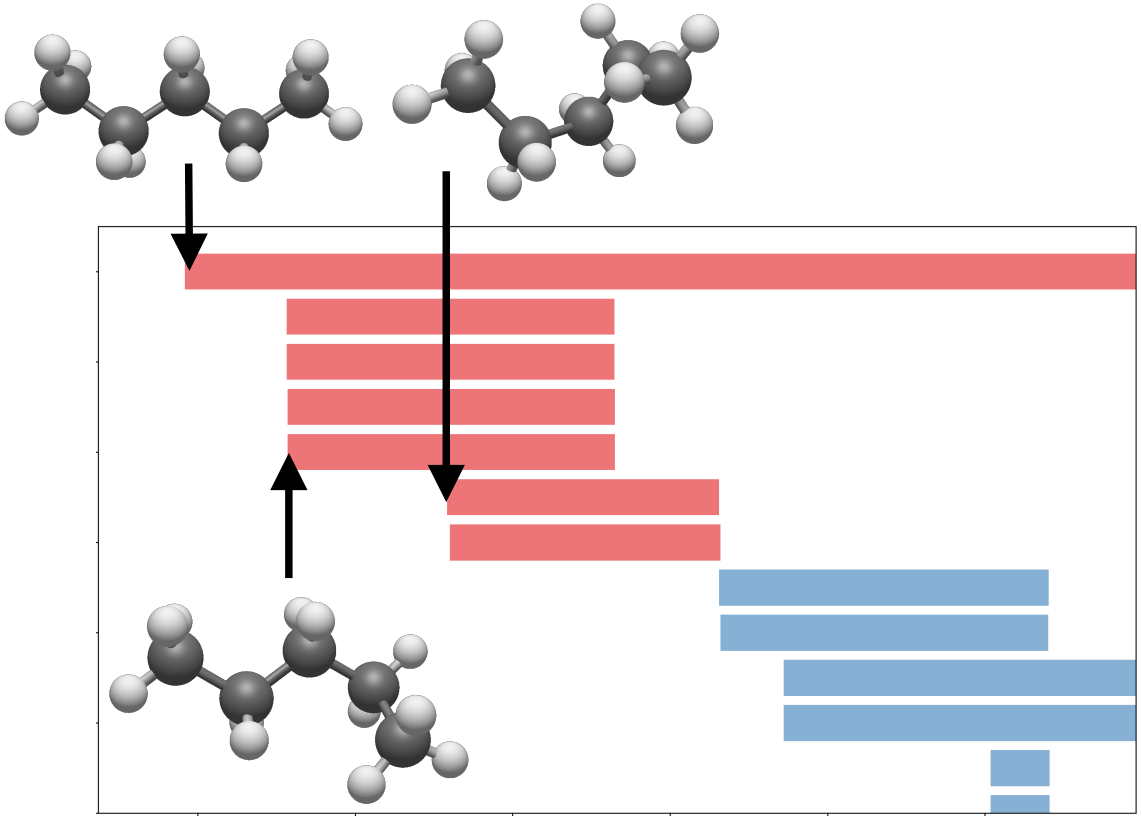
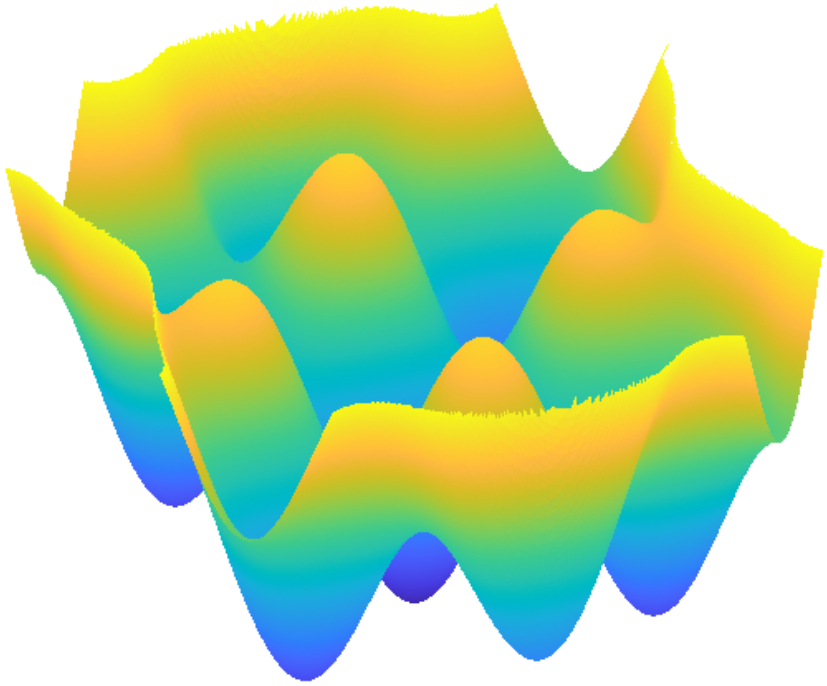
# Descriptors of Energy Landscapes using Topological Analysis

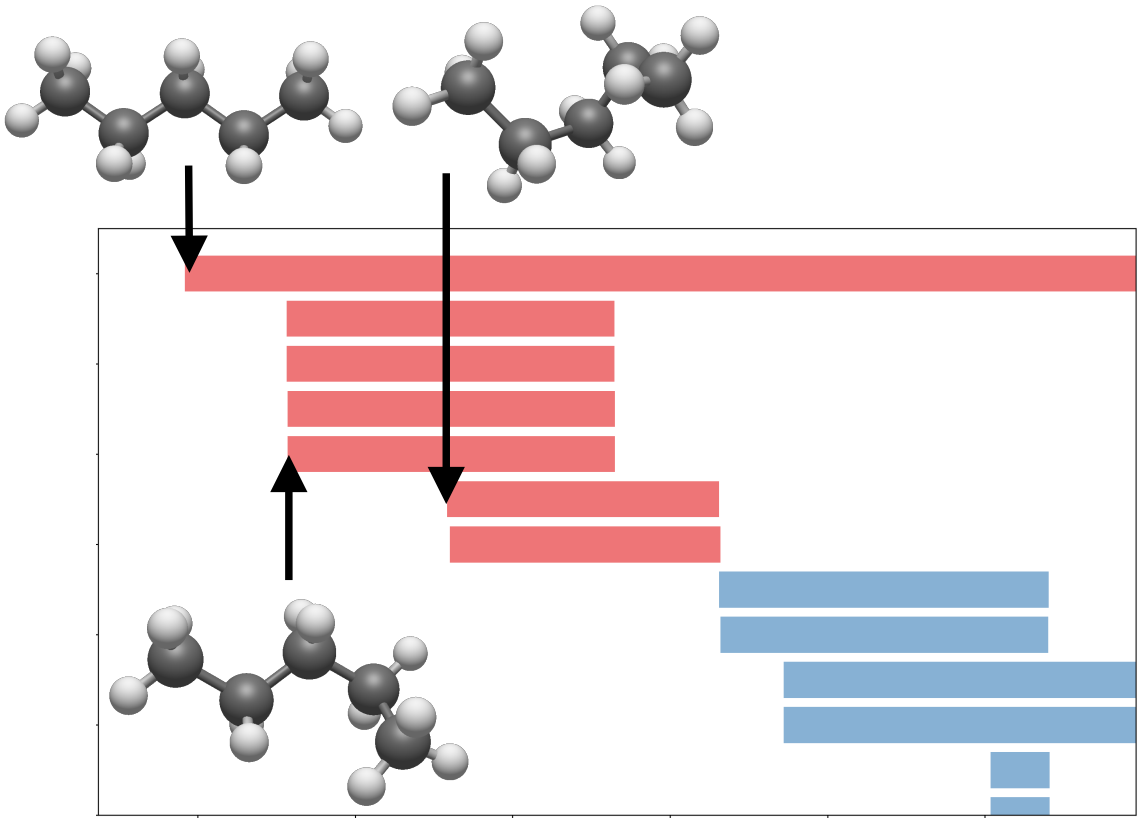


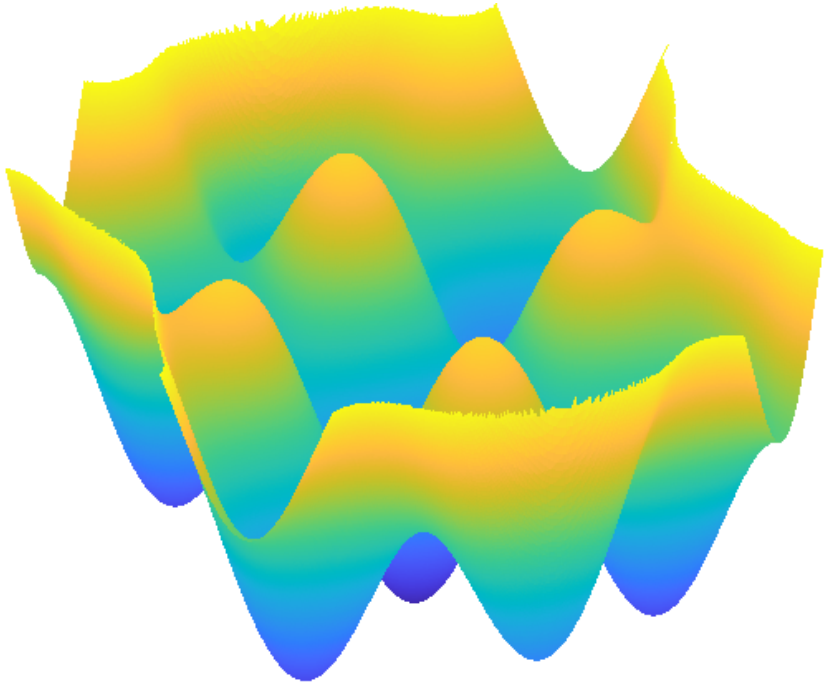


Henry Adams, Colorado State University

DELTA NSF #1934725

# Descriptors of Energy Landscapes using Topological Analysis

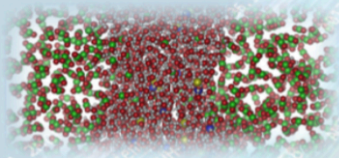




- An introduction to persistent homology
- The n-alkane energy landscapes
- Machine learning tasks

# DELTA

# Descriptors of Energy Landscapes Using Topological Analysis




3N Energy Landscape (Simulation/ Experiment)

Dimensionality Reduction

- PCA
- Non-linear Methods
- Generalized Collective Coordinates

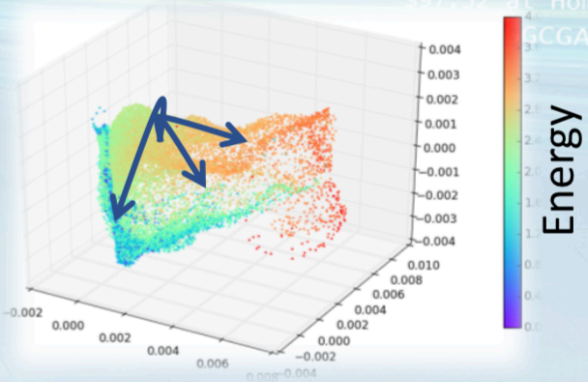
Topology of Reduced Energy Landscapes

- Morse Theory
- Persistent Homology
- Catastrophe Theory
- Singularity Theory

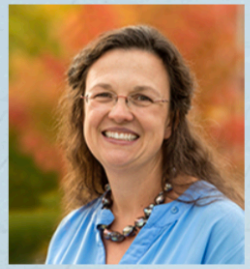
Predictive Machine Learning

- Optimized Synthetic Conditions
- Phase Behavior
- Tuning Catalytic Pathways

Accelerated Sampling

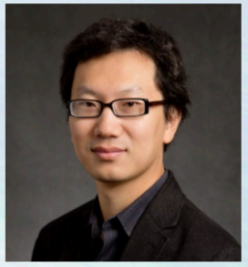



NSF #1934725









WASHINGTON STATE UNIVERSITY

Colorado State University

Rensselaer

AUBURN UNIVERSITY

University of Colorado Boulder

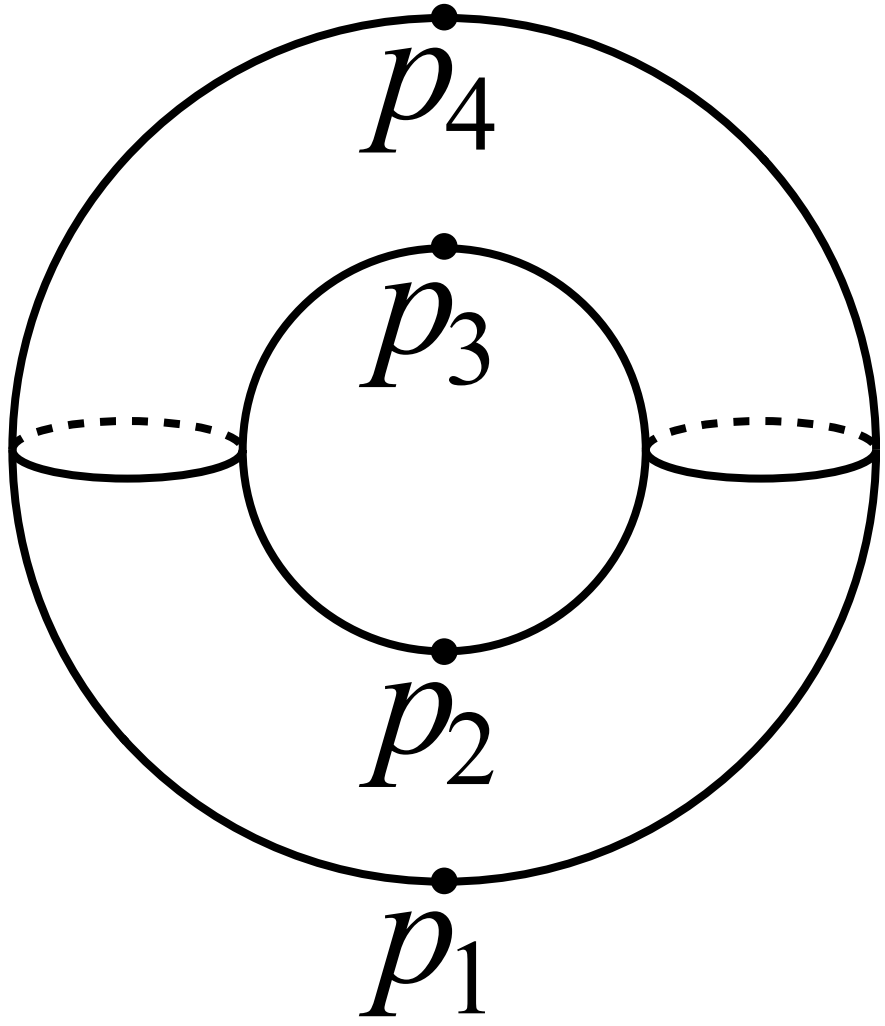
UNIVERSITY OF ARKANSAS

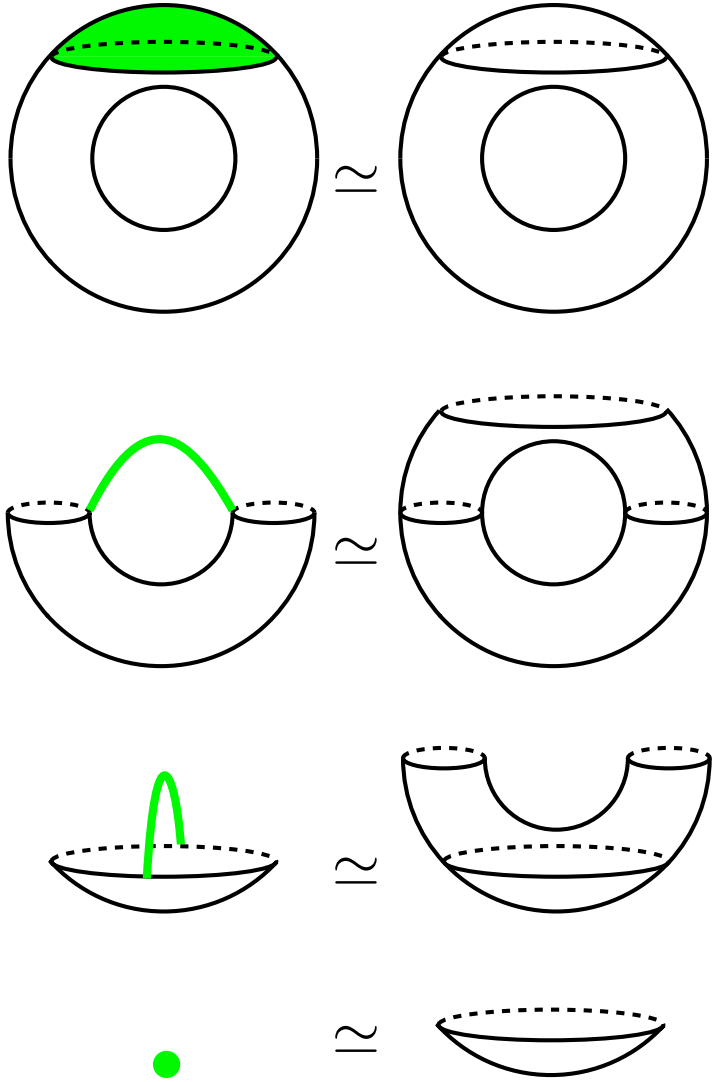
ILLINOIS

# Morse Theory and Persistent Homology

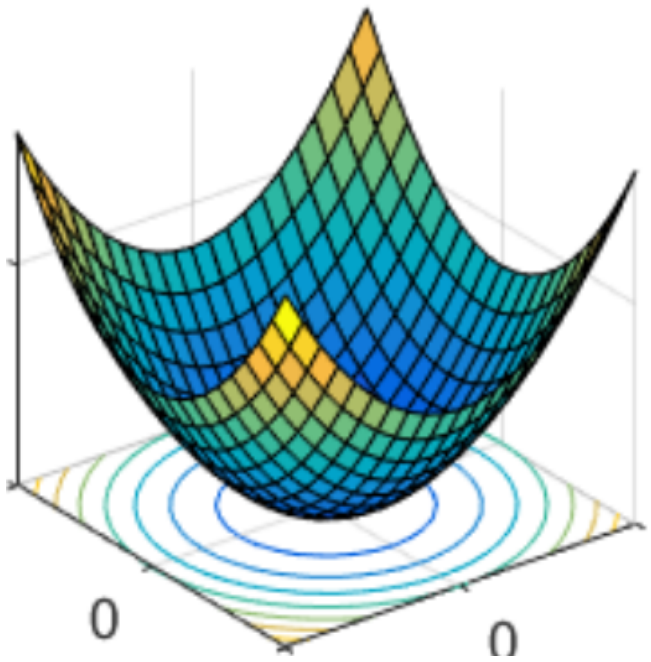
Abstract: Many of the properties of a chemical system are described by its energy landscape, a real-valued function defined on a high-dimensional domain. I will explain how topology, and in particular persistent homology, can be used in order to describe some of the pertinent features of an energy landscape. As a motivating example, the low-energy conformation space of an aluminate molecular anion interacting with a simple potassium cation can be well-modeled by the permutohedron (joint work with Aurora Clark and Hung Le). In a recently funded NSF Harnessing the Data Revolution project, the DELTA team is learning how to identify and leverage changing topological features of energy landscapes across a range of chemical conditions in order to predict reactivity.

# Morse Theory Overview

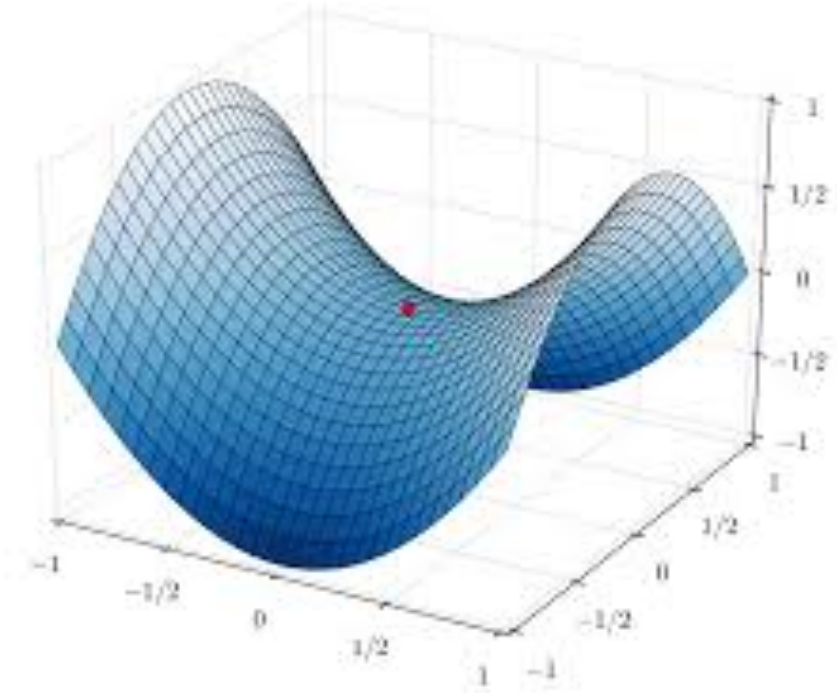


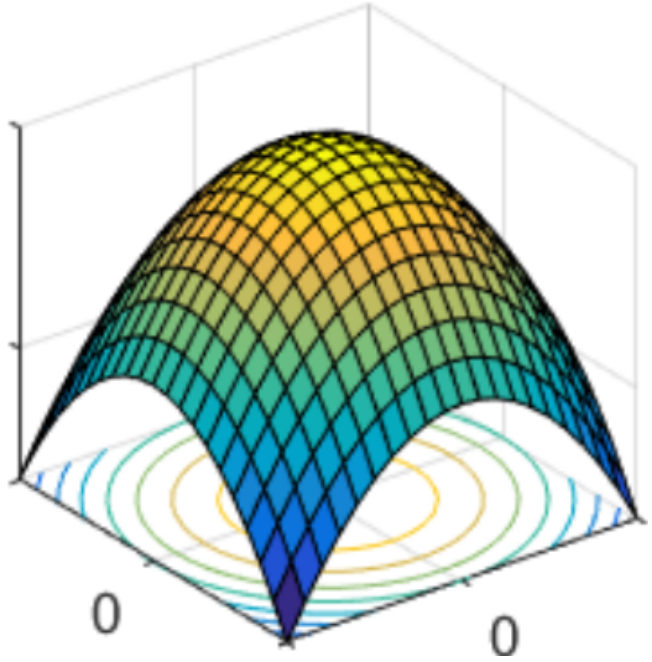
# Morse Theory Overview




**Index 0 critical point**




**Index 1 critical point**




**Index 2 critical point**

Images from Wikipedia

# Morse Theory Overview

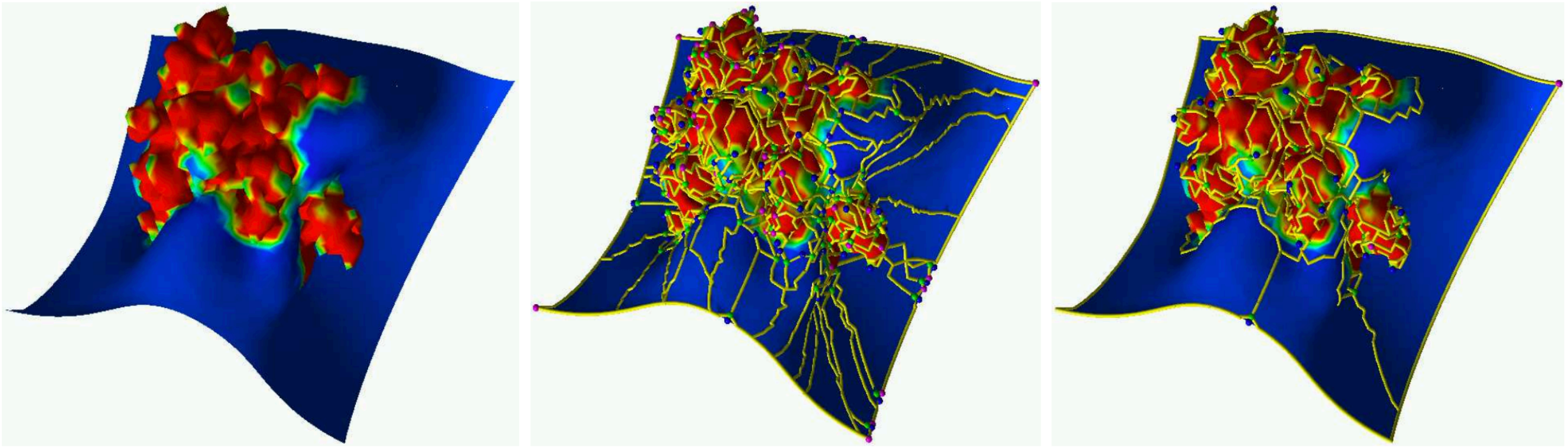


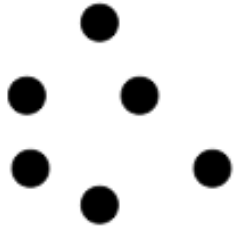
Fig. 16. Interaction energy between glucose and ethane under the three translational degrees of freedom. Left: isosurface of the electrostatic interaction pseudo-colored with the corresponding van der Waals potential. Middle: full MS complex with 564 critical points. Right: simplified MS complex with 166 critical points highlighting good candidate binding sites.

*Topological hierarchy for functions on triangulated surfaces* by Peer Timo-Brener, Herbert Edelsbrunner, Bernd Hamann, and Valerio Pasucci

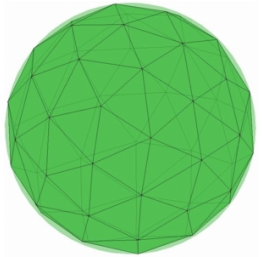
# An Introduction to Persistent Homology

# Homology

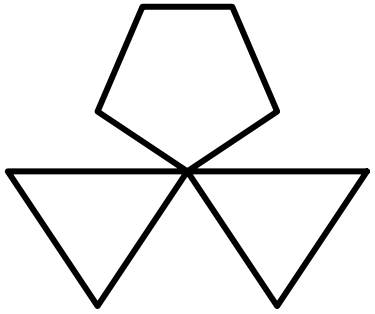
*i*-dimensional homology $H_i$ "counts the number of *i*-dimensional holes"



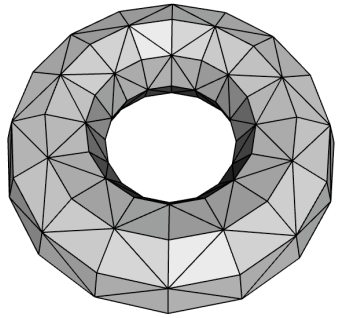
Six connected components
No 1-dimensional holes
No 2-dimensional holes



One connected component
No 1-dimensional holes
One 2-dimensional hole

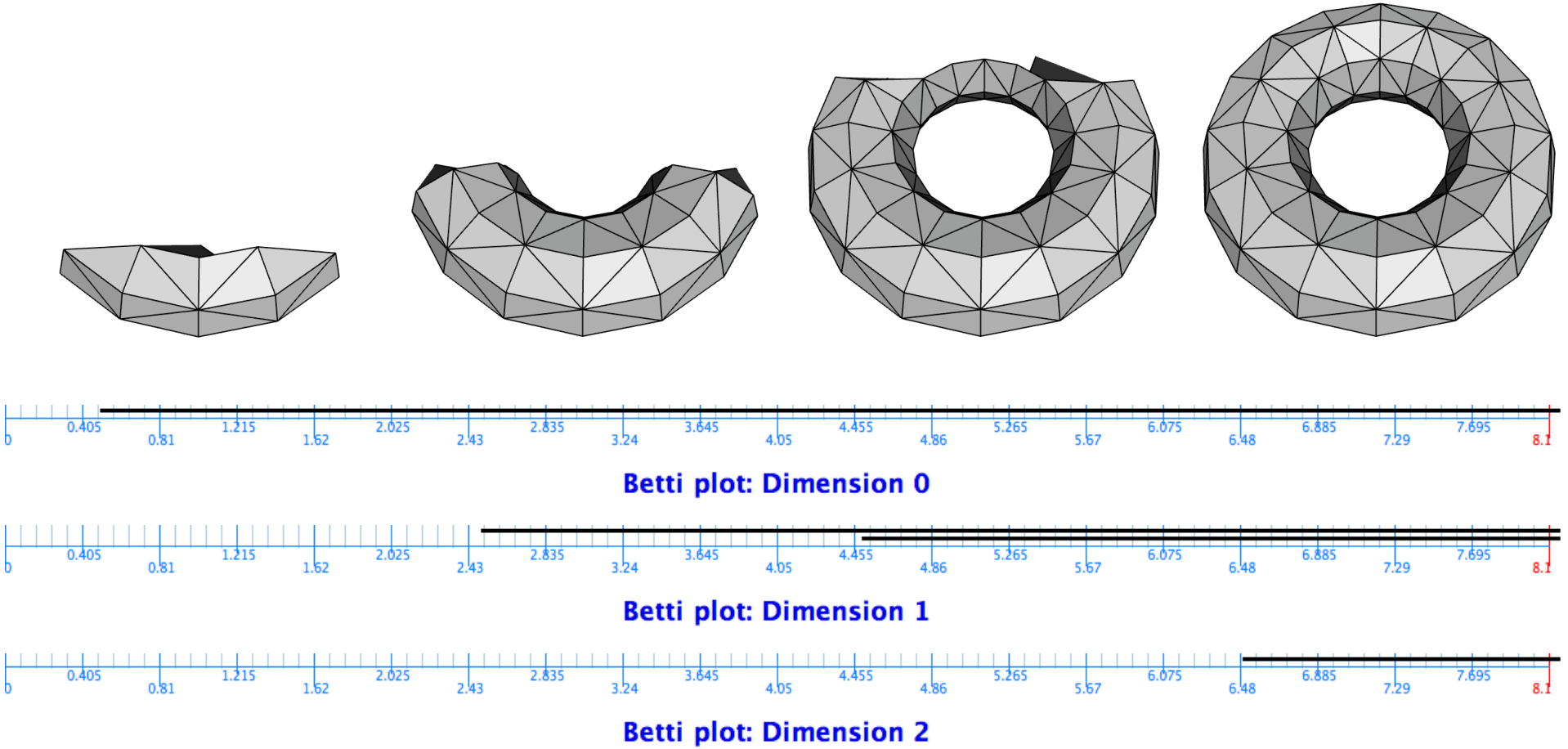


One connected component
Three 1-dimensional holes
No 2-dimensional holes



One connected component
Two 1-dimensional holes
One 2-dimensional hole

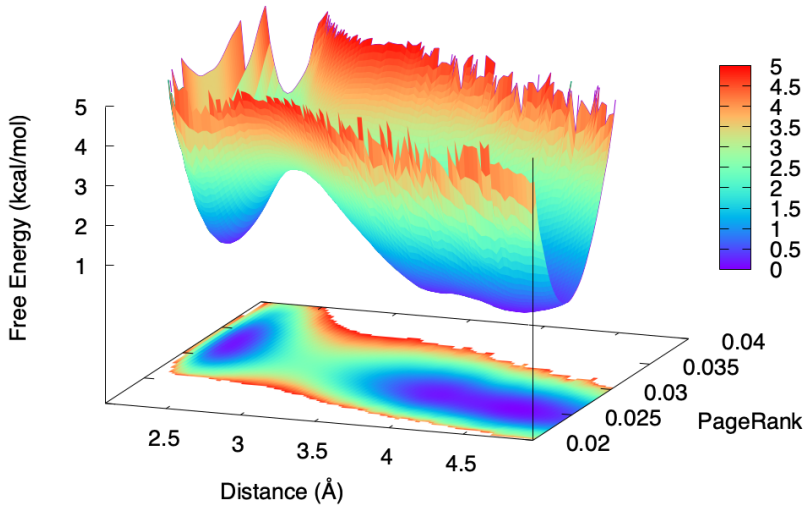
# Sublevelset persistent homology

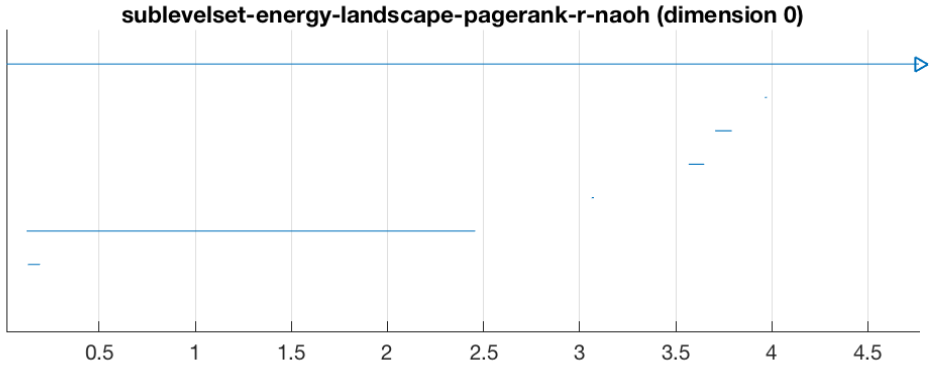


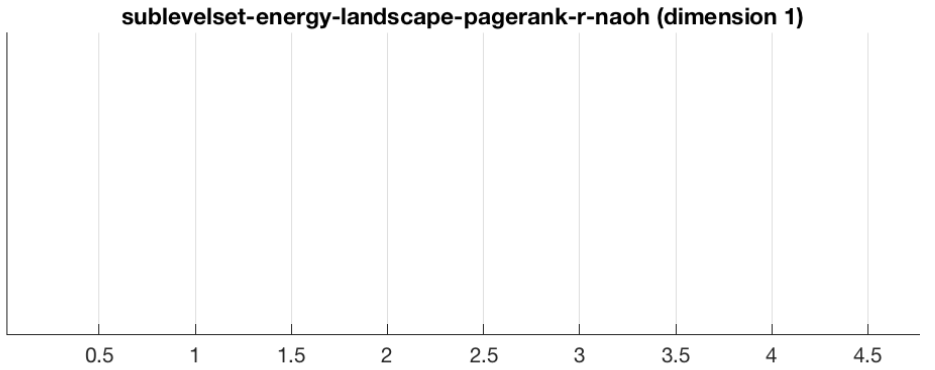

Input: Real-valued function on a space. Output: barcode.

# Sublevelset persistent homology

EL (PageRank) of ion-pair formation of $Na^+$ and $OH^-$: superlevel sets



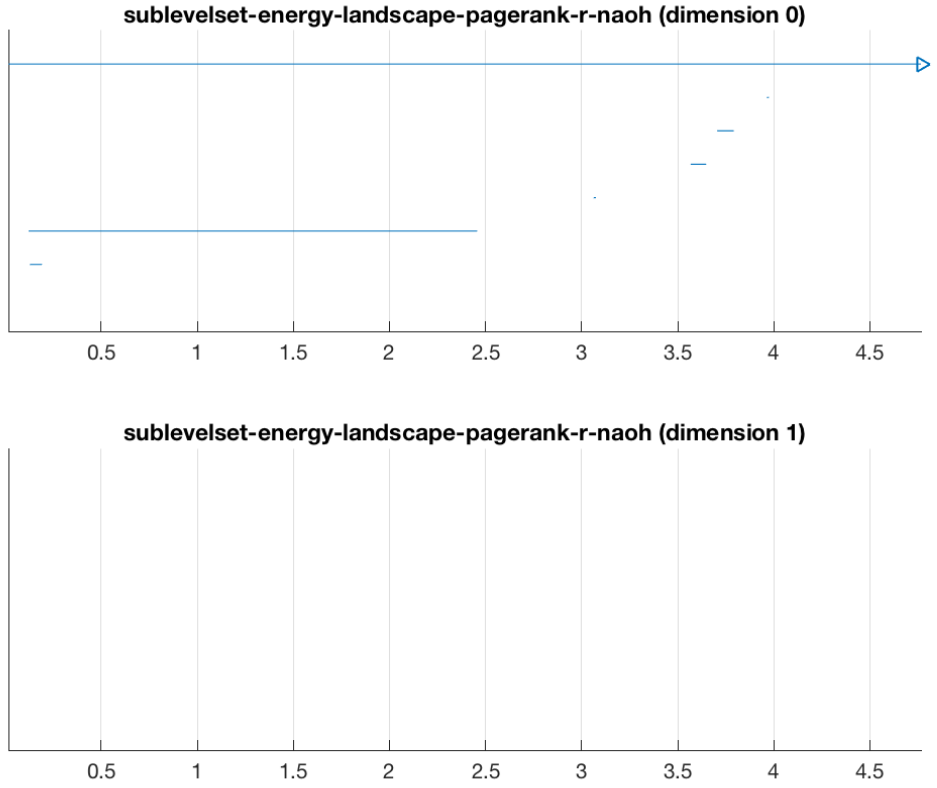




Input: Real-valued function on a space. Output: barcode.

# Sublevelset persistent homology

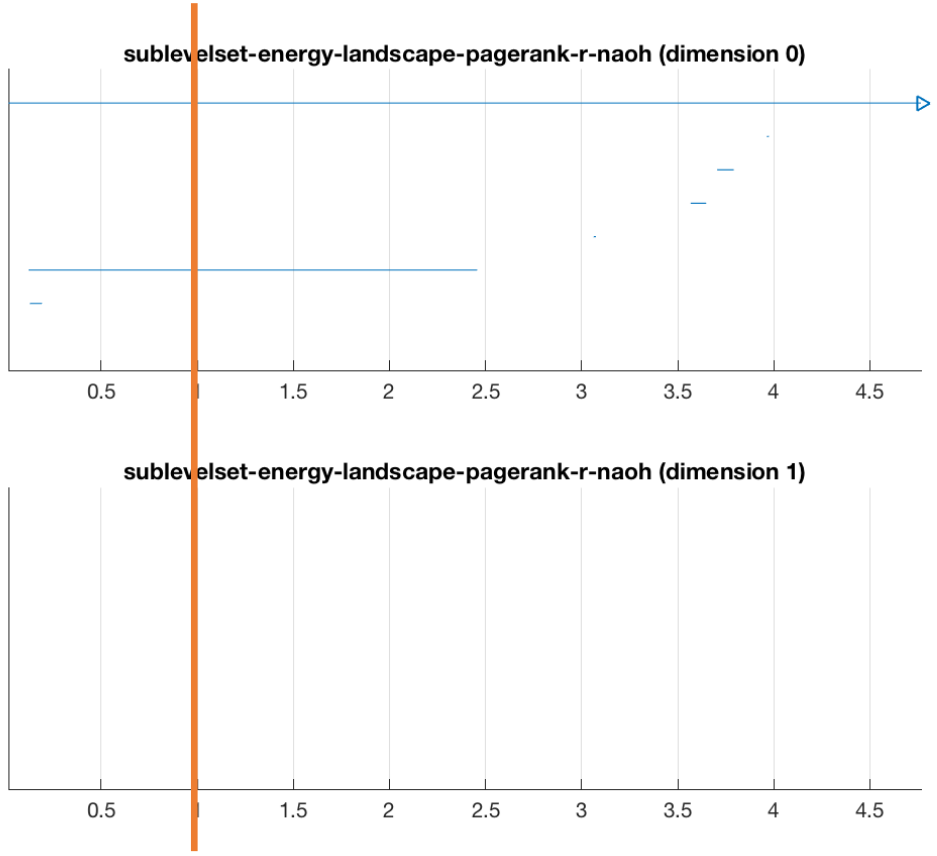
EL (PageRank) of ion-pair formation of $Na^+$ and $OH^-$: sublevel sets




Input: Real-valued function on a space. Output: barcode.

# Sublevelset persistent homology

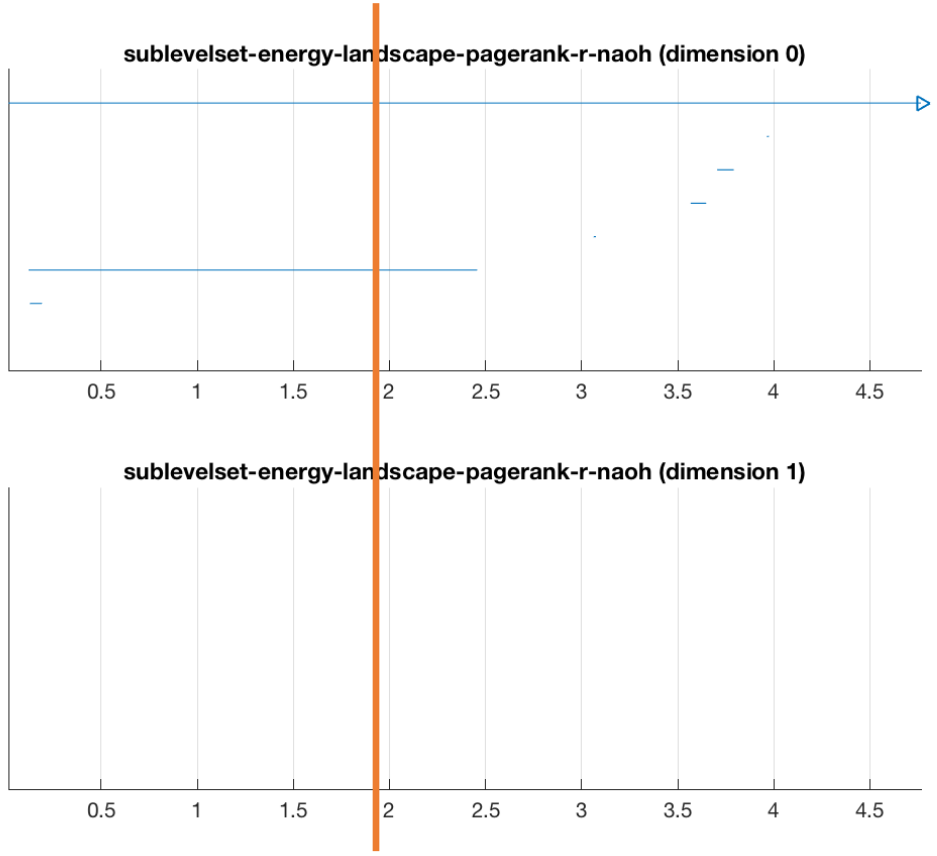
EL (PageRank) of ion-pair formation of $Na^+$ and $OH^-$: sublevel sets




Input: Real-valued function on a space. Output: barcode.

# Sublevelset persistent homology

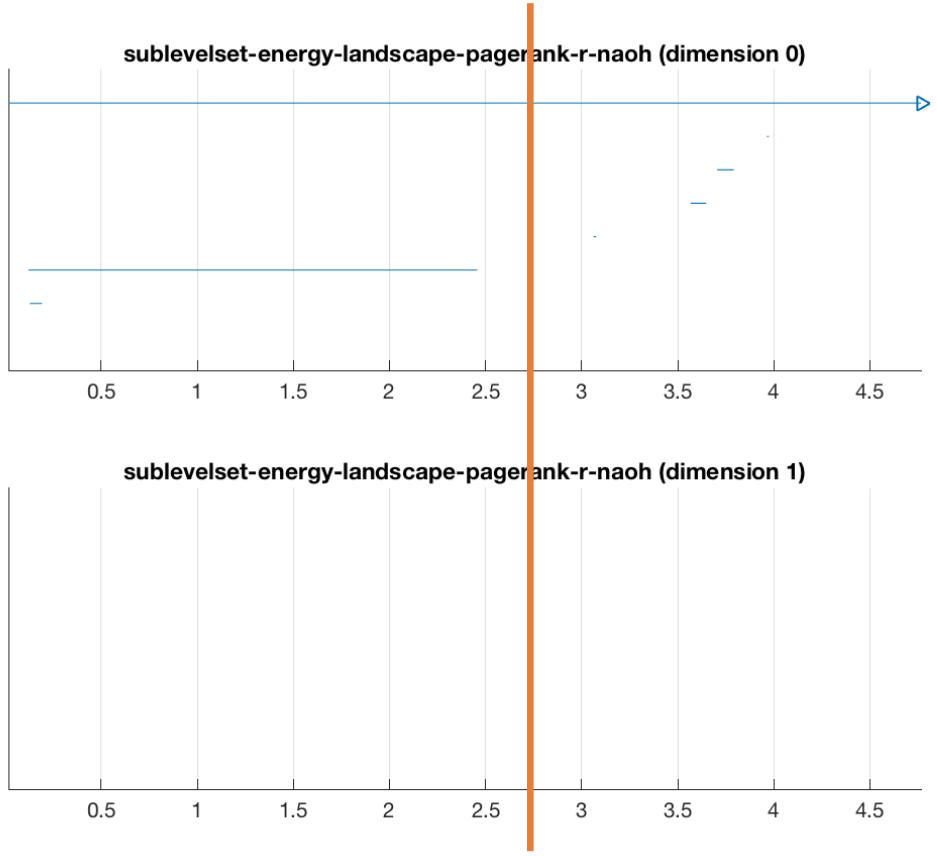
EL (PageRank) of ion-pair formation of $Na^+$ and $OH^-$: sublevel sets




Input: Real-valued function on a space. Output: barcode.

# Sublevelset persistent homology

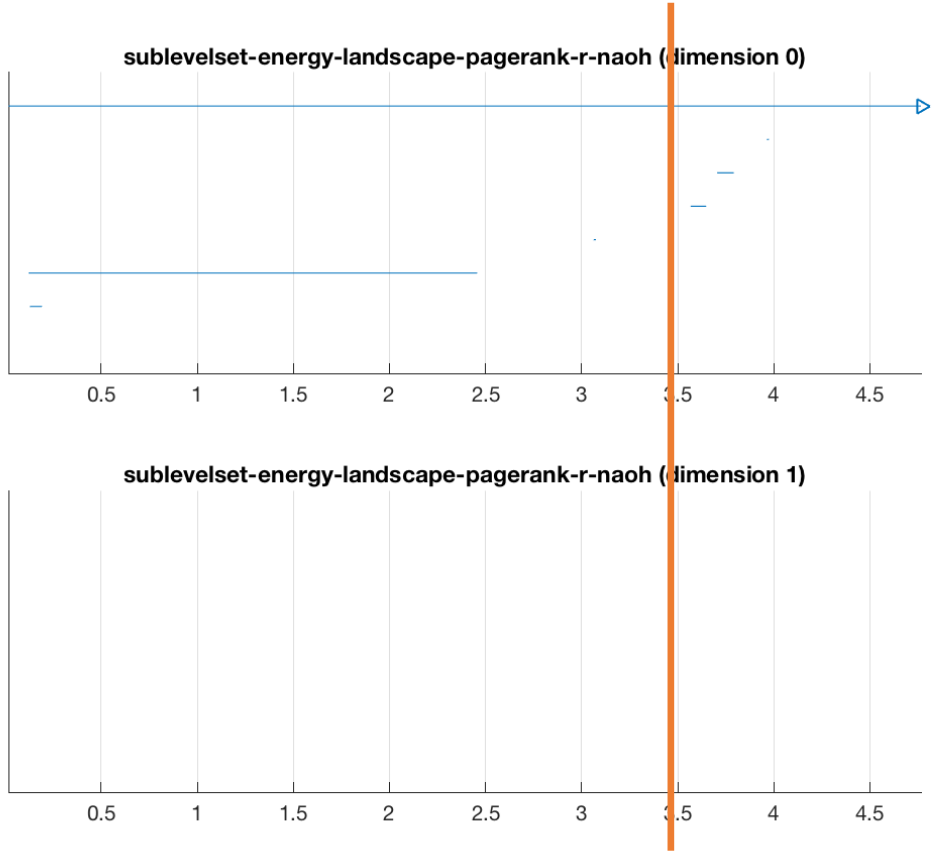
EL (PageRank) of ion-pair formation of $Na^+$ and $OH^-$: sublevel sets




Input: Real-valued function on a space. Output: barcode.

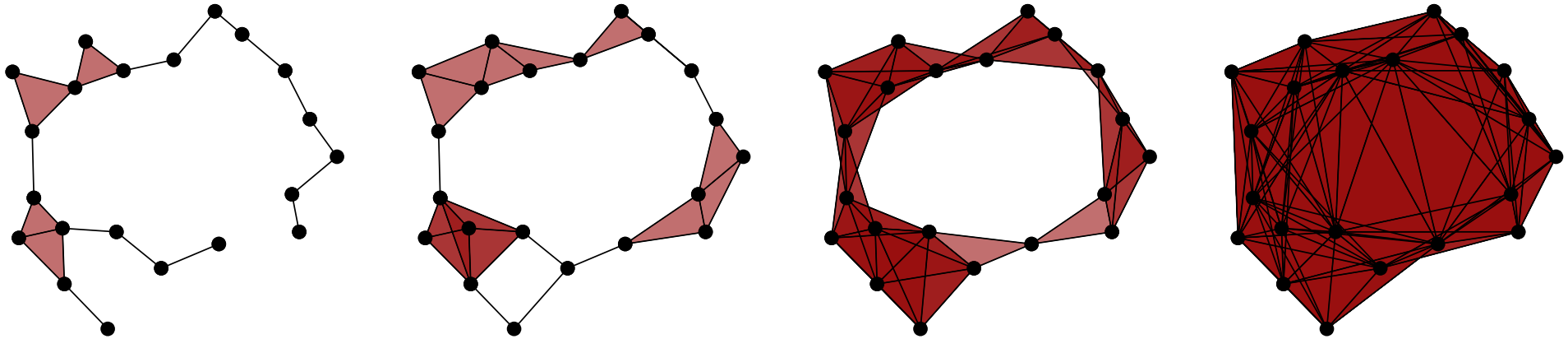
# Sublevelset persistent homology

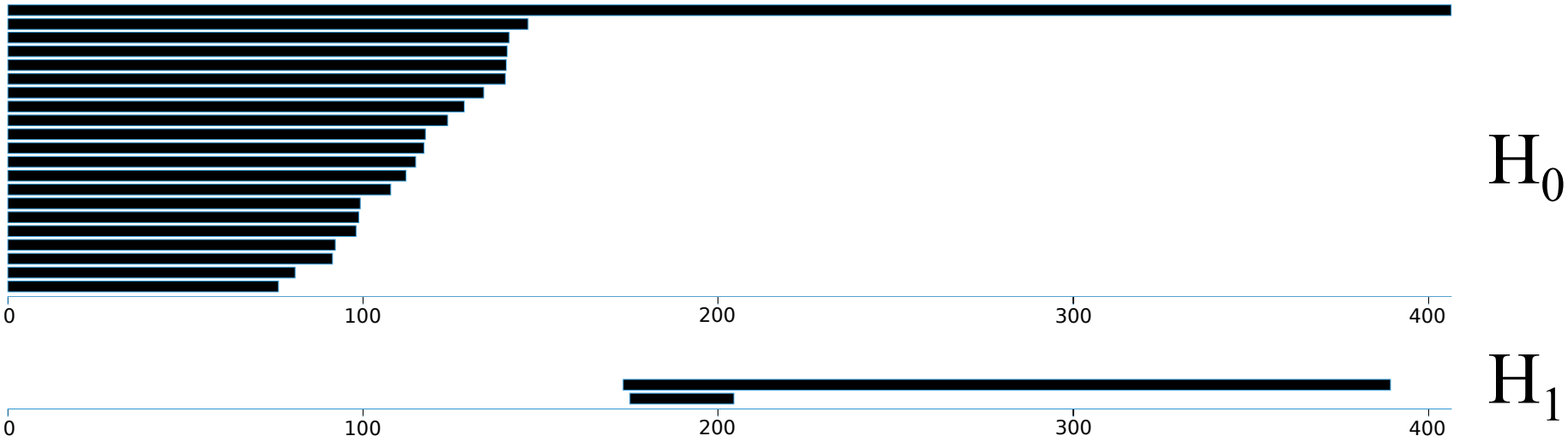
EL (PageRank) of ion-pair formation of $Na^+$ and $OH^-$: sublevel sets




Input: Real-valued function on a space. Output: barcode.

# Persistent homology

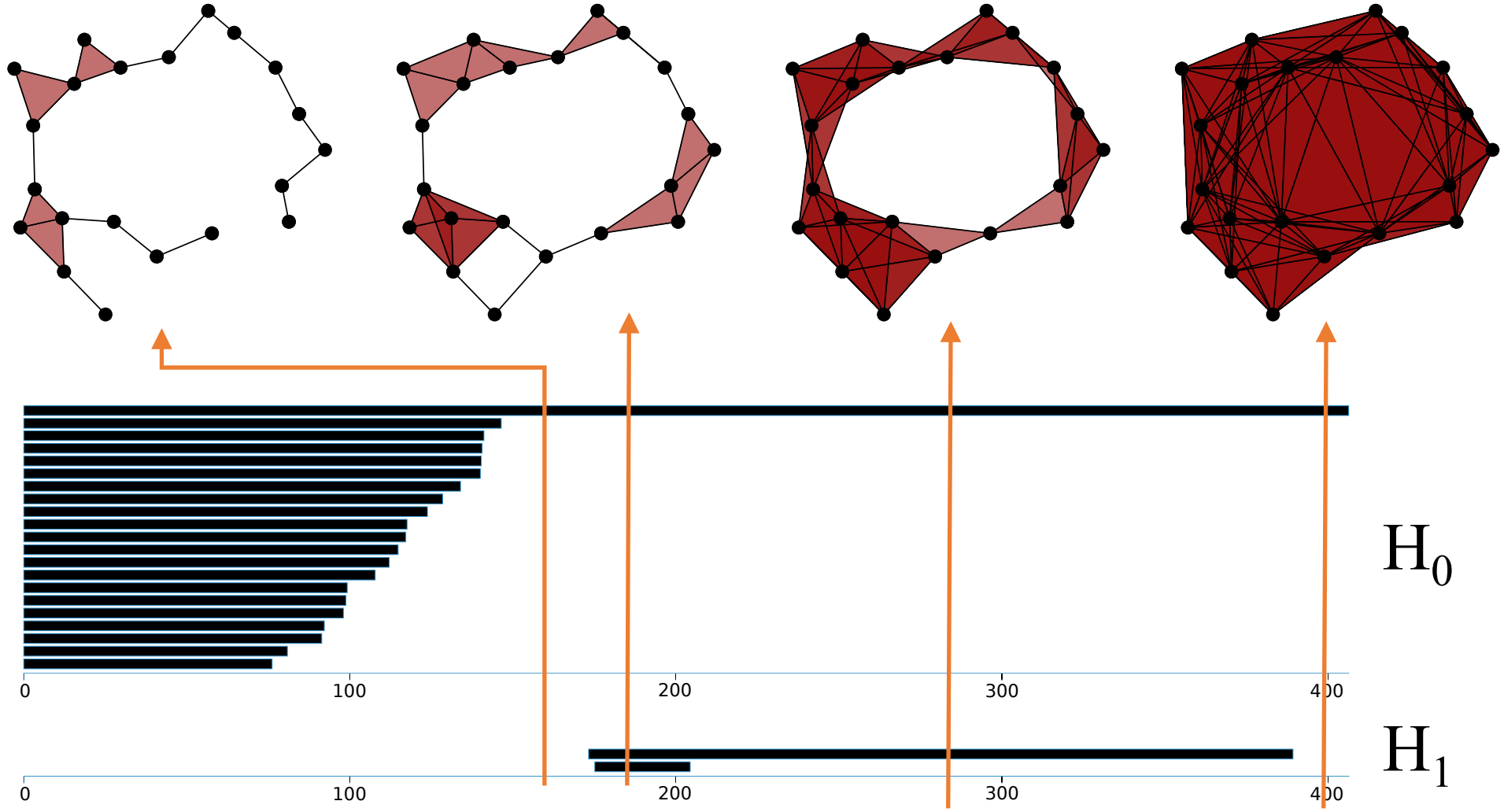





Input: Increasing spaces. Output: barcode.

Significant features persist.

# Persistent homology



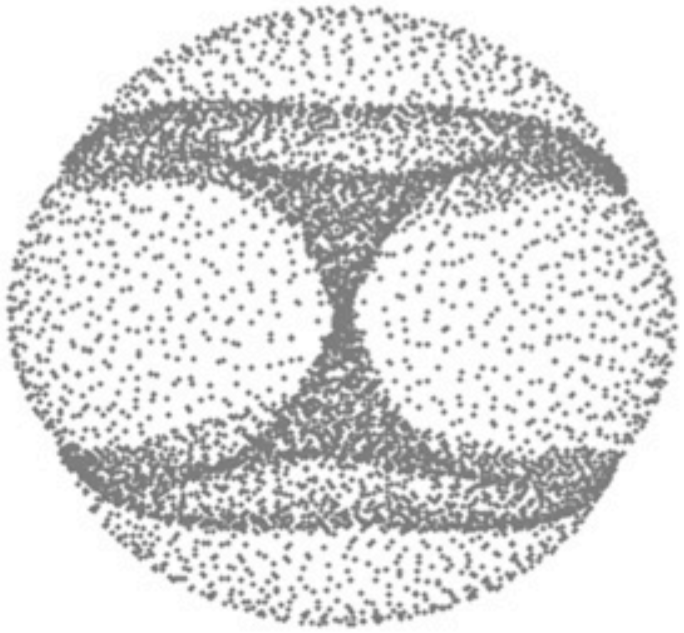

Input: Increasing spaces. Output: barcode.

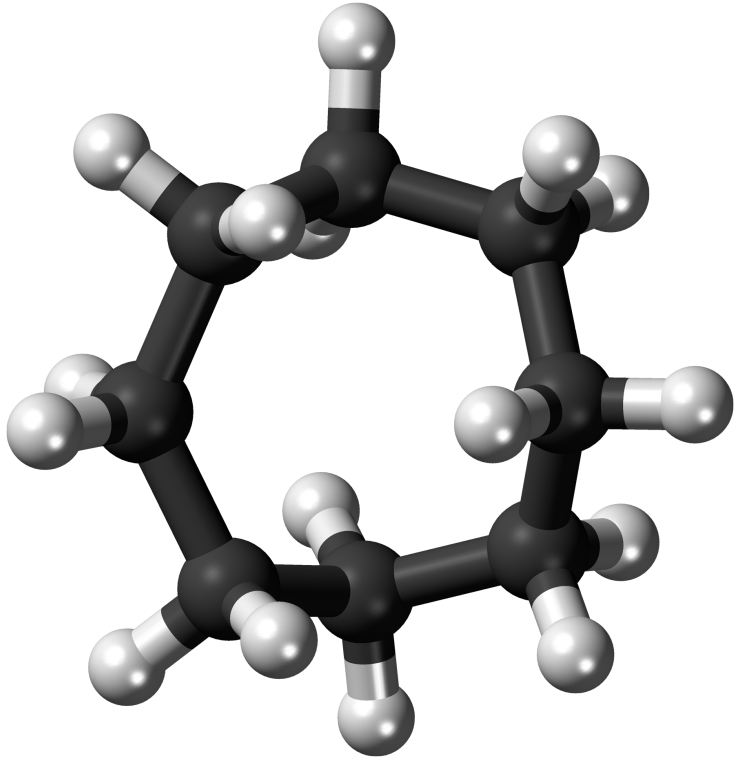
Significant features persist.

# Persistent homology applied to data

Example: Cyclo-Octane ($C_8H_{16}$) data

6040 points in 24-dimensional space



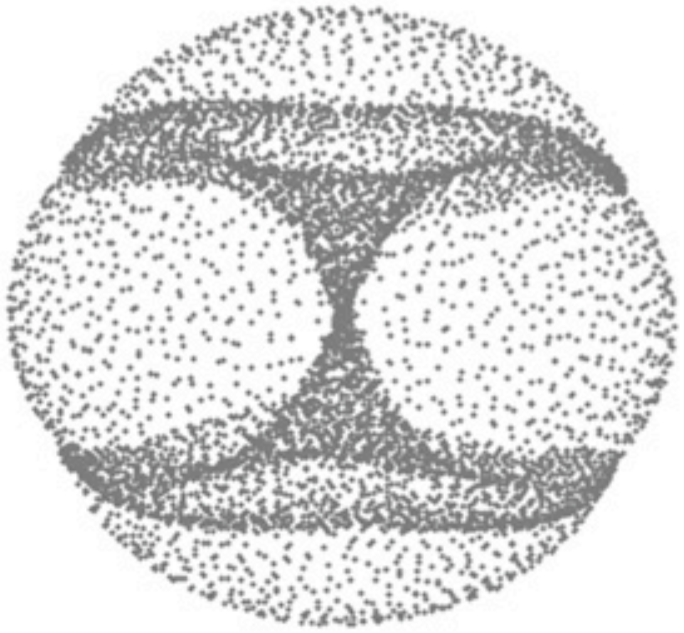


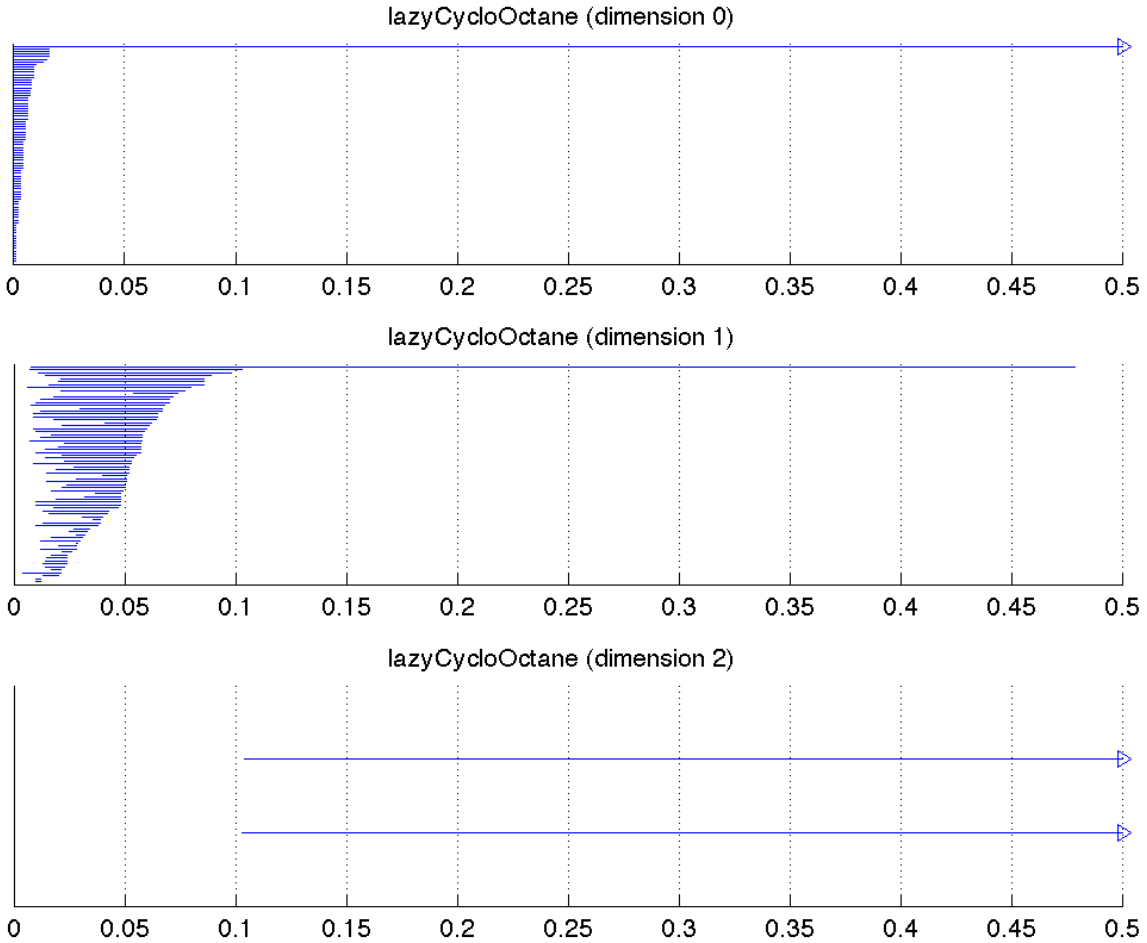
*Non-Manifold Surface Reconstruction from High Dimensional Point Cloud Data*
by Shawn Martin and Jean-Paul Watson, 2010.

# Persistent homology applied to data

## Example: Cyclo-Octane ($C_8H_{16}$) data

## 6040 points in 24-dimensional space



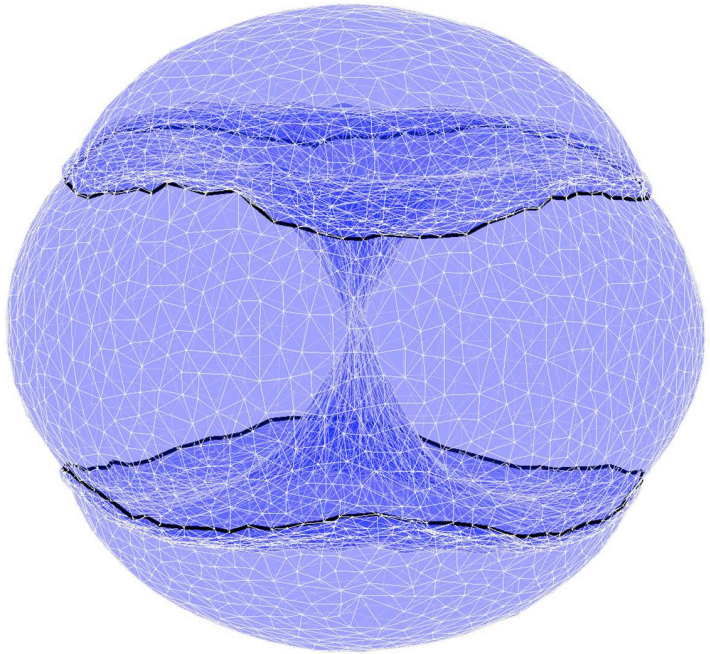


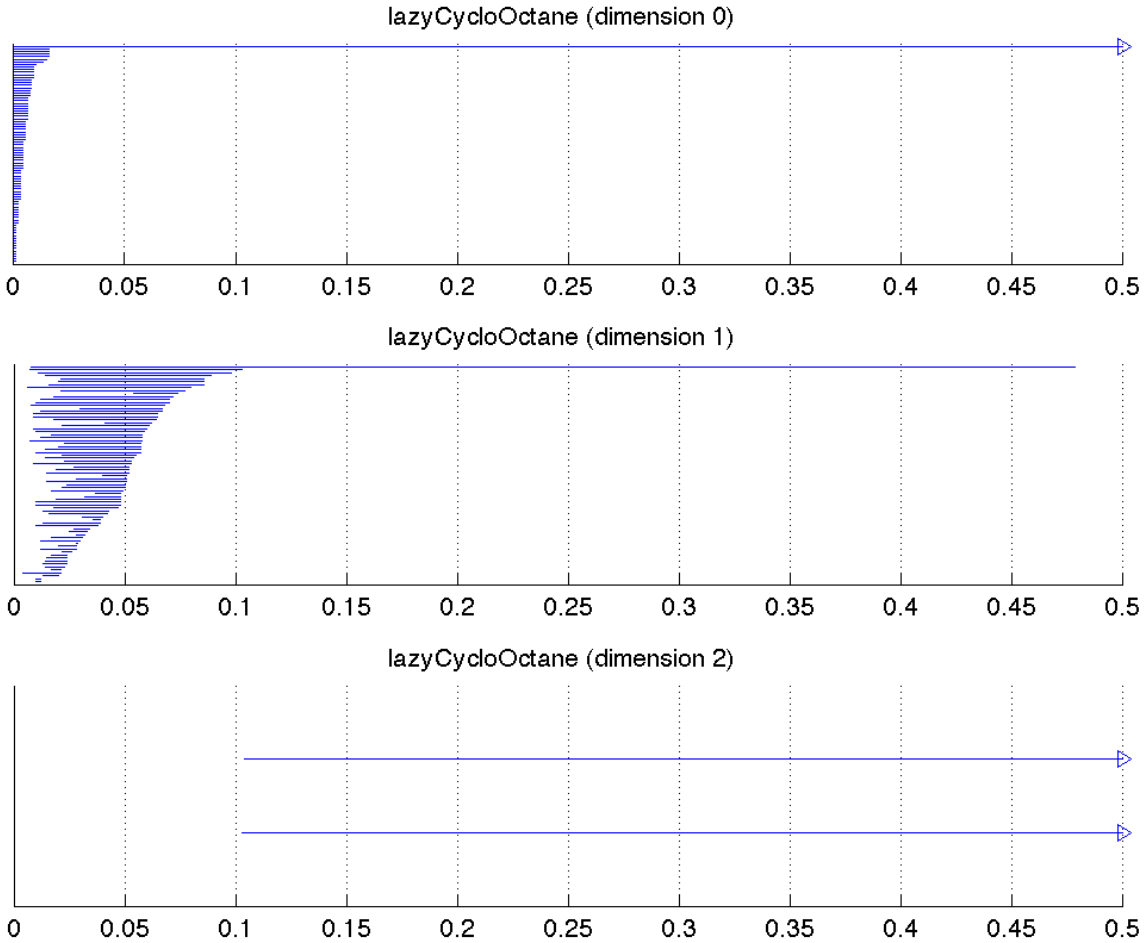

*Non-Manifold Surface Reconstruction from High Dimensional Point Cloud Data*
by Shawn Martin and Jean-Paul Watson, 2010.

# Persistent homology applied to data

Example: Cyclo-Octane ($C_8H_{16}$) data

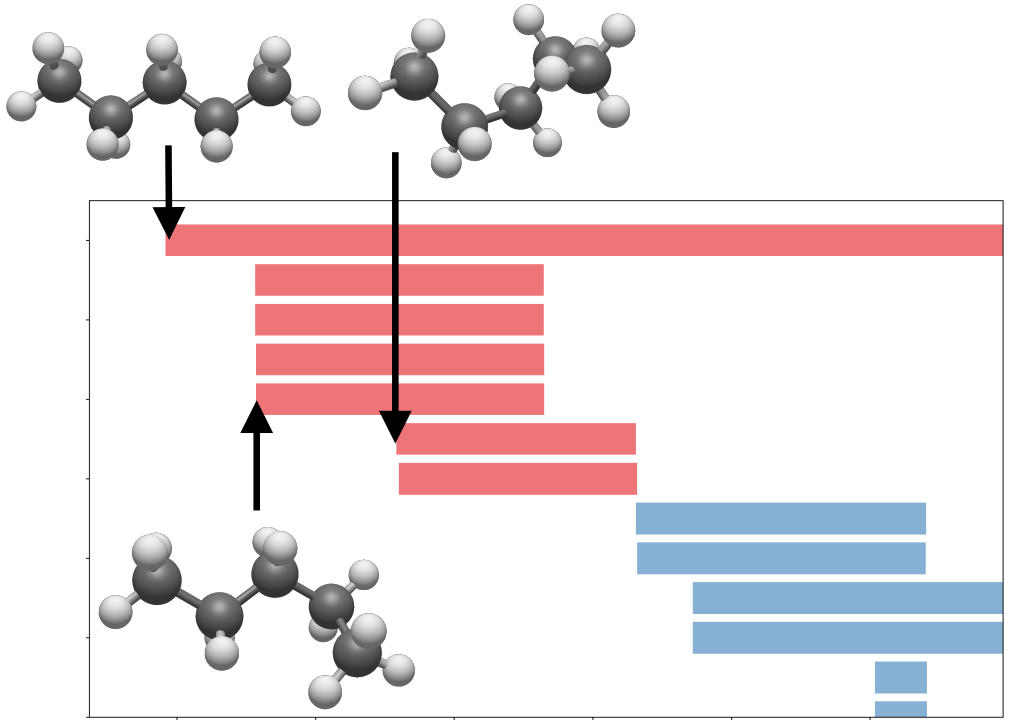
6040 points in 24-dimensional space

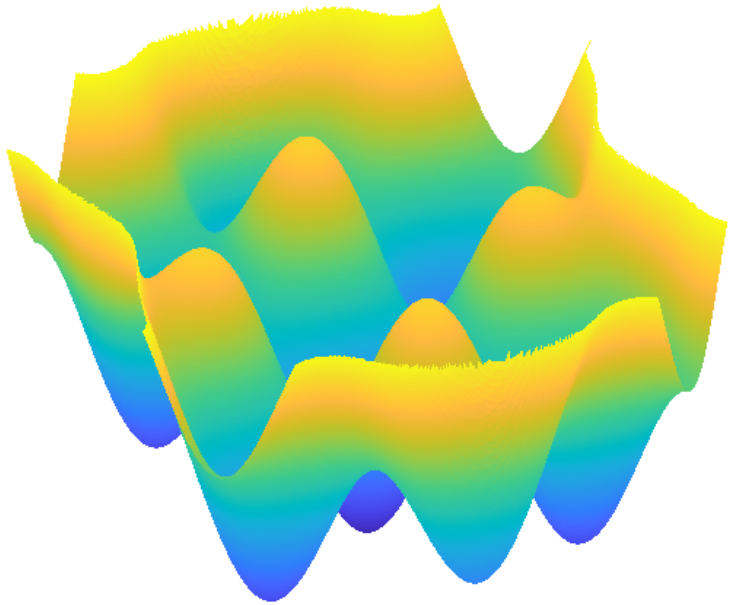





*Non-Manifold Surface Reconstruction from High Dimensional Point Cloud Data*
by Shawn Martin and Jean-Paul Watson, 2010.

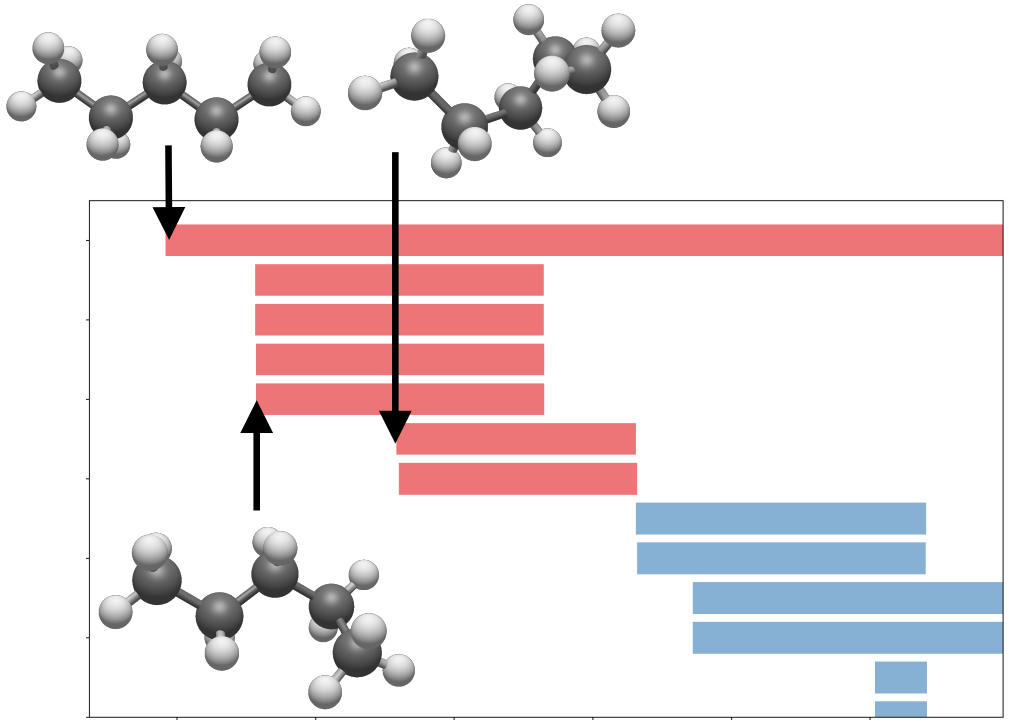
# Sublevelset persistent homology of the *n*-alkanes

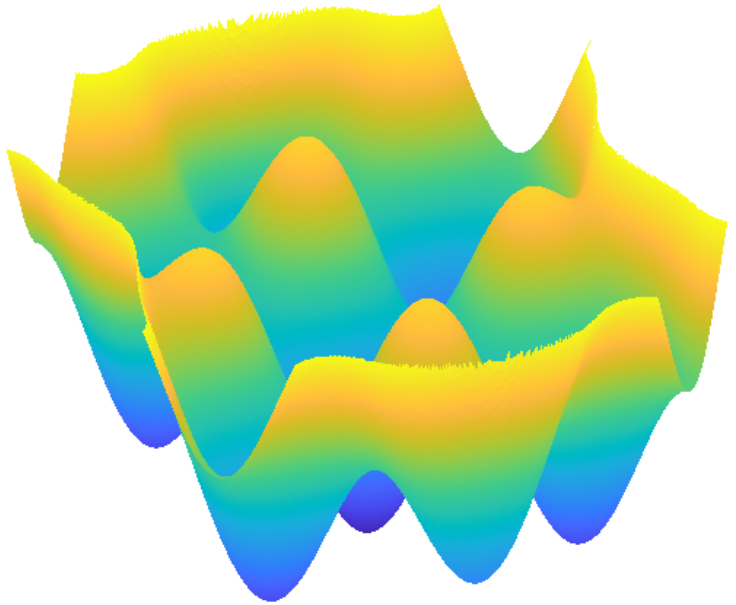


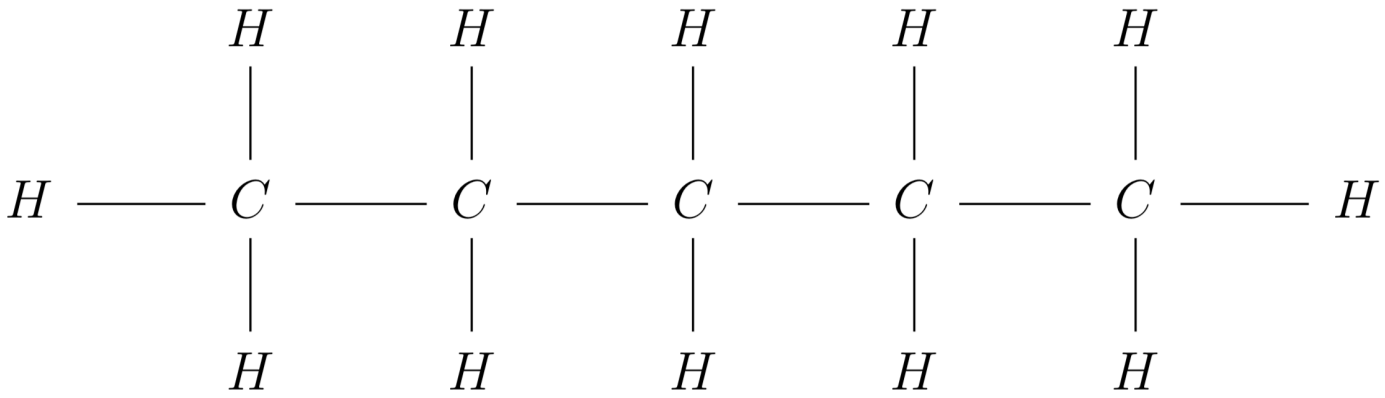


Joint with Johnathan Bush, Mark Heim, Joshua Mirth,
Yang Zang, Yanqin Zhai, and the DELTA team

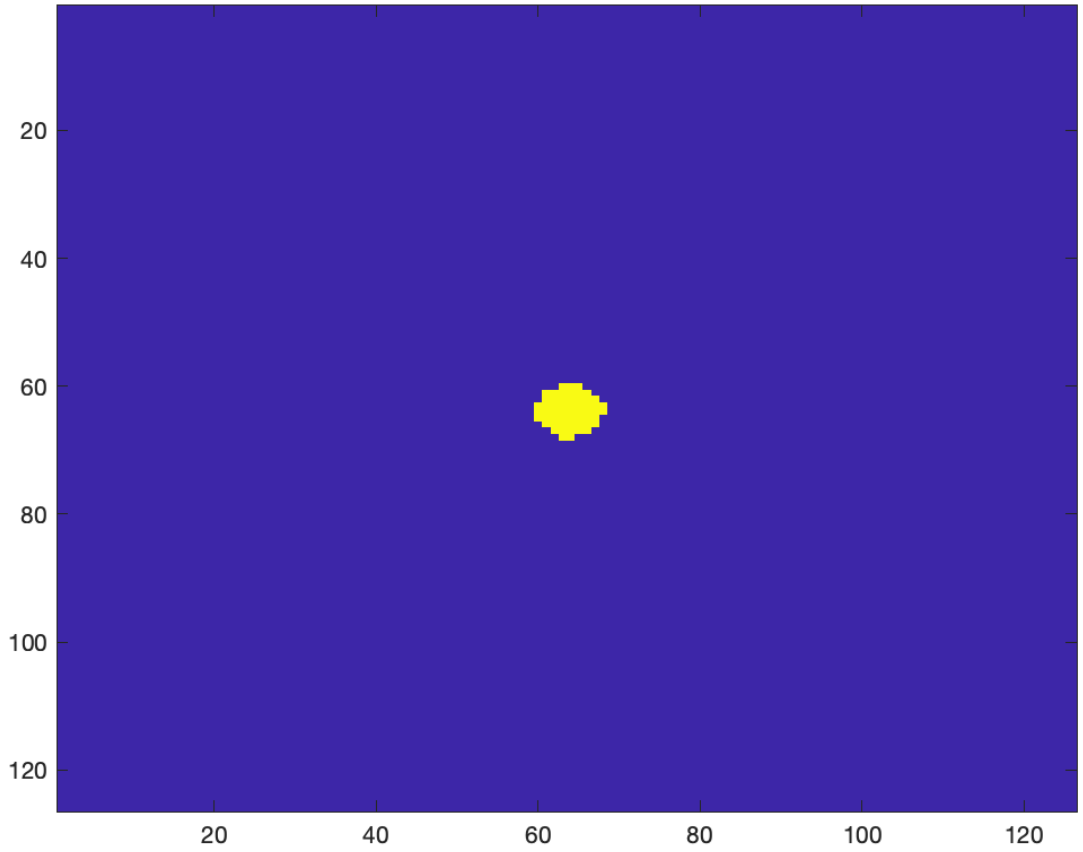
# Sublevelset persistent homology of the *n*-alkanes

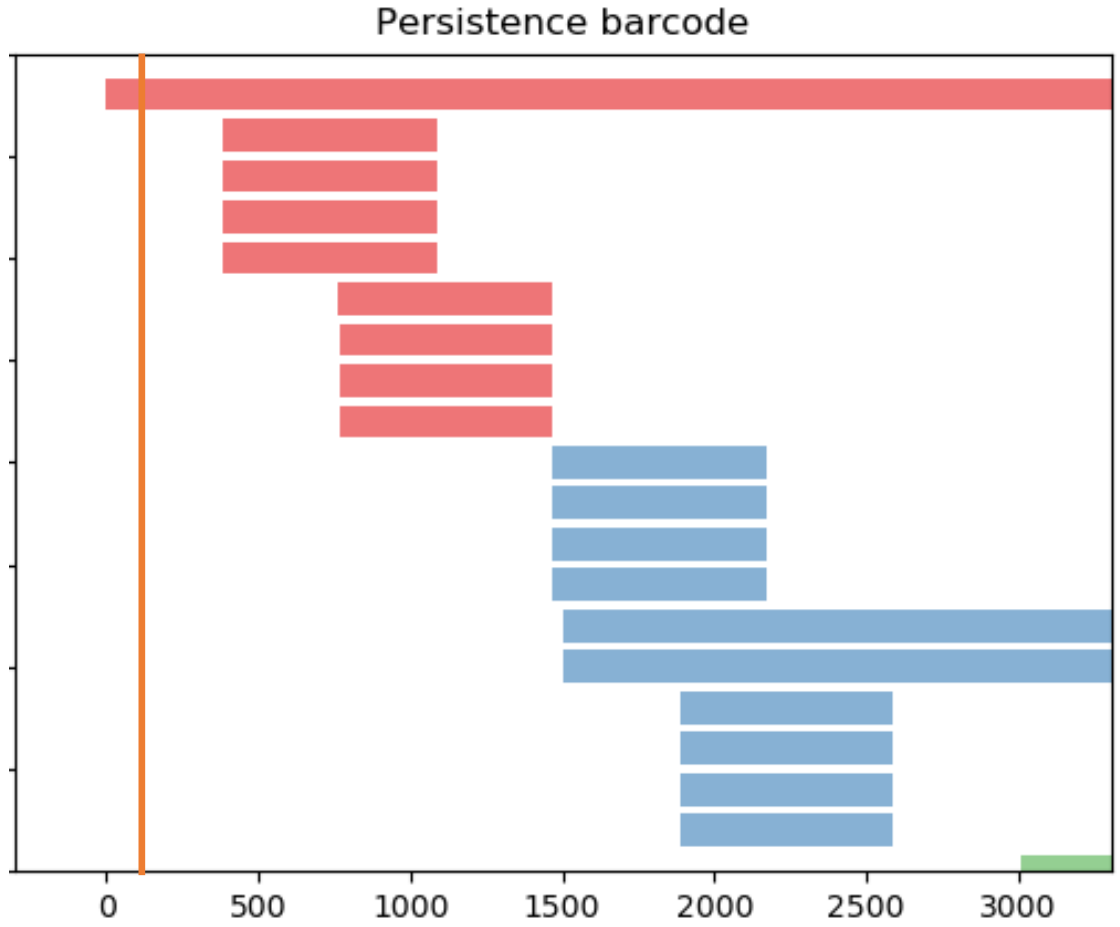






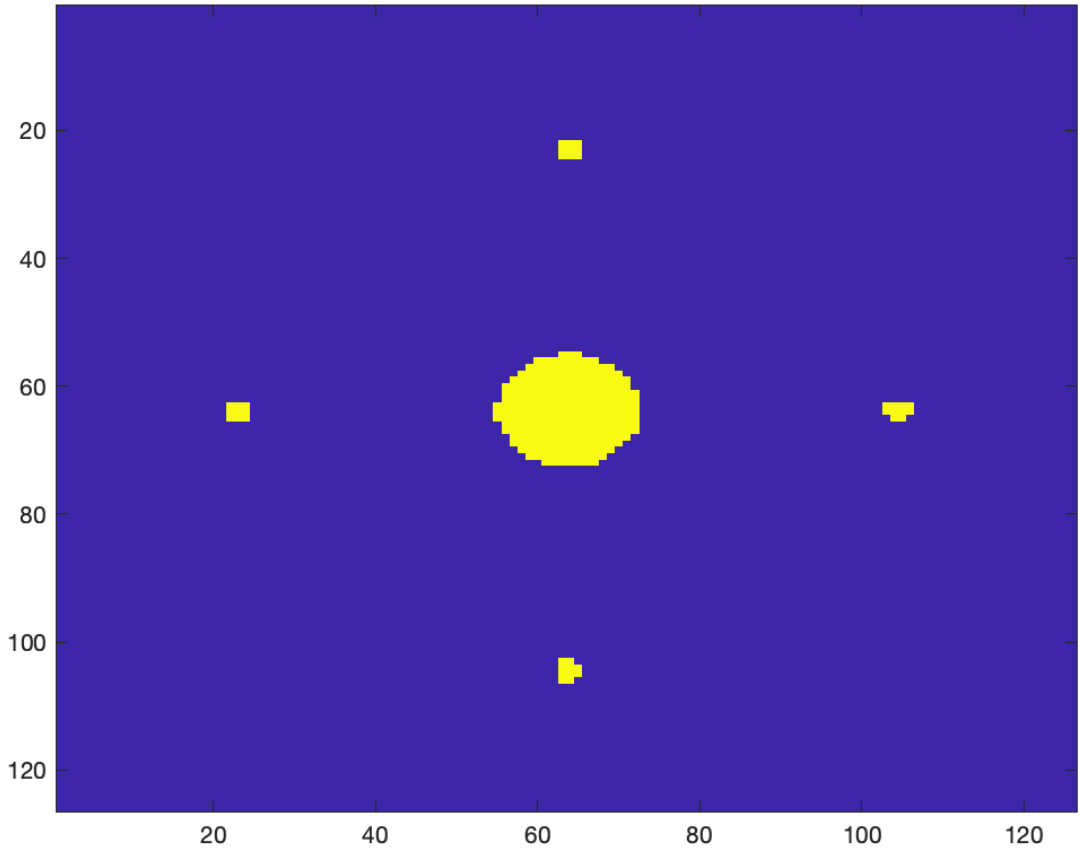
# Sublevelset persistent homology of the *n*-alkanes

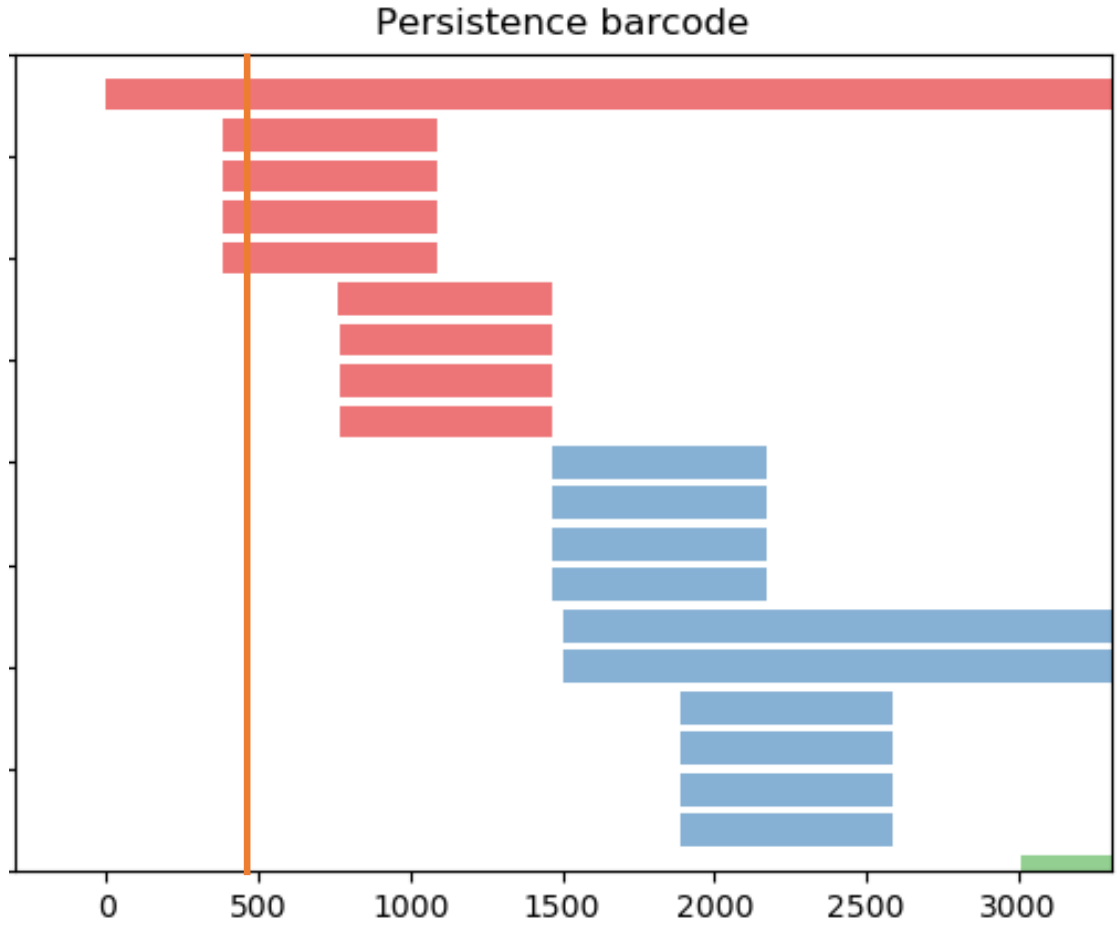



We have a complete characterization of the persistence barcodes of all *n*-alkanes

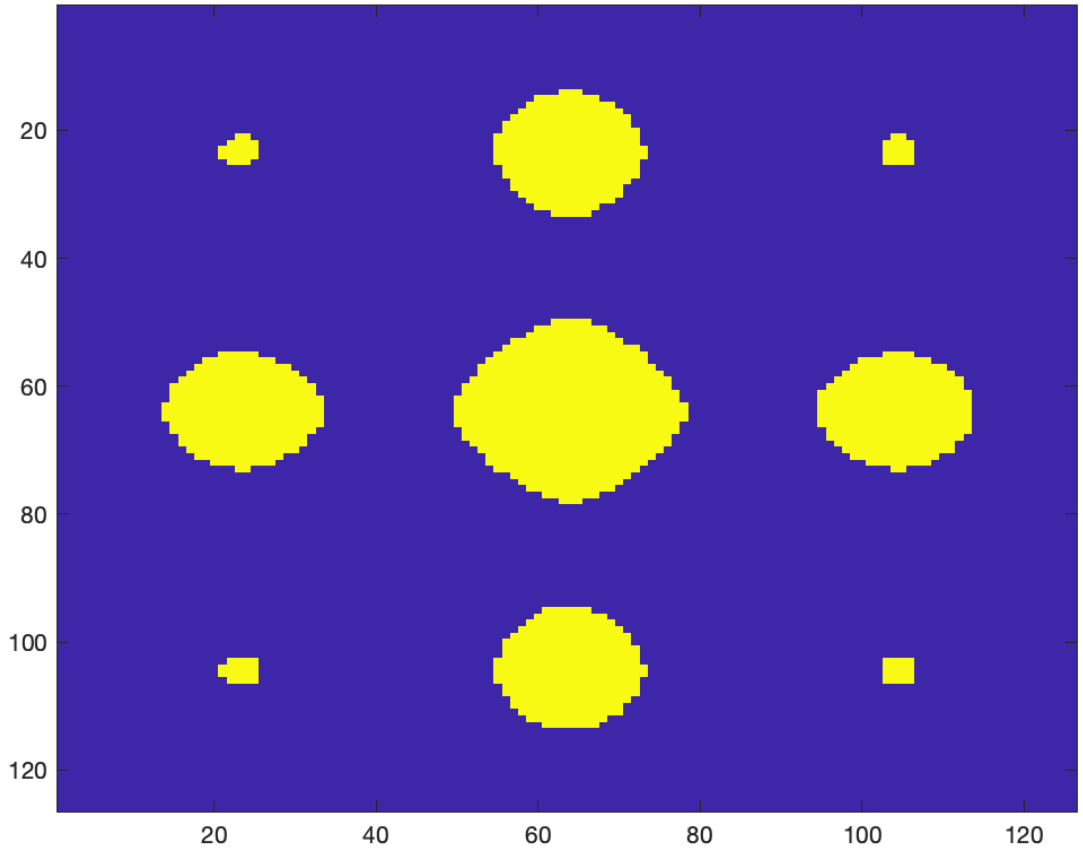
# Sublevelset persistent homology of the *n*-alkanes

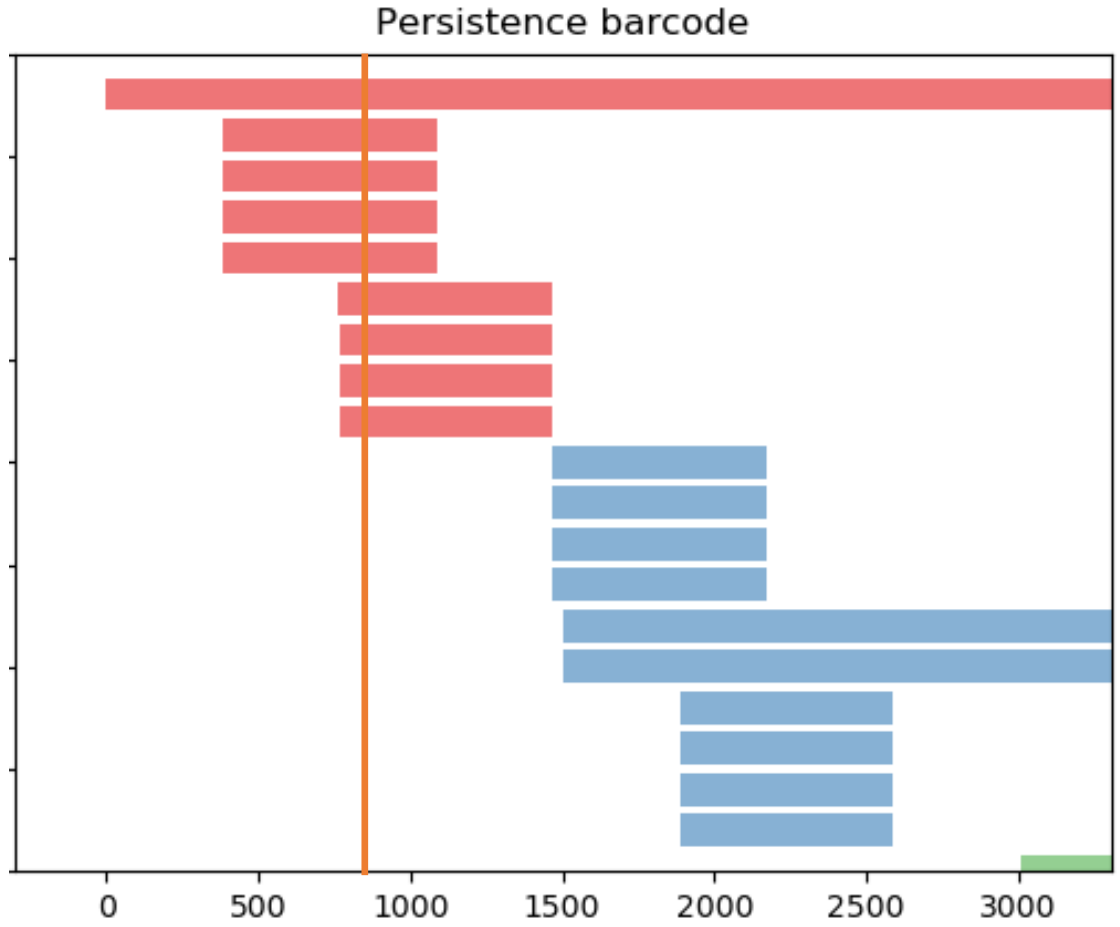



We have a complete characterization of the persistence barcodes of all *n*-alkanes

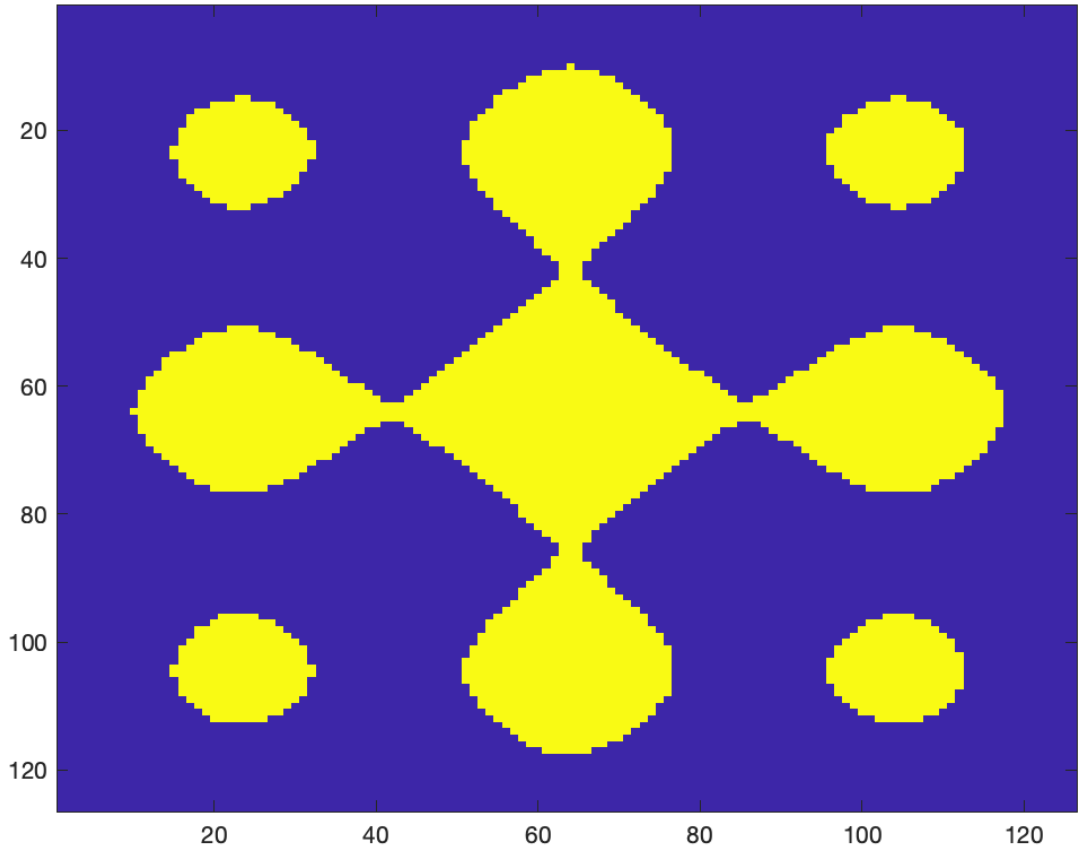
# Sublevelset persistent homology of the *n*-alkanes

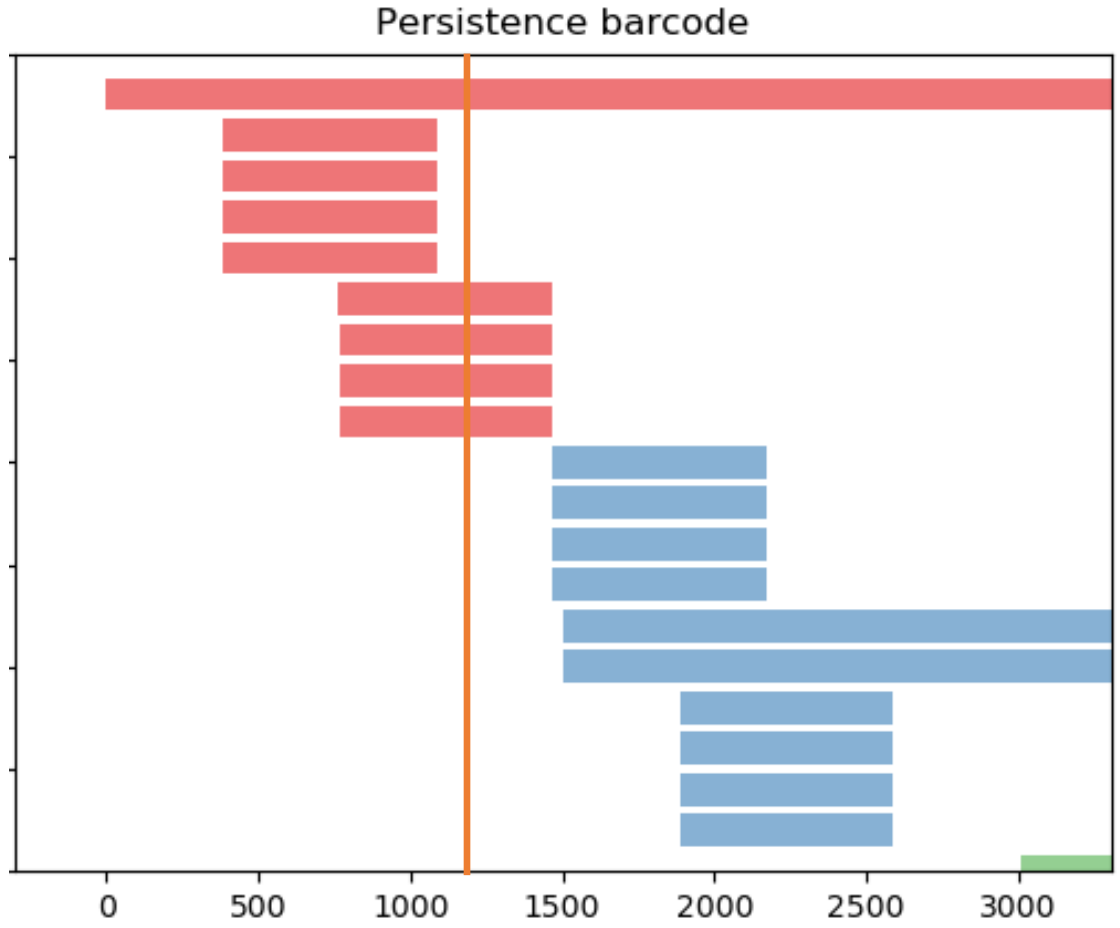



We have a complete characterization of the persistence barcodes of all *n*-alkanes

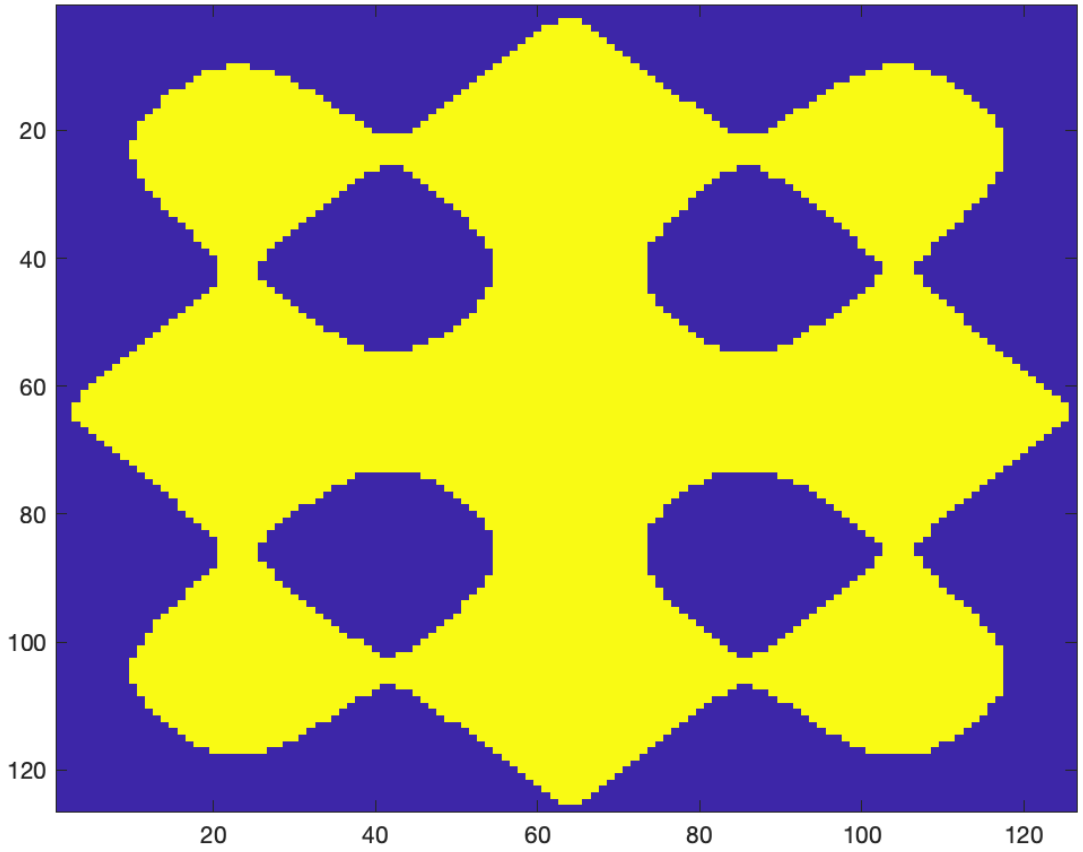
# Sublevelset persistent homology of the *n*-alkanes

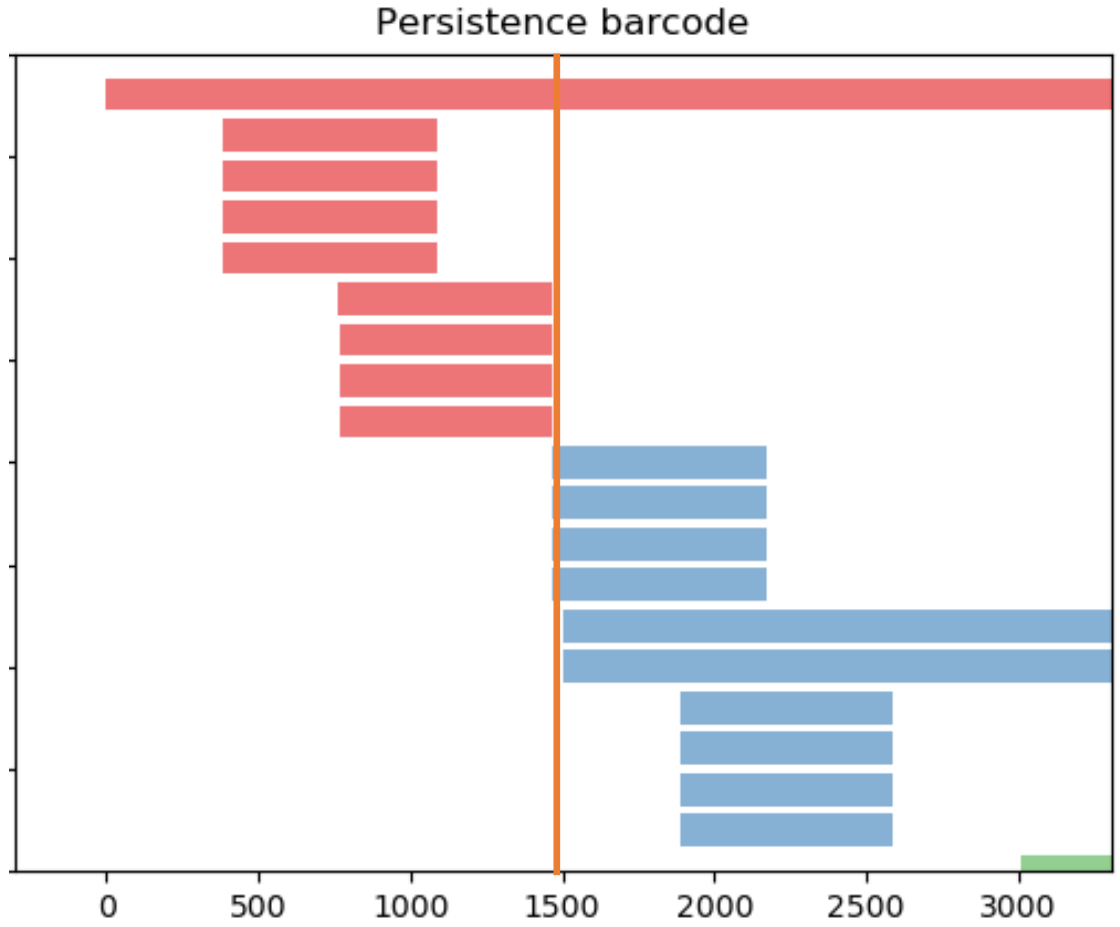



We have a complete characterization of the persistence barcodes of all *n*-alkanes

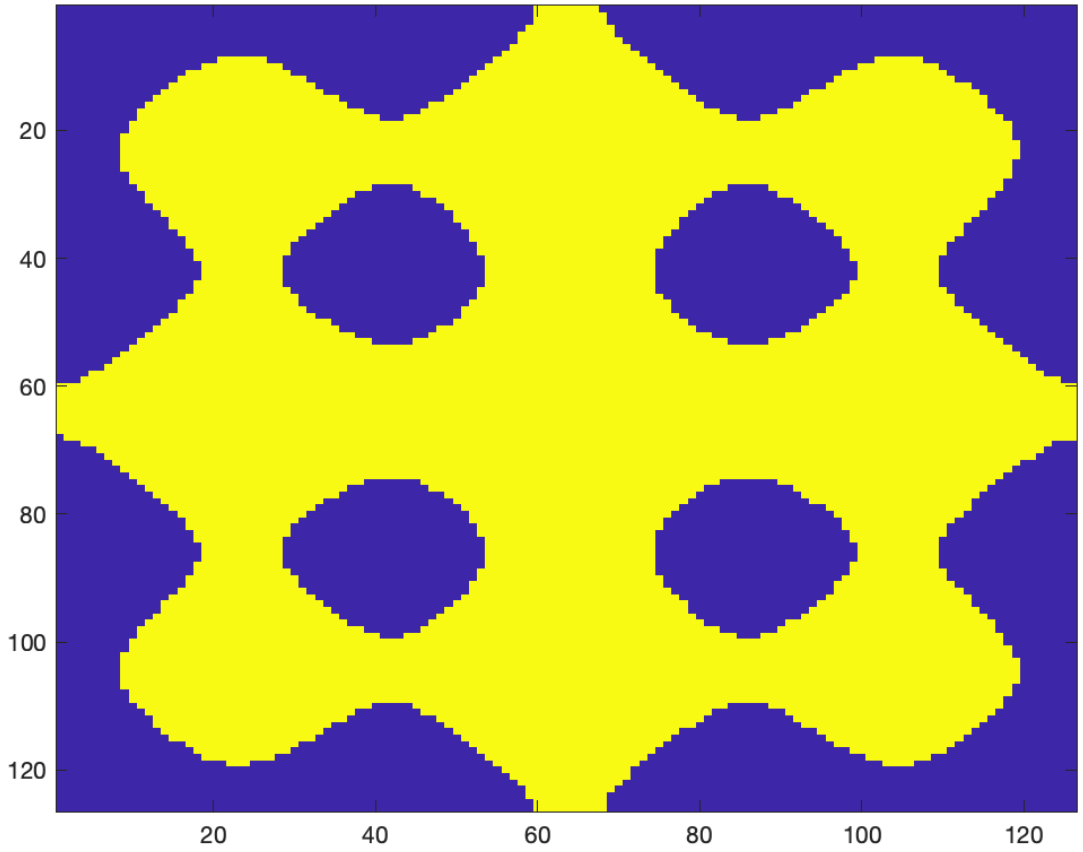
# Sublevelset persistent homology of the *n*-alkanes

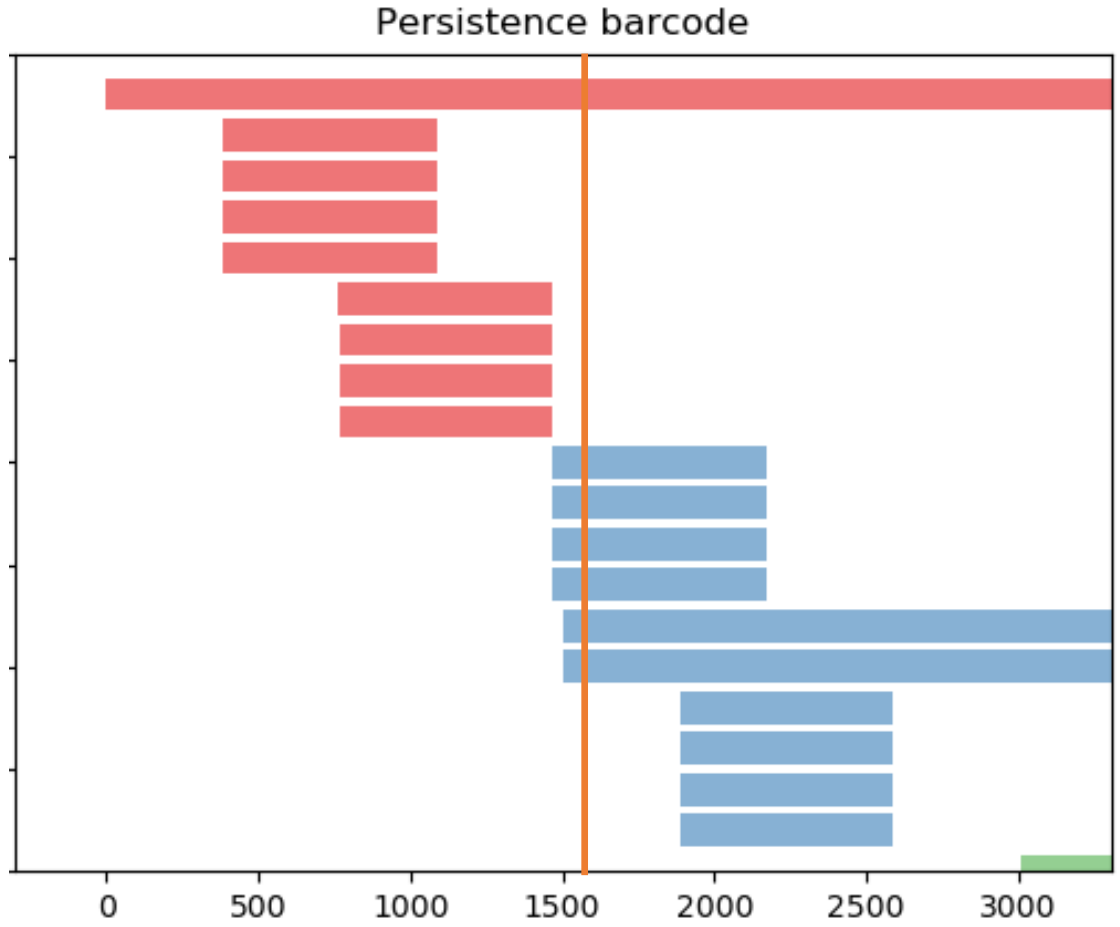



We have a complete characterization of the persistence barcodes of all *n*-alkanes

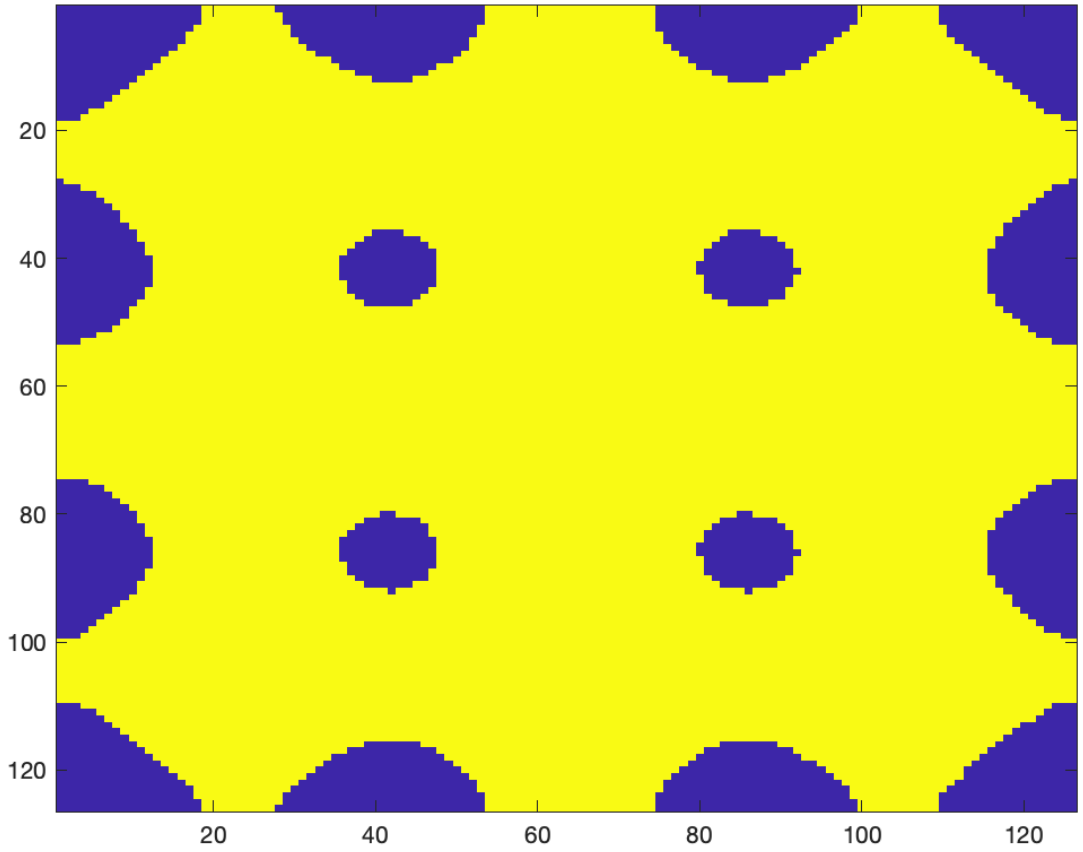
# Sublevelset persistent homology of the *n*-alkanes

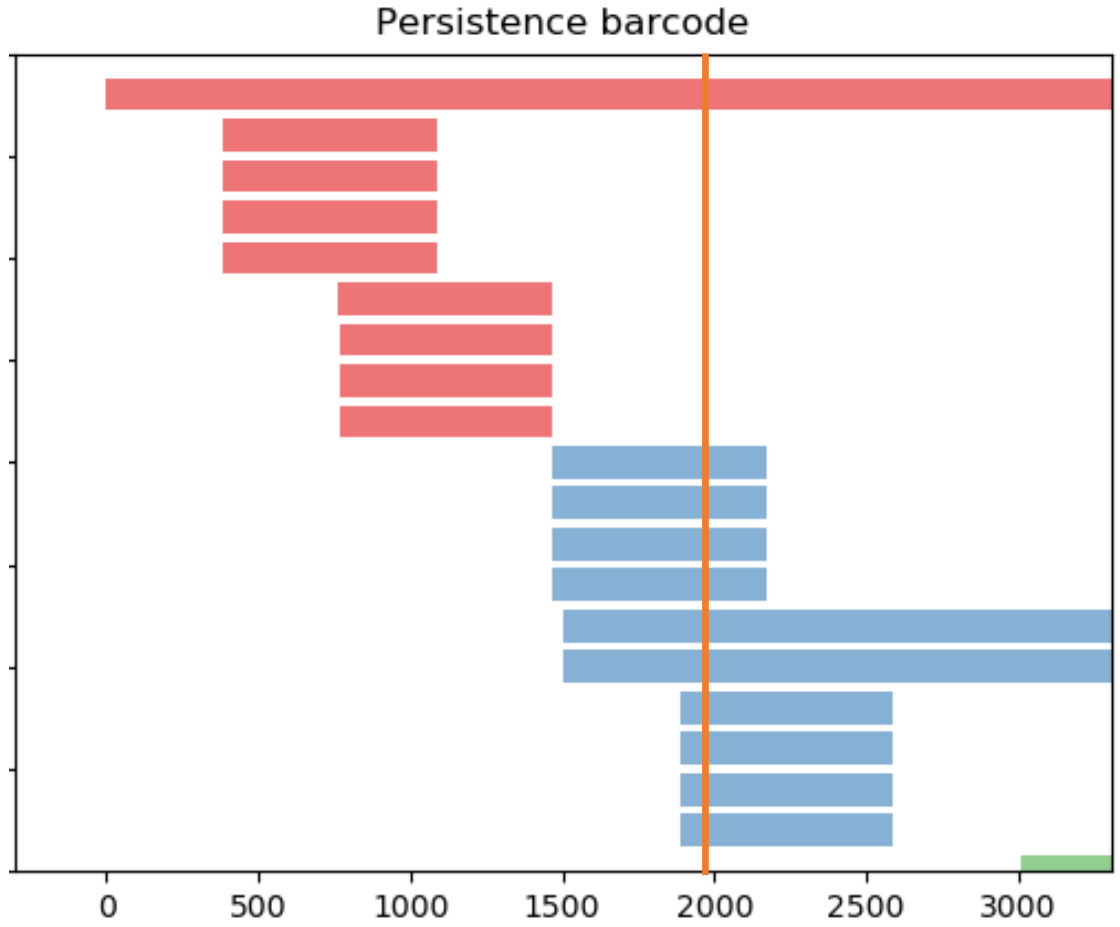



We have a complete characterization of the persistence barcodes of all *n*-alkanes

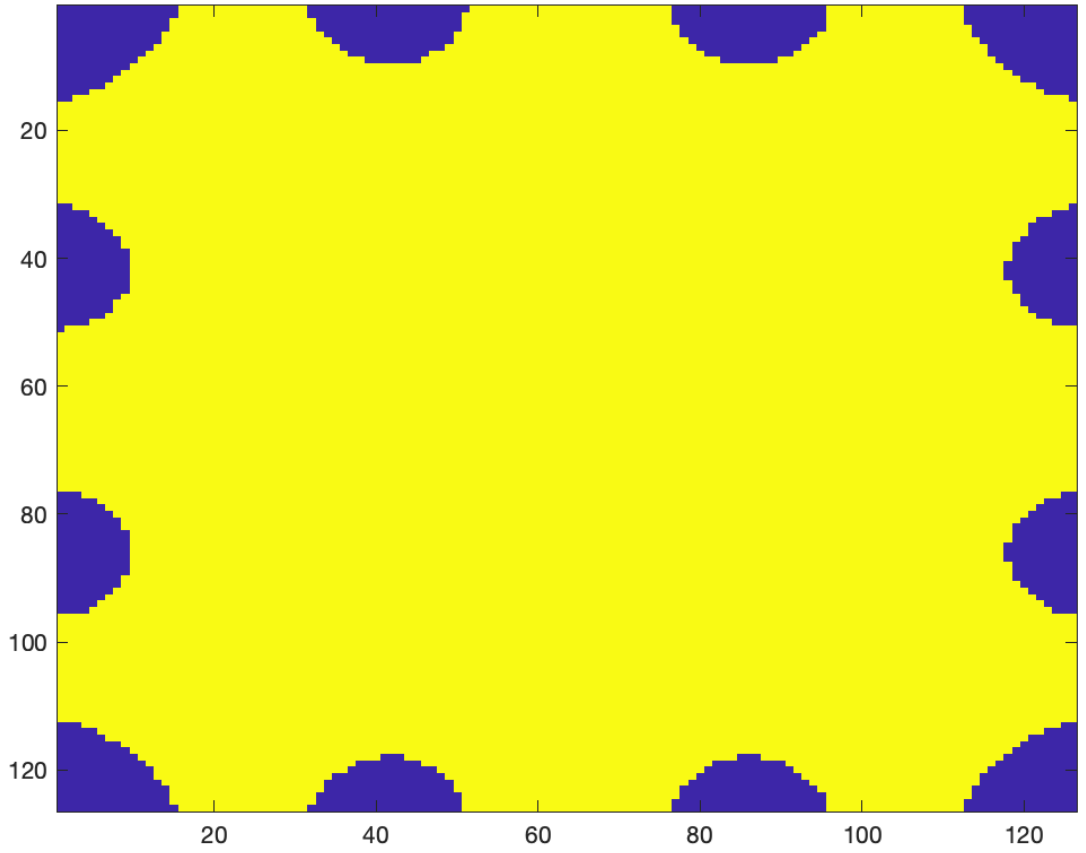
# Sublevelset persistent homology of the *n*-alkanes

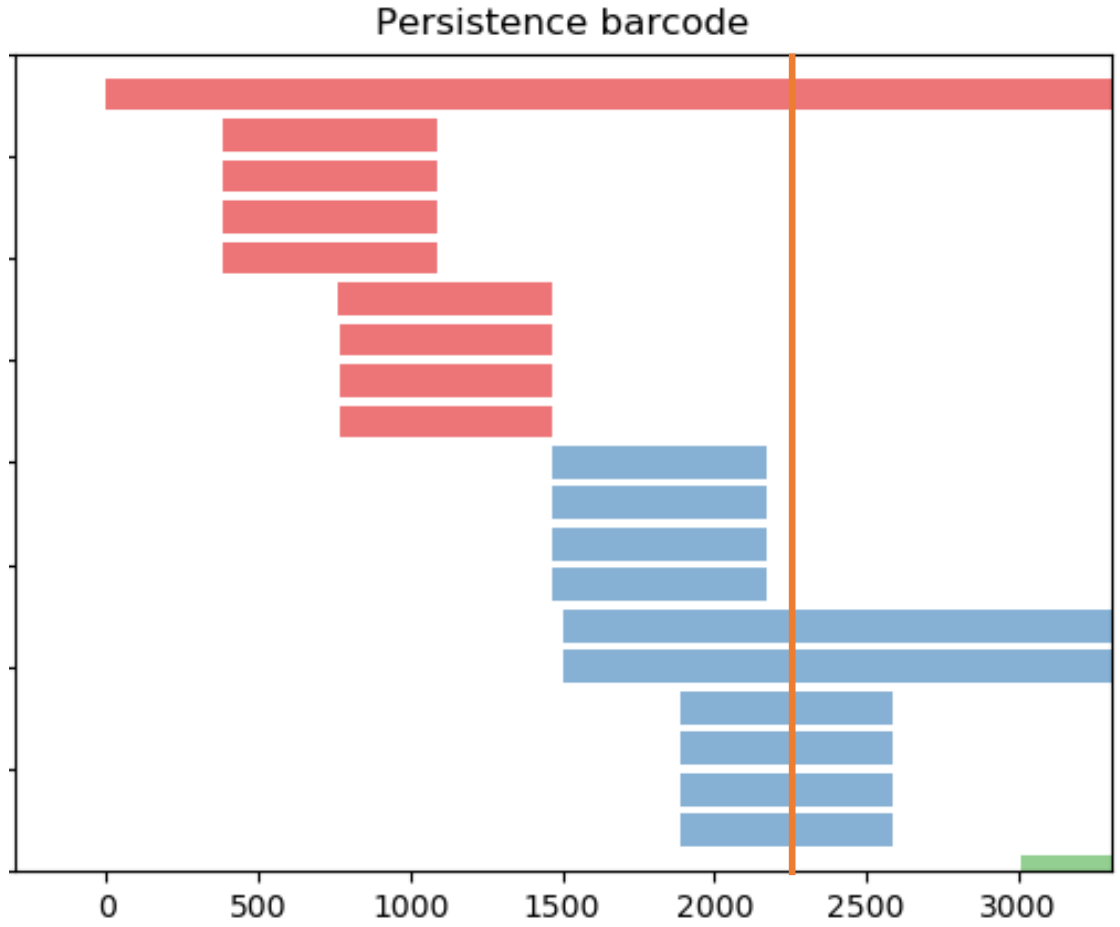



We have a complete characterization of the persistence barcodes of all *n*-alkanes

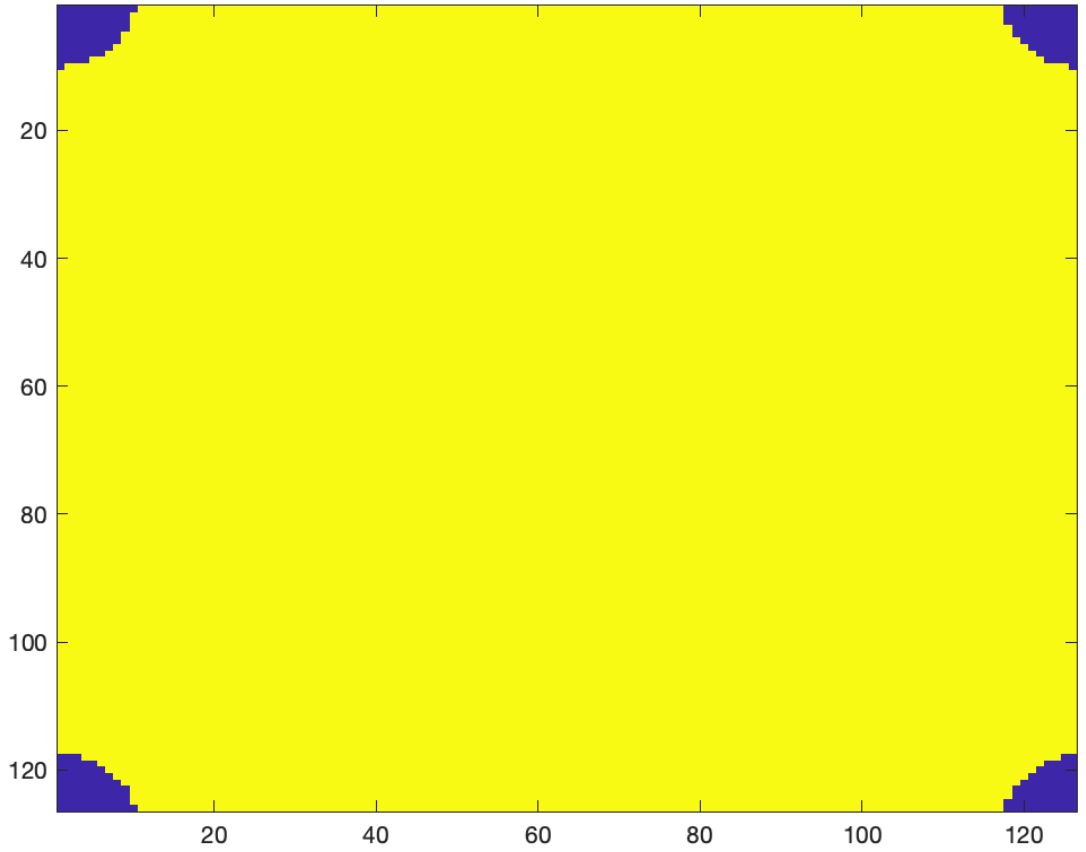
# Sublevelset persistent homology of the *n*-alkanes

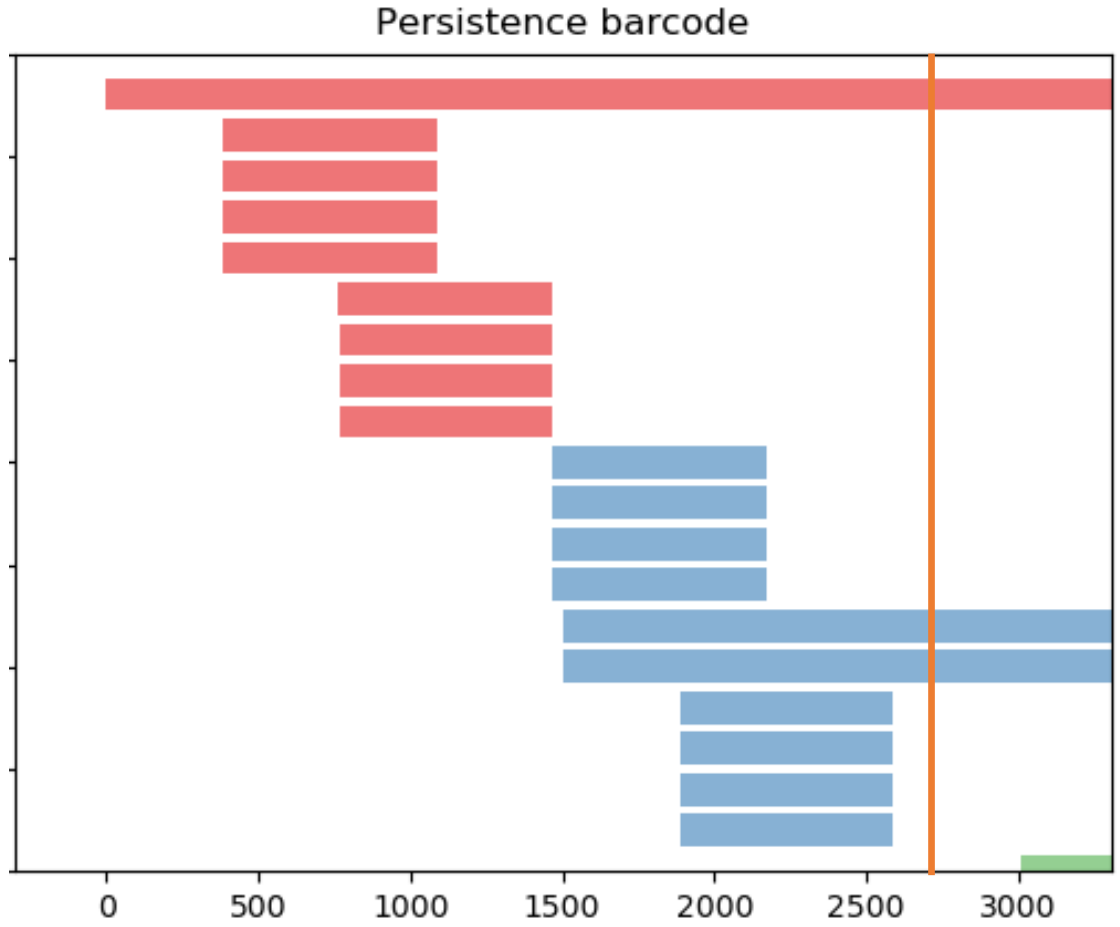



We have a complete characterization of the persistence barcodes of all *n*-alkanes

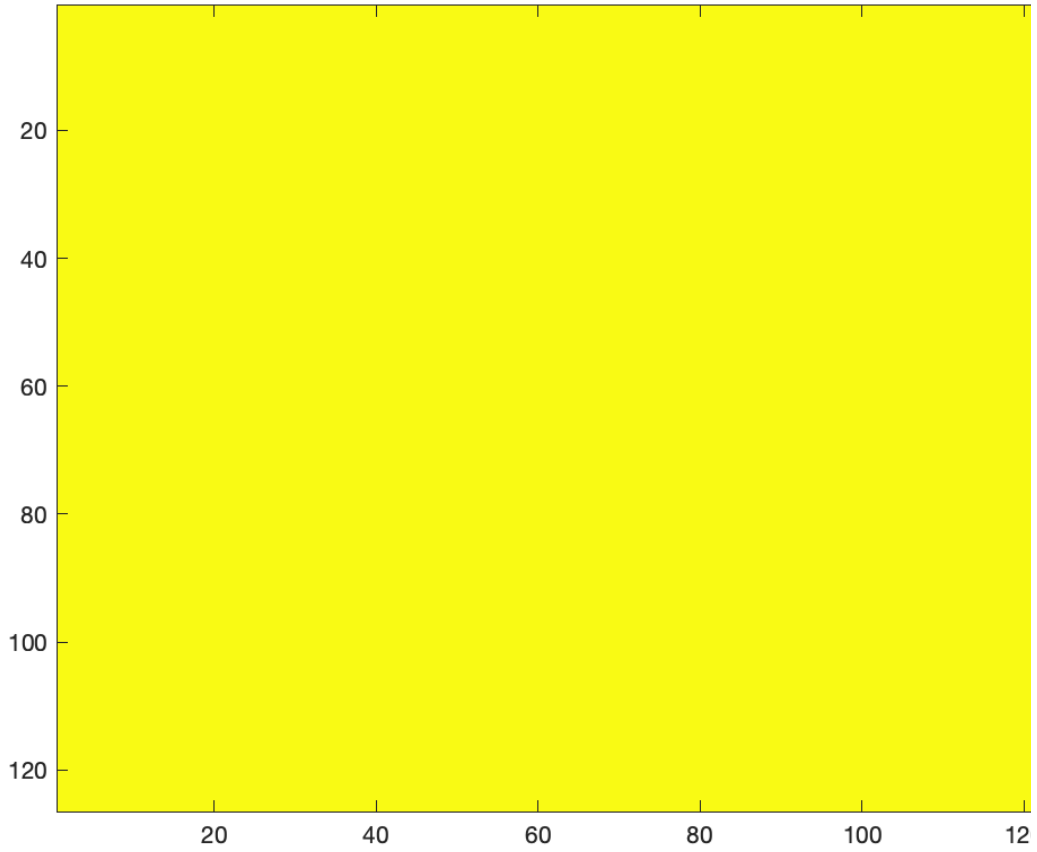
# Sublevelset persistent homology of the *n*-alkanes

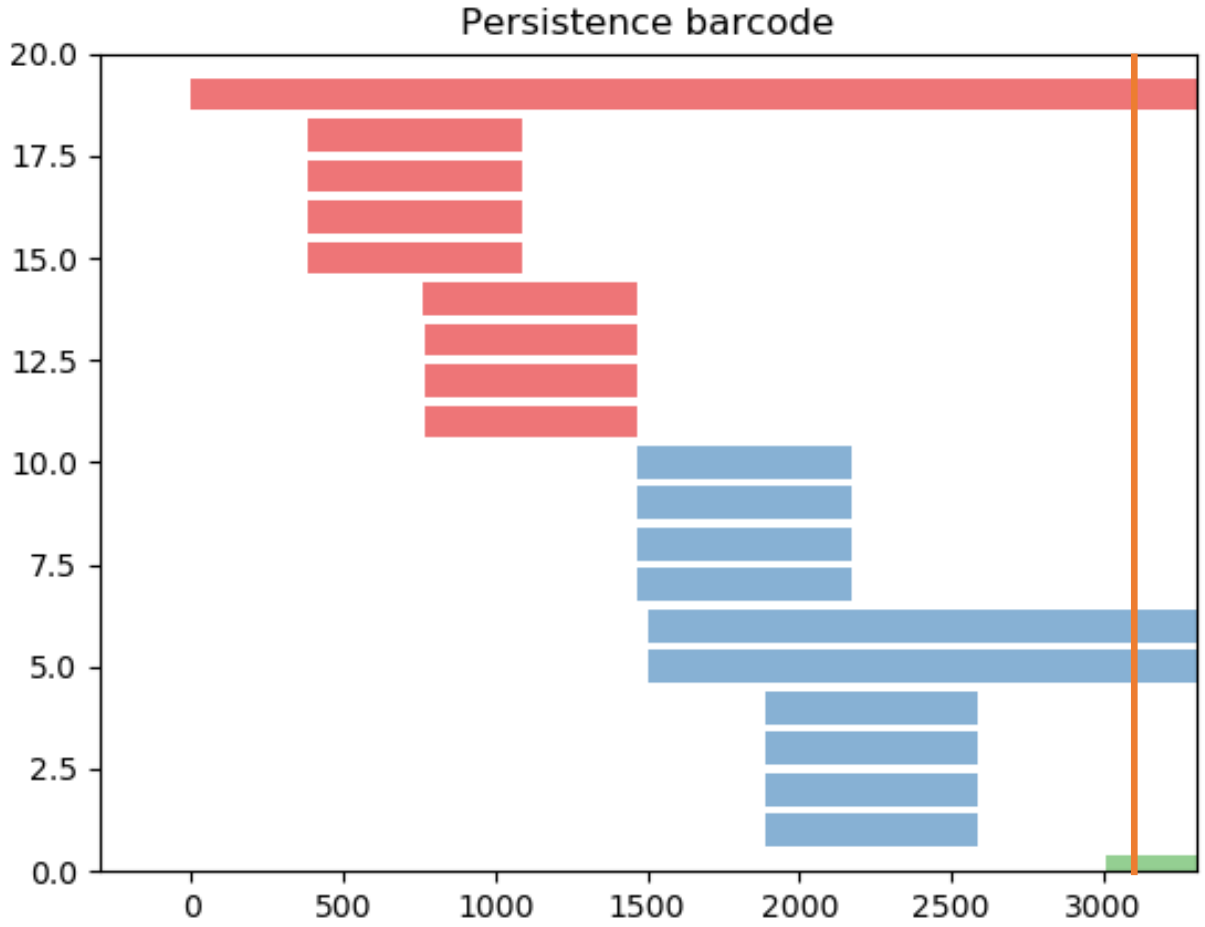



We have a complete characterization of the persistence barcodes of all *n*-alkanes

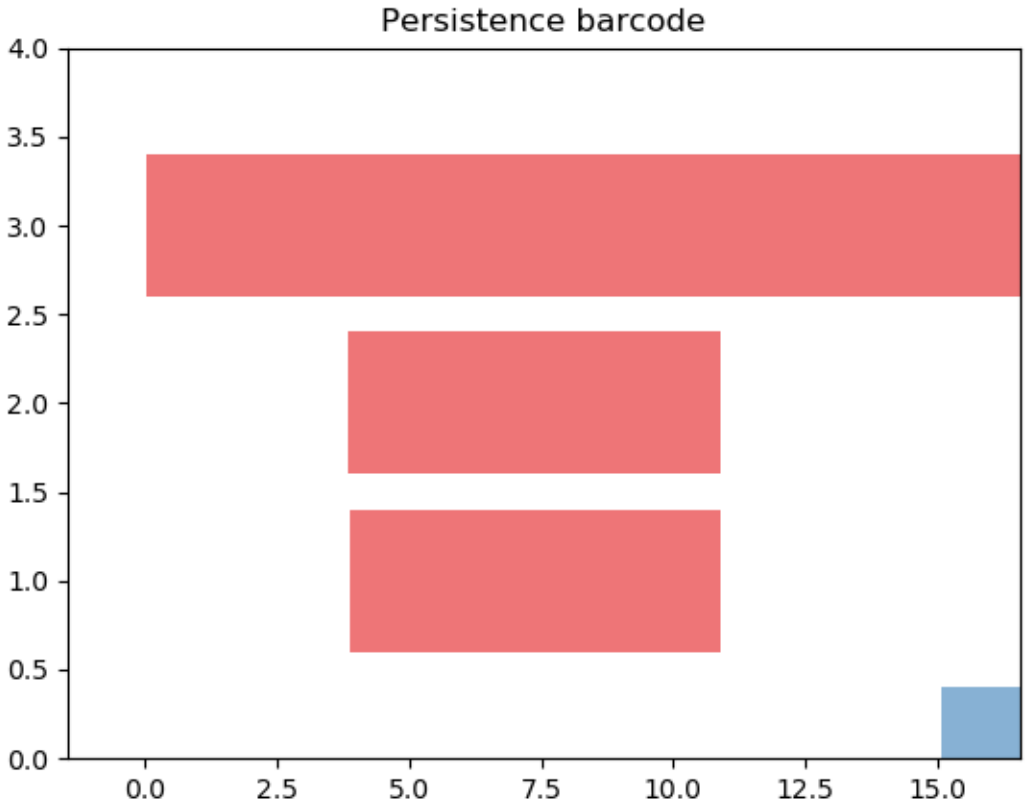
# Sublevelset persistent homology of the *n*-alkanes



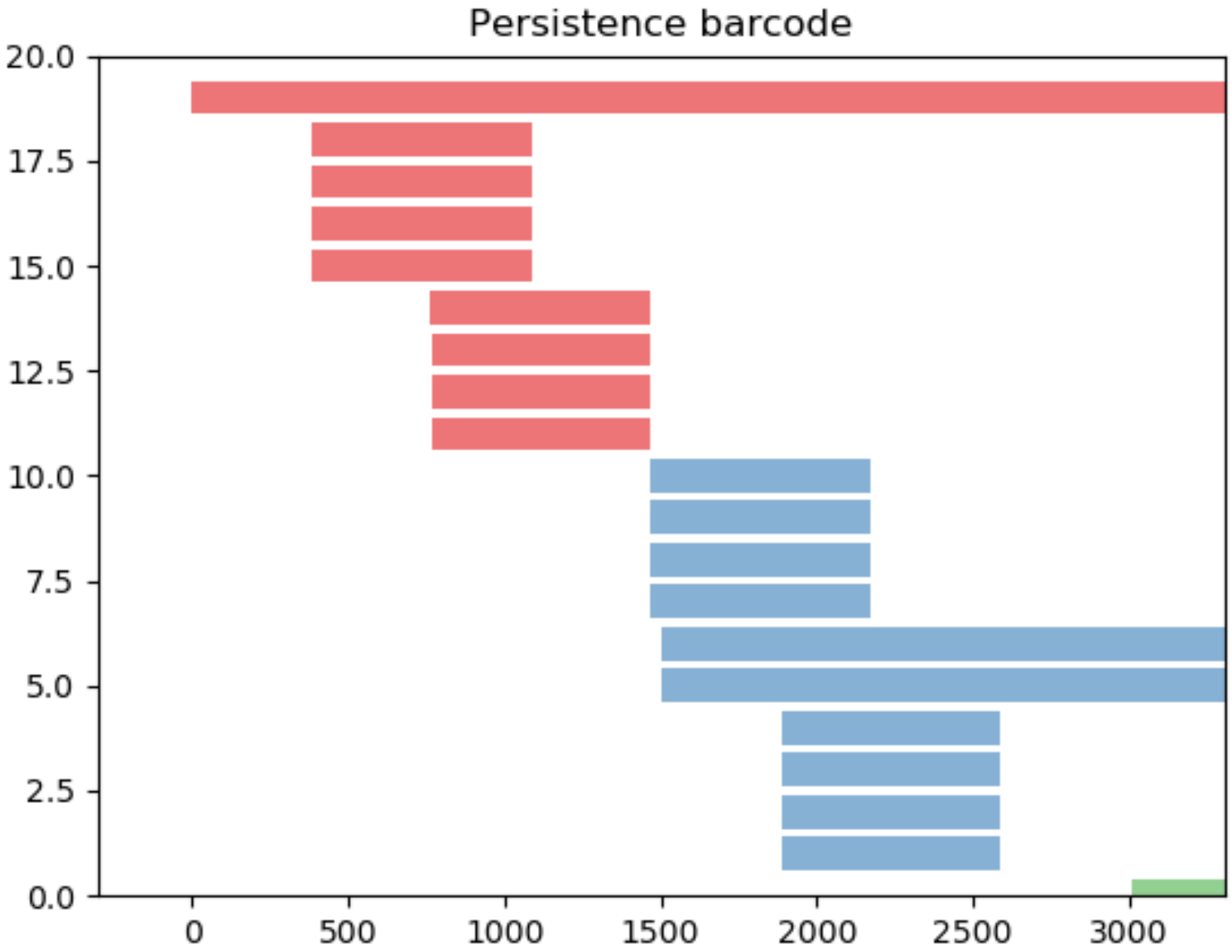

We have a complete characterization of the persistence barcodes of all *n*-alkanes

# Sublevelset persistent homology of butane



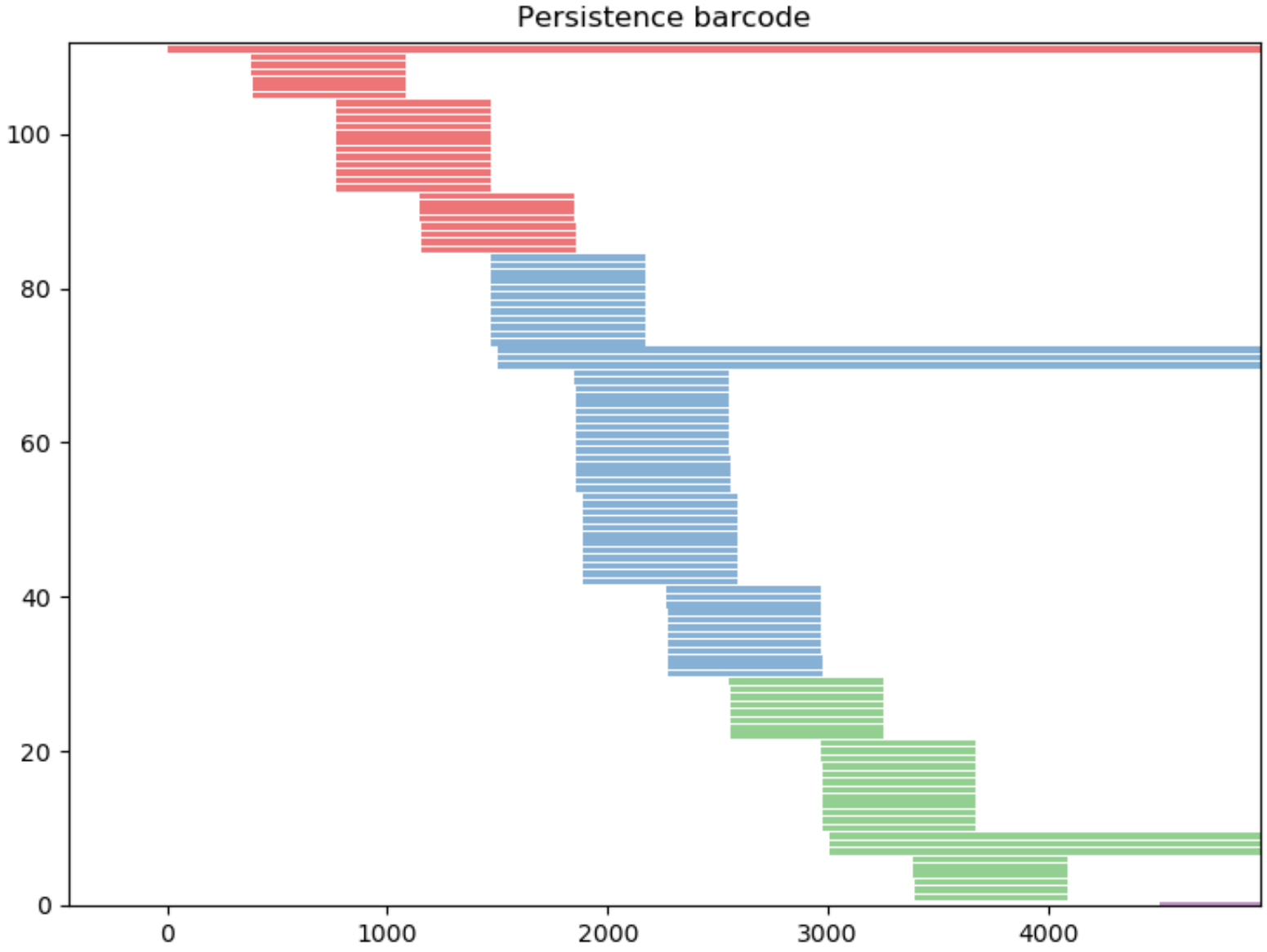

We have a complete characterization of the persistence barcodes of all *n*-alkanes

# Sublevelset persistent homology of pentane



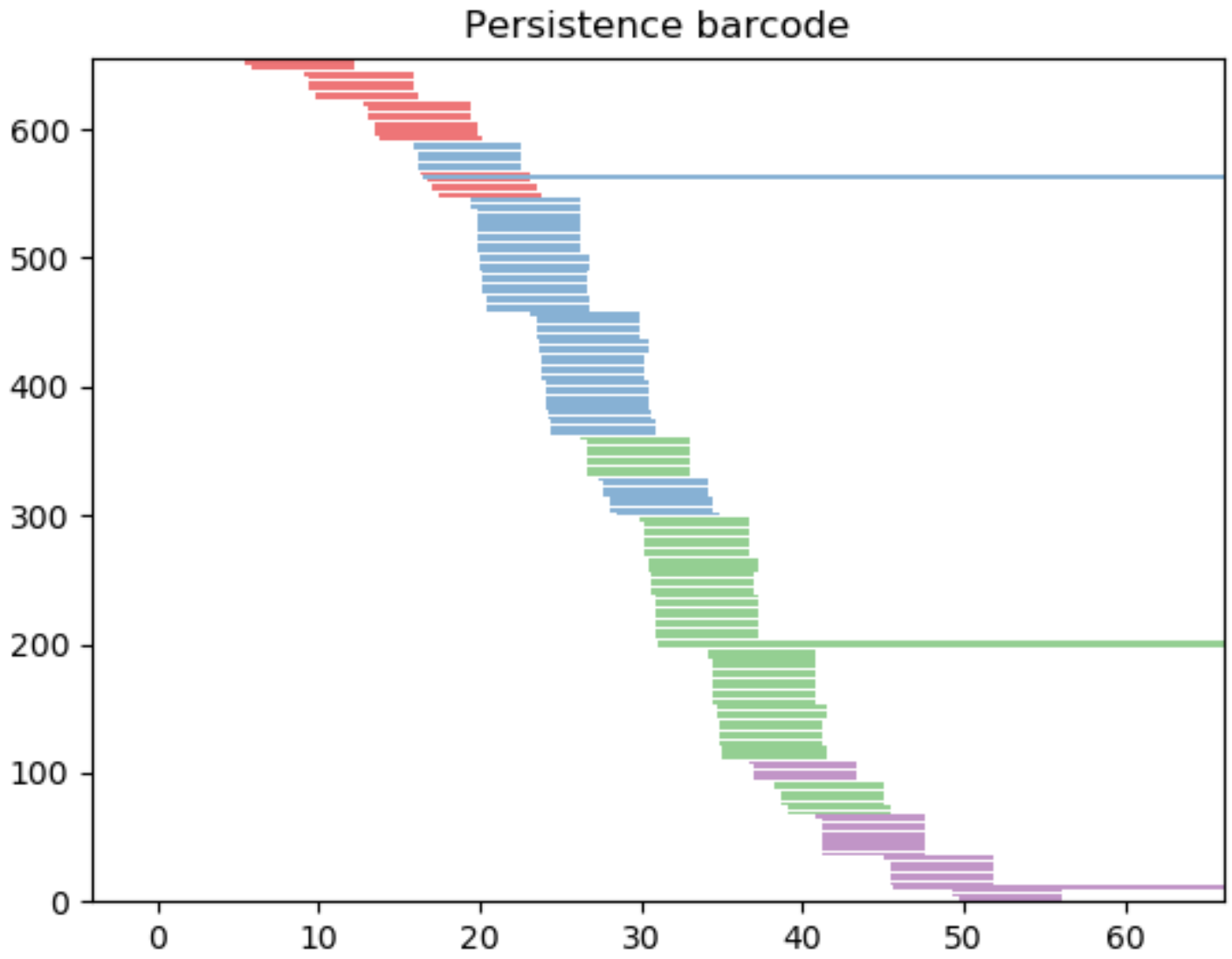

We have a complete characterization of the persistence barcodes of all *n*-alkanes

# Sublevelset persistent homology of hexane



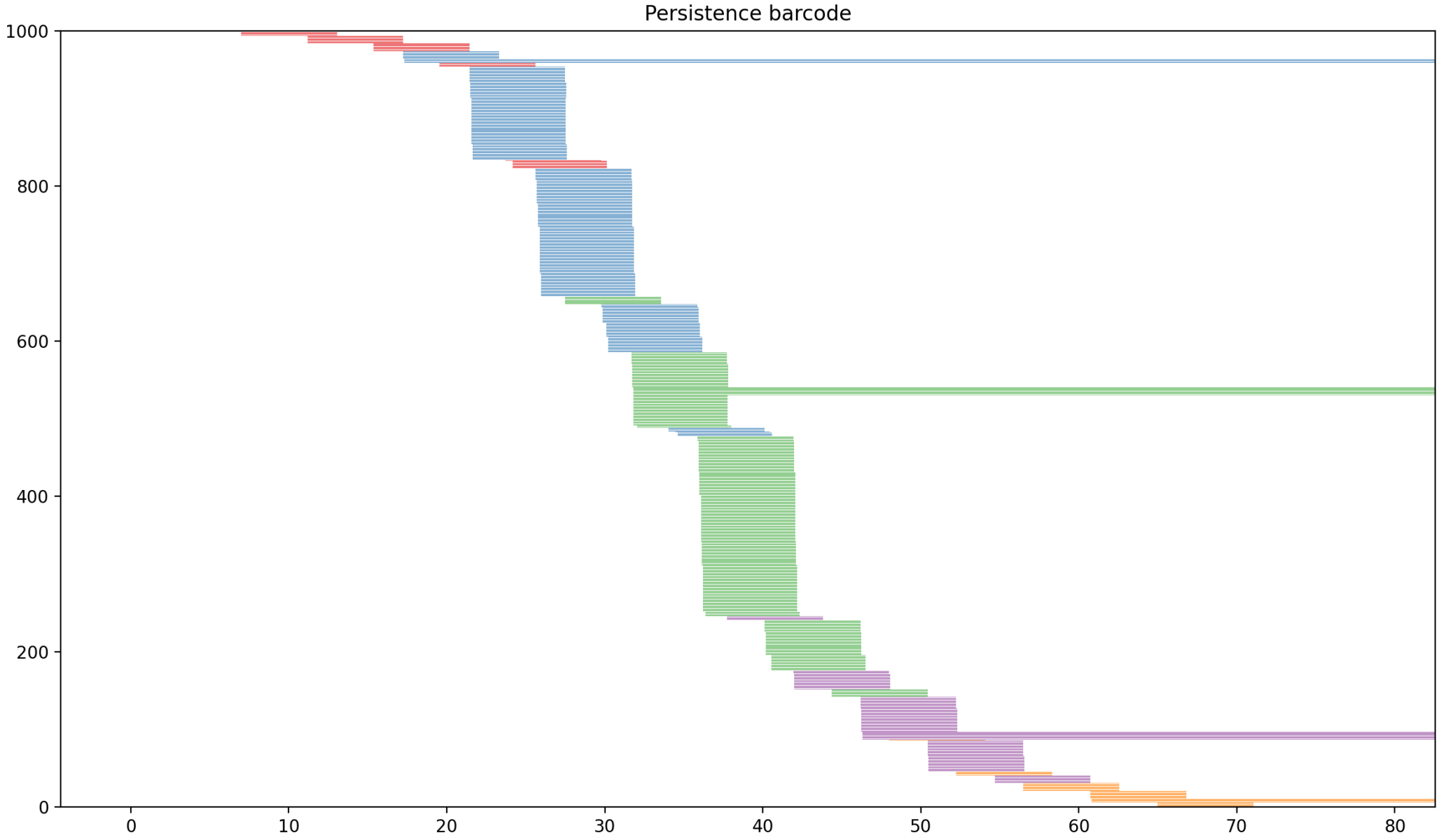

We have a complete characterization of the persistence barcodes of all *n*-alkanes

# Sublevelset persistent homology of heptane



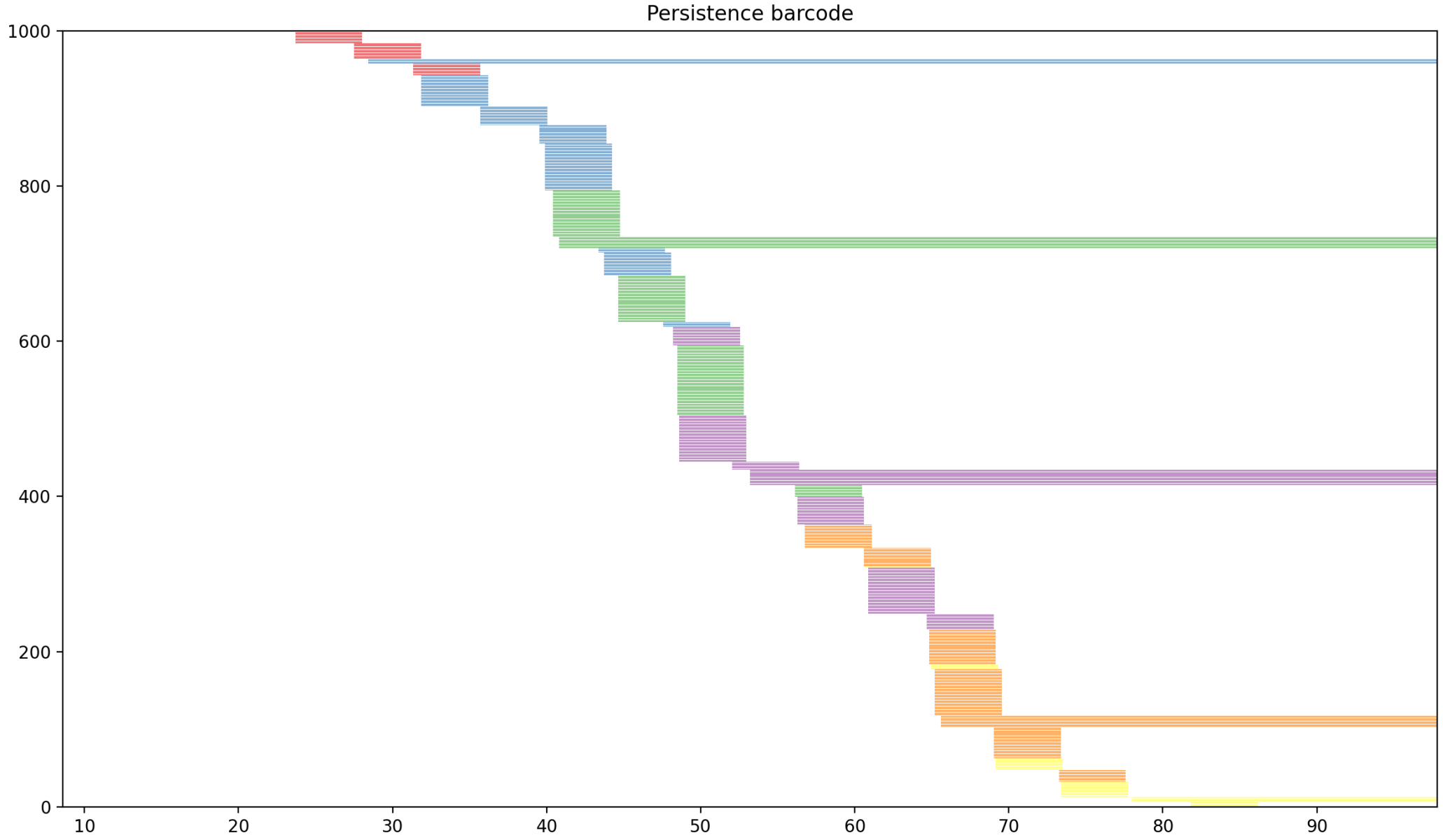

We have a complete characterization of the persistence barcodes of all *n*-alkanes

# Sublevelset persistent homology of octane




We have a complete characterization of the persistence barcodes of all $n$-alkanes

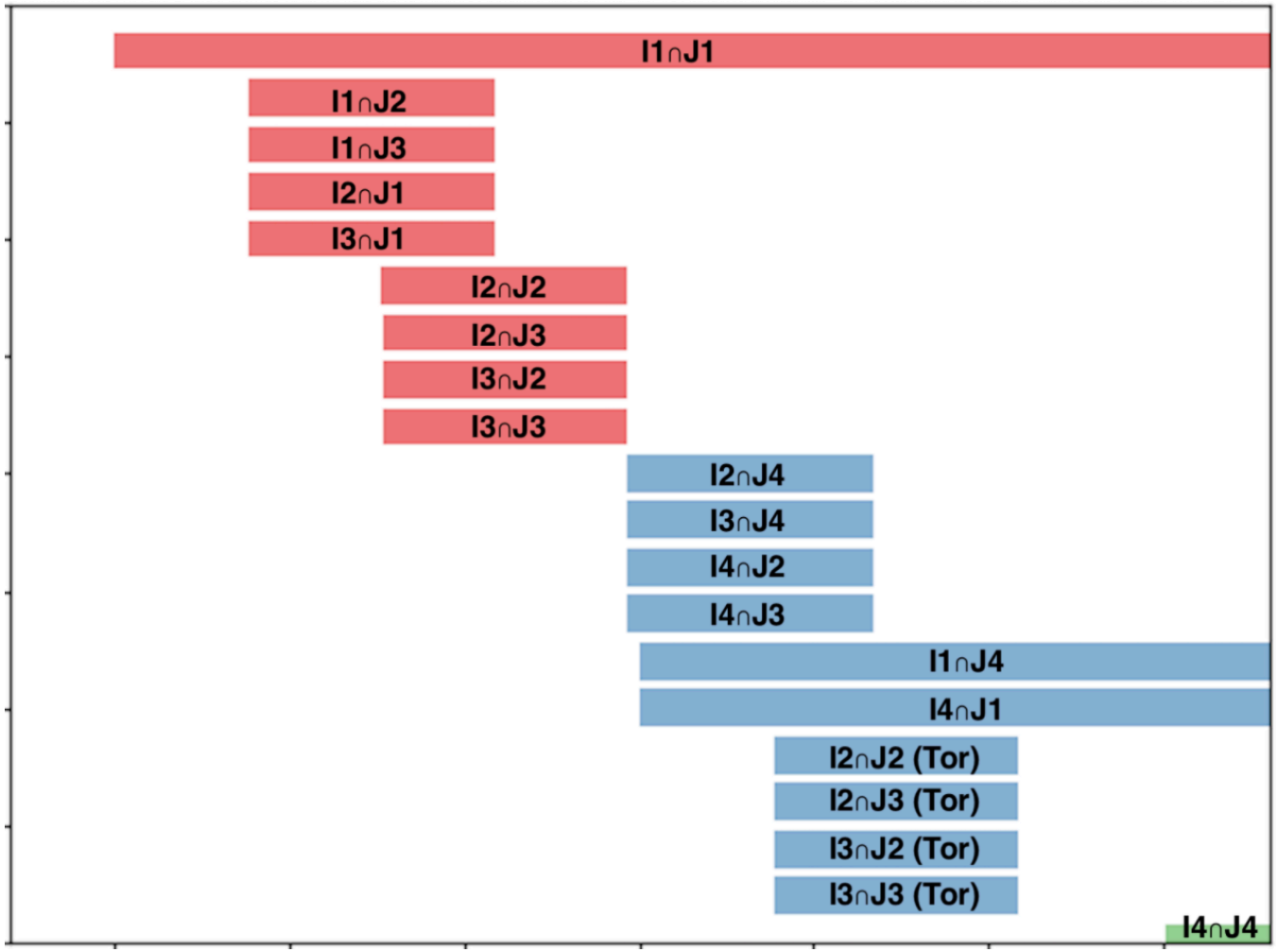
# Sublevelset persistent homology of nonane




We have a complete characterization of the persistence barcodes of all *n*-alkanes

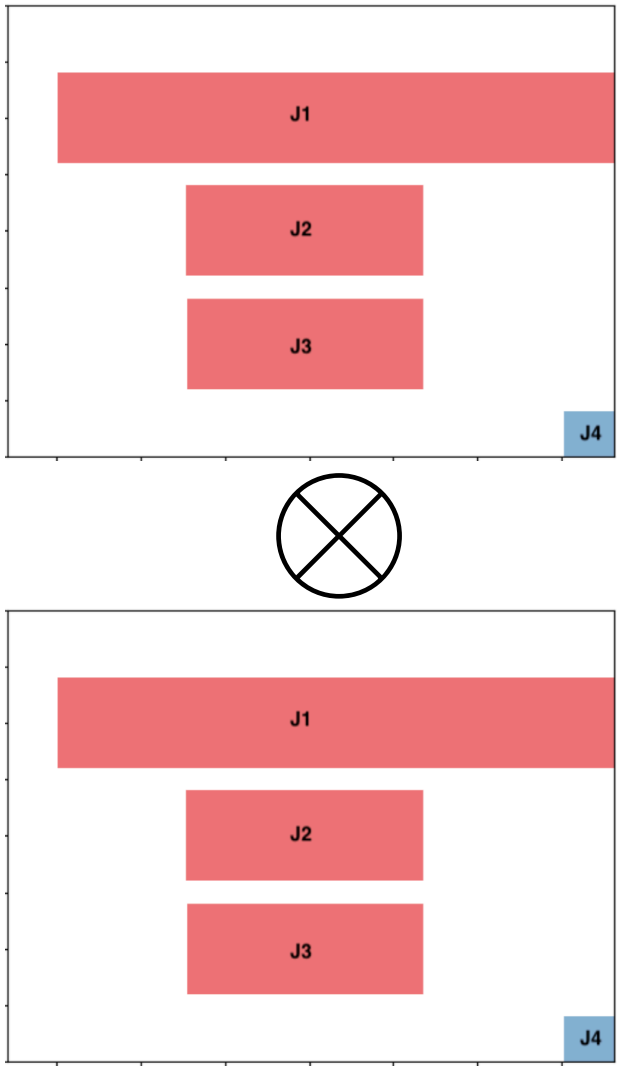
# Sublevelset persistent homology of the *n*-alkanes

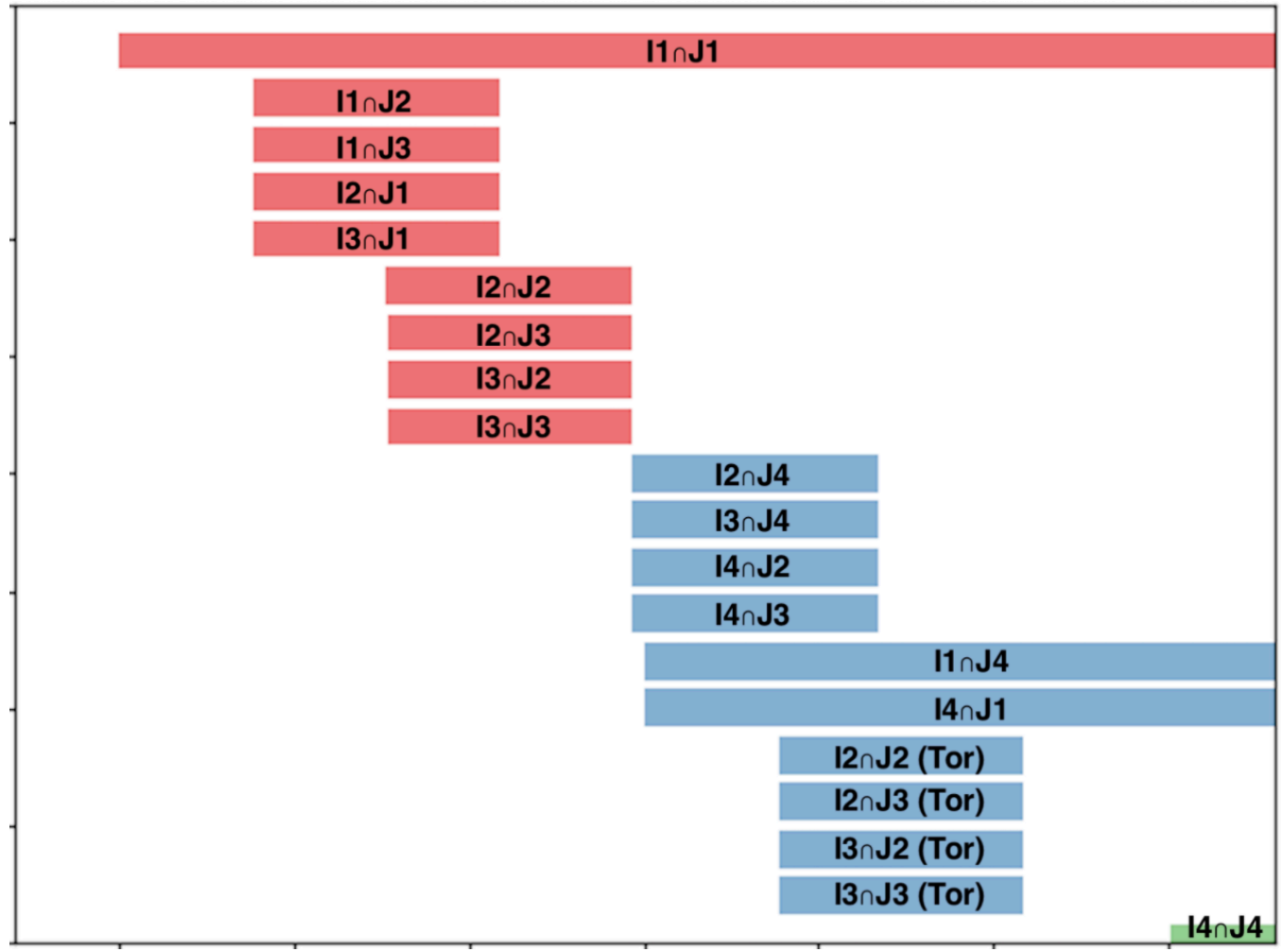
- We have a formula for the birth and death time of each bar in every homological dimension
- The proof uses a Künneth formula for persistent homology by Jose Perea and Hitesh Gakhar




We have a complete characterization of the persistence barcodes of all *n*-alkanes

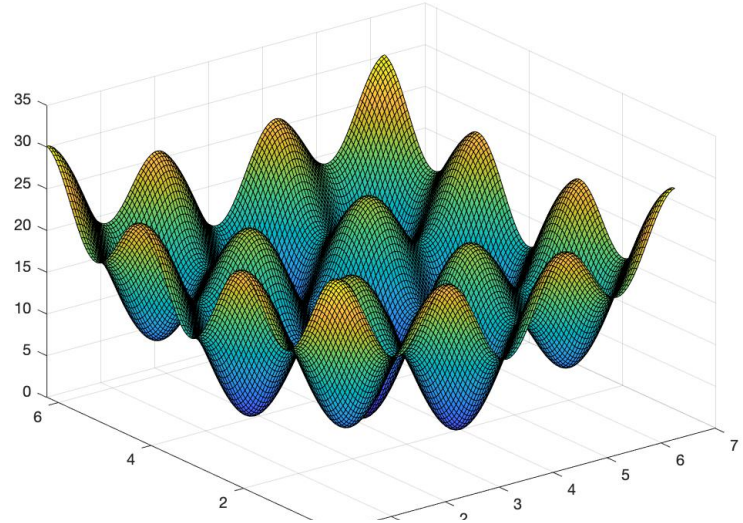
# Sublevelset persistent homology of the *n*-alkanes

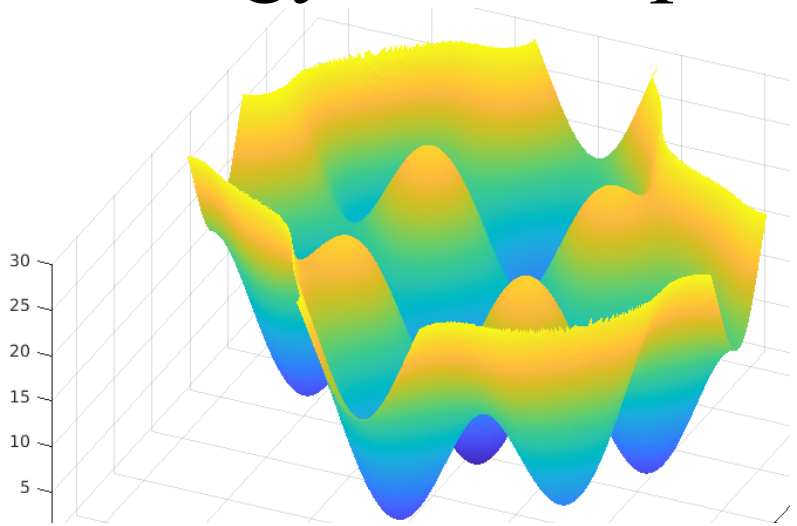


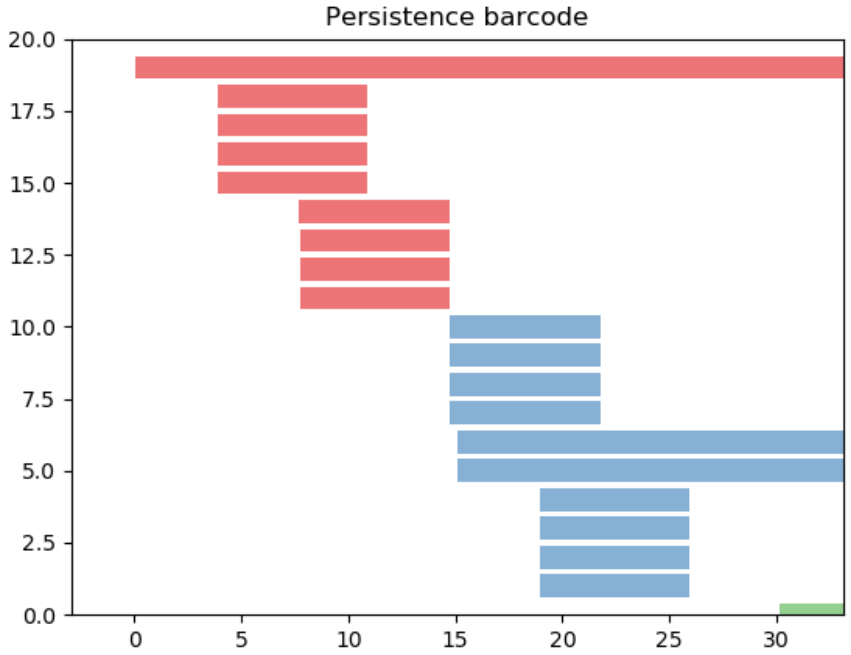



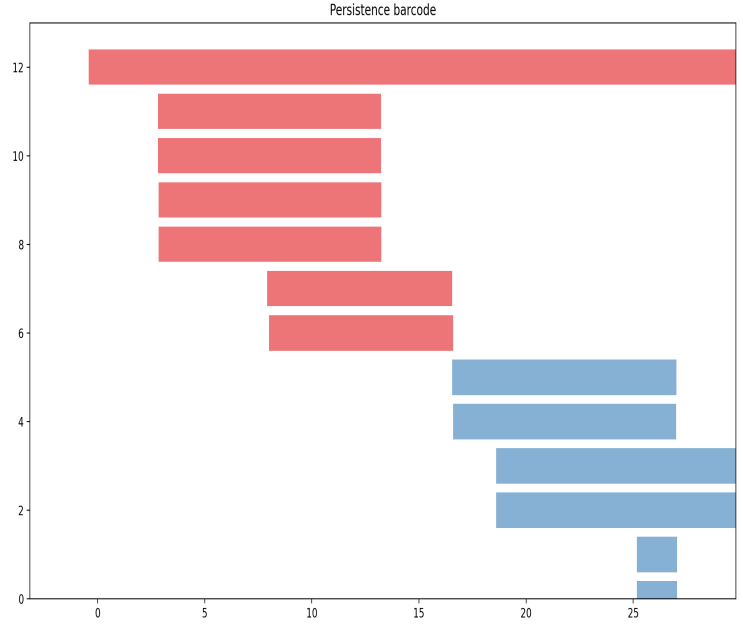

We have a complete characterization of the persistence barcodes of all *n*-alkanes

# A notion of distance between energy landscapes



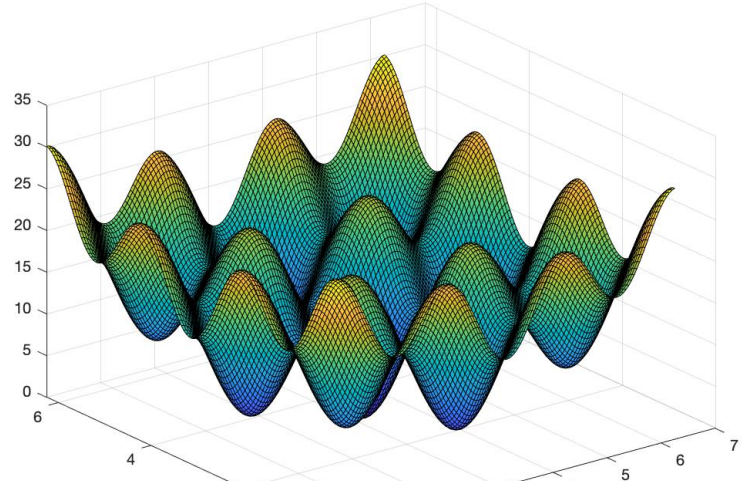



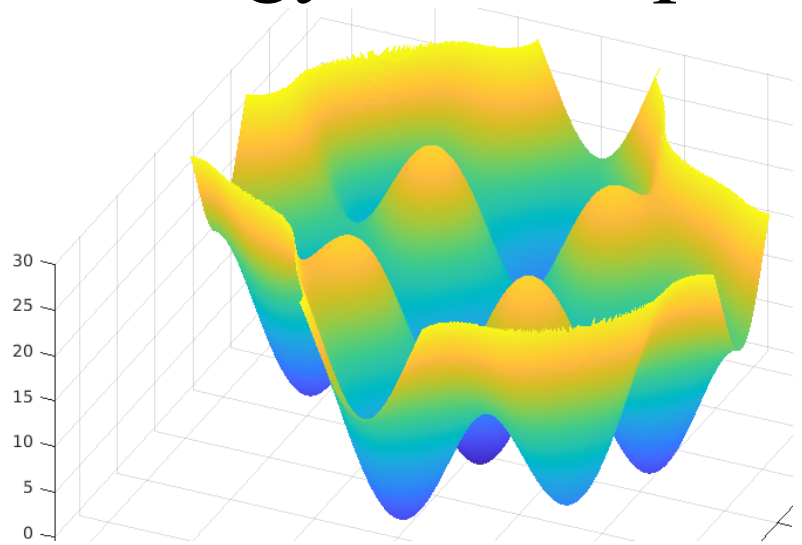



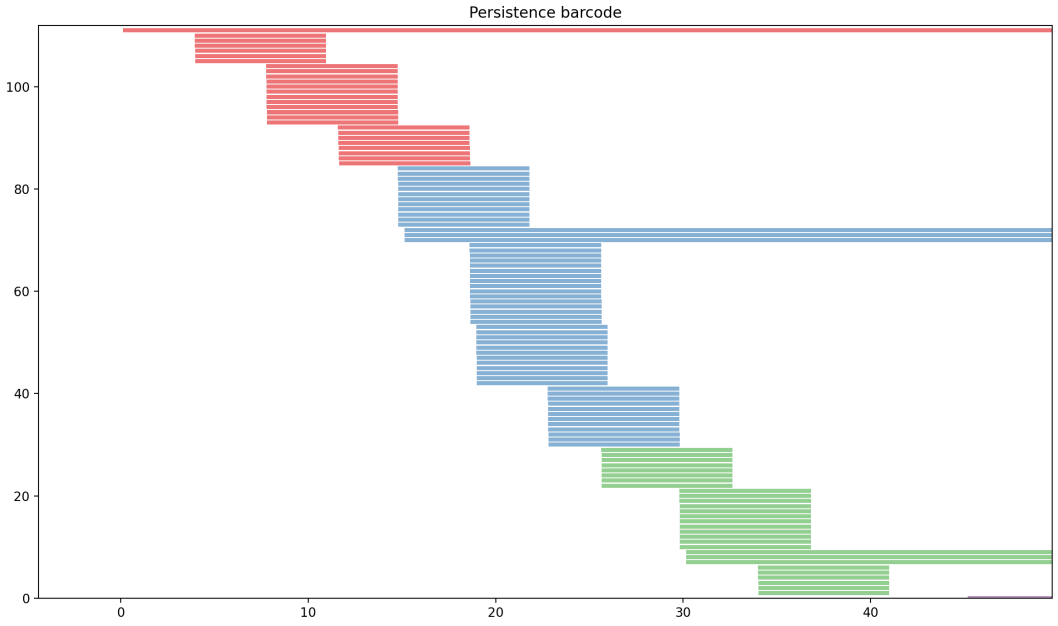



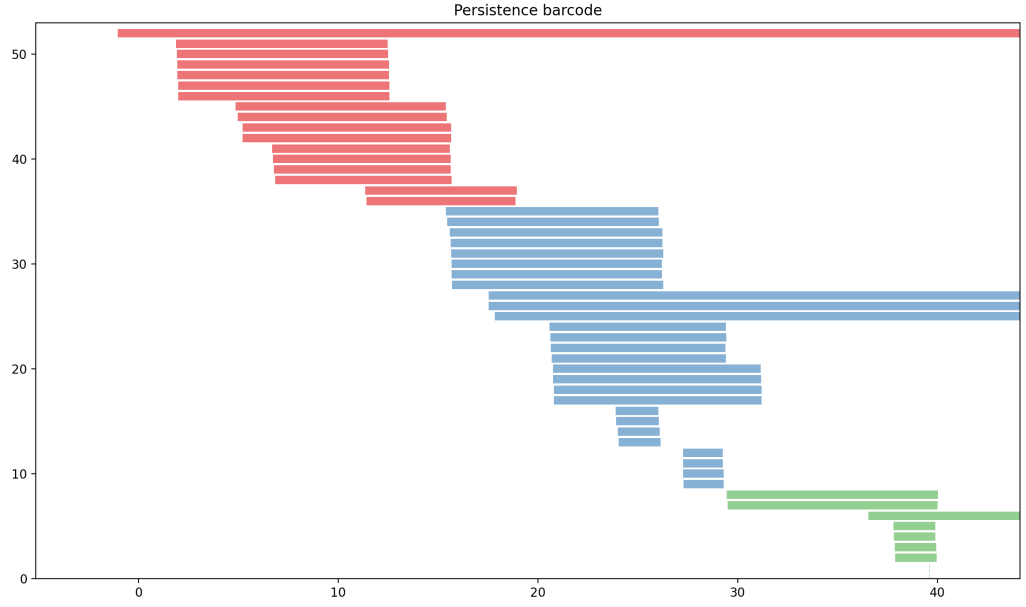

Bottleneck and Wasserstein distances are “edit” distances between barcodes

# A notion of distance between energy landscapes



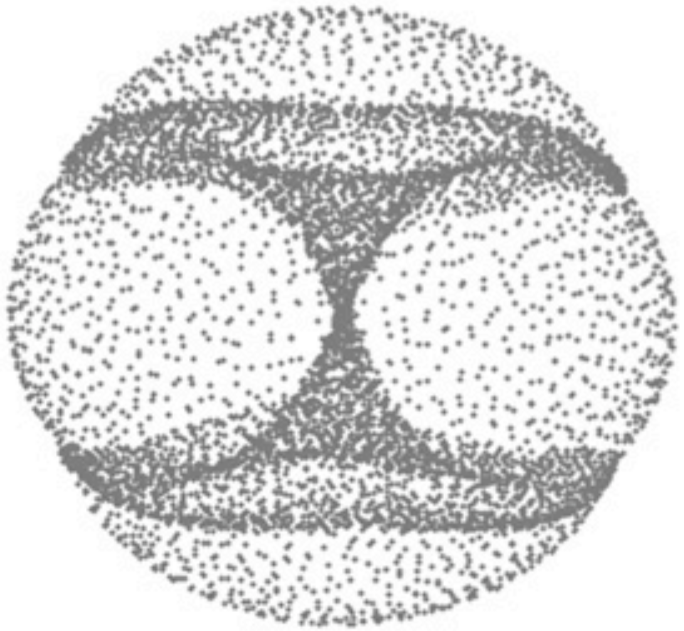



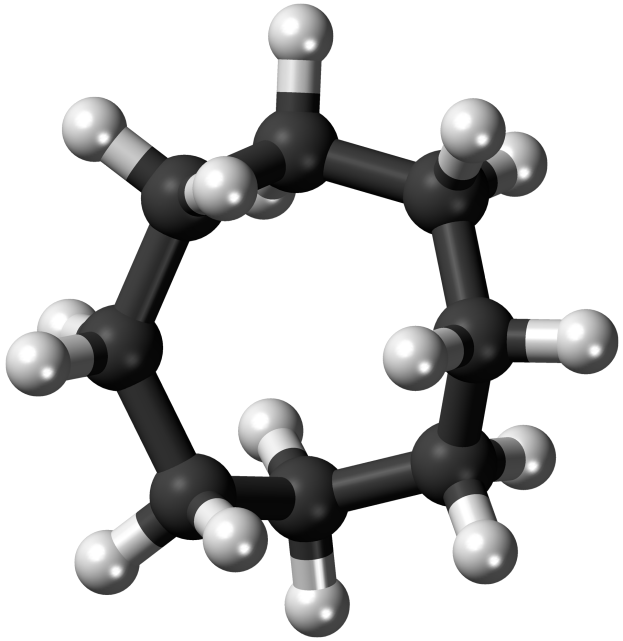




Bottleneck and Wasserstein distances are “edit” distances between barcodes

# The cyclo-alkanes

- We have a plan in place for computing the sublevelset persistent homology of cyclo-octane, $C_8H_{16}$, using Vietoris-Rips complexes and the lower star filtration.

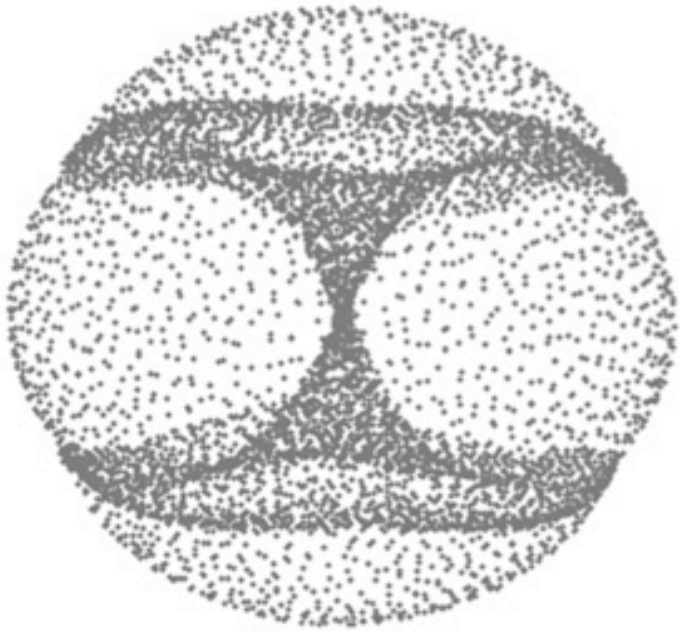


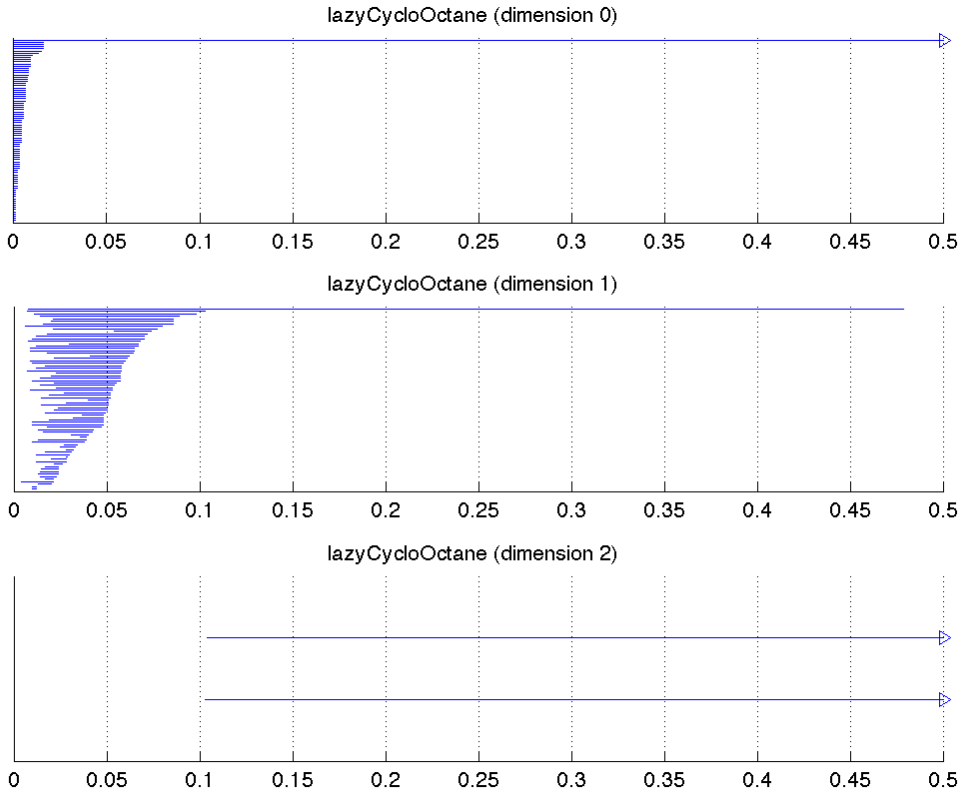


*Non-Manifold Surface Reconstruction from High Dimensional Point Cloud Data*
by Shawn Martin and Jean-Paul Watson, 2010.

# The cyclo-alkanes

- We have a plan in place for computing the sublevelset persistent homology of cyclo-octane, $C_8H_{16}$, using Vietoris-Rips complexes and the lower star filtration.

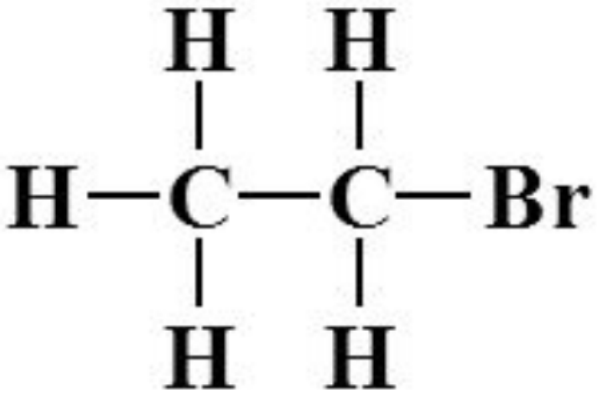





*Non-Manifold Surface Reconstruction from High Dimensional Point Cloud Data*
by Shawn Martin and Jean-Paul Watson, 2010.

# Machine learning tasks

- Big goal: Predict how the energy landscape of a chemical system evolves as a result to changes in salinity, temperature, pressure, etc.
- Related subgoal: Predict how the sublevelset persistent homology evolves as a result to changes in salinity, temperature pressure, etc.
  - Halide task




- Big goal: Given a partial sampling of a restricted energy landscape, predict the energy value at a new sample point.
- Related subgoal: Given a partial sampling of a restricted energy landscape, predict the sublevelset persistent homology.