papers with different methods they are 86.2% and 95.5%.

As a warm-up, our research group has reproduced their results using the original data and open-source TDA programming package.

An application: detection of wheeze in medical science (pulmonology)



Original sound signals

Realized topological shapes embedded in 2D Euclidean space

“Persistence barcodes” as representations of the algebraic invariant (1D homology group)

A pipeline 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 $(x_1, x_{1+\tau}), (x_2, x_{2+\tau}), (x_3, x_{3+\tau}), \ldots$

delay

Topological feature extraction → Persistent homology

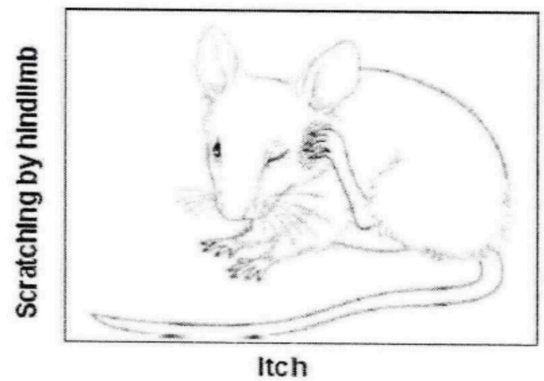
Algebraic invariants: homology groups, persistence barcodes, ...

Statistical inference → Machine learning, persistence scoring, etc.

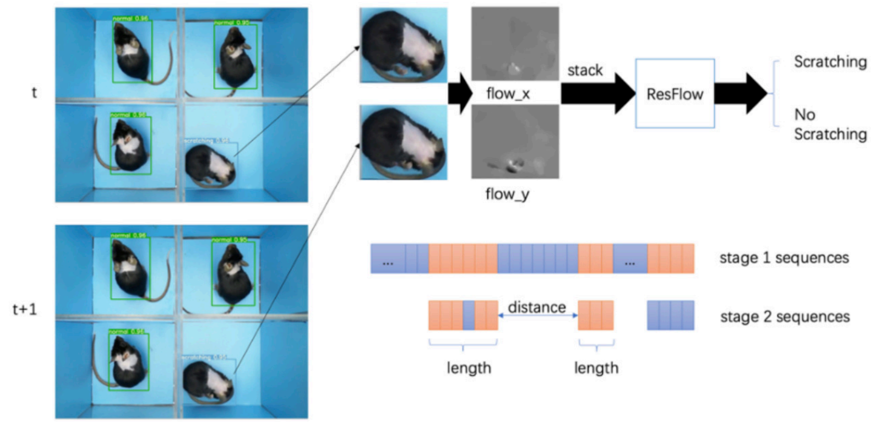
Characteristics conclusions

Application I: detection of mouse scratching behavior

Joint with the biomedical engineering group led by Fangyi Chen and the data science group led by Zhen Zhang, both at SUSTech, we applied topological methods to the problem of automated and real-time detection of mouse scratching behavior, with motivations from pharmacology.



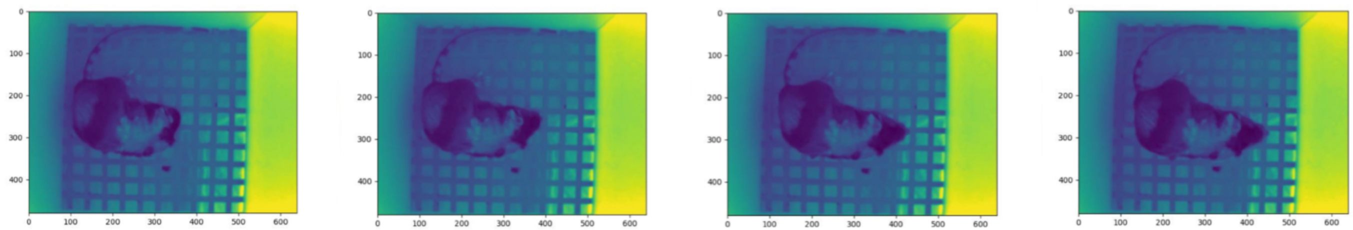
Prior to our group's involvement, machine learning via neural networks was applied with satisfactory accuracy (https://yifeizhu.github.io/scratch.mp4).



However, the learning process was time consuming, which is impractical for time-sensitive purposes and lab efficiency.

Application I: detection of mouse scratching behavior

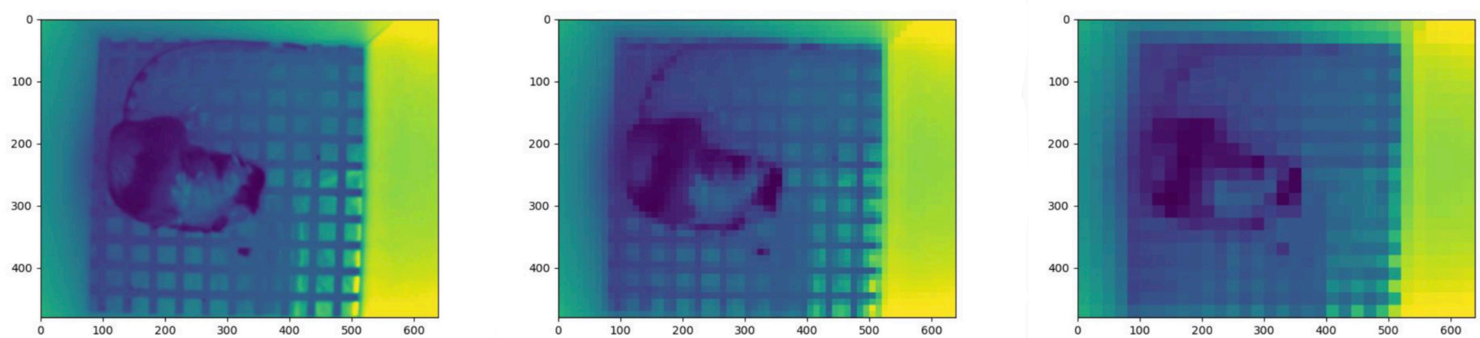
We observed that the scratching behavior exhibits periodicity. In the meantime, however, global movements of a mouse may significantly reduce the pattern.



To resolve this issue, we adopted the following approaches:

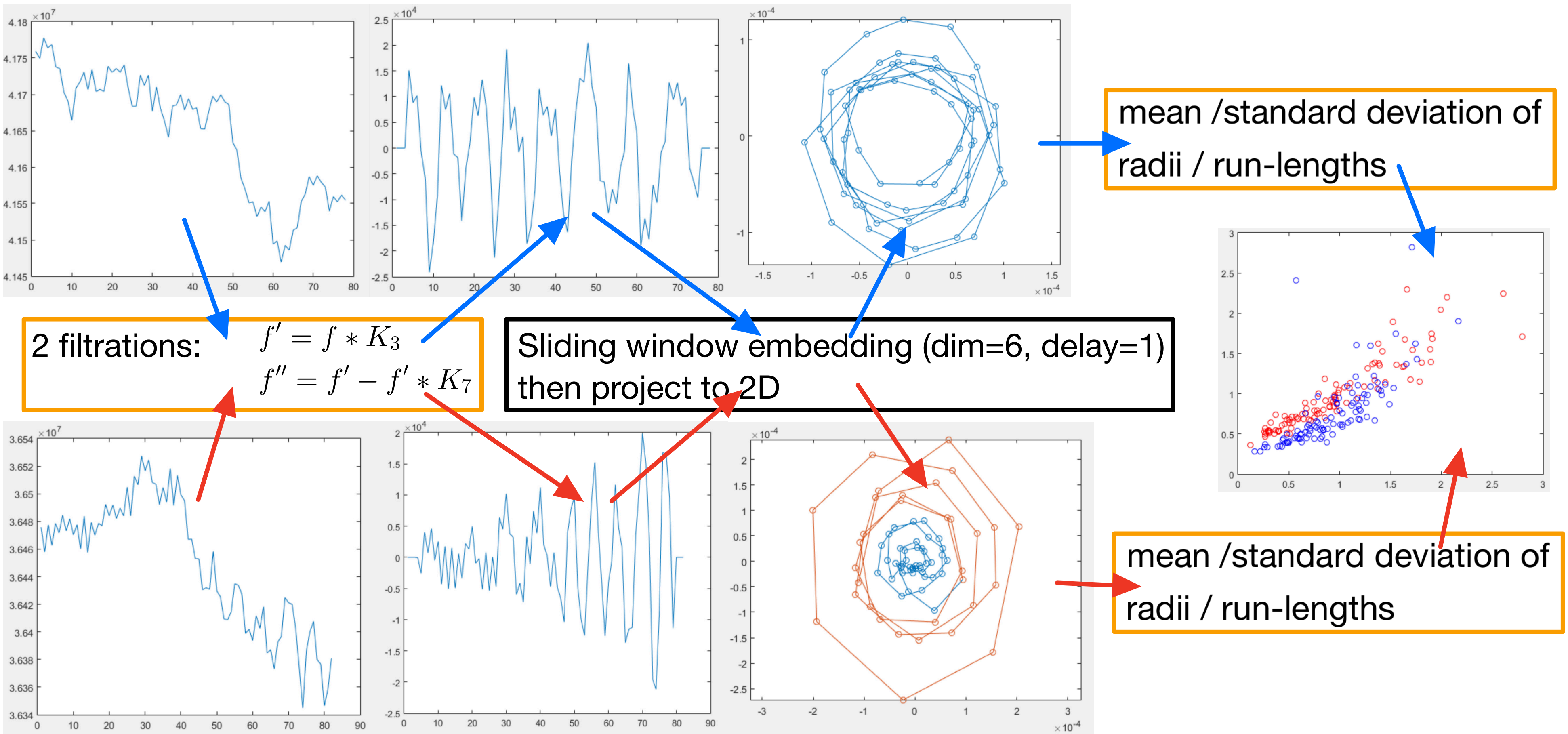
<u>Approach 1</u> Sum up all 460 x 640 pixels to extract a series of 1D data which ignores differences caused by global movements. Too coarse?

<u>Approach 2</u> Blur the images by pooling, and feed the topological pipeline with reduced 100-dimensional data. Still too refined?



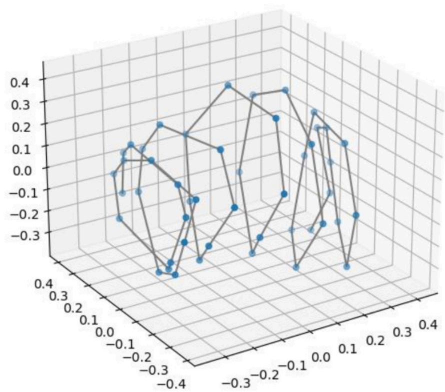
Application I: detection of mouse scratching behavior

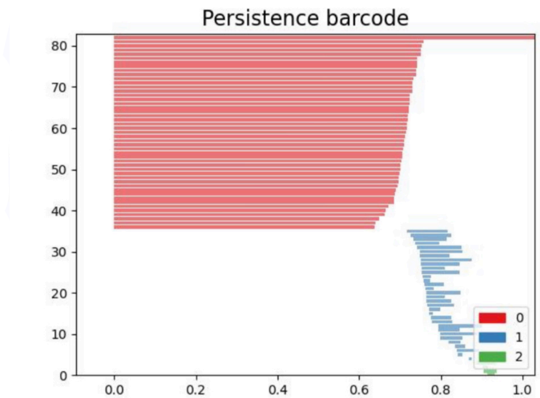
Approach 1 (1D data), combined with carefully designed filtration for wave signals + suitably chosen geometric statistics, yielded a close-to-real-time, decently accurate detection performance.

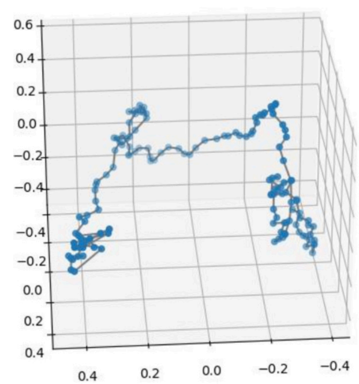


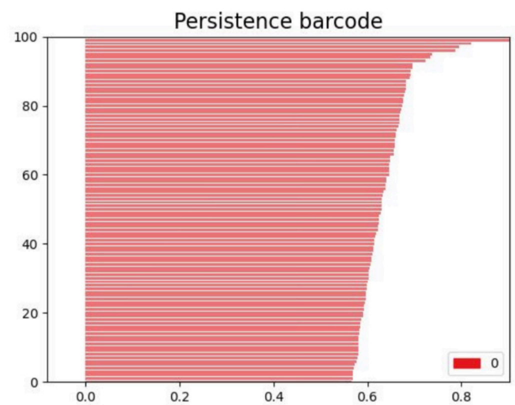
Application I: detection of mouse scratching behavior

Approach 2 (multi-dimensional data), combined with persistent homology and its representations, yielded recognizable characteristics but required considerable computational time.









Application II: classification of voiced and unvoiced speech signals

Joint with Meng Yu of Tencent AI Lab, we applied topological methods to classify voiced/unvoiced and vowel/consonant speech data, with motivations from industrial applications.

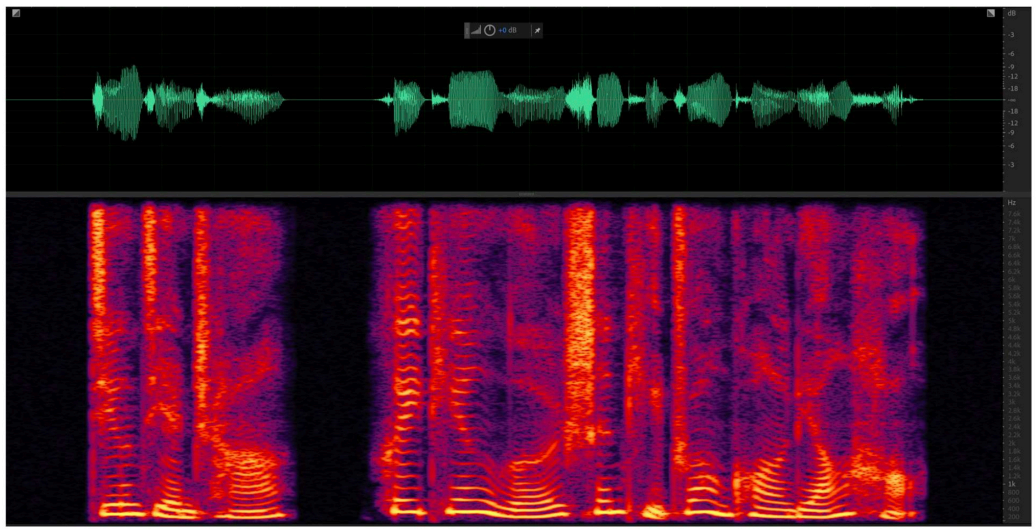
We were inspired by Carlsson et al.'s discovery of the Klein-bottle distribution of local natural images, as well as their subsequent recent work of topological convolutional neural networks learning video data. We would like to understand an analogous "moduli space" for speech data and how its input may enable smarter learning.

Display of speech signals

There are speech signal processing softwares for professional use.

Time

Time-Frequency



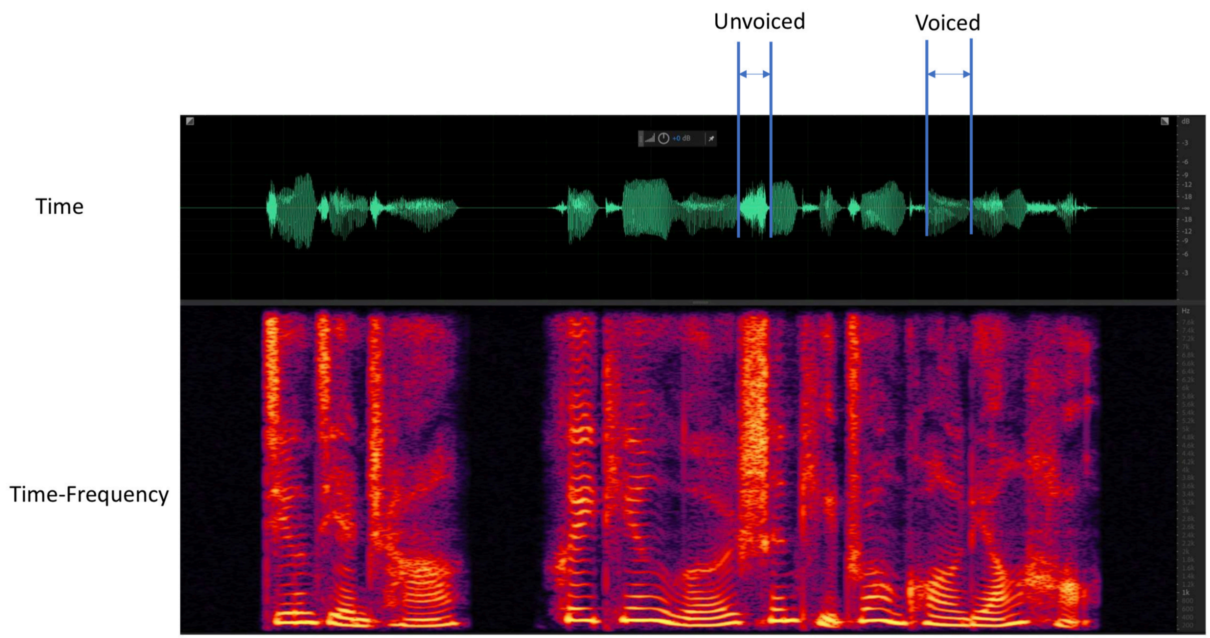
Application II: classification of voiced and unvoiced speech signals

Joint with Meng Yu of Tencent AI Lab, we applied topological methods to classify voiced/unvoiced and vowel/consonant speech data, with motivations from industrial applications.

We were inspired by Carlsson et al.'s discovery of the Klein-bottle distribution of local natural images, as well as their subsequent recent work of topological convolutional neural networks learning video data. We would like to understand an analogous "moduli space" for speech data and how its input may enable smarter learning.

Display of speech signals

There are speech signal processing softwares for professional use.

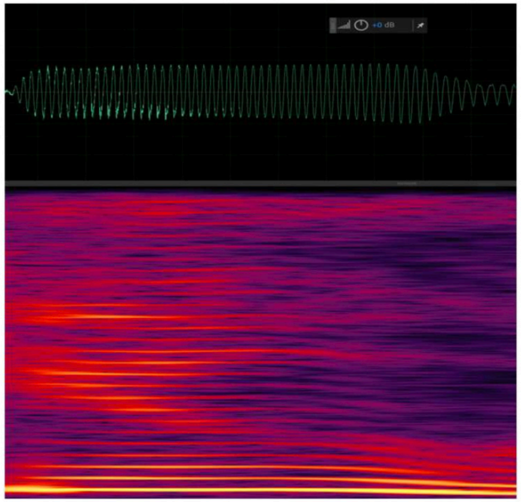


Application II: classification of voiced and unvoiced speech signals

Voiced

Sinusoid in time domain

Harmonics in frequency domain

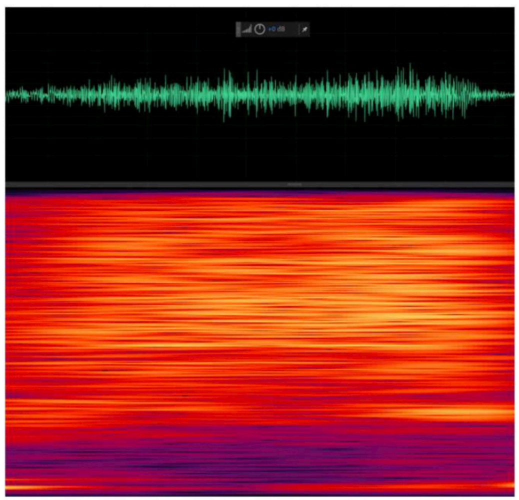


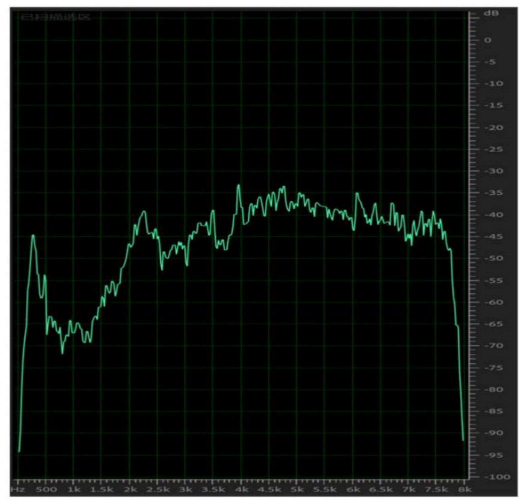
Time and Time-Frequency domain

Frequency response

Unvoiced

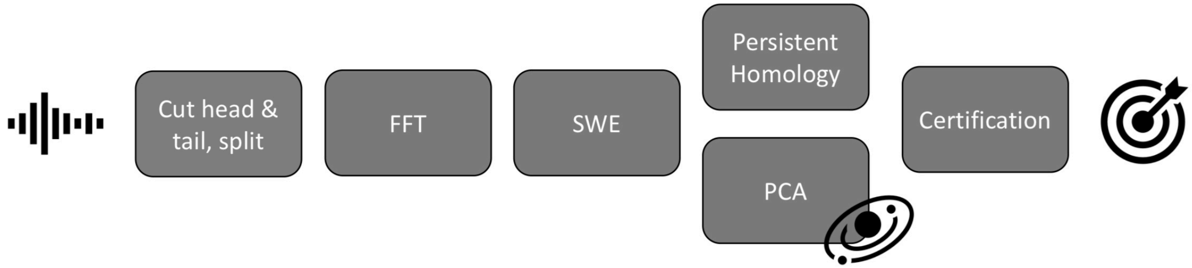
Like a white noise



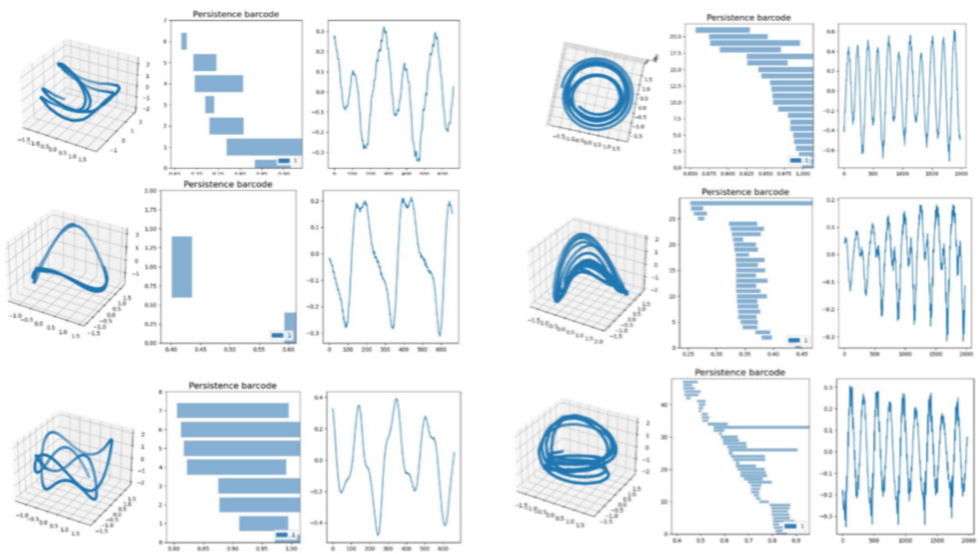


Application II: classification of voiced and unvoiced speech signals

Here is a flowchart for our topological approach:

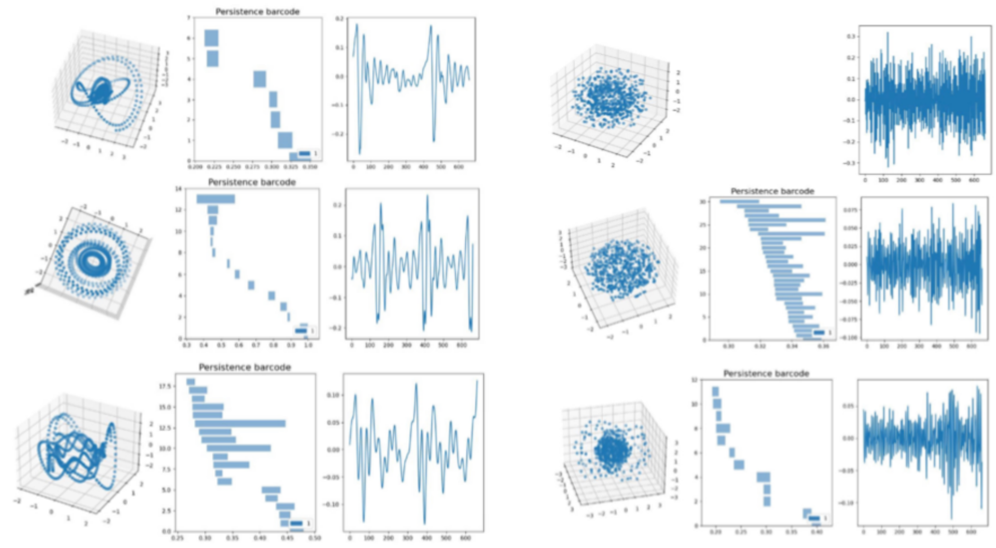


Topological profiles for vowels and consonants



Features for vowels

Left: frame size: 15ms, frame shift: 5ms; Right: frame size: 45ms, frame shift: 22.5ms

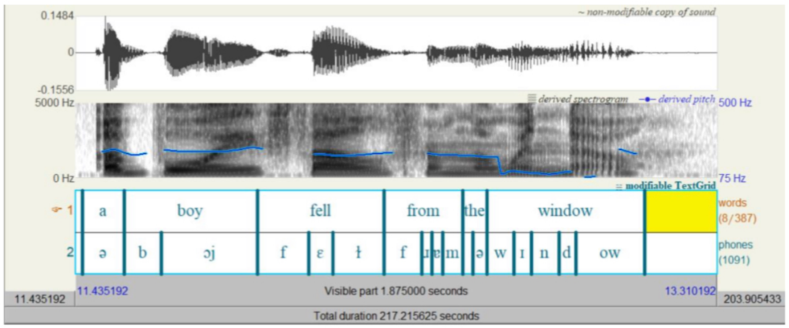


Features for consonants

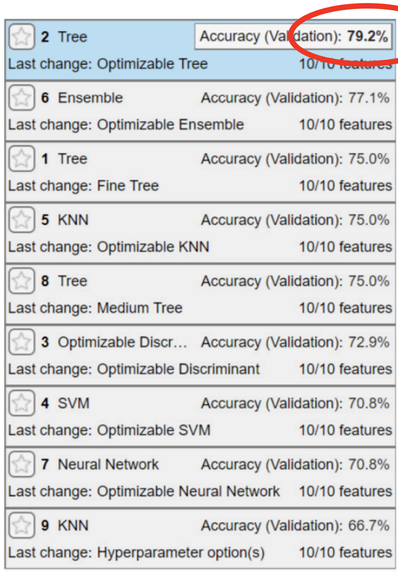
Left: pulmonic consonant; Right: non-pulmonic consonant

Application II: classification of voiced and unvoiced speech signals

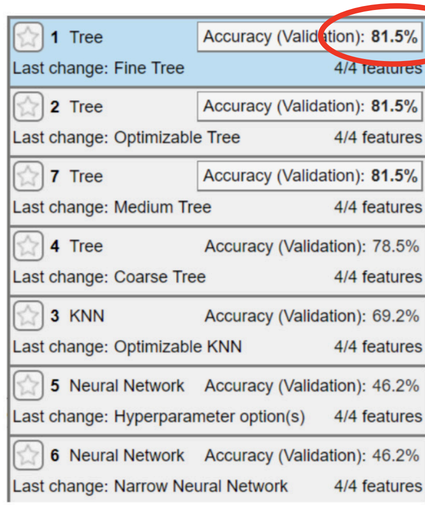
Using real-world speech data from the MFA aligner, we further fed the topological features for machine learning, and obtained positive preliminary results for classification.




```
vowel_phones=['ɔj','ɛ','ə','ɪ','aj','ɑ','æ','i','o','ʊ','aw','e','u','a']
consonant_phones=['b','f','m','ɹ','ð','w','h','p','t','z','n','g','dʒ','s','ʃ','v','l','ŋ','k','θ','j','tʃ','ʒ','d']
```

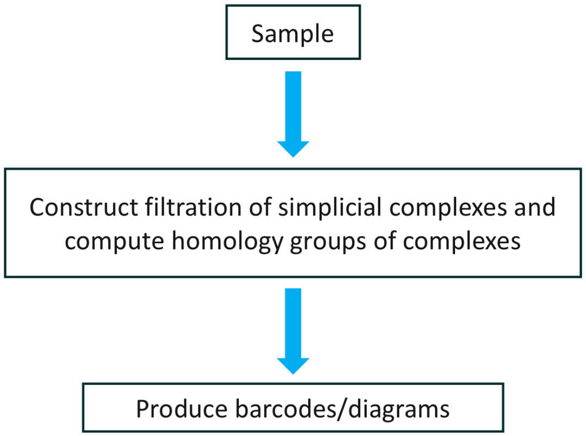


32 vowels, 16 consonants. 10 features: 5 are barcodes number of 5 diag, other 5 are number of barcodes that reaches inf(both consider barcode of 1 dimension for only)

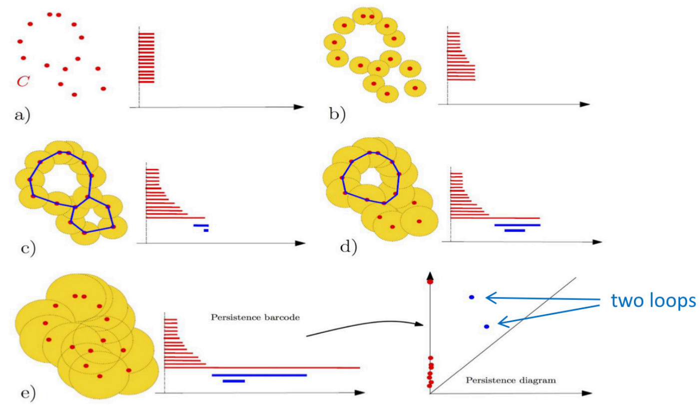


32 vowels, 33 consonants. 4 features: bottleneck distance between neighborhood barcode(currently the best result)

Persistent homology



How filtration through varying distance measure reveals essential topological features



Sliding window embedding

Euclidean embedding of time series 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. Given pairs (ϕ, y) with $\Phi: M \to M$ a smooth diffeomorphism and $y: M \to \mathbb{R}$ a smooth function, it is a generic property that the map $\Phi_{(\phi,\,y)}: M \to \mathbb{R}^{2n+1}$ defined by

$$\Phi_{(\varphi,y)}(x) = \Big(y\,(x)\,, y\big(\varphi(x)\big), \ldots, y\big(\varphi^{2n}(x)\big)\Big)$$

is an embedding.

Thank you.