Algebraic topology

Introduction

Algebraic Topology at a glance

Algebraic Topology

Topology is the study of the geometric properties of objects.

Algebraic topology is the study of the algebraic properties of spaces that are associated with their geometry.

Algebraic Topology

Imagine a map of a city. The topology of the city would describe the geometric properties of the map, such as the locations of streets, buildings, and landmarks, and *how* they are connected to each other. Topology focuses on the "raw data" of the map, which has information about the city's layout and structure.

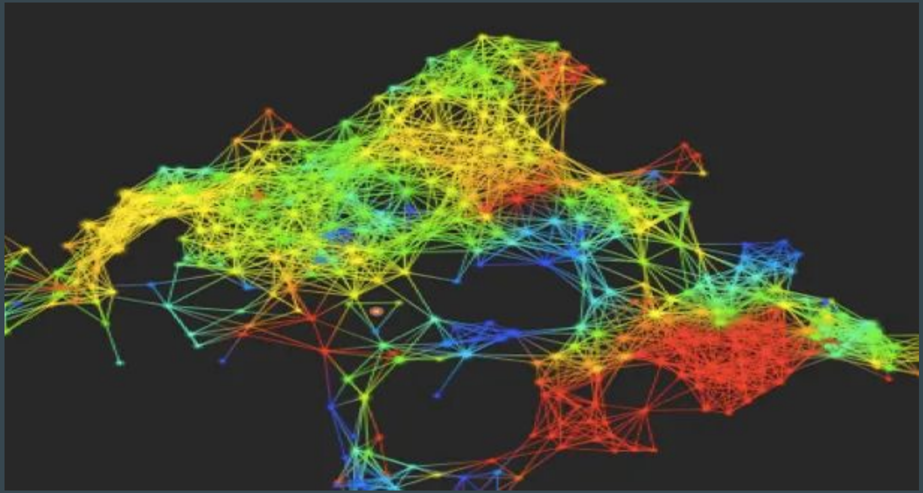
Algebraic topology would assign *numbers* to certain features of the map, such as the number of streets that intersect at a particular point or the distance between two landmarks. These numbers would form an algebraic structure, which captures certain properties of the city's geometry in a more abstract way.

Algebraic topology is using this algebraic structure to study the city, rather than the raw data of the map itself.

Topological Data Analysis

Topological Data Analysis studies data sets that are represented as point clouds, where each data point represents an observation or measurement. Topological information are used to extract information about the underlying structure of the data: for example about the number of connected components, and the relationships between them.

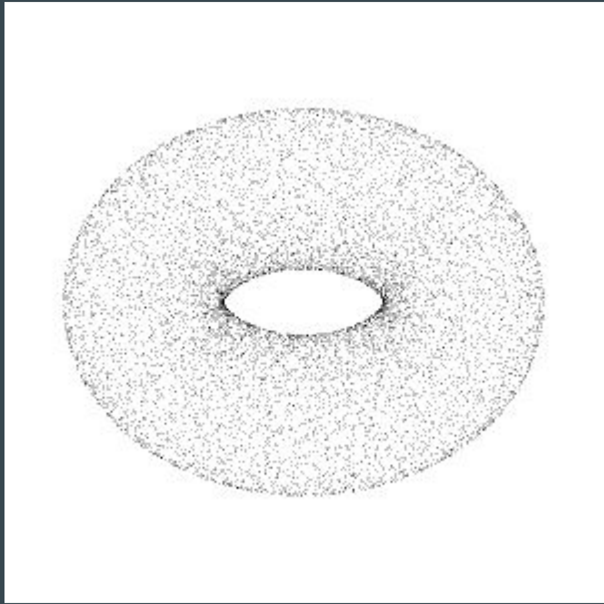
https://www.datanami.com/2015/03/25/mapping-the-shape-of-complex-data-with-ayasdi/



I. Definitions

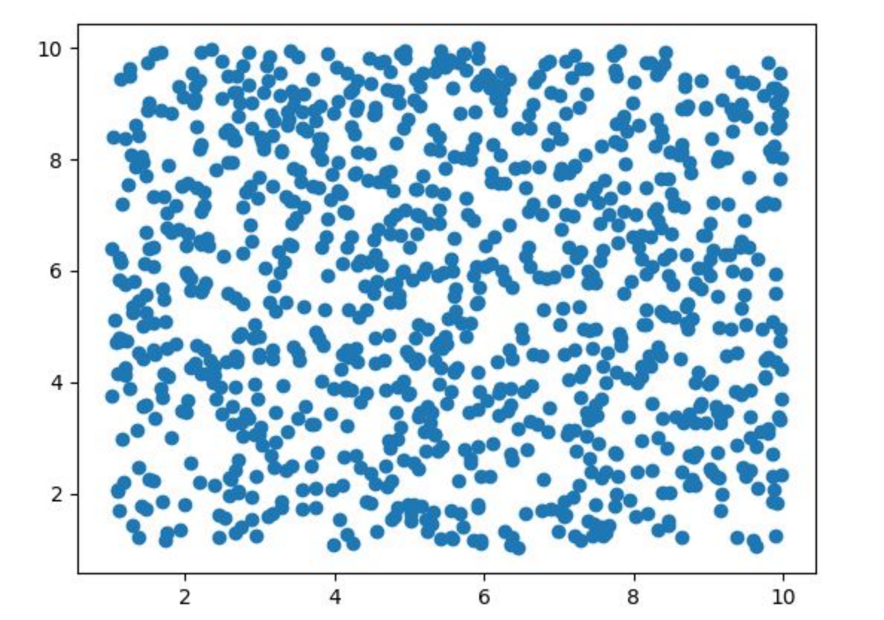
Point cloud

- Data can be represented as an unordered sequence of points in a Euclidean n-dimensional space
- A collection of points that are unorderly distributed in n-dimensional space is a point cloud
- An example of a point cloud is a set of points uniformly distributed on a torus



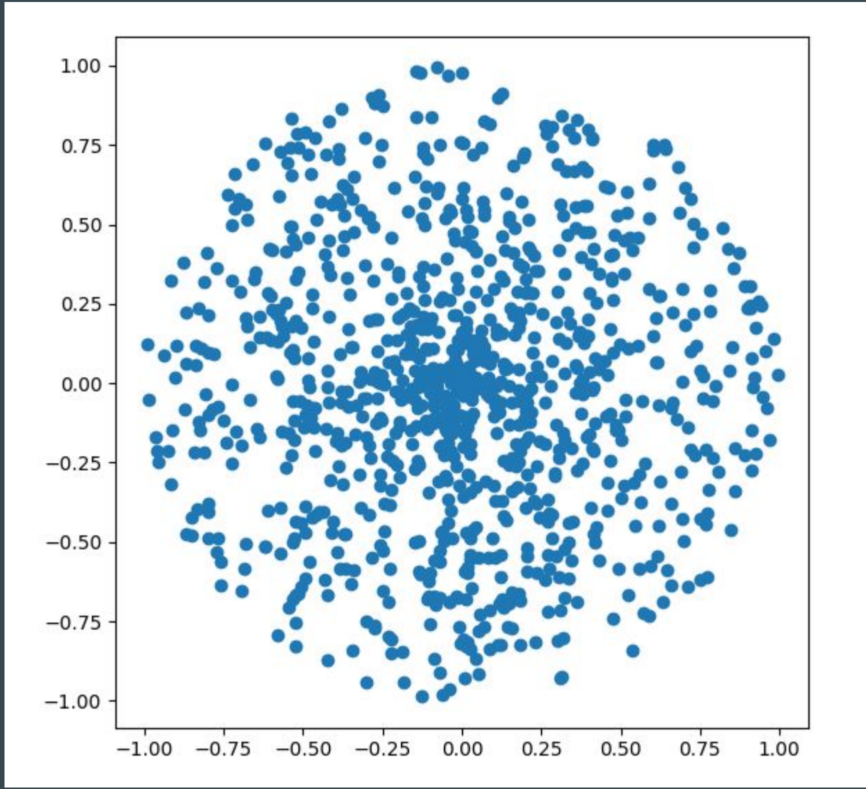
Point cloud

- An example of a point cloud is a set of points uniformly distributed on a torus
- Other examples include randomly distributed points in a box, gaussian distribution of points in a circle



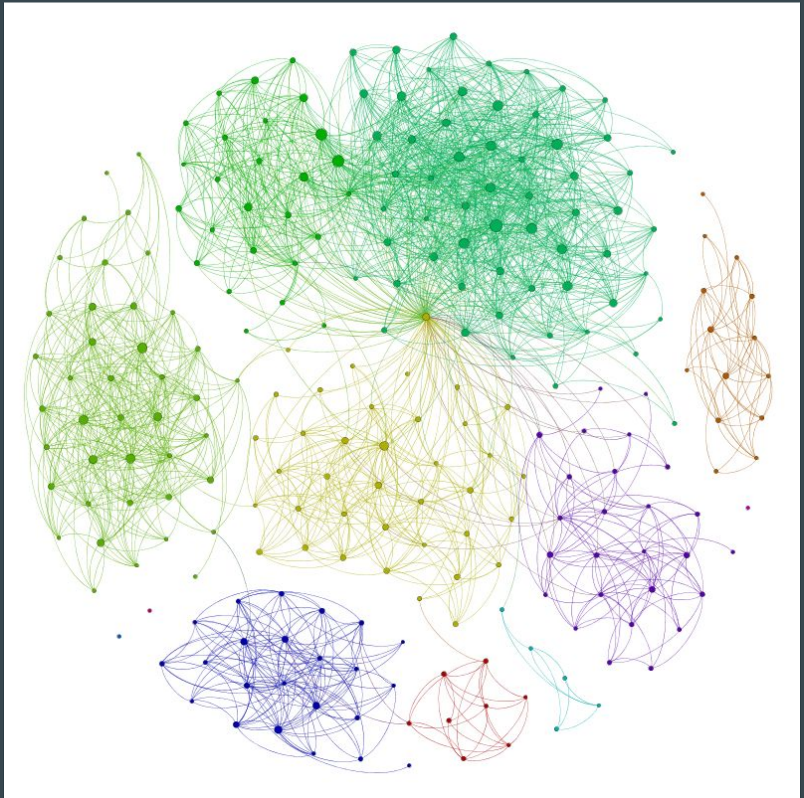
Point cloud

- An example of a point cloud is a set of points uniformly distributed on a torus
- Other examples include randomly distributed points in a box, gaussian distribution of points in a circle



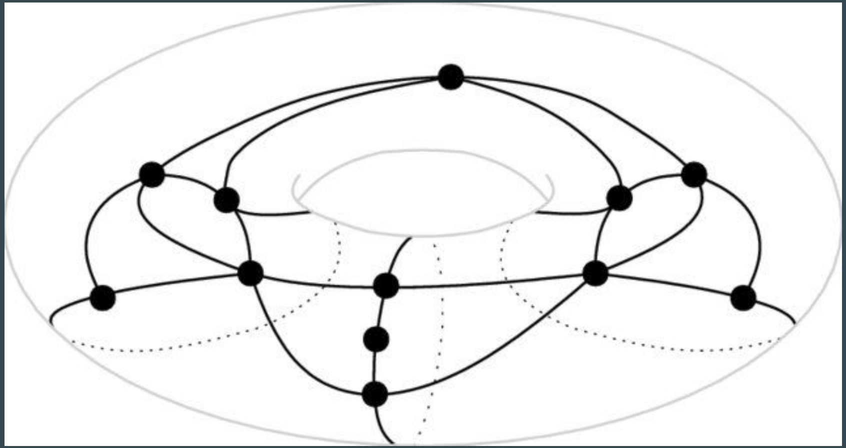
Graphs

- Graphs consist of a set of vertices that are connected by edges
- In discrete mathematics vertices can be abstract, and edges are pairs of these vertices that do not need to be connected. For example a graph can be composed by people and edges are defined between pairs of people that know each other



Graphs

- In **topology** a graph as a 1-dimensional geometric object, vertices are points and edges are curves connecting these points in pairs.



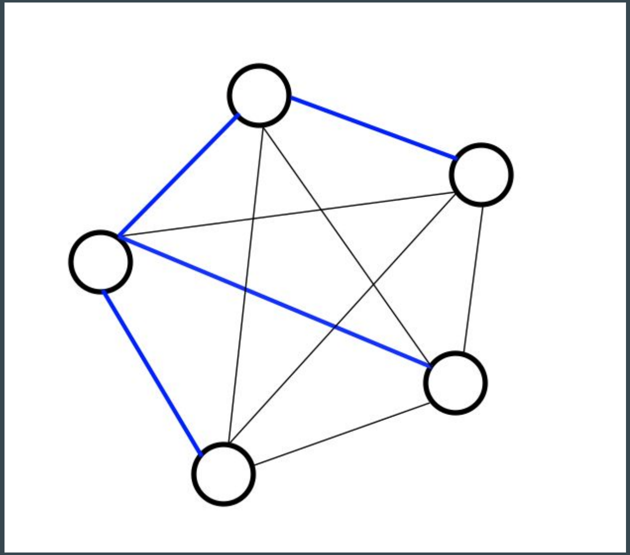
Sneddon 2004

Simple Graphs

A graph is *simple* if the edge set is a subset of the set of unordered pairs, which means that

- no two edges connect the same two vertices
- no edge joins a vertex to itself

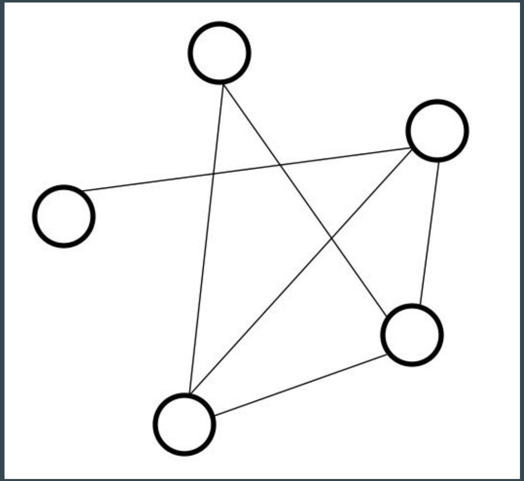
Every simple graph with n vertices is a subgraph of the *complete graph*, Kn, that contains an edge for every pair of vertices:



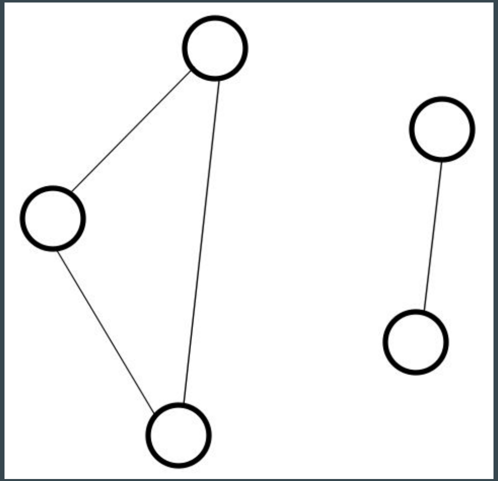
Edelsbrunner and Harer
Blue: example of simple graph, subset of complete graph, K5

Connected Graphs

- A path between vertices *u* and *v* is a sequence of vertices that are connected by edges: *u, u1, u2, ..., v*
- A simple graph is *connected* if there is a path between every pair of vertices



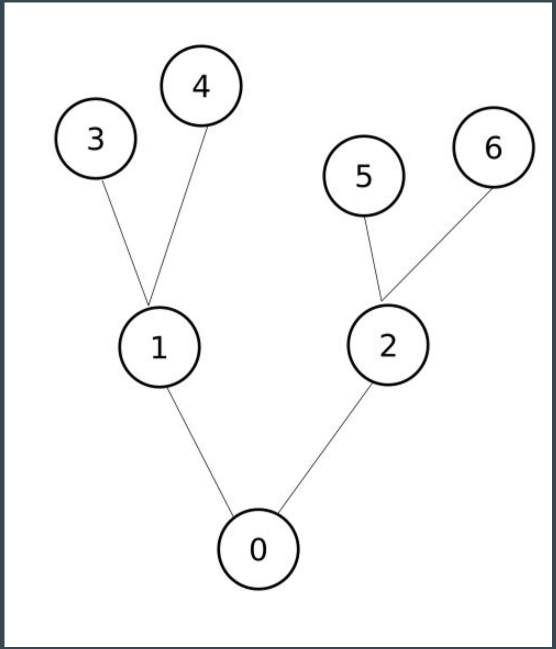
Connected graph



Disconnected graph

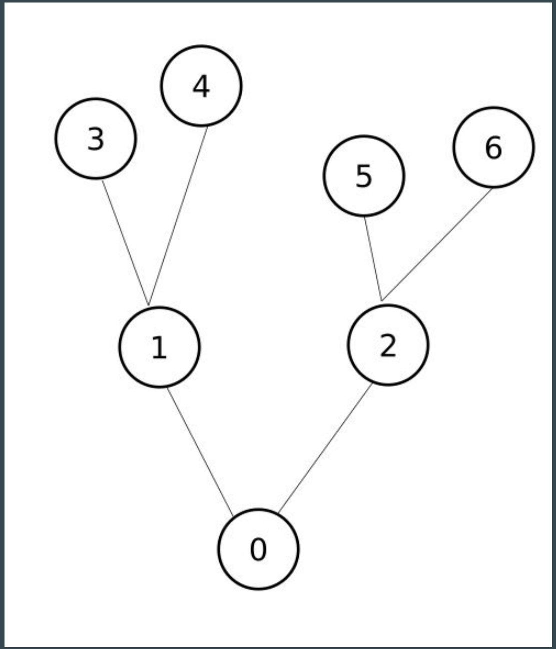
Trees

- The smallest connected graph is a tree: there is a unique simple path between every pair of vertices.
- Removing any one edge from the tree, disconnects the graph



Trees

- In graph theory, a *tree* is an (un)directed graph in which any two vertices are connected by exactly one path, which is equivalently a connected acyclic (un)directed graph.
- A tree with a defined root is directed and all paths point away from the root.
- A *forest* is a disjoint union of trees.



Topological Spaces

A topology on a point set X is a collection U of subsets of X, called open sets, such that:

- X is open and the empty set ∅ is open.
- If U1 and U2 are open, then the intersection of the two subsets U1 ∩ U2 is open. Equivalently, intersections of finitely many open sets are open.
- If Ui is open every i, then the union of all Ui is open.

The pair (X, U) is a called a *topological space*, but we will usually refer to X as a topological space.

Connected / Disjoint spaces

A function from one topological space to another is *continuous* if the preimage of every open set is open. Let f: A→ B be a map from A to B, and Y a subset of B. Then the preimage $Y^{-1}(f)$ is the set of all elements in A that map to elements in B.

For example: the function $f : R \to R$ is *not* continuous

$(-\infty, 0] \to 0$

$(0, \infty) \to 1$

because for any $0 < \varepsilon < 1$, $(-\varepsilon, \varepsilon)$ is open, but $f^{-1}((-\varepsilon, \varepsilon))$ is not.

Connected / Disjoint spaces

A path is a *continuous* function from the unit interval, $\gamma : [0, 1] \to X$.

It connects the point $\gamma(0)$ to the point $\gamma(1)$ in X.

Similar to paths in graphs we allow for self-intersections:

$\gamma(s) = \gamma(t)$, with s not equal to t

If there are no self-intersections then the function is *injective* and the path is *simple*

Connected / Disjoint spaces

Definition I.

A topological space X is *path-connected* if every pair of points in X is connected by a path.

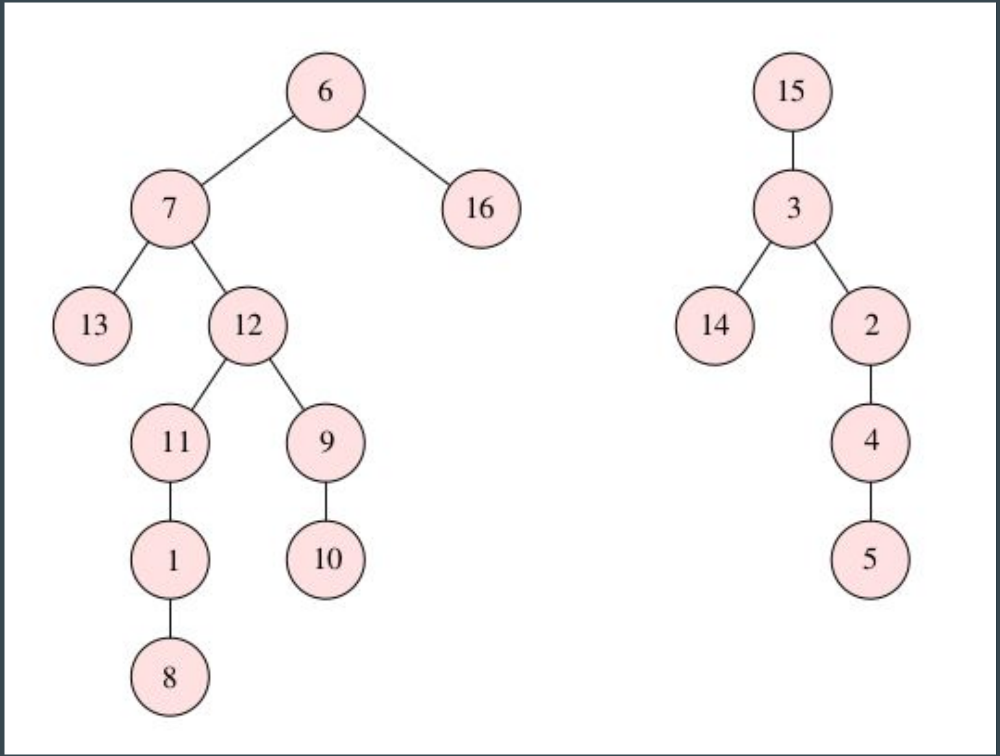
Definition II.

A *separation* of a topological space X is a partition of $X = U \cup W$ into two non-empty, open subsets. A topological space is *connected* if it has no separation.

Connectedness is strictly weaker than path-connectedness, but typically a path-connected space is also connected.

Connected / Disjoint spaces

The topological space that has a separation into two trees is not path-connected because there is no path connecting the vertices that belong to different trees.



Disjoint sets: Algorithm

An example of a *connectedness detection algorithm* to detect if a set is connected is the following:

Let's take a graph with n vertices: [n] = {1, 2, ..., n}

We store each connected component of the graph as a subset of [n] = {1, 2, . . . , n}. Adding one edge at a time and maintaining the sets of edges representing the components, we find that the graph is connected iff in the end there is only one set left, namely [n].

Disjoint sets: Algorithm

Connectedness detection algorithm:

Let's take a graph with n vertices: [n] = {1, 2, ..., n}

- Starting from n single sets, each containing one of the vertices, we replace two sets with their union if there is an edge in the graph connecting them
- For each successful union the number of components is reduced by one

We need two operations for this:

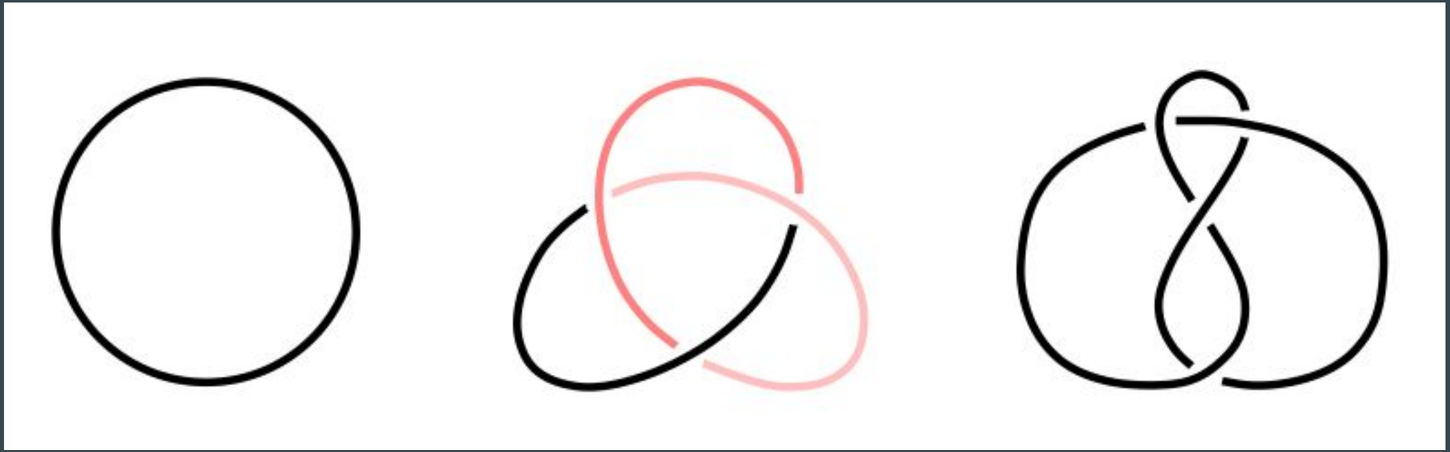
- Find (i) : returns the set that contains i: {i}
- Replace {i}, {j} with {i,j}: if (i, j) in set of edges replace the two sets by their union

Knots / Links

A closed curve embedded in R^3 does not decompose the space but it can be tangled up in inescapable ways. The field of mathematics that studies such tangles is knot theory. Its prime subject is a knot which is an embedding $\kappa : S^1 \rightarrow R^3$, that is, an injective, continuous function that is a homeomorphism onto its image. It turns out that any injective, continuous function from the $S^1 \rightarrow R^3$ is an embedding, but this is not true for general domains. A knot is equivalent to κ if it can be continuously deformed into κ without crossing itself during this process.

Knots / Links

The simplest knot is a circle, also known as a trivial knot.



The “unknot”, the “trefoil” knot, and the “figure-eight” knot.

Knots / Links (Reidemeister moves)

A knot can be deformed, by drawing its projections to a plane, keeping track of the under- and over-passes at crossings. For example, using the three fundamental Reidemeister moves one can prove that the trefoil cannot transform into the unknot:

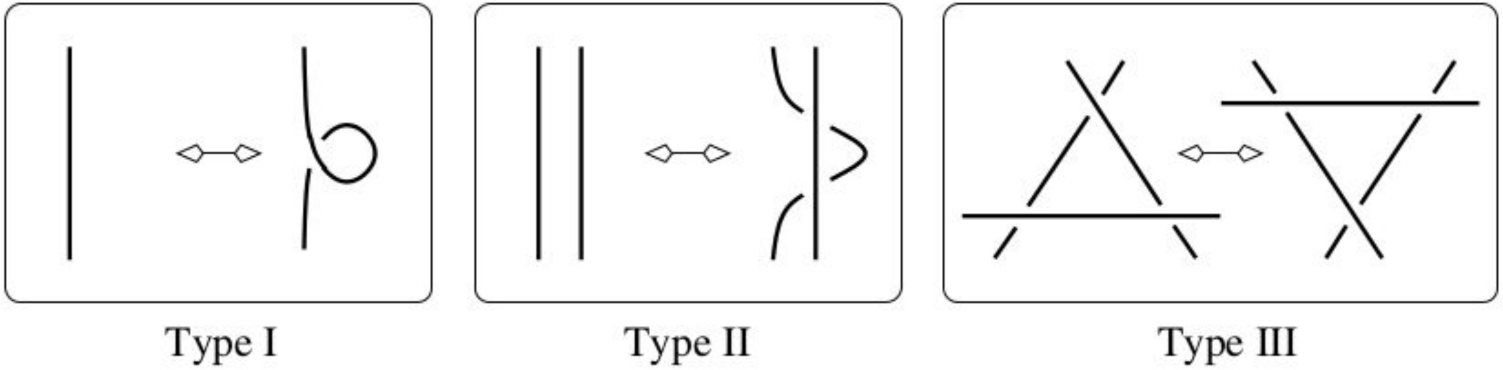
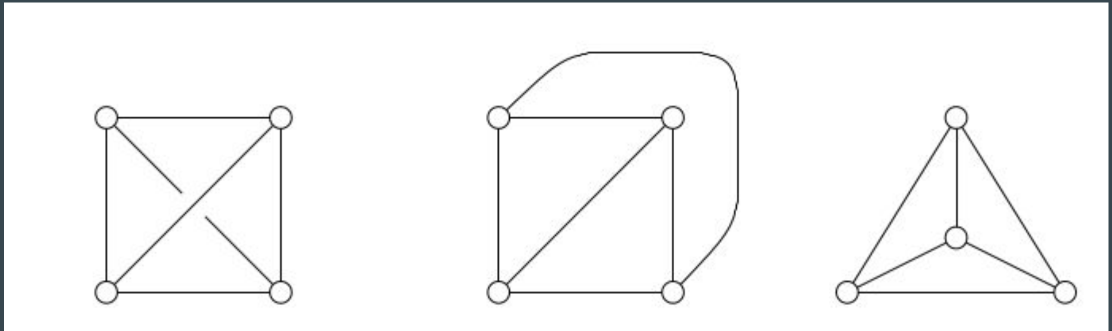


Figure I.8: The three types of Reidemeister moves.

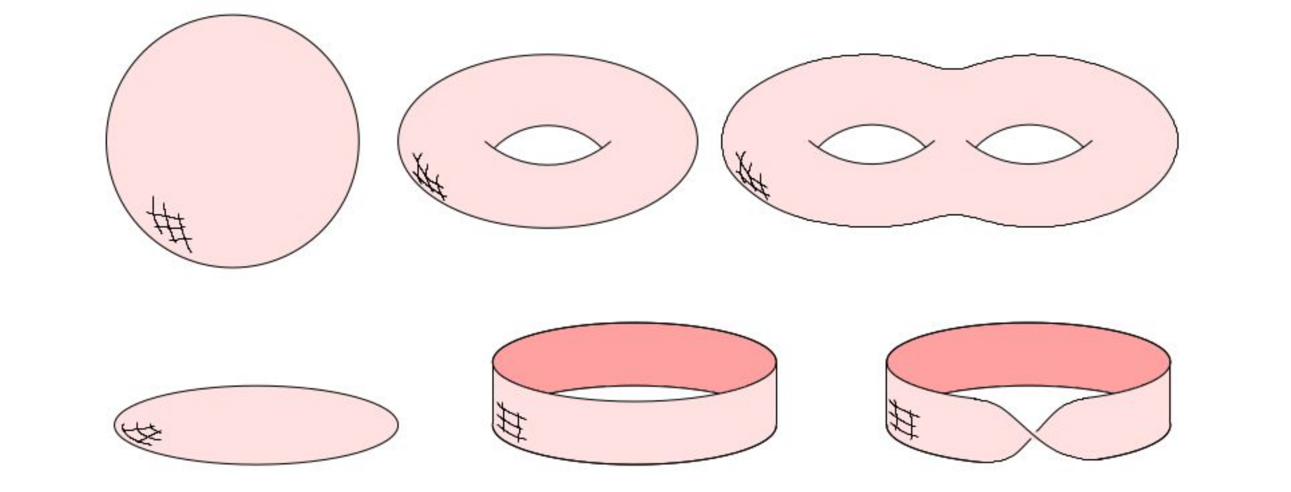
Planar graphs

Only graphs with relatively few edges can be drawn without crossings in the plane. Let $G = (V, E)$ be a simple, undirected graph. A *drawing* maps every vertex $u \in V$ to a point $f(u)$ in R2 and every edge $uv \in E$ to a path with endpoints $f(u)$ and $f(v)$. The drawing is an embedding if the points are distinct, the paths are simple and do not cross each other, and incidences are limited to endpoints. A graph is planar if it has an embedding in the plane, or if it can be drawn without crossings.



Surfaces

Consider the open disk of points at distance less than one from the origin, D. We will call any subset of a topological space that is homemorphic to D an open disk. A 2-manifold (without boundary) is a topological space M whose points all lie in open disk. Examples of compact 2-manifolds are shown below:



II. Complexes

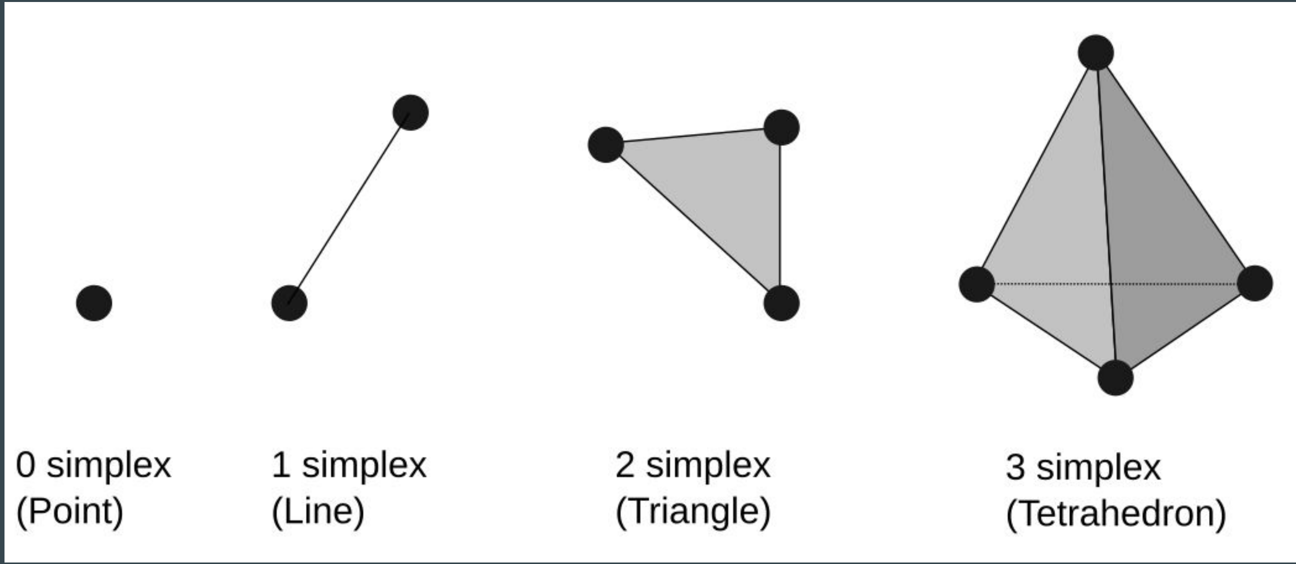
Simplices

In Topological Data Analysis (TDA), a simplex is a geometric object that generalizes the notion of a triangle to higher dimensions. A simplex of dimension k is defined as the convex hull of (k+1) affinely independent points in Euclidean space.

Simplices are fundamental building blocks in TDA, and are used to define simplicial complexes, which are collections of simplices that can be glued together to form a topological space.

Simplices

For example, a 0-dimensional simplex is a point, a 1-simplex is a line, a 2-simplex is a triangle and a 3-simplex is a tetrahedron.



Simplices

Simplices have well-understood topological properties (a dimension of the simplex fully defines its properties) that can be used to study the shape of complex objects. The number and dimension of simplices in a simplicial complex can provide information about the topology of the underlying space.

Simplices

In simplicial complexes, a face of a simplex is a subset of its vertices. For a k-simplex, which is defined as the convex hull of (k+1) affinely independent points, a face of the simplex is any subset of its vertices that can be obtained by removing one or more vertices.

For example, if we have a 2-simplex (i.e., a triangle) with vertices {v1, v2, v3}, then its faces are:

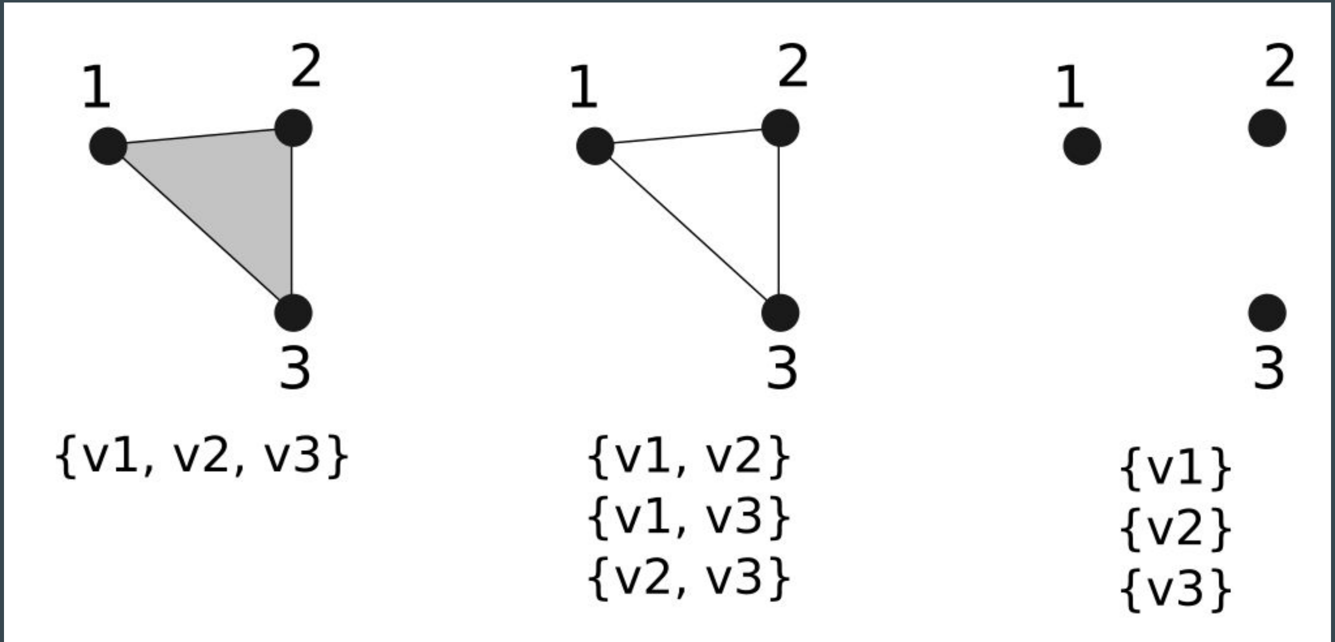
Its vertices: {v1}, {v2}, and {v3}.

Its edges: {v1, v2}, {v2, v3}, and {v3, v1}.

The whole simplex itself: {v1, v2, v3}.

Simplices

For example, if we have a 2-simplex (i.e., a triangle) with vertices {v1, v2, v3}, then its faces are:



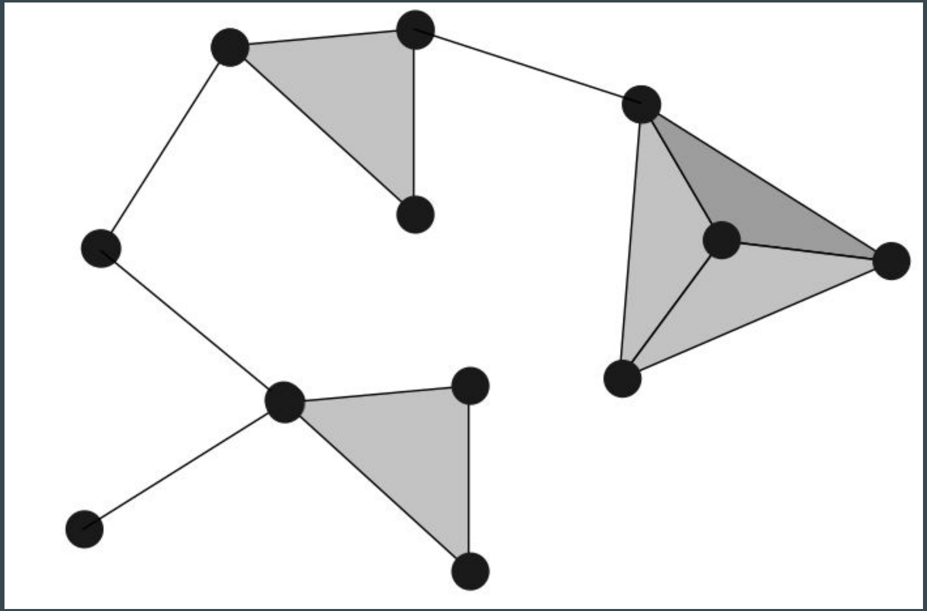
Simplices

Faces of simplices allow us to define the boundary of a simplex and the incidence relations between simplices in a simplicial complex. This relation allows us to compute the homology of a simplicial complex, which is a fundamental topological invariant.

Simplicial complexes

Definition. A simplicial complex is a finite collection of simplices K such that $\sigma \in K$ and $\tau \leq \sigma$ implies $\tau \in K$, and σ, $\sigma_0 \in K$ implies $\sigma \cap \sigma_0$ is either empty or a face of both.

Intuitively, a simplicial complex is a collection of simple building blocks (the simplices) that are glued together in a way that preserves their combinatorial structure.



Simplicial complexes

A simplicial complex K is a set of simplices that satisfies the following conditions:

1. Every face of a simplex from K is also a face in K.

2. The non-empty intersection of any two simplices σ_1 , σ_2 $\in$ K is a face of both σ_1 and σ_2.

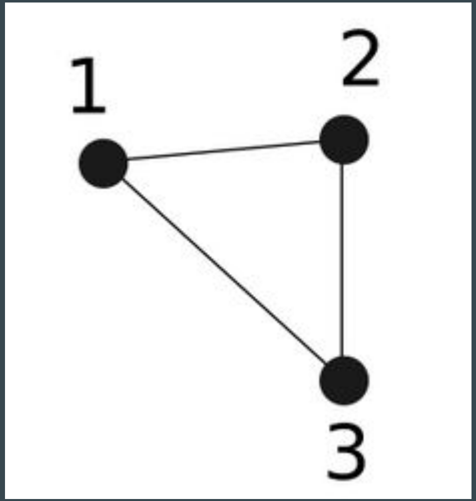
Simplicial complexes

Using simplicial complexes allows the computation of homology, a fundamental topological invariant that measures the number and dimension of the "holes" in a space. Homology can provide insight into the structure and connectivity of a data set, and can be used to identify topological features such as clusters and loops.

Simplicial complexes also provide a framework for defining and computing other topological invariants such as persistent homology, which measures the stability of topological features across different scales of the data.

Homology

Homology is a mathematical formalism about how a space is connected. Homology groups provide a mathematical language for the holes in a topological space. In fact, holes are described indirectly, by focusing on their boundaries. For example, a one-dimensional hole is described by a set of lines and the absence of their interior:



{v1, v2}, {v1, v3} and {v2, v3} in K

{v1, v2, v3} not in K

Homology

The homology of a simplicial complex is a fundamental topological invariant that measures the number and dimension of the "holes" in the complex. In algebraic topology, homology is a way to assign a sequence of homology groups to a topological space or a simplicial complex.

The k-th homology group of a simplicial complex measures the number of k-dimensional "holes" in the complex that cannot be filled by a (k+1)-dimensional surface or "bubble".

Homology

For example, the 0th homology group counts the number of connected components in the complex, the 1st homology group counts the number of independent loops or circles, and so on.

Homotopy

Homotopy is a relation between continuous maps from one space to another, which captures the idea of deformation or continuous transformation. Two maps are said to be homotopic equivalent if one can be continuously deformed into the other without tearing or cutting the space. Homotopy theory studies the properties of spaces that are invariant under homotopy, such as their fundamental group, which measures the number of loops that can be continuously shrunk to a point.

Homotopy

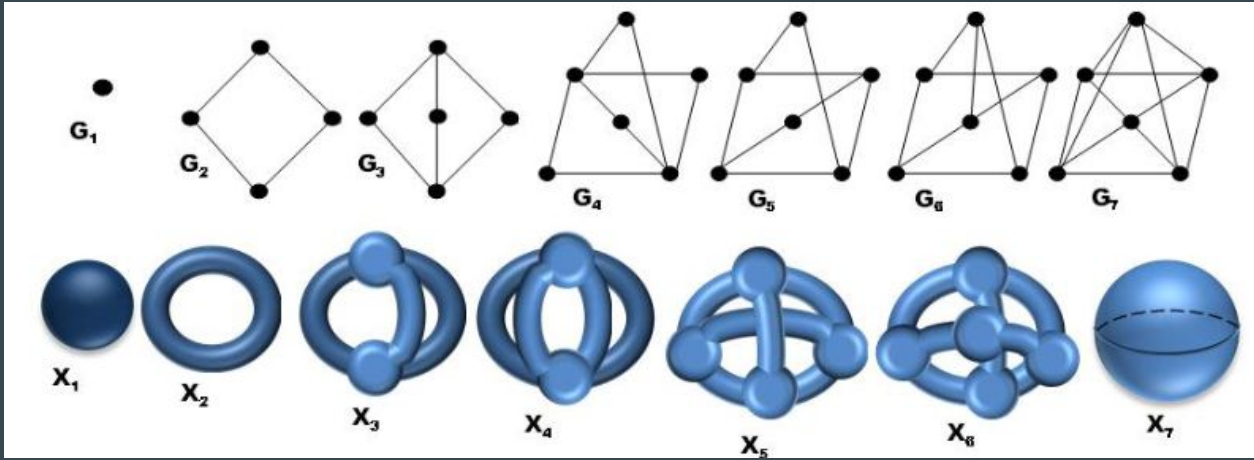
In TDA, homotopy equivalence relates two spaces or simplicial complexes that can be continuously deformed into each other.

Two spaces X and Y are said to be homotopy equivalent if there exist continuous maps f: X -> Y and g: Y -> X such that the compositions $g \circ f: X \to X$ and $f \circ g: Y \to Y$ are homotopic to the identity maps id_X and id_Y, respectively.

This gives a homotopy equivalence relation and we write $X \simeq Y$ if they have the same homotopy type.

Homotopy

Intuitively, a homotopy equivalence between two spaces means that they have the same topological structure. Any topological property that is invariant under continuous deformation is the same for both spaces. For example, homotopy equivalent spaces have the same fundamental group, which measures the number of loops that can be continuously shrunk to a point.



Evako 2015

Nerve theorem

Given a finite collection of sets, F, and without assuming that the sets are convex, we define the nerve to consist of all non-empty subcollections whose sets have a non-empty common intersection

$$\mathrm{Nrv}\, F \quad = \quad \{X \subseteq F \mid \bigcap X \neq \emptyset\}.$$

The nerve is always an abstract simplicial complex, no matter what sets we have in F.

Nerve theorem

Nerve Theorem.

Let F be a finite collection of closed, convex sets in Euclidean space. Then the nerve of F and the union of the sets in F have the same homotopy type.

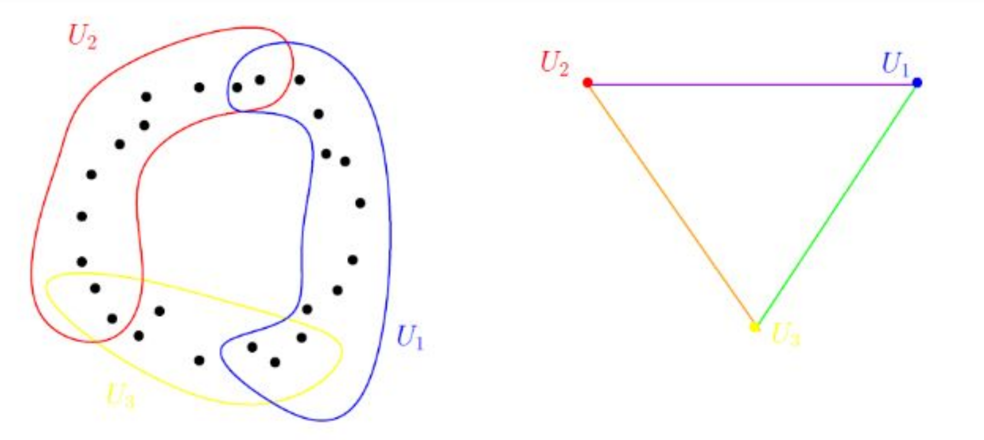


Figure 3: The nerve of a cover of a set of sampled points in the plane.

A collection of cover sets whose union is a triangle with one holes in the plane. The nerve is the boundary complex of the triangle which is homotopical equivalent

Čech complexes

The Čech complex is a simplicial complex constructed from a finite set of points in a metric space. It is used to approximate the underlying topological structure of the data set and to compute its persistent homology.

Čech complexes

The Čech complex is defined as follows:

For each subset of points in the data set whose pairwise distances are less than or equal to a fixed radius epsilon, we add a simplex to the complex.

The simplices correspond to the subsets of points of all dimensions, including the empty set and individual points.

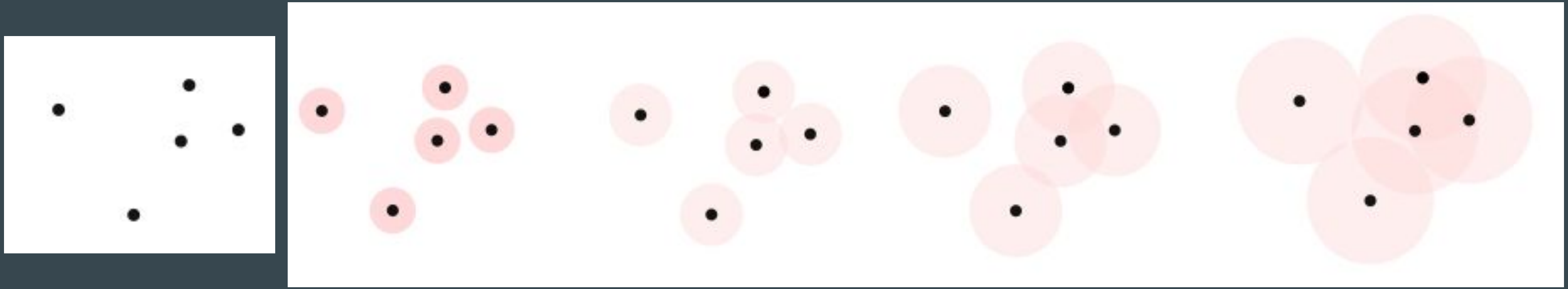
Čech complexes

Let S be a finite set of points in R^d and write $B_x(r) = x + r\,B^d$ for the closed ball with center x and radius r. The Čech complex of S and radius r is isomorphic to the nerve of this collection of balls:

$$\check{\mathrm{C}}\mathrm{ech}(r) \quad = \quad \{\sigma \subseteq S \mid \bigcap_{x \in \sigma} B_x(r) \neq \emptyset\}.$$

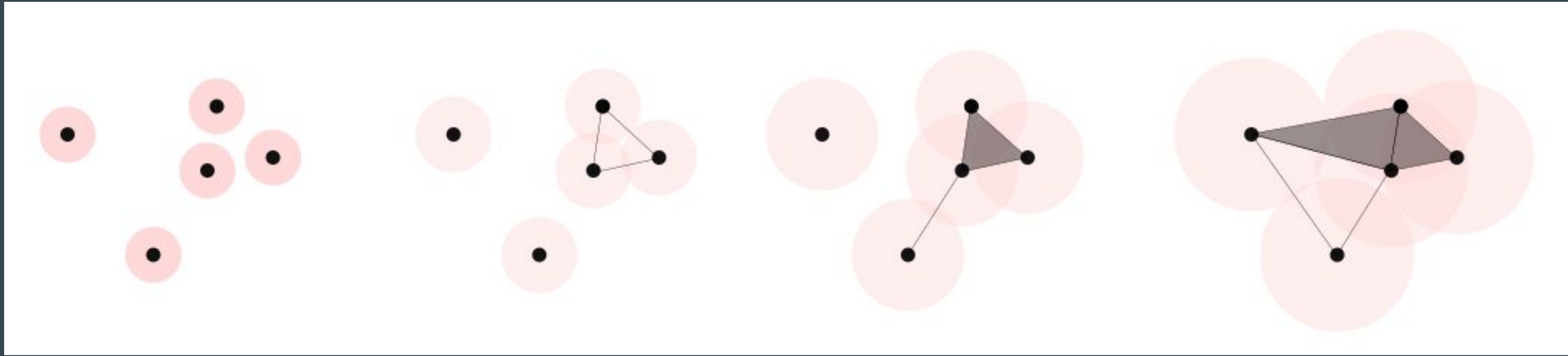
Čech complexes

For a collection of points (in 2D), we add balls around each point of radius ε, for increasing values of ε.



Čech complexes

In Čech complex, two points are connected if their pairwise distance is less than epsilon. Similarly, for three points to be connected and form a triangle, all epsilon circles should intersect.



Vietoris - Rips complexes

The Vietoris-Rips complex is defined as follows:

For each pair of points in the data set whose pairwise distance is less than or equal to a fixed radius epsilon, we add an edge to the complex.

For each subset of points of size k whose pairwise distances are less than or equal to epsilon, we add a k-simplex to the complex.

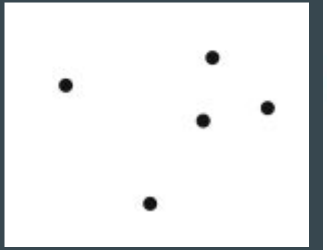
Vietoris - Rips complexes

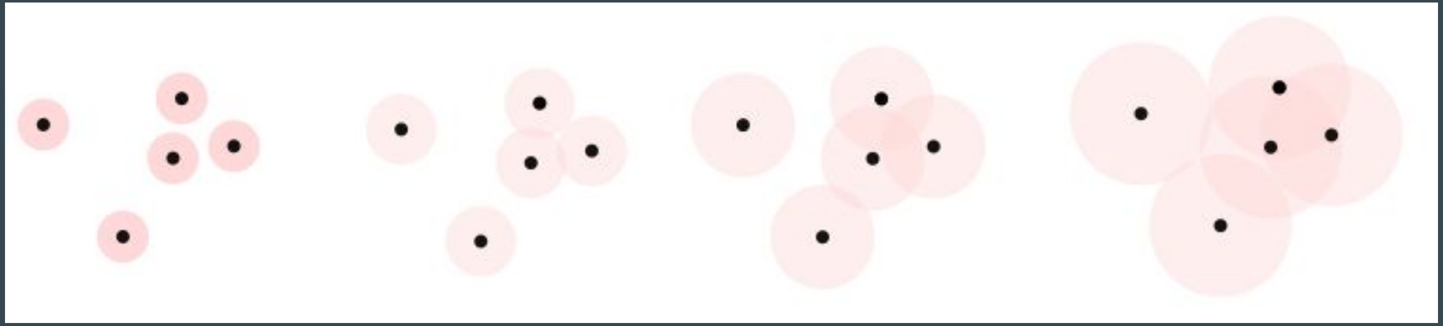
Instead of checking all subcollections, we just check pairs and add 2- and higher-dimensional simplices whenever we can. This simplification leads to the Vietoris-Rips complex of S and r consisting of all subsets of diameter at most 2r:

$$\text{Vietoris-Rips}(r) = \{\sigma \subseteq S \mid \operatorname{diam} \sigma \leq 2r\}.$$

Vietoris - Rips complexes

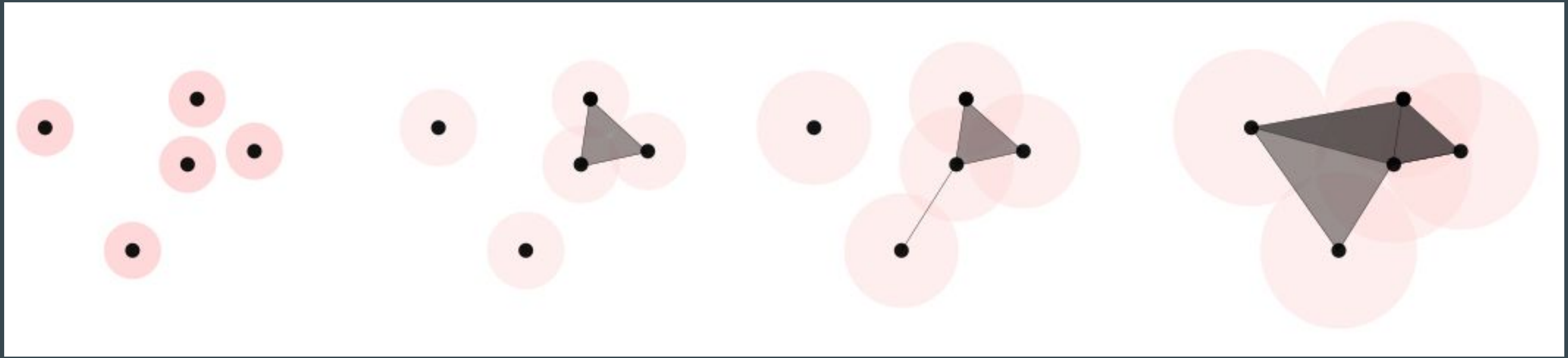
For a collection of points (in 2D), we add balls around each point of radius ε, for increasing values of ε.





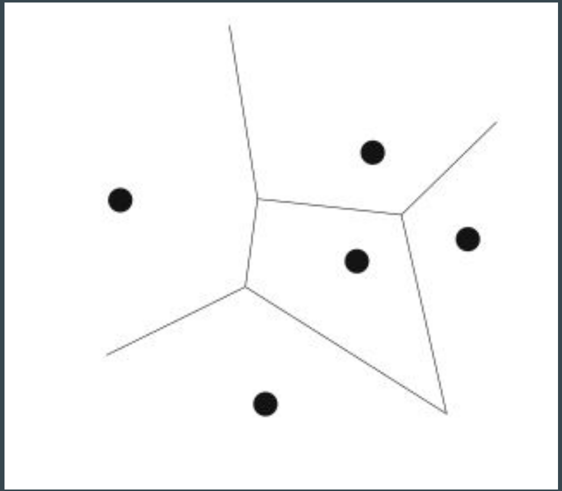
Vietoris - Rips complexes

In Vietoris-Rips complex, two points are connected if their pairwise distance is less than epsilon. Each higher dimensional simplex is added to the simplicial complex, as long as all the points are already connected by lines.



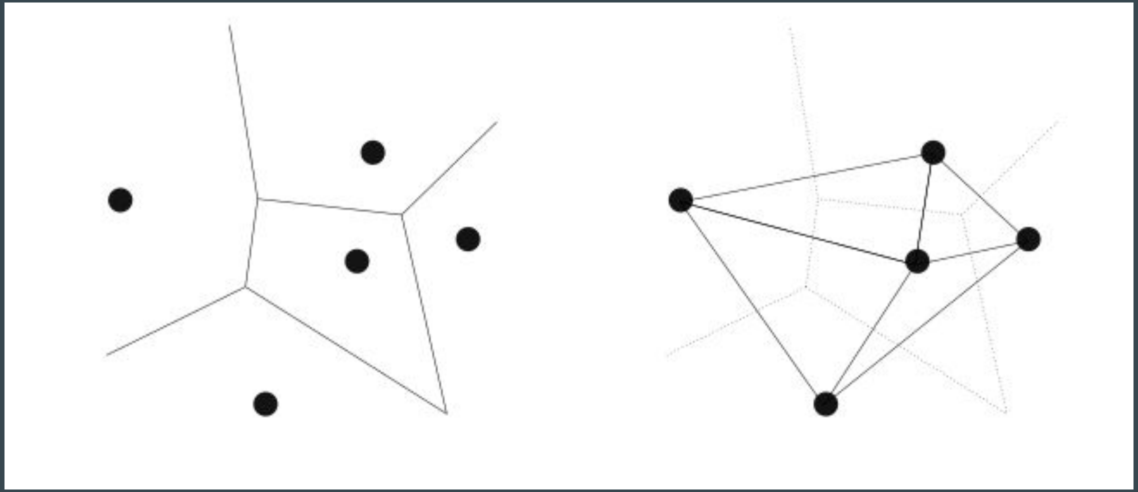
Delaunay complexes

To compute the Delaunay complex, one first constructs the Voronoi diagram, which is a simplicial complex formed by connecting points that are equidistant from each other:



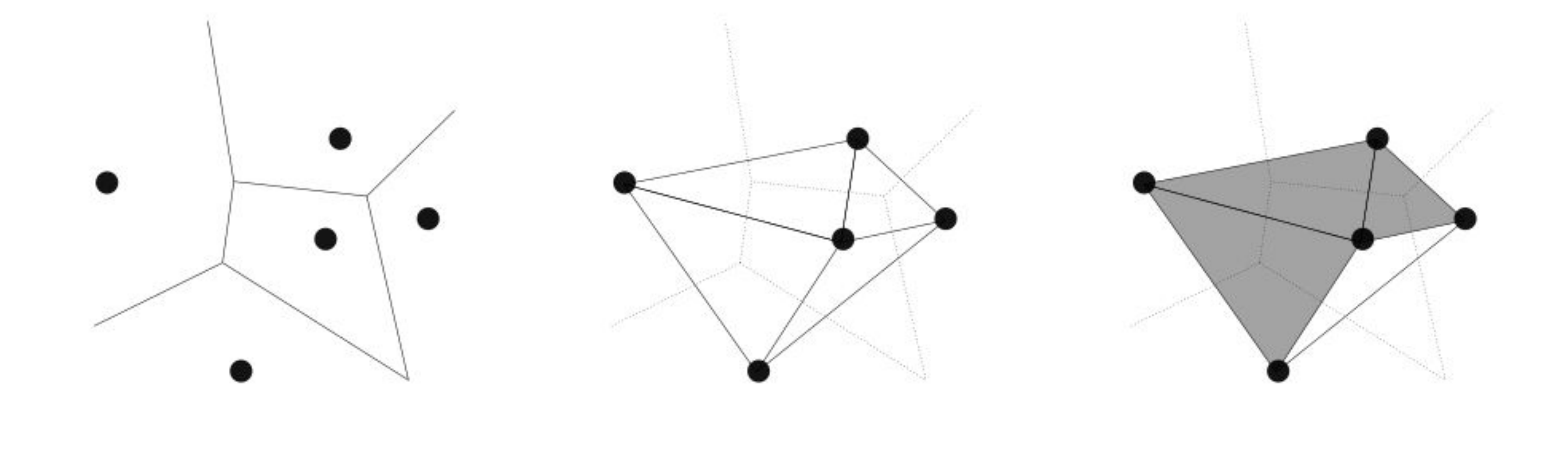
Delaunay complexes

The second step is to generate the Delaunay triangulation of the data set by connecting points that have at least one intersecting face:



Delaunay complexes

The Delaunay complex is generated by filling in the connected components of the Delaunay triangulation, with higher order intersections:



Delaunay complexes

The Delaunay complex of a finite set $S \subseteq R^d$ is isomorphic to the nerve of the Voronoi diagram:

$$\text{Delaunay} \quad = \quad \{\sigma \subseteq S \mid \bigcap_{u \in \sigma} V_u \neq \emptyset\}.$$

Alpha complexes

The Alpha complex is a simplicial complex constructed from the finite cells of a Delaunay Triangulation. The Delaunay complex is an upper bound for the alpha - complex.

The alpha complex can be seen as a combination of the Delaunay complex and the Vietoris-Rips complex: it is a subset of the Vietoris-Rips complex formed by removing simplices whose circum-spheres contain other points of the data set, and a subset of the Delaunay complex formed by removing simplices whose vertices do not form a convex set.

Alpha complexes

The Alpha complex is defined based on a radius alpha that is used as a constraint to generate a family of subcomplexes from the Delaunay complex.

For each subset of points whose intersection with the union of the open balls centered at the points of the subset is non-empty and contractible, we add a simplex to the complex. The simplices correspond to the subsets of points of all dimensions, including the empty set and individual points.

Alpha complexes

Let S be a finite set of points in R^d and r a non-negative real number. For each $u \in S$, we let $B_u(r) = u + rB^d$ be the closed ball with center u and radius r. The union of these balls is the set of points at distance at most r from at least one of the points in S.

To decompose the union, we intersect each ball with the corresponding Voronoi cell, $R_u(r) = B_u(r) \cap V_u$.

Since balls and Voronoi cells are convex, the $R_u(r)$ are also convex. Any two of them are disjoint or overlap along a common piece of their boundaries, and together the $R_u(r)$ covers the entire union.

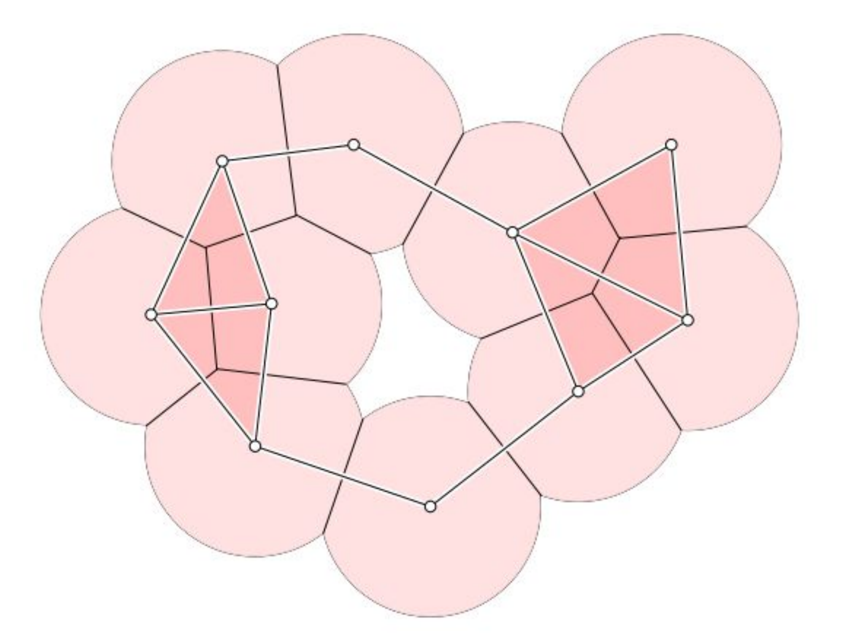
Alpha complexes

The Alpha complex is isomorphic to the nerve of this cover:

$$\mathrm{Alpha}(r) \quad = \quad \{\sigma \subseteq S \mid \bigcap_{u \in \sigma} R_u(r) \neq \emptyset\}$$

Alpha complexes

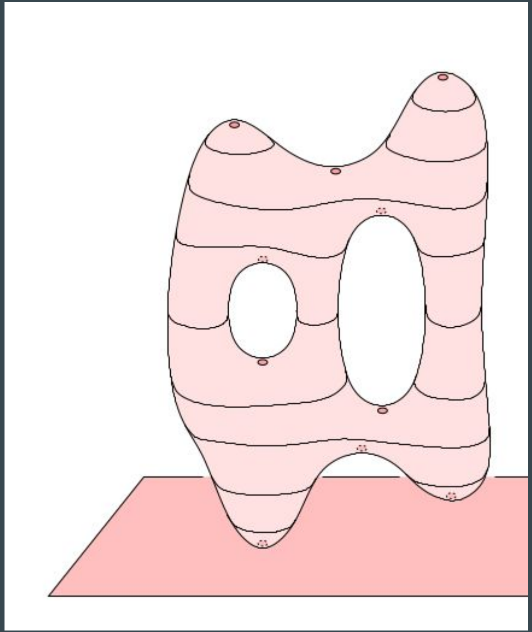
The union of disks is decomposed into convex regions by the Voronoi diagram. The corresponding alpha complex is superimposed.



IV. Morse functions

Morse function

Definition. A Morse function is a smooth function on a manifold, $f : M \to \mathbb{R}$, such that (i) all critical points are non-degenerate, and (ii) the critical points have distinct function values.

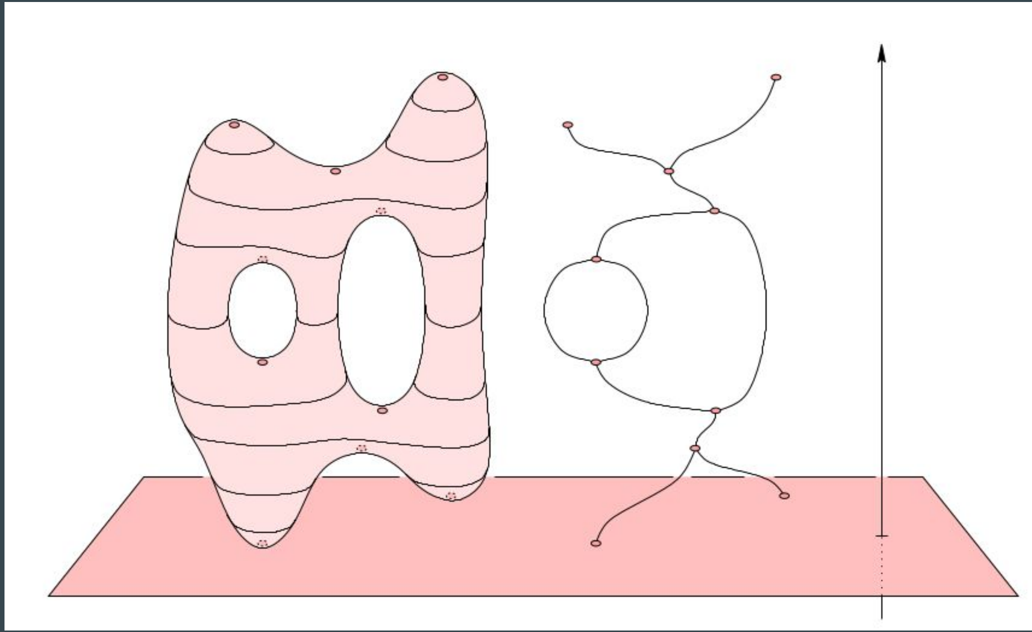


Reebs graphs

Reeb graphs of Morse functions. More can be said if $X = M$ is a manifold of dimension $d \geq 2$ and $f : M \to R$ is a Morse function. Recall that each point $u \in R(f)$ is the image of a contour in M. We call u a node of the Reeb graph if $\psi^{-1}(u)$ contains a critical point or, equivalently, if u is the image of a critical point under ψ. By definition of Morse function, the critical points have distinct function values, which implies a bijection between the critical points of f and the nodes of $R(f)$.

Reebs graphs

The Reebs graph of a Morse function is a collection of critical points and their connectivity is define by their ordering according to a predefined direction



V. Persistence

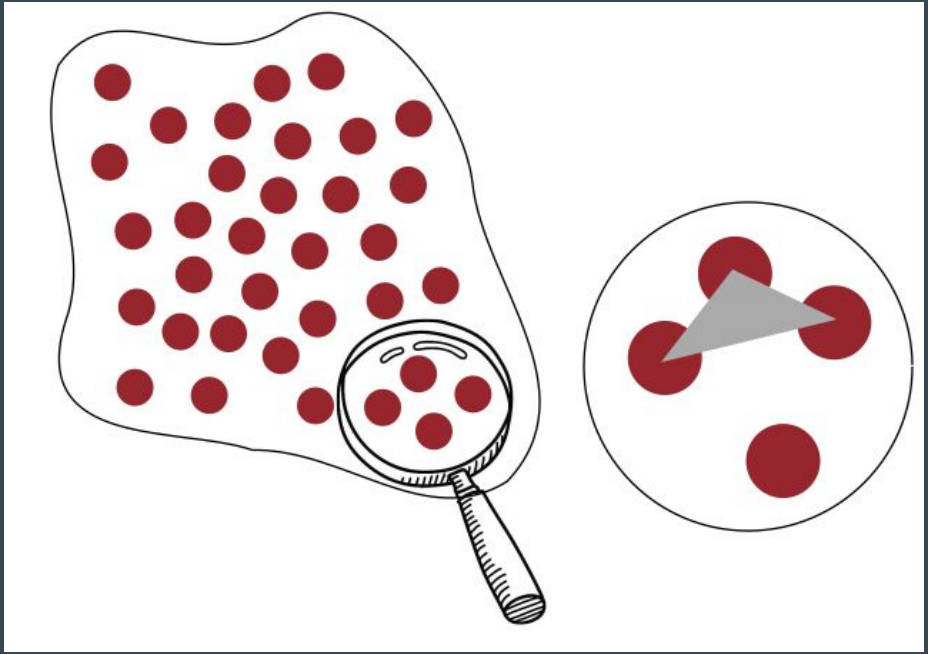
Persistence homology

The central concept of persistence is motivated by the practical need to cope with noise in data. This includes defining, recognizing, and possibly eliminating noise. However, the distinction between noise and feature is not well-defined but rather a subjective notion. For any particular case, the focus is on a range of scales and it is desired to ignore everything that is smaller or larger.

Persistence homology

You can imagine TDA as a magnifying glass that focuses on a specific aspect of the data. This can mean a focus on:

- A length scale
- The skeleton of the data
- The largest components in a structure



Elder rule

The Elder rule is a heuristic in TDA that is used to filter out simplices from a simplicial complex that are unlikely to be significant in terms of the underlying topology of the data set.

The rule is based on the intuition that simplices that appear early in a filtration are more likely to be topologically significant than simplices that appear later.

Elder rule

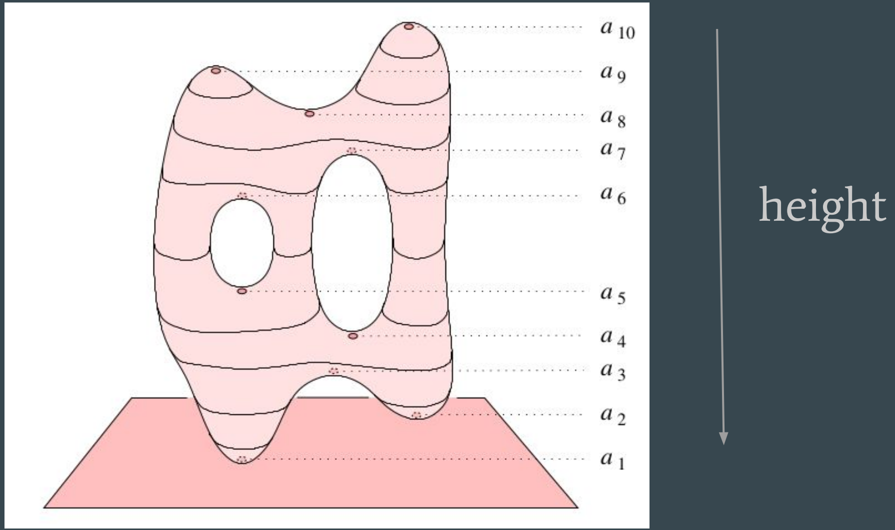
Formally, the Elder rule is implemented as follows:

At a juncture, the older of the two merging paths continues and the younger path ends.

By applying the Elder rule to a filtration (for example The Vietoris-Rips complex), the homology of each simplex in the complex is computed. For each simplex, we record the filtration level of the first complex at which the simplex appears.

Filtrations

In TDA, a filtration is a sequence of simplicial complexes that is used to track the evolution of the topological features of a data set. The simplicial complexes in the filtration are ordered by inclusion, meaning that each complex in the sequence contains all the simplices of the previous complexes.



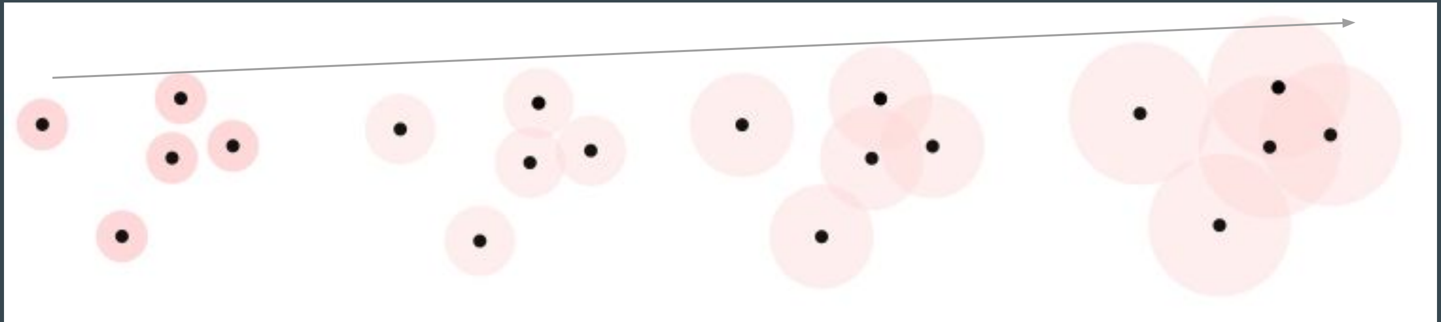
Filtrations

We obtain persistence by formulating the Elder Rule for the homology groups of all dimensions. Consider a simplicial complex, K, and a function $f : K \to R$. We require that f be monotonic by which we mean it is non-decreasing along increasing chains of faces, that is, $f(\sigma) \leq f(\tau)$ whenever σ is a face of τ. Monotonicity implies that the sublevel set, $K(a) = f^{-1}(-\infty, a]$, is a subcomplex of K for every $a \in R$. Letting m be the number of simplices in K, we get $n + 1 \leq m + 1$ different subcomplexes, which we arrange as an increasing sequence:

$$\emptyset = K_0 \subseteq K_1 \subseteq \ldots \subseteq K_n = K.$$

Filtrations

For example, for Vietoris-Rips and Čech complexes, the filtration is the increasing spheres' radii:



Persistence barcodes

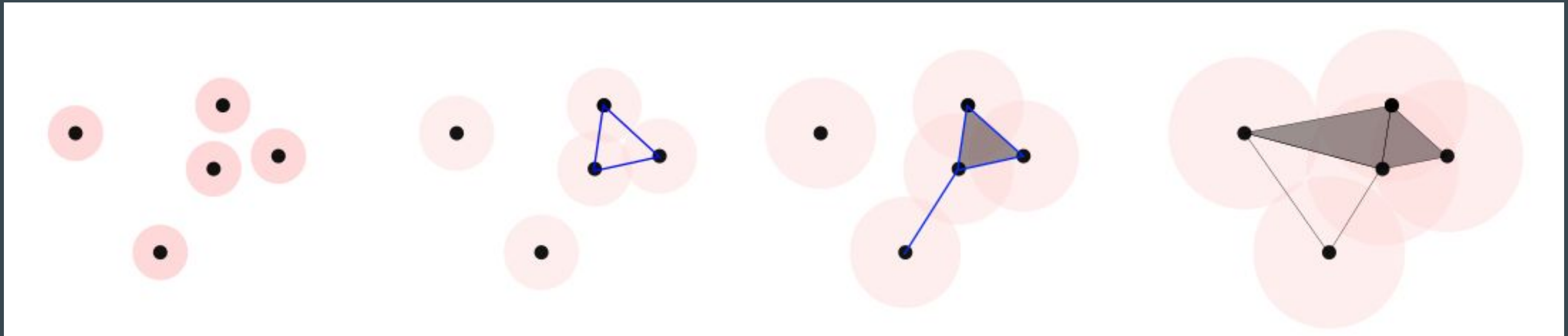
Persistence keeps track of the first time (birth time) a simplex is observed in a simplicial complex, and the time (death time) it merges with a larger component, according to a reference parameter (for example, epsilon in Vietoris-Rips).

A persistence barcode is a graphical representation of the persistence of topological features in a data set and it provides a summary of the birth and death times of each topological feature, as well as its lifetime, which reflects its topological significance.

Persistence barcodes

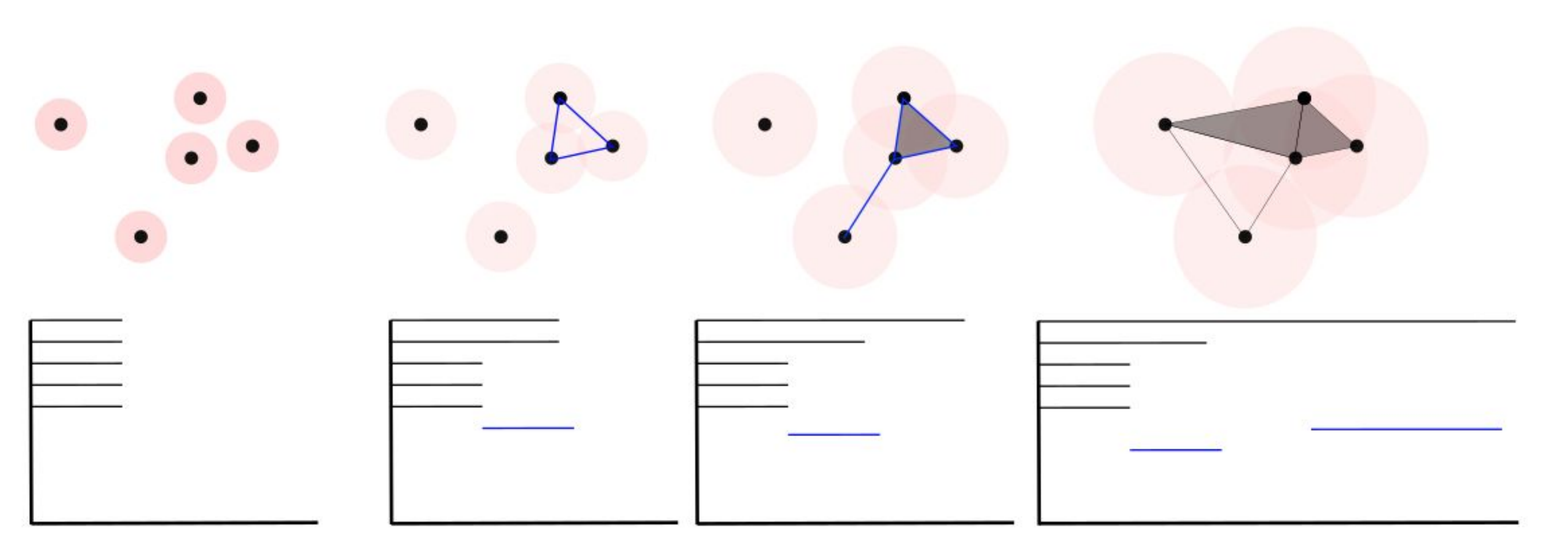
For a concrete example, assume that $R = (R_{\varepsilon i})^N$ is a sequence of complexes associated to a fixed point cloud for an increasing sequence of parameter values $(\varepsilon_i)^N$. There are natural inclusion maps:

$$\mathcal{R}_1 \overset{\iota}{\hookrightarrow} \mathcal{R}_2 \overset{\iota}{\hookrightarrow} \cdots \overset{\iota}{\hookrightarrow} \mathcal{R}_{N-1} \overset{\iota}{\hookrightarrow} \mathcal{R}_N$$



Persistence barcodes

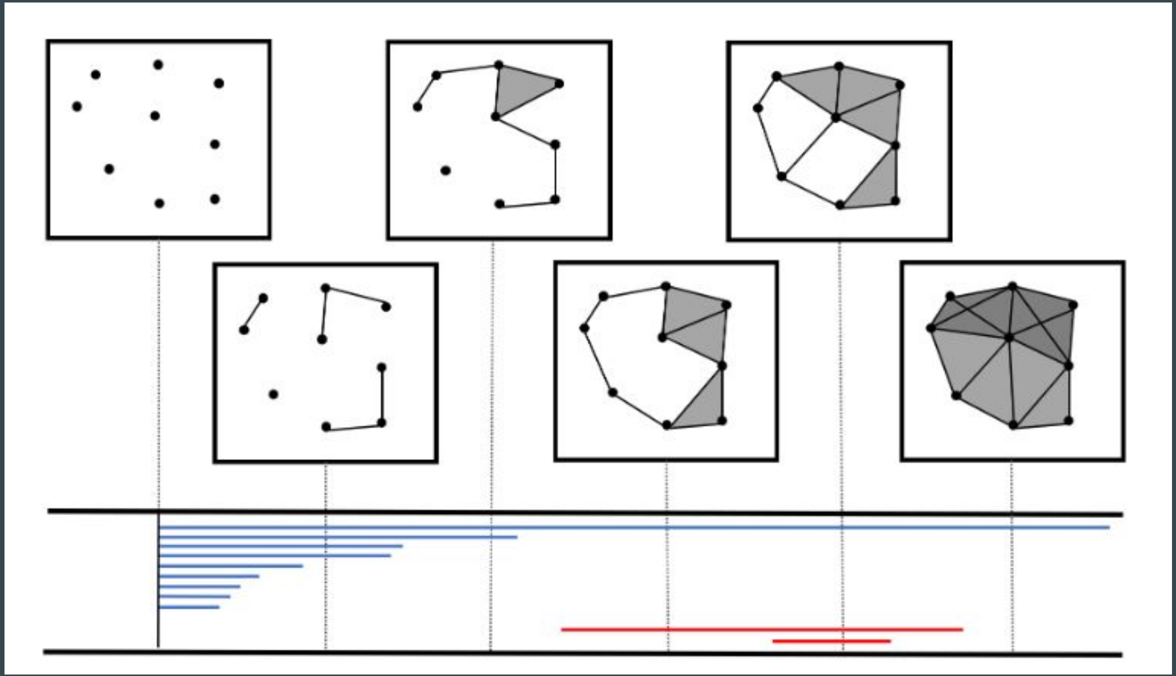
Keeping a record of the birth and death times of each component we can generate a persistence barcode that describes the lifetime of each topological component



Persistence barcodes

Point clouds have been analyzed for a filtration that generates a sequence of simplicial complexes. A connected component (blue) that persists infinitely is generated at the first step. At 4th step a cycle (in red) is formed that turns into two (5th step) and disappears when all points are connected.

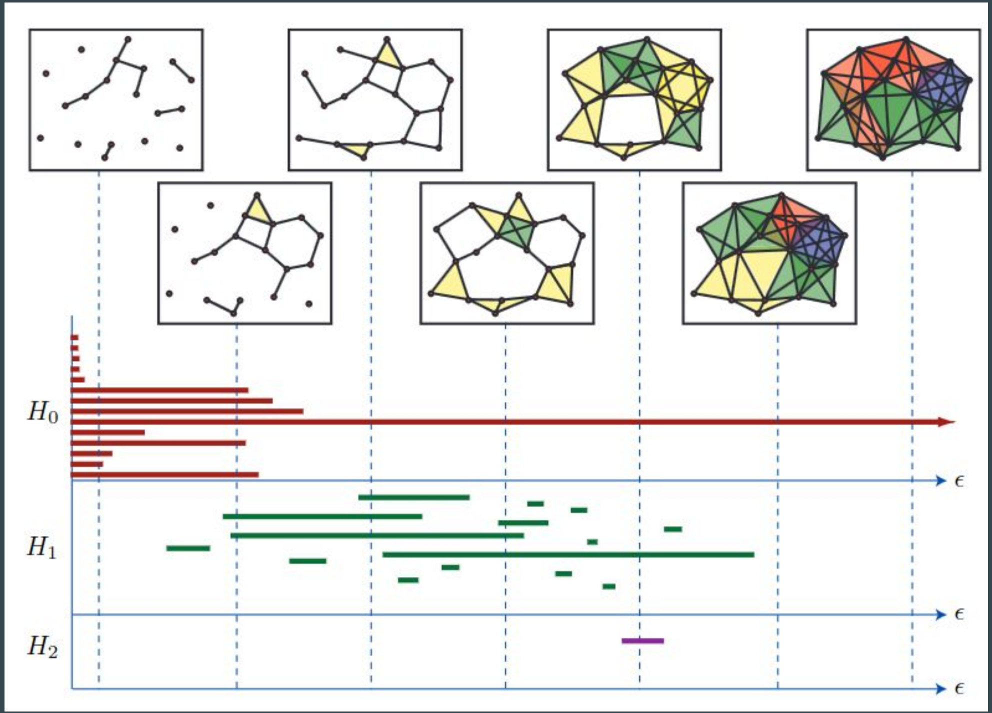
Chevyrev et al. 2018



Persistence barcodes

A more advanced example of the barcodes for a filtration that generates a sequence of simplicial complexes.

The rank of $H_k(R_{\varepsilon i})$ (top), equals the number of intervals in the barcode (bottom) for $H_k(R)$ intersecting the (dashed) line



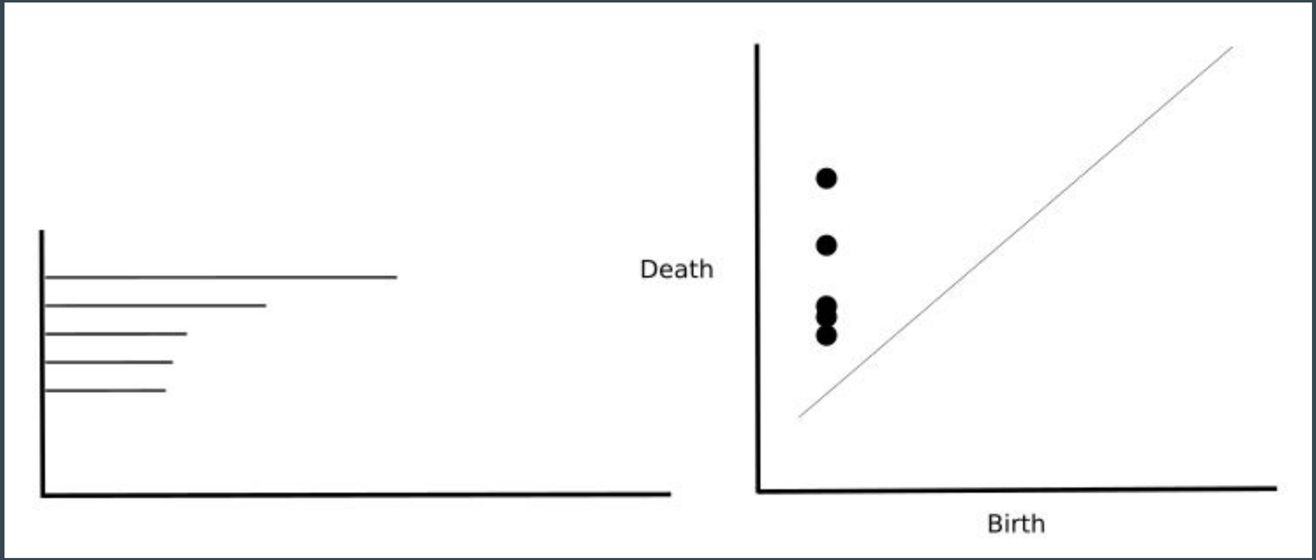
Robert Ghrist 2008

Persistence diagrams

Equivalently to the persistence barcodes, the persistence diagram is a graphical representation of the persistence of topological features in a data set. The persistence barcode presents the lifetime of the components as a line (bar), while the persistence diagram presents the birth and death times of each component in a two dimensional plane.

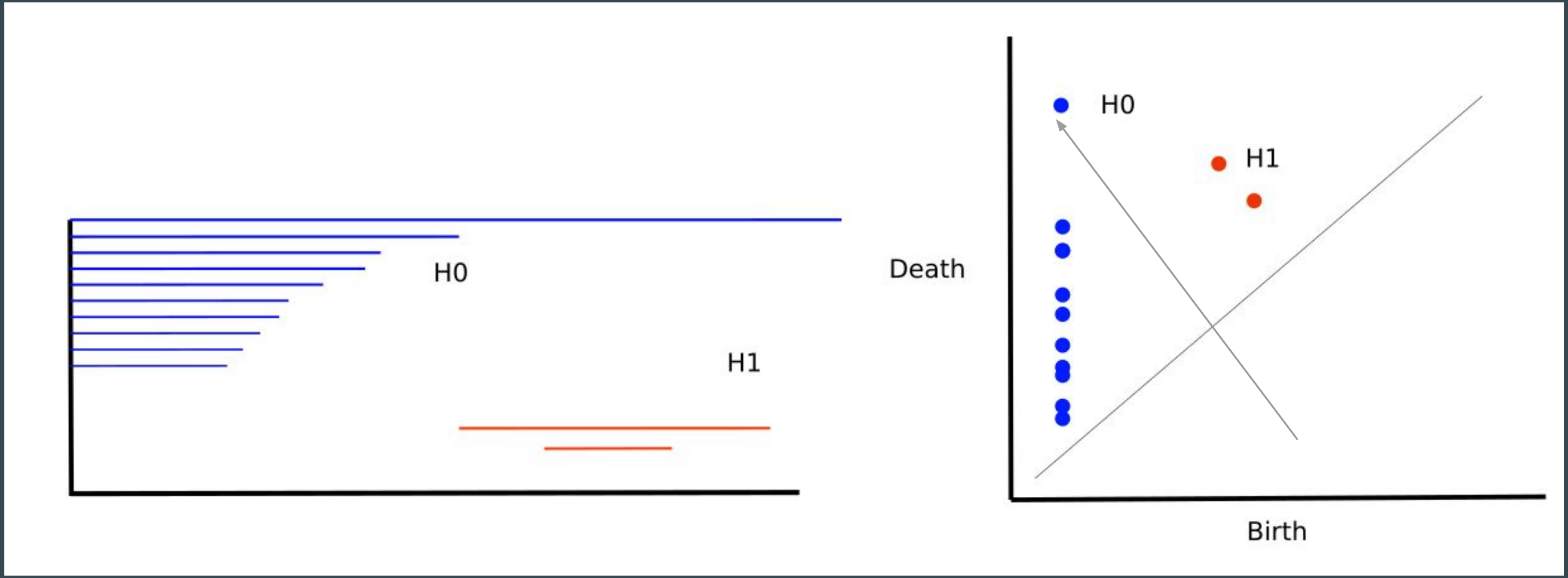
Persistence diagrams

For example, the lifetimes of a four bar barcode can be equivalently represented in a 2D diagram which encodes start - end times.



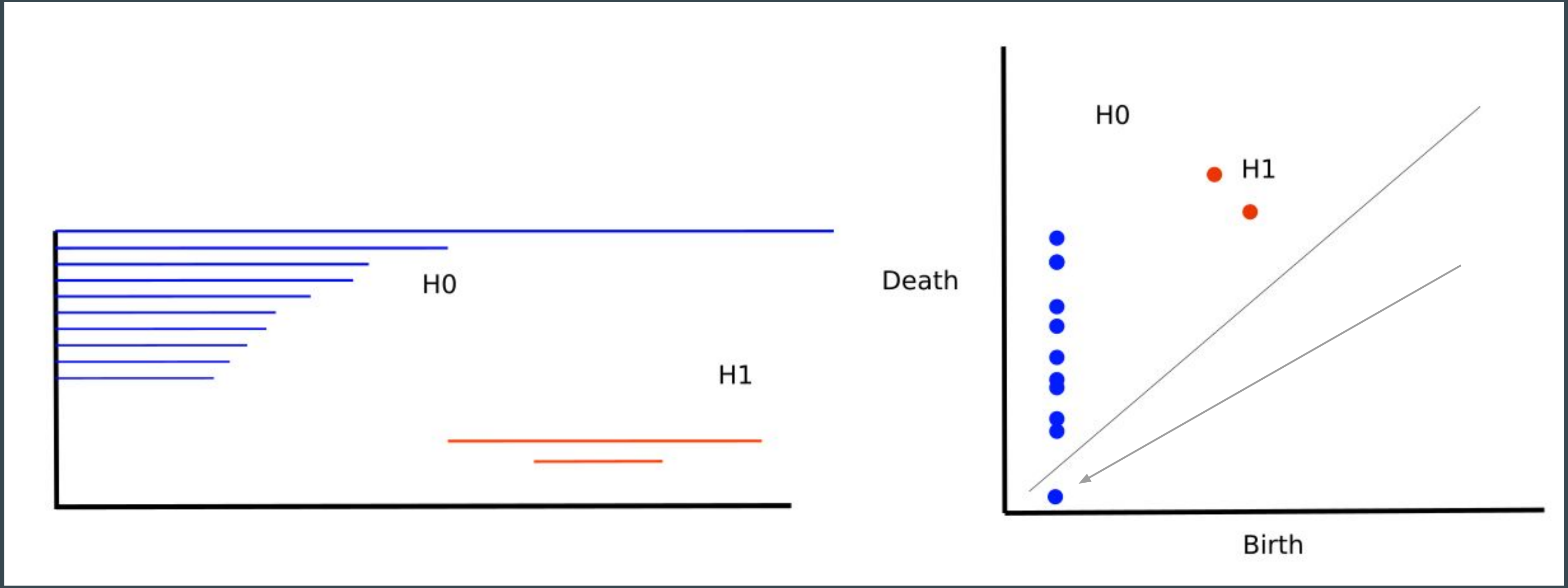
Persistence diagrams

For barcodes of data with multiple dimensions (blue: components, red: loops) the infinite component can be either a large number “inf” or be set to “-1”



Persistence diagrams

For barcodes of data with multiple dimensions (blue: components, red: loops) the infinite component can be either a large number “inf” or be set to “-1”



Extracting insights from the shape of complex data using topology

Lum et al. 2013

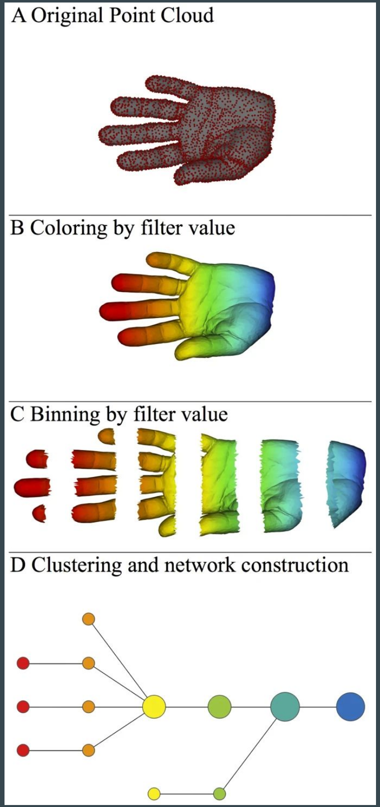
Extracting insights from the shape of complex data using topology

This paper applies topological methods to study complex high dimensional data sets by extracting shapes (patterns) and obtaining insights about them. Our method combines the best features of existing standard methodologies such as principal component and cluster analyses to provide a geometric representation of complex data sets. Through this hybrid method, we often find subgroups in data sets that traditional methodologies fail to find. Our method also permits the analysis of individual data sets as well as the analysis of relationships between related data sets. We illustrate the use of our method by applying it to three very different kinds of data, namely gene expression from breast tumors, voting data from the United States House of Representatives and player performance data from the NBA, in each case finding stratifications of the data which are more refined than those produced by standard methods.

Extracting insights from the shape of complex data using topology

Mathematical underpinnings of topological data analysis (TDA)

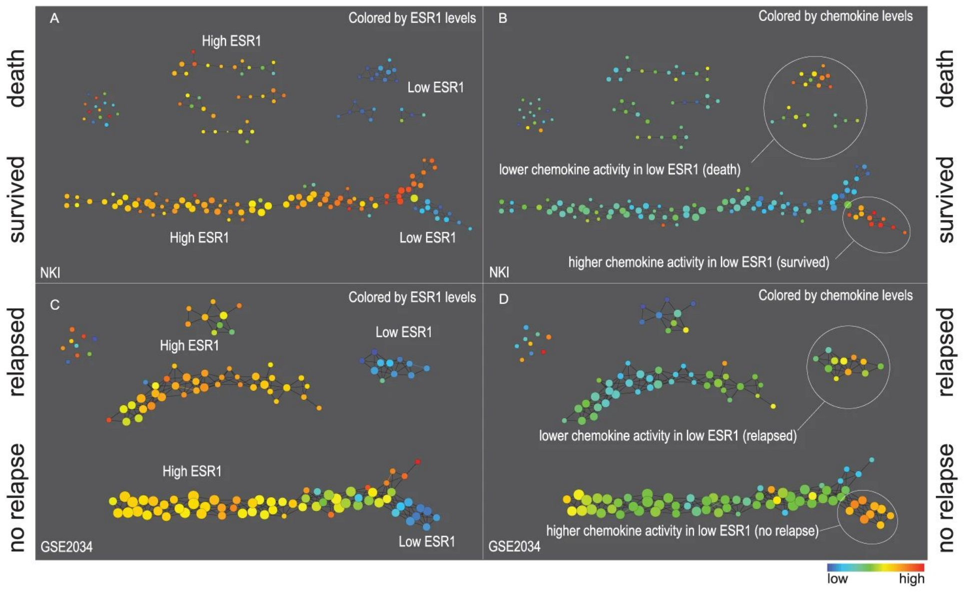
The approach as applied to a data set in our analysis pipeline. A) A 3D object (hand) represented as a point cloud. B) A filter value is applied to the point cloud and the object is now colored by the values of the filter function. C) The data set is binned into overlapping groups. D) Each bin is clustered and a network is built.



Extracting insights from the shape of complex data using topology

Identifying patient outcome in breast cancer

Two filter functions, L-Infinity centrality and survival or relapse were used to generate the networks. The top half of panels A and B are the networks of patients who didn't survive, the bottom half are the patients who survived. Panels A and C are colored by the average expression of the ESR1 gene. Panels B and D are colored by the average expression of the genes in the KEGG chemokine pathway.

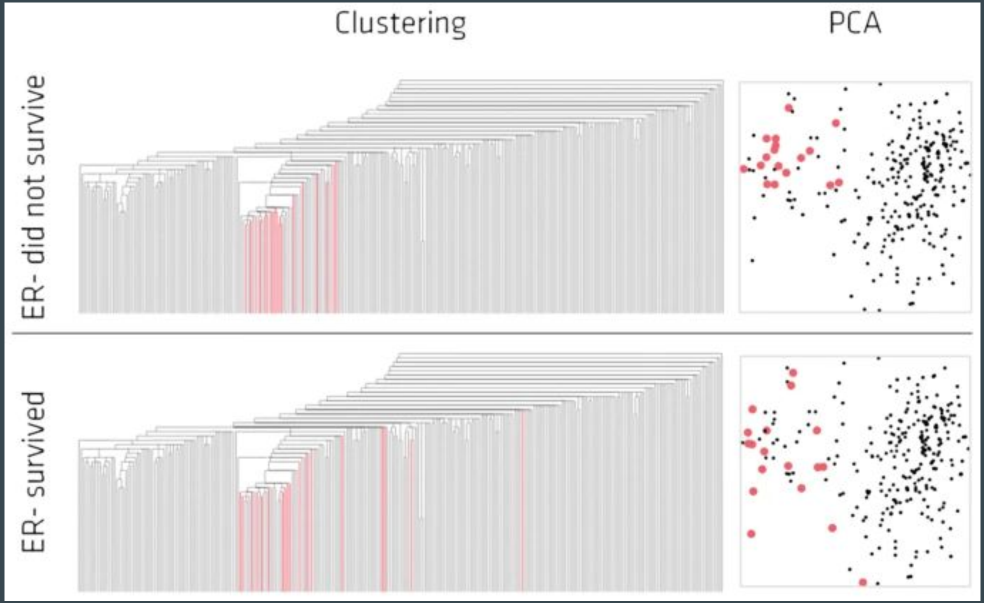


Extracting insights from the shape of complex data using topology

Identifying patient outcome in breast cancer

Single linkage hierarchical clustering and PCA of the NKI data set.

Highlighted in red are the lowERNS (top panel) and the lowERHS (bottom panel) patient subgroups.

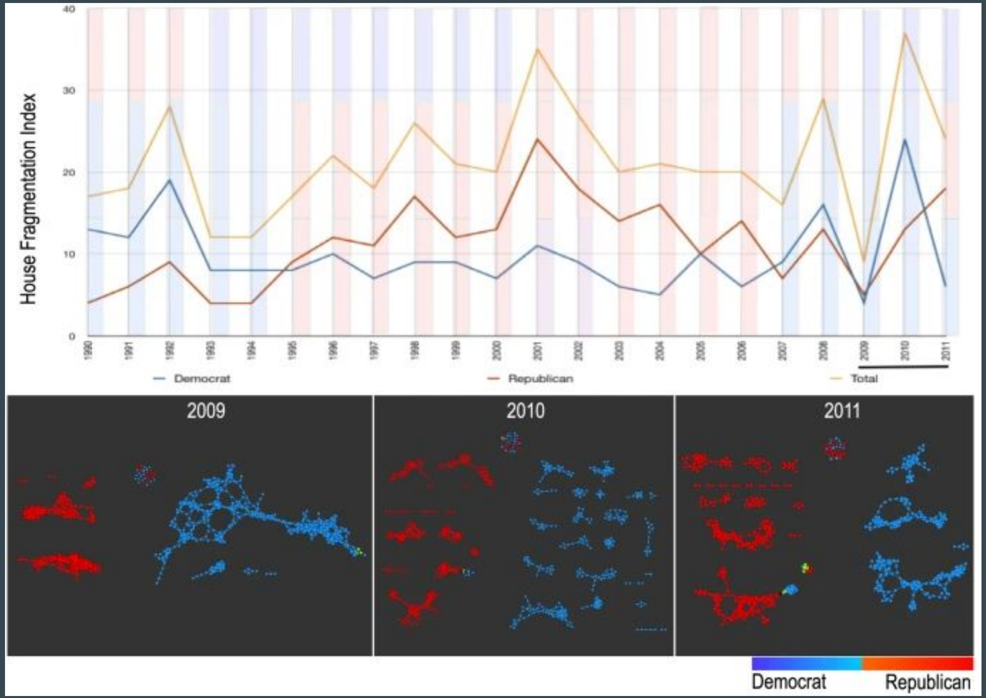


Extracting insights from the shape of complex data using topology

US House of Representatives based on voting behavior

Top panel is the fragmentation index calculated from the number of sub-networks formed each year per political party.

The bottom 3 panels are the topological networks for the members. Networks are constructed from voting behavior of the member of the house, (1,0,-1).



Extracting insights from the shape of complex data using topology

Basketball team stratification

A) Low resolution map at 20 intervals for each filter

B) High resolution map at 30 intervals for each filter.

The graphs are colored by points per game and a variance normalized Euclidean distance metric is applied.