Topology for statistical analysis of brain artery images

Ezra Miller

Duke University, Department of Mathematics

ezra@math.duke.edu

joint with

Paul Bendich & Alex Pieloch (Duke Math)
J.S. Marron & Sean Skwerer (Chapel Hill Stat/Oper.Res.)

Young Mathematicians Conference
The Ohio State University

12 August 2017

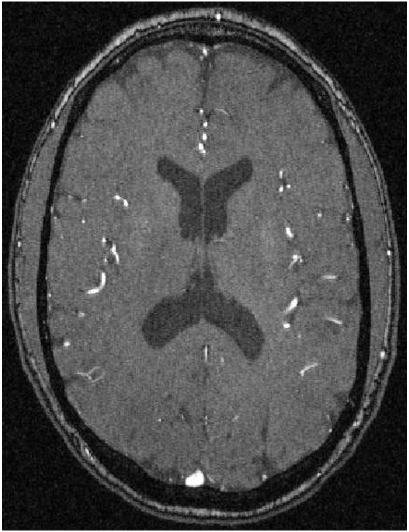
Outline

Brain artery trees

Goal: Statistical analysis taking 3D geometry into account

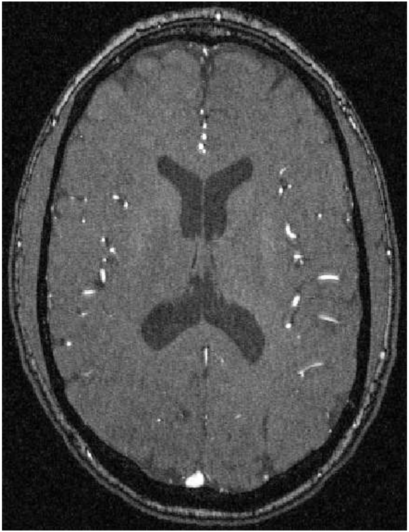
- predict stroke tendency
- screen for loci of pathology, such as tumors
- explore how age affects vascularization

Magnetic Resonance Angiography (MRA)



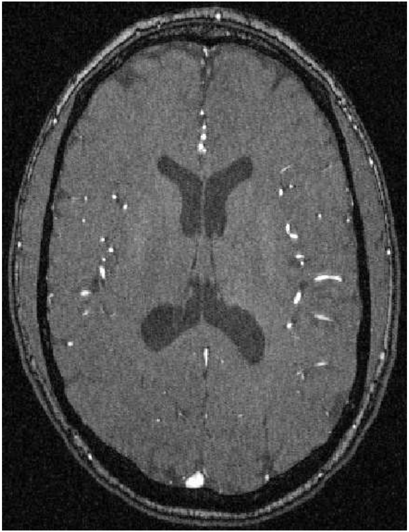
from Elizabeth Bullitt, Dept. of Neurosurgery, UNC-CH

Magnetic Resonance Angiography (MRA)



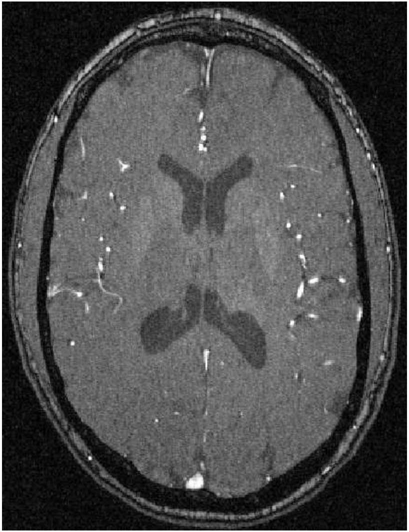
from Elizabeth Bullitt, Dept. of Neurosurgery, UNC-CH

Magnetic Resonance Angiography (MRA)



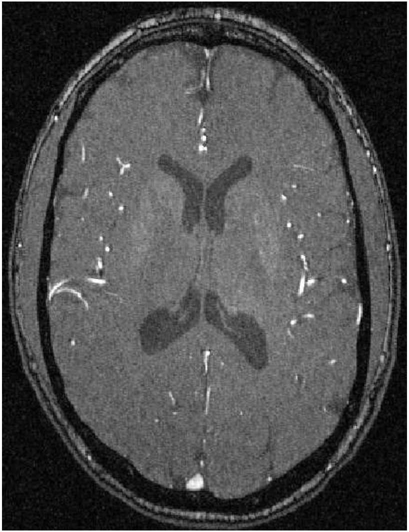
from Elizabeth Bullitt, Dept. of Neurosurgery, UNC-CH

Magnetic Resonance Angiography (MRA)



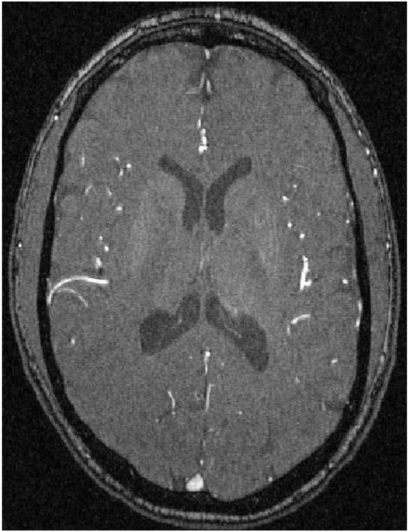
from Elizabeth Bullitt, Dept. of Neurosurgery, UNC-CH

Magnetic Resonance Angiography (MRA)



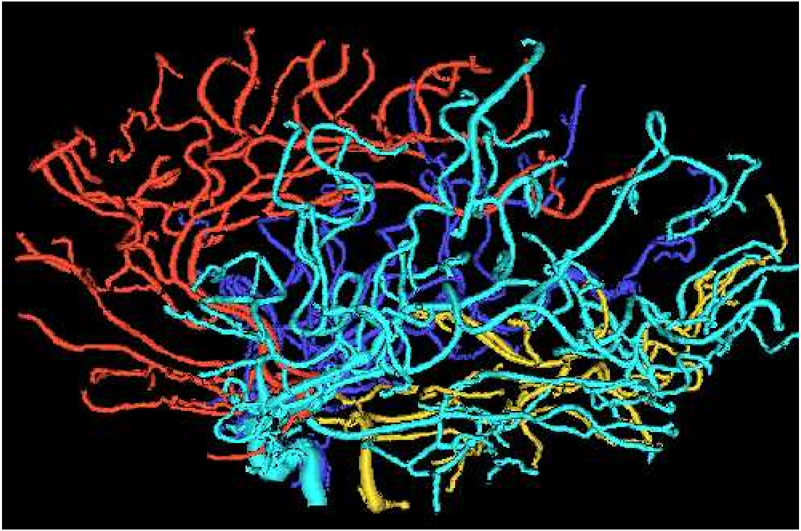
from Elizabeth Bullitt, Dept. of Neurosurgery, UNC-CH

Magnetic Resonance Angiography (MRA)



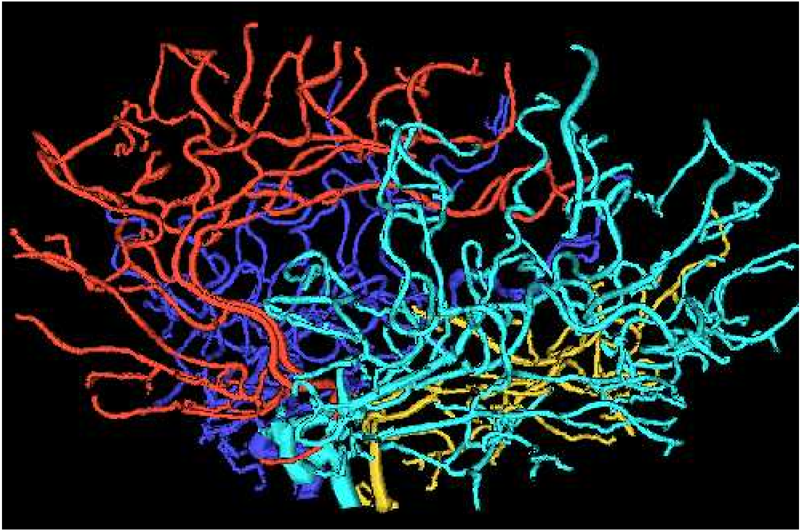
from Elizabeth Bullitt, Dept. of Neurosurgery, UNC-CH

Tube tracking



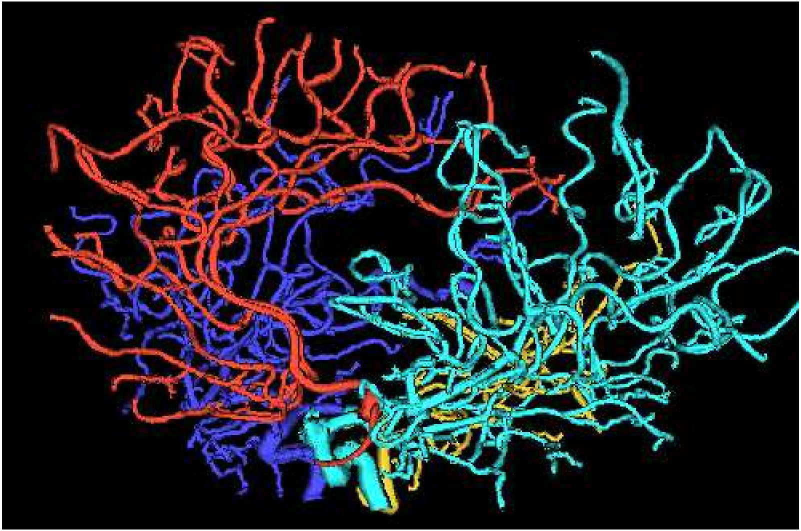
[Bullitt and Aylward, 2002]

Tube tracking



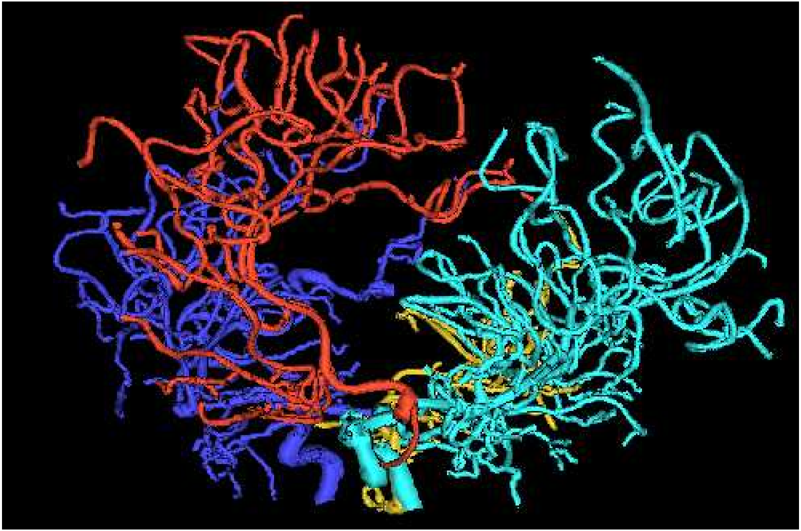
[Bullitt and Aylward, 2002]

Tube tracking



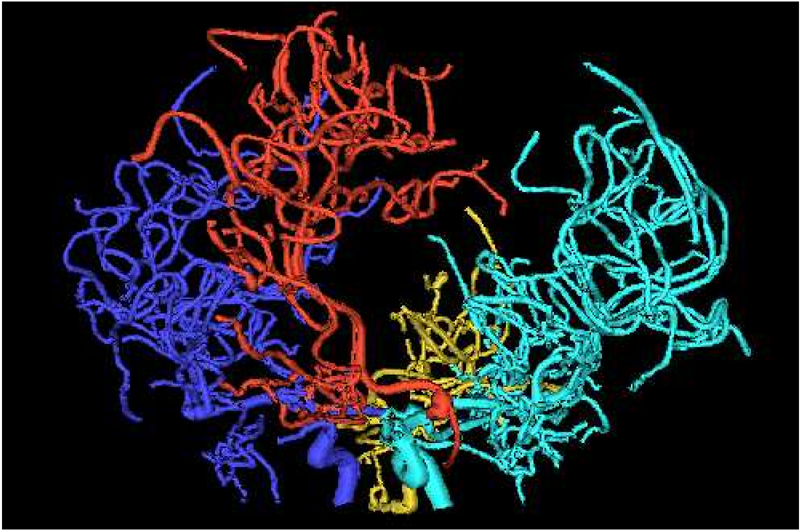
[Bullitt and Aylward, 2002]

Tube tracking



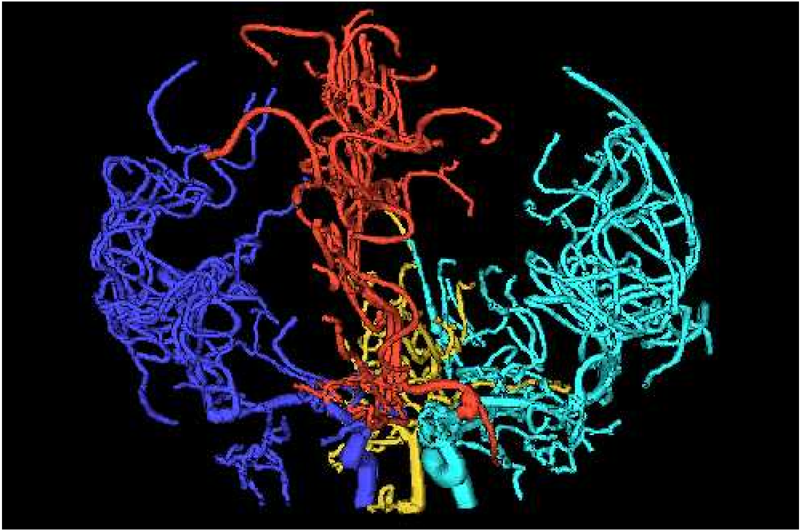
[Bullitt and Aylward, 2002]

Tube tracking



[Bullitt and Aylward, 2002]

Tube tracking



[Bullitt and Aylward, 2002]

Brain artery trees

Goal: Statistical analysis taking 3D geometry into account

- predict stroke tendency
- screen for loci of pathology, such as tumors
- explore how age affects vascularization

Brain artery trees

Goal: Statistical analysis taking 3D geometry into account

- predict stroke tendency
- screen for loci of pathology, such as tumors
- explore how age affects vascularization

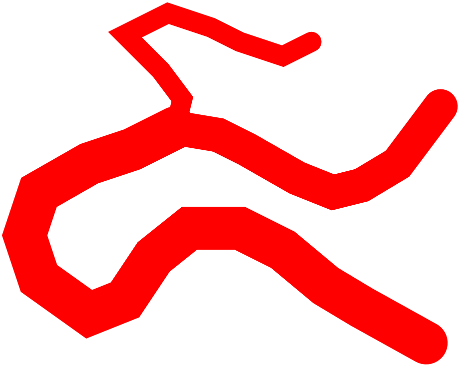
The images:

Brain artery trees

Goal: Statistical analysis taking 3D geometry into account

- predict stroke tendency
- screen for loci of pathology, such as tumors
- explore how age affects vascularization

The images:

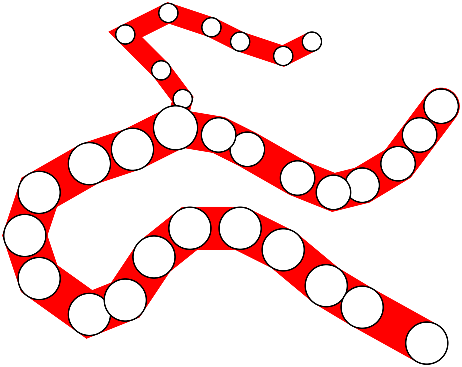


Brain artery trees

Goal: Statistical analysis taking 3D geometry into account

- predict stroke tendency
- screen for loci of pathology, such as tumors
- explore how age affects vascularization

The images:

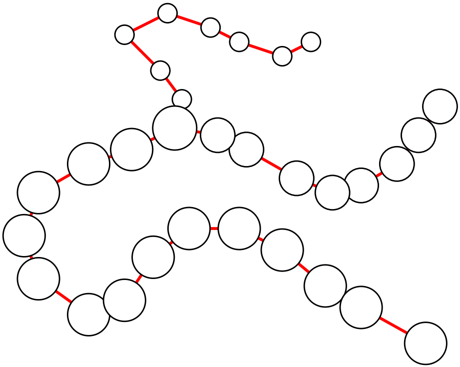


Brain artery trees

Goal: Statistical analysis taking 3D geometry into account

- predict stroke tendency
- screen for loci of pathology, such as tumors
- explore how age affects vascularization

The data structure:



Brain artery trees

Goal: Statistical analysis taking 3D geometry into account

- predict stroke tendency
- screen for loci of pathology, such as tumors
- explore how age affects vascularization

The data structure:

Brain artery trees

Goal: Statistical analysis taking 3D geometry into account

- predict stroke tendency
- screen for loci of pathology, such as tumors
- explore how age affects vascularization

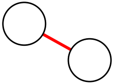
The data structure:

Brain artery trees

Goal: Statistical analysis taking 3D geometry into account

- predict stroke tendency
- screen for loci of pathology, such as tumors
- explore how age affects vascularization

The data structure:

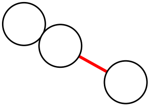


Brain artery trees

Goal: Statistical analysis taking 3D geometry into account

- predict stroke tendency
- screen for loci of pathology, such as tumors
- explore how age affects vascularization

The data structure:

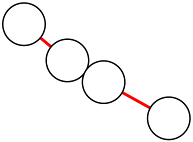


Brain artery trees

Goal: Statistical analysis taking 3D geometry into account

- predict stroke tendency
- screen for loci of pathology, such as tumors
- explore how age affects vascularization

The data structure:

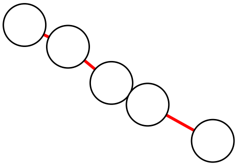


Brain artery trees

Goal: Statistical analysis taking 3D geometry into account

- predict stroke tendency
- screen for loci of pathology, such as tumors
- explore how age affects vascularization

The data structure:

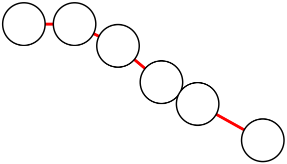


Brain artery trees

Goal: Statistical analysis taking 3D geometry into account

- predict stroke tendency
- screen for loci of pathology, such as tumors
- explore how age affects vascularization

The data structure:

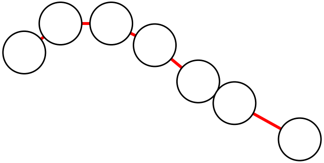


Brain artery trees

Goal: Statistical analysis taking 3D geometry into account

- predict stroke tendency
- screen for loci of pathology, such as tumors
- explore how age affects vascularization

The data structure:

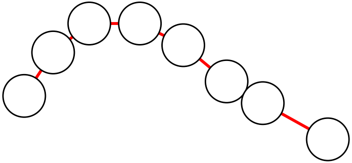


Brain artery trees

Goal: Statistical analysis taking 3D geometry into account

- predict stroke tendency
- screen for loci of pathology, such as tumors
- explore how age affects vascularization

The data structure:

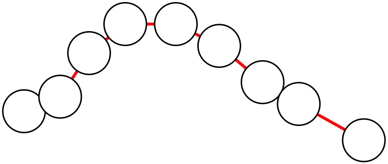


Brain artery trees

Goal: Statistical analysis taking 3D geometry into account

- predict stroke tendency
- screen for loci of pathology, such as tumors
- explore how age affects vascularization

The data structure:

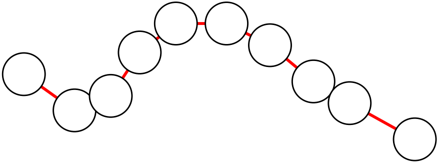


Brain artery trees

Goal: Statistical analysis taking 3D geometry into account

- predict stroke tendency
- screen for loci of pathology, such as tumors
- explore how age affects vascularization

The data structure:

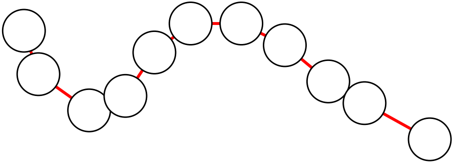


Brain artery trees

Goal: Statistical analysis taking 3D geometry into account

- predict stroke tendency
- screen for loci of pathology, such as tumors
- explore how age affects vascularization

The data structure:

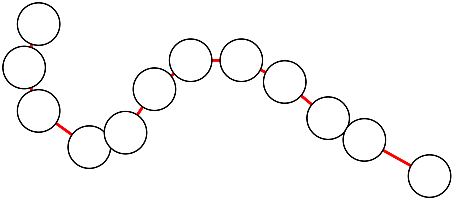


Brain artery trees

Goal: Statistical analysis taking 3D geometry into account

- predict stroke tendency
- screen for loci of pathology, such as tumors
- explore how age affects vascularization

The data structure:

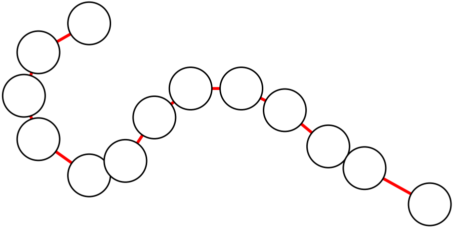


Brain artery trees

Goal: Statistical analysis taking 3D geometry into account

- predict stroke tendency
- screen for loci of pathology, such as tumors
- explore how age affects vascularization

The data structure:

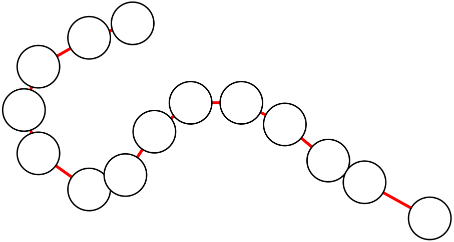


Brain artery trees

Goal: Statistical analysis taking 3D geometry into account

- predict stroke tendency
- screen for loci of pathology, such as tumors
- explore how age affects vascularization

The data structure:

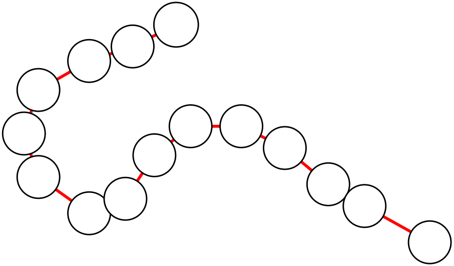


Brain artery trees

Goal: Statistical analysis taking 3D geometry into account

- predict stroke tendency
- screen for loci of pathology, such as tumors
- explore how age affects vascularization

The data structure:

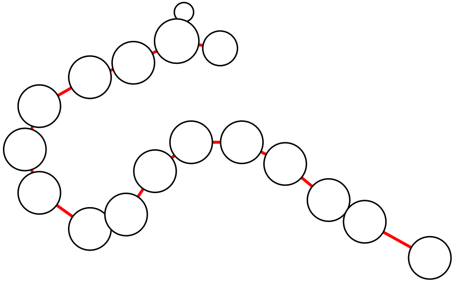


Brain artery trees

Goal: Statistical analysis taking 3D geometry into account

- predict stroke tendency
- screen for loci of pathology, such as tumors
- explore how age affects vascularization

The data structure:

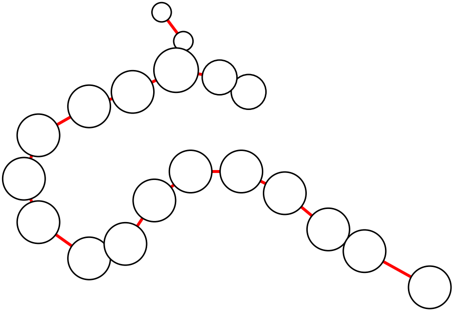


Brain artery trees

Goal: Statistical analysis taking 3D geometry into account

- predict stroke tendency
- screen for loci of pathology, such as tumors
- explore how age affects vascularization

The data structure:

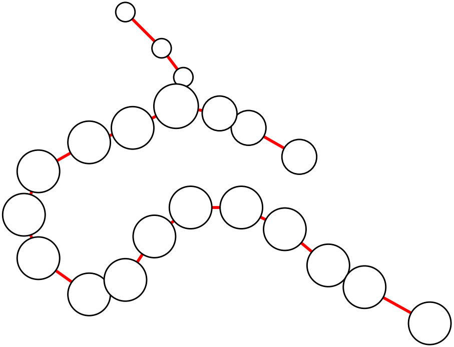


Brain artery trees

Goal: Statistical analysis taking 3D geometry into account

- predict stroke tendency
- screen for loci of pathology, such as tumors
- explore how age affects vascularization

The data structure:

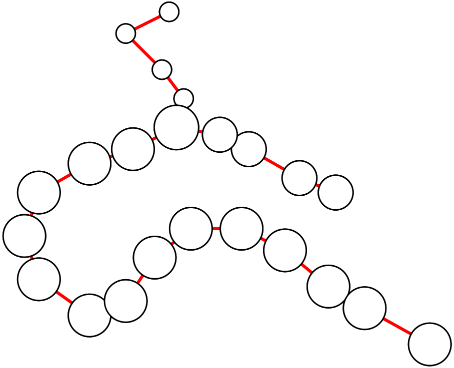


Brain artery trees

Goal: Statistical analysis taking 3D geometry into account

- predict stroke tendency
- screen for loci of pathology, such as tumors
- explore how age affects vascularization

The data structure:

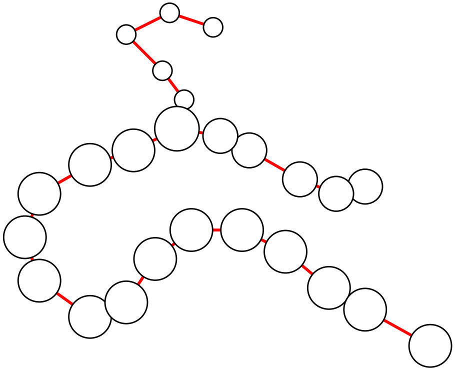


Brain artery trees

Goal: Statistical analysis taking 3D geometry into account

- predict stroke tendency
- screen for loci of pathology, such as tumors
- explore how age affects vascularization

The data structure:

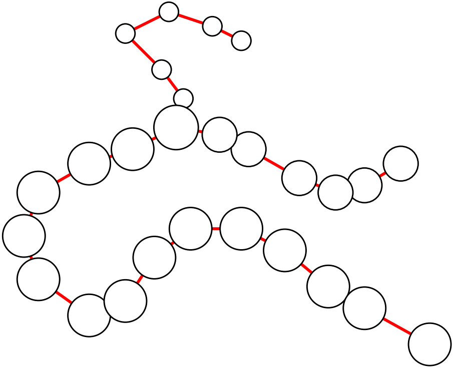


Brain artery trees

Goal: Statistical analysis taking 3D geometry into account

- predict stroke tendency
- screen for loci of pathology, such as tumors
- explore how age affects vascularization

The data structure:

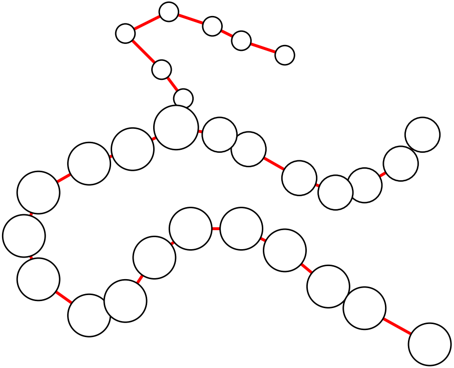


Brain artery trees

Goal: Statistical analysis taking 3D geometry into account

- predict stroke tendency
- screen for loci of pathology, such as tumors
- explore how age affects vascularization

The data structure:

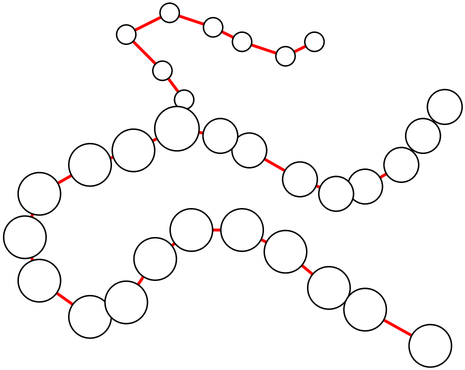


Brain artery trees

Goal: Statistical analysis taking 3D geometry into account

- predict stroke tendency
- screen for loci of pathology, such as tumors
- explore how age affects vascularization

The data structure:

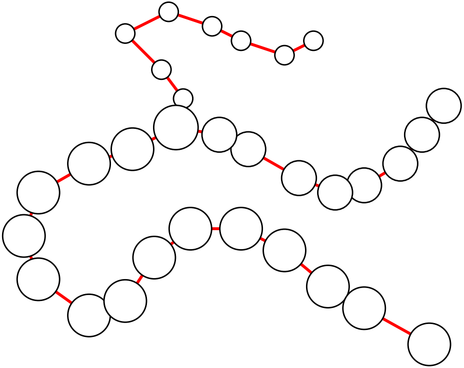


Brain artery trees

Goal: Statistical analysis taking 3D geometry into account

- predict stroke tendency
- screen for loci of pathology, such as tumors
- explore how age affects vascularization

The data structure:

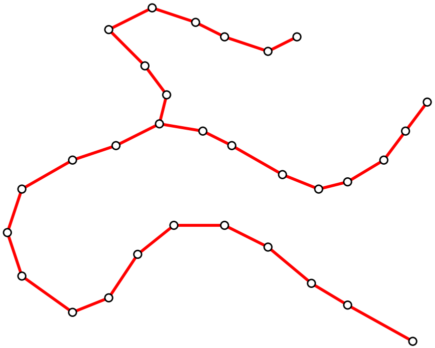


Brain artery trees

Goal: Statistical analysis taking 3D geometry into account

- predict stroke tendency
- screen for loci of pathology, such as tumors
- explore how age affects vascularization

The data structure:

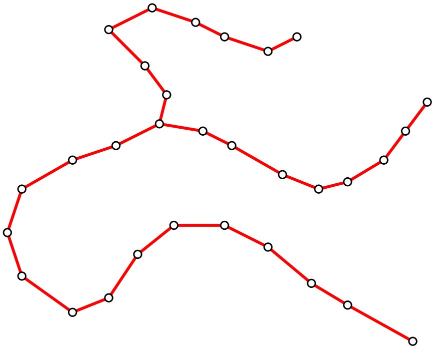


Brain artery trees

Goal: Statistical analysis taking 3D geometry into account

- predict stroke tendency
- screen for loci of pathology, such as tumors
- explore how age affects vascularization

The data structure:



Homology

Topological space $X \rightsquigarrow$ homology $H_i X$ for each dimension i.

- vector space that measures "i-dimensional holes" in X

Homology

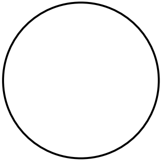
Topological space $X \rightsquigarrow$ homology $H_i X$ for each dimension i.

- vector space that measures "i-dimensional holes" in X

Homology

Topological space $X \rightsquigarrow$ homology $H_i X$ for each dimension i.

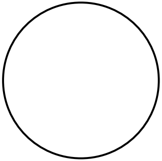
- vector space that measures "i-dimensional holes" in X



Homology

Topological space $X \rightsquigarrow$ homology $H_i X$ for each dimension i.

- vector space that measures "i-dimensional holes" in X

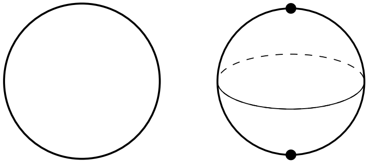


$\dim(H_1) = 1$

Homology

Topological space $X \rightsquigarrow$ homology $H_i X$ for each dimension i.

- vector space that measures "i-dimensional holes" in X

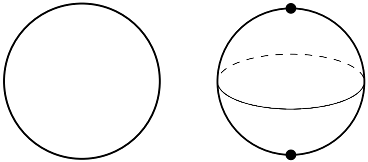


$\dim(H_1) = 1$

Homology

Topological space $X \rightsquigarrow$ homology $H_i X$ for each dimension i.

- vector space that measures "i-dimensional holes" in X

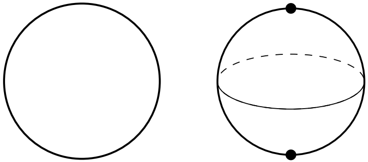


$\dim(H_1) = 1$ $\dim(H_1) = 0$

Homology

Topological space $X \rightsquigarrow$ homology $H_i X$ for each dimension i.

- vector space that measures "i-dimensional holes" in X



$\dim(H_1) = 1$

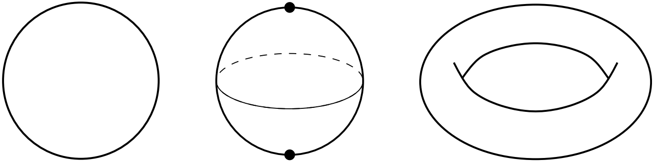
$\dim(H_1) = 0$

$\dim(H_2) = 1$

Homology

Topological space $X \rightsquigarrow$ homology $H_i X$ for each dimension i.

- vector space that measures "i-dimensional holes" in X



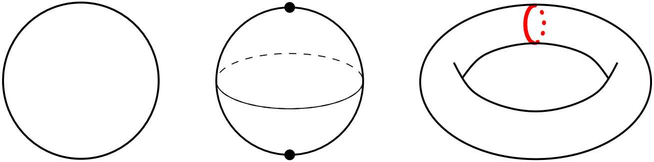
$\dim(H_1) = 1$ $\dim(H_1) = 0$

$\dim(H_2) = 1$

Homology

Topological space $X \rightsquigarrow$ homology $H_i X$ for each dimension i.

- vector space that measures "i-dimensional holes" in X



$\dim(H_1) = 1$

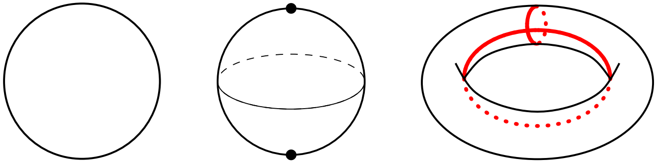
$\dim(H_1) = 0$

$\dim(H_2) = 1$

Homology

Topological space $X \rightsquigarrow$ homology $H_i X$ for each dimension i.

- vector space that measures "i-dimensional holes" in X



$\dim(H_1) = 1$

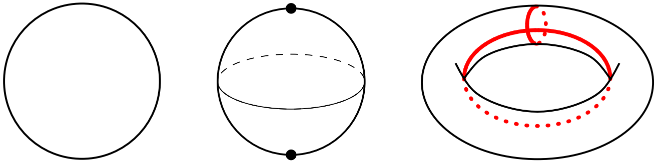
$\dim(H_1) = 0$

$\dim(H_2) = 1$

Homology

Topological space $X \rightsquigarrow$ homology $H_i X$ for each dimension i.

- vector space that measures "i-dimensional holes" in X



$\dim(H_1) = 1$

$\dim(H_1) = 0$

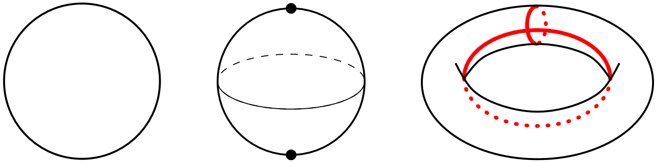
$\dim(H_2) = 1$

$\dim(H_1) = 2$

Homology

Topological space $X \rightsquigarrow$ homology $H_i X$ for each dimension i.

- vector space that measures "i-dimensional holes" in X



$\dim(H_1) = 1$

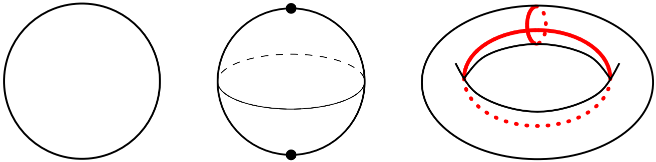
$\dim(H_1) = 0$
$\dim(H_2) = 1$

$\dim(H_1) = 2$
$\dim(H_2) = 1$

Homology

Topological space $X \rightsquigarrow$ homology $H_i X$ for each dimension i.

- vector space that measures "i-dimensional holes" in X



$\dim(H_1) = 1$ $\dim(H_1) = 0$ $\dim(H_1) = 2$

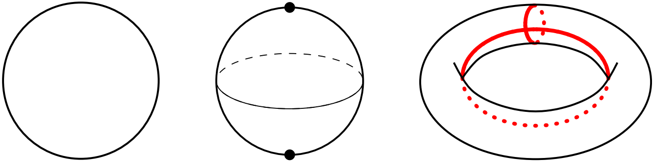
$\dim(H_2) = 1$ $\dim(H_2) = 1$

- $i = 0$ case: basis for $H_i \leftrightarrow$ connected components of X

Homology

Topological space $X \rightsquigarrow$ homology $H_i X$ for each dimension i.

- vector space that measures "i-dimensional holes" in X



$\dim(H_0) = 1$	$\dim(H_0) = 1$	$\dim(H_0) = 1$
$\dim(H_1) = 1$	$\dim(H_1) = 0$	$\dim(H_1) = 2$
	$\dim(H_2) = 1$	$\dim(H_2) = 1$

- $i = 0$ case: basis for $H_i \leftrightarrow$ connected components of X

Persistent homology

Build X step by step

- measure evolving topology.

Def. Suppose X is a filtered space, meaning X is a union of an increasing sequence of subspaces: $\varnothing = X_0 \subset X_1 \subset \cdots \subset X_m = X$.

- The persistent homology of this filtration is $H_i X_1 \to H_i X_2 \to \cdots \to H_i X_m$, a sequence of vector space homomorphisms.
- A feature persists from j to k if it appears first in $H_i X_j$ and last in $H_i X_k$.

Examples:

1. Given a function $f : X \to \mathbb{R}$, let $X_t = \{x \in X \mid f(x) \leq t\}$. Good choice of $t_0, \ldots, t_m \in \mathbb{R}$: the values of t across which $H_i X_t$ changes.
2. Any simplicial complex: build it simplex by simplex in some order.

History. invented by [Frosini, Landi 1999], [Robins 1999];
[Edelsbrunner, Letscher, Zomorodian 2002]: includes efficient computation;
[Carlsson, Zomorodian 2009]: multiparameter persistence;
[Knudson and many others]: further theoretical developments, applications

Persistent homology

Build X step by step

- measure evolving topology.

Def. Suppose X is a filtered space, meaning X is a union of an increasing sequence of subspaces: $\varnothing = X_0 \subset X_1 \subset \cdots \subset X_m = X$.

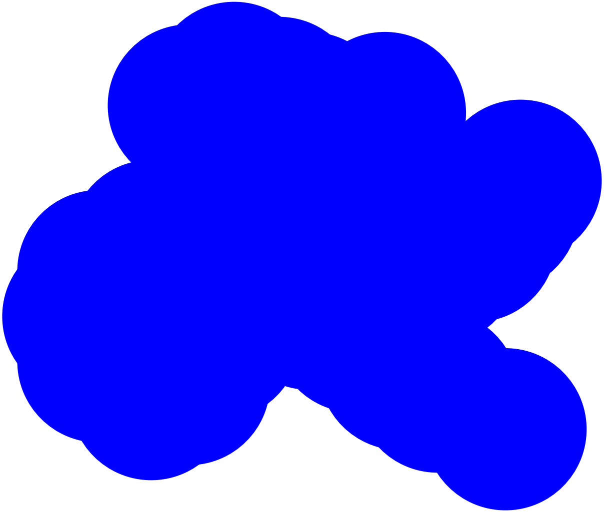
- The persistent homology of this filtration is $H_i X_1 \to H_i X_2 \to \cdots \to H_i X_m$, a sequence of vector space homomorphisms.
- A feature persists from j to k if it appears first in $H_i X_j$ and last in $H_i X_k$.

Examples:

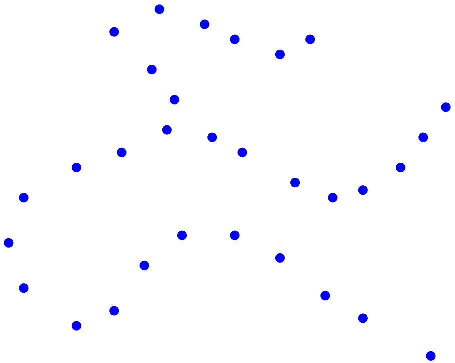
1. Given a function $f : X \to \mathbb{R}$, let $X_t = \{x \in X \mid f(x) \leq t\}$. Good choice of $t_0, \ldots, t_m \in \mathbb{R}$: the values of t across which $H_i X_t$ changes.
2. Any simplicial complex: build it simplex by simplex in some order.

History. invented by [Frosini, Landi 1999], [Robins 1999];
[Edelsbrunner, Letscher, Zomorodian 2002]: includes efficient computation;
[Carlsson, Zomorodian 2009]: multiparameter persistence;
[Knudson and many others]: further theoretical developments, applications

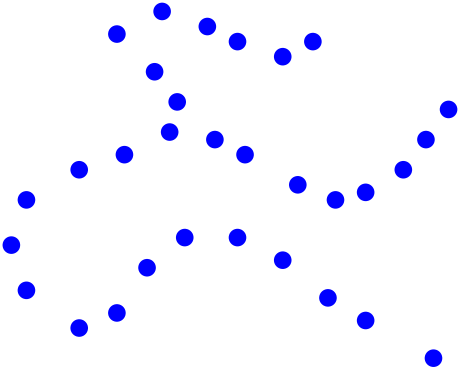
Example: expanding balls



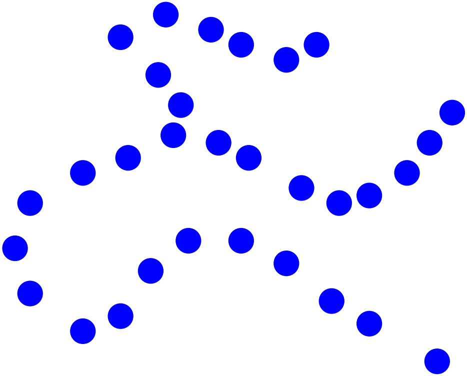
Example: expanding balls



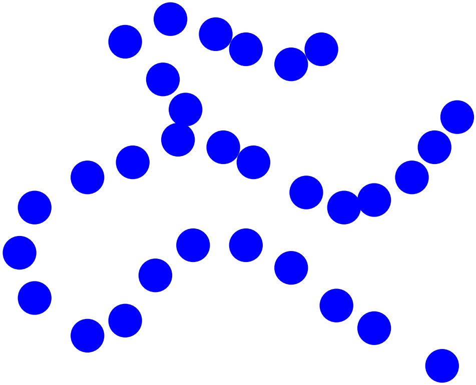
Example: expanding balls



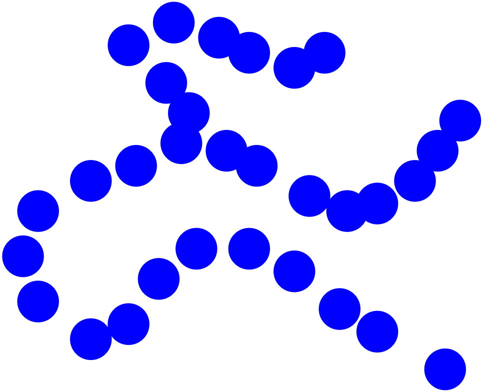
Example: expanding balls



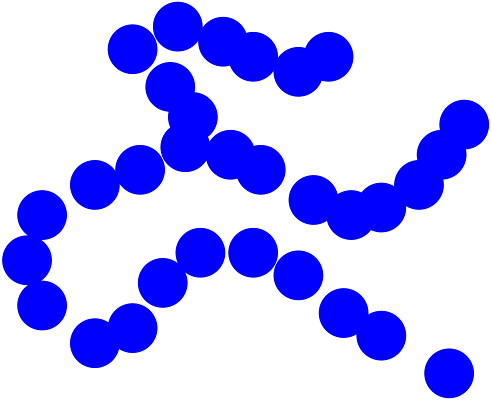
Example: expanding balls



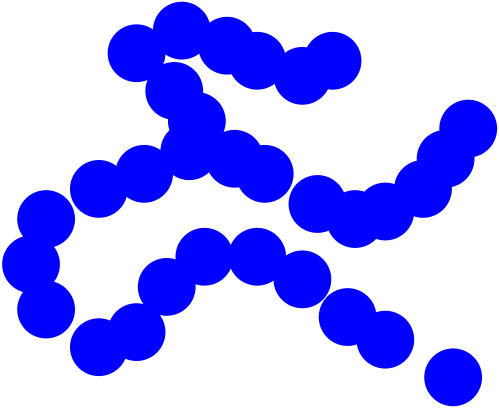
Example: expanding balls



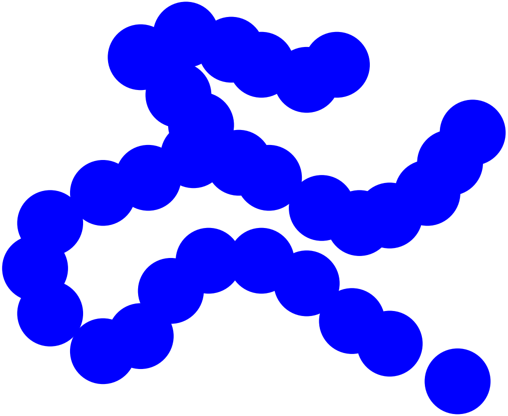
Example: expanding balls



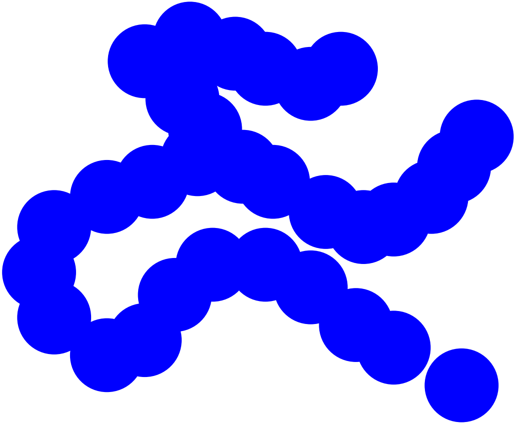
Example: expanding balls



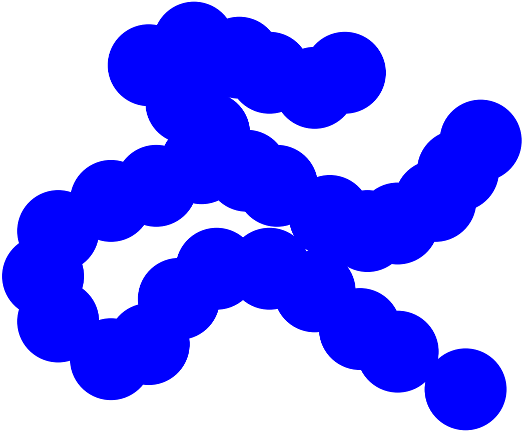
Example: expanding balls



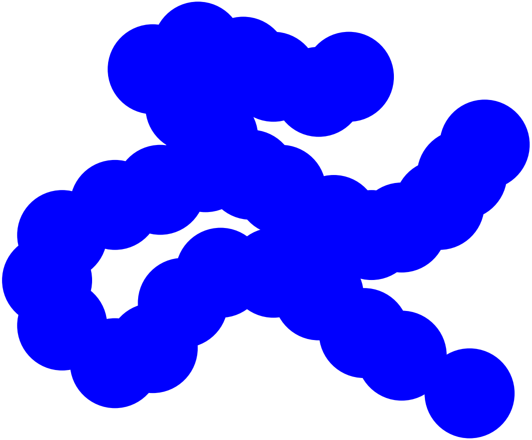
Example: expanding balls



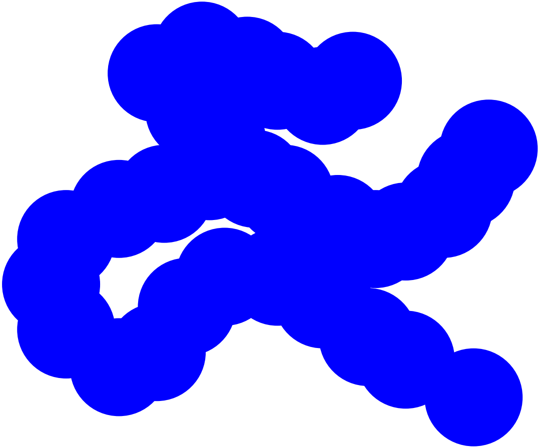
Example: expanding balls



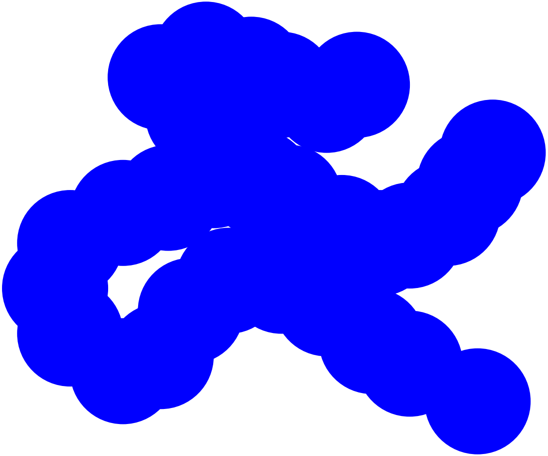
Example: expanding balls



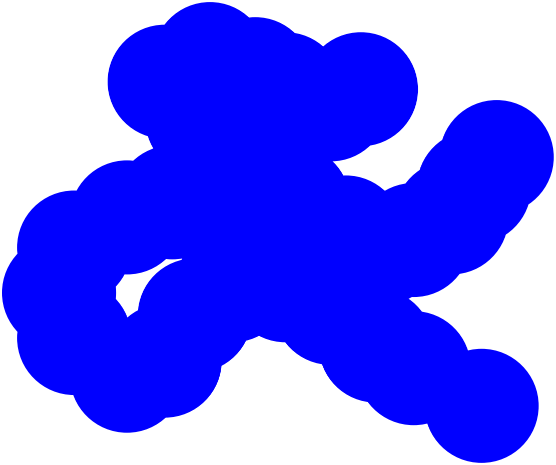
Example: expanding balls



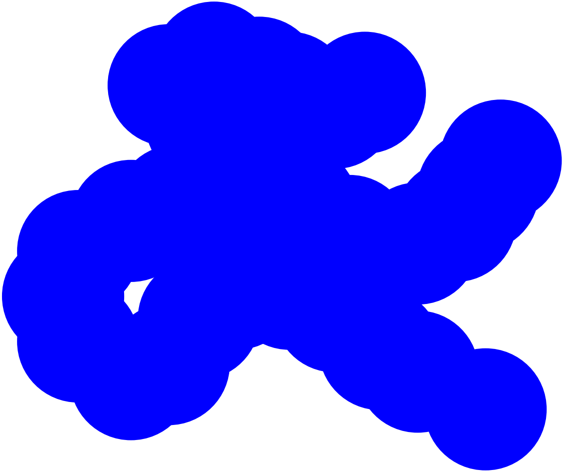
Example: expanding balls



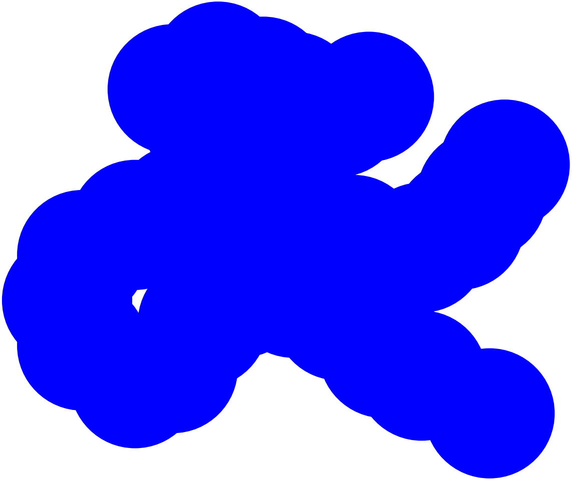
Example: expanding balls



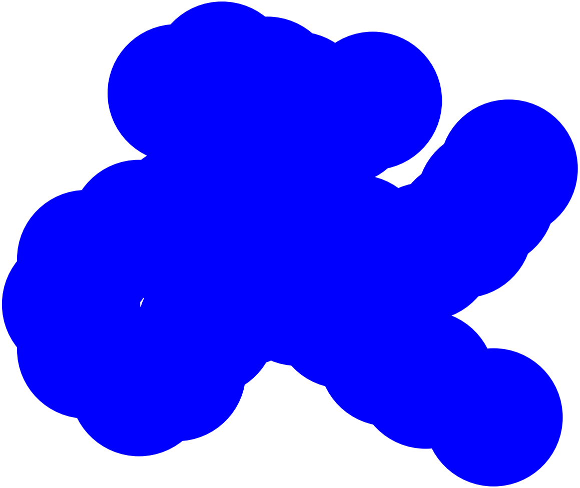
Example: expanding balls



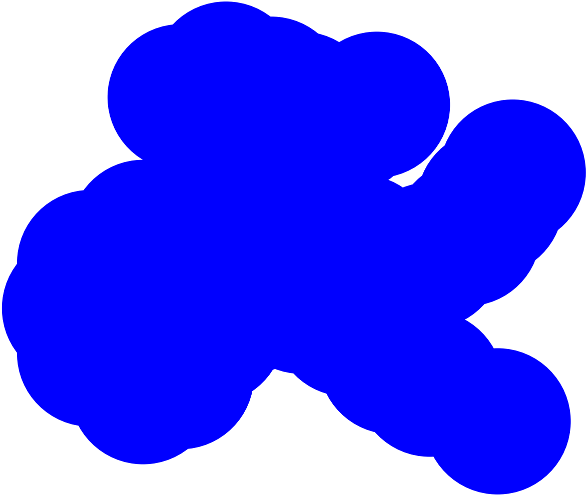
Example: expanding balls



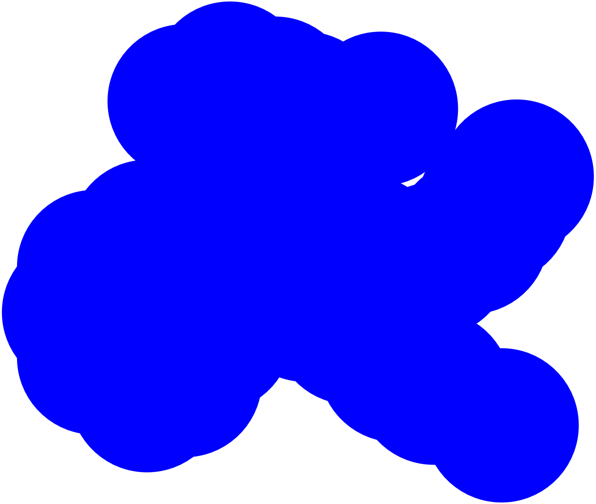
Example: expanding balls



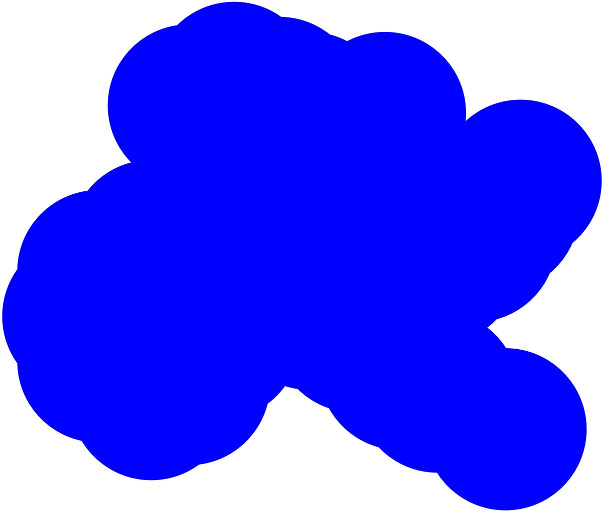
Example: expanding balls



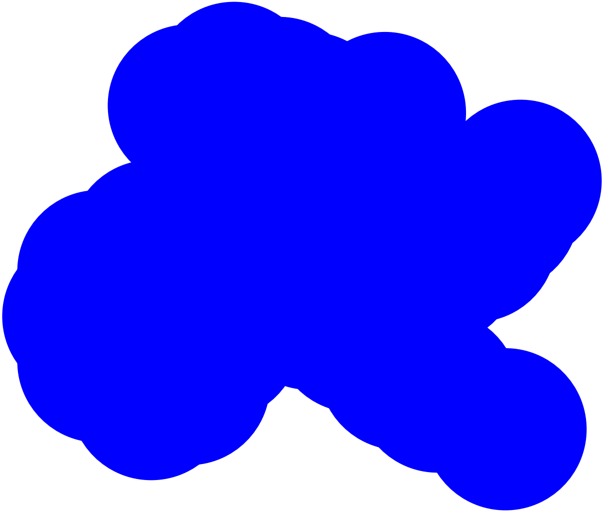
Example: expanding balls



Example: expanding balls



Example: expanding balls



Persistent homology

Build X step by step

- measure evolving topology.

Def. Suppose X is a filtered space, meaning X is a union of an increasing sequence of subspaces: $\varnothing = X_0 \subset X_1 \subset \cdots \subset X_m = X$.

- The persistent homology of this filtration is $H_i X_1 \to H_i X_2 \to \cdots \to H_i X_m$, a sequence of vector space homomorphisms.
- A feature persists from j to k if it appears first in $H_i X_j$ and last in $H_i X_k$.

Examples:

1. Given a function $f : X \to \mathbb{R}$, let $X_t = \{x \in X \mid f(x) \le t\}$. Good choice of $t_0, \ldots, t_m \in \mathbb{R}$: the values of t across which $H_i X_t$ changes.
2. Any simplicial complex: build it simplex by simplex in some order.

History. invented by [Frosini, Landi 1999], [Robins 1999];
[Edelsbrunner, Letscher, Zomorodian 2002]: includes efficient computation;
[Carlsson, Zomorodian 2009]: multiparameter persistence;
[Knudson and many others]: further theoretical developments, applications

Persistent homology

Build X step by step

- measure evolving topology.

Def. Suppose X is a filtered space, meaning X is a union of an increasing sequence of subspaces: $\varnothing = X_0 \subset X_1 \subset \cdots \subset X_m = X$.

- The persistent homology of this filtration is $H_i X_1 \to H_i X_2 \to \cdots \to H_i X_m$, a sequence of vector space homomorphisms.
- A feature persists from j to k if it appears first in $H_i X_j$ and last in $H_i X_k$.

Examples:

1. Given a function $f : X \to \mathbb{R}$, let $X_t = \{x \in X \mid f(x) \leq t\}$. Good choice of $t_0, \ldots, t_m \in \mathbb{R}$: the values of t across which $H_i X_t$ changes.
2. Any simplicial complex: build it simplex by simplex in some order.

History. invented by [Frosini, Landi 1999], [Robins 1999];
[Edelsbrunner, Letscher, Zomorodian 2002]: includes efficient computation;
[Carlsson, Zomorodian 2009]: multiparameter persistence;
[Knudson and many others]: further theoretical developments, applications

Persistent homology

Build X step by step

- measure evolving topology.

Def. Suppose X is a filtered space, meaning X is a union of an increasing sequence of subspaces: $\varnothing = X_0 \subset X_1 \subset \cdots \subset X_m = X$.

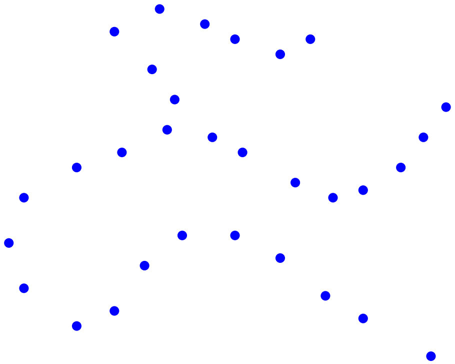
- The persistent homology of this filtration is $H_i X_1 \to H_i X_2 \to \cdots \to H_i X_m$, a sequence of vector space homomorphisms.
- A feature persists from j to k if it appears first in $H_i X_j$ and last in $H_i X_k$.

Examples:

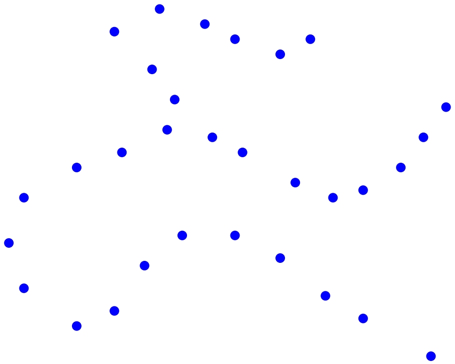
1. Given a function $f : X \to \mathbb{R}$, let $X_t = \{x \in X \mid f(x) \leq t\}$. Good choice of $t_0, \ldots, t_m \in \mathbb{R}$: the values of t across which $H_i X_t$ changes.
2. Any simplicial complex: build it simplex by simplex in some order.

History. invented by [Frosini, Landi 1999], [Robins 1999];
[Edelsbrunner, Letscher, Zomorodian 2002]: includes efficient computation;
[Carlsson, Zomorodian 2009]: multiparameter persistence;
[Knudson and many others]: further theoretical developments, applications

Example: expanding balls

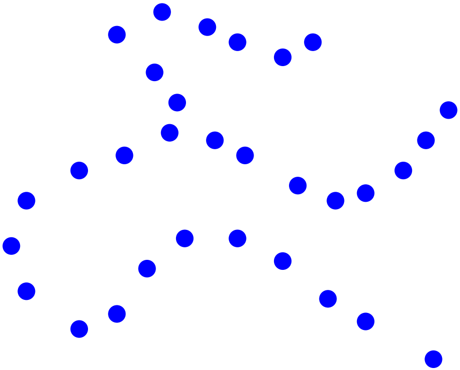


Example: expanding balls



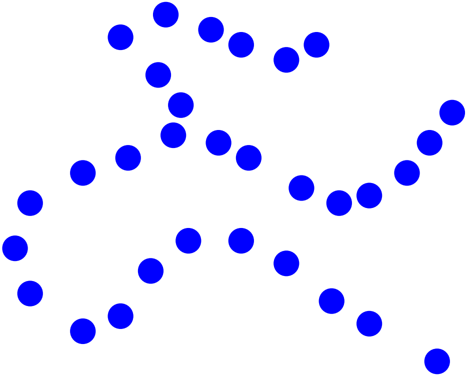
$\dim(H_0) = 31$

Example: expanding balls



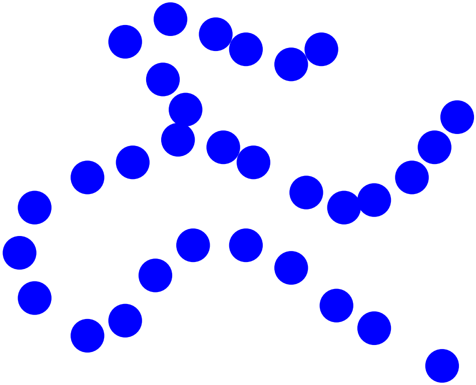
$\dim(H_0) = 31$

Example: expanding balls



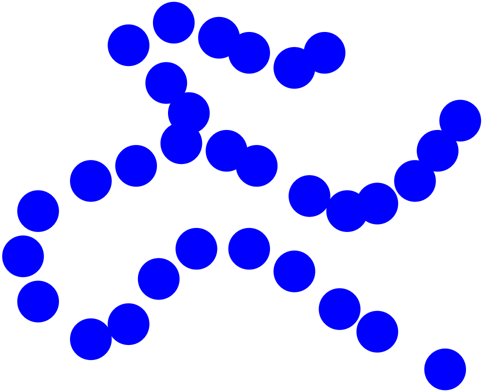
$\dim(H_0) = 31$

Example: expanding balls



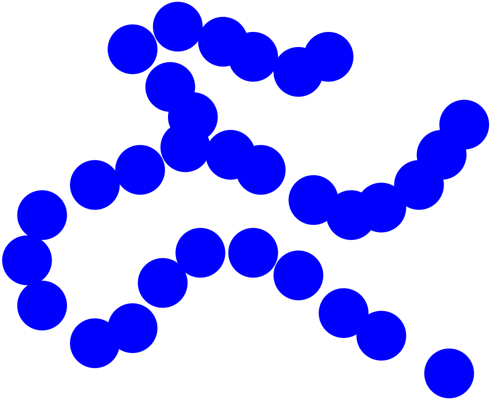
$\dim(H_0) = 26$

Example: expanding balls



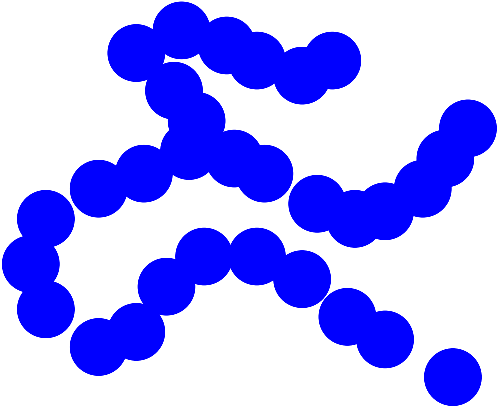
$\dim(H_0) = 21$

Example: expanding balls



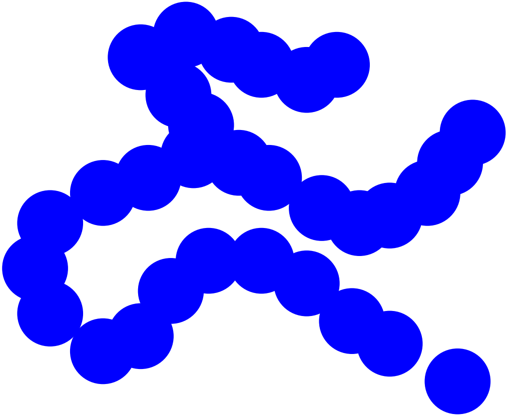
$\dim(H_0) = 12$

Example: expanding balls



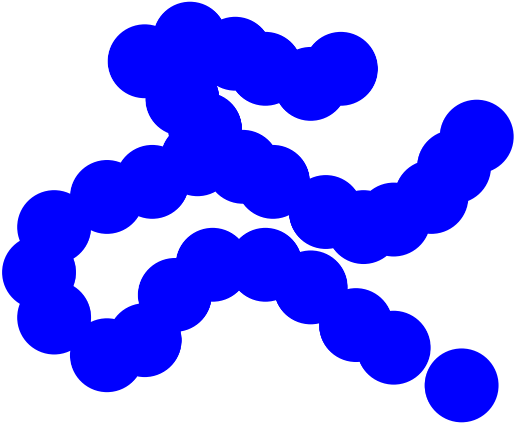
$\dim(H_0) = 6$

Example: expanding balls



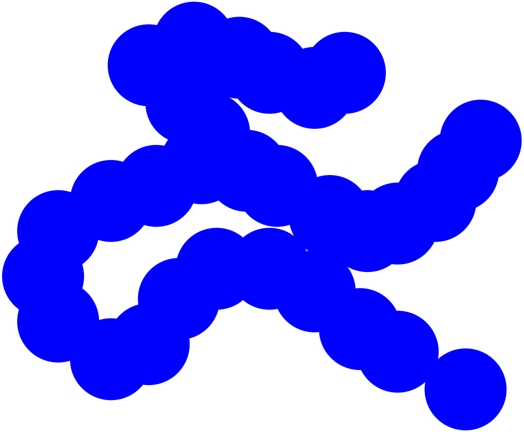
$\dim(H_0) = 2$

Example: expanding balls



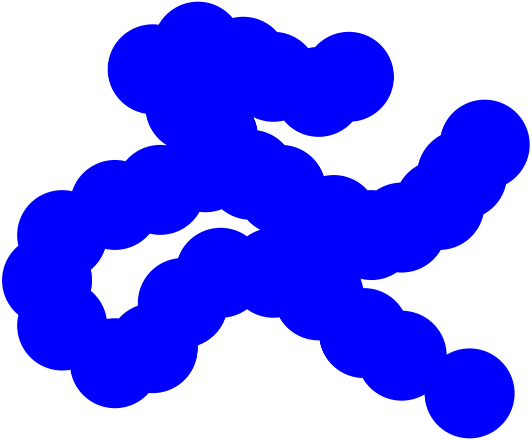
$\dim(H_0) = 2$

Example: expanding balls



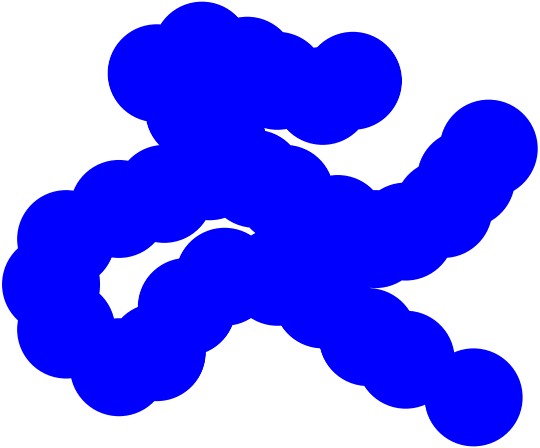
$\dim(H_0) = 1 \qquad \dim(H_1) = 2$

Example: expanding balls



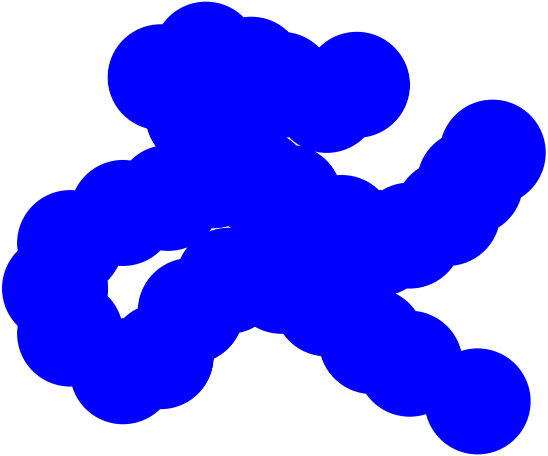
$\dim(H_0) = 1 \qquad \dim(H_1) = 1$

Example: expanding balls



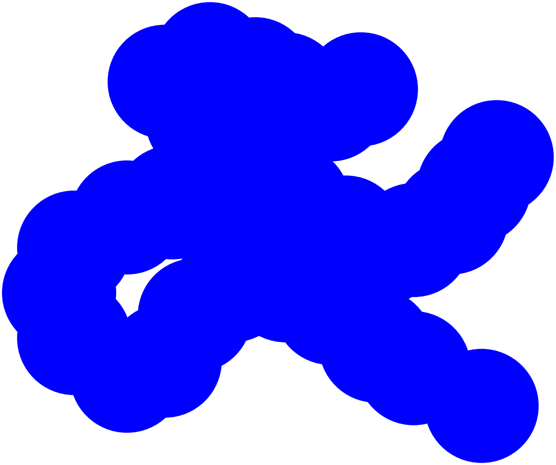
$\dim(H_0) = 1 \qquad \dim(H_1) = 1$

Example: expanding balls



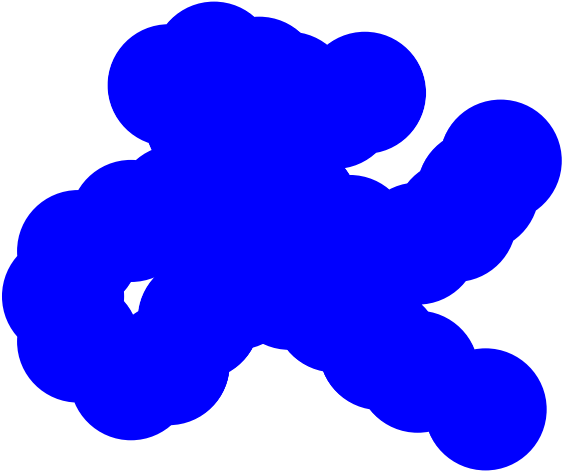
$\dim(H_0) = 1 \qquad \dim(H_1) = 3$

Example: expanding balls



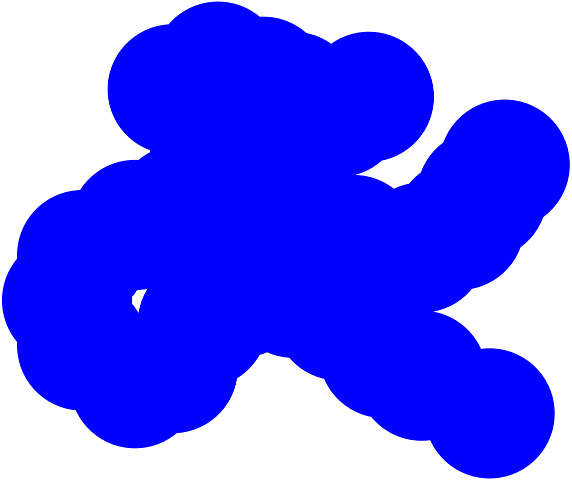
$\dim(H_0) = 1 \qquad \dim(H_1) = 1$

Example: expanding balls



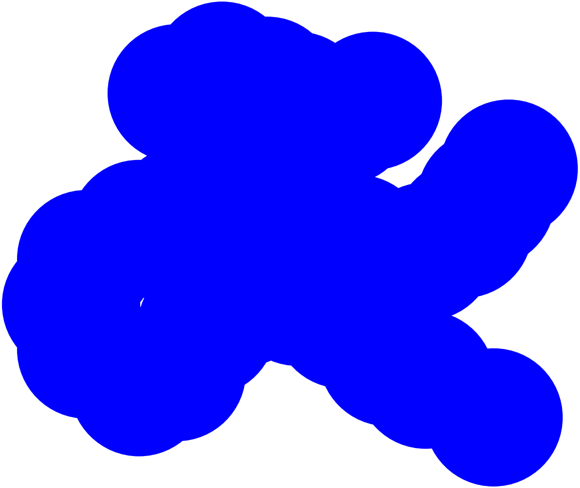
$\dim(H_0) = 1 \qquad \dim(H_1) = 1$

Example: expanding balls



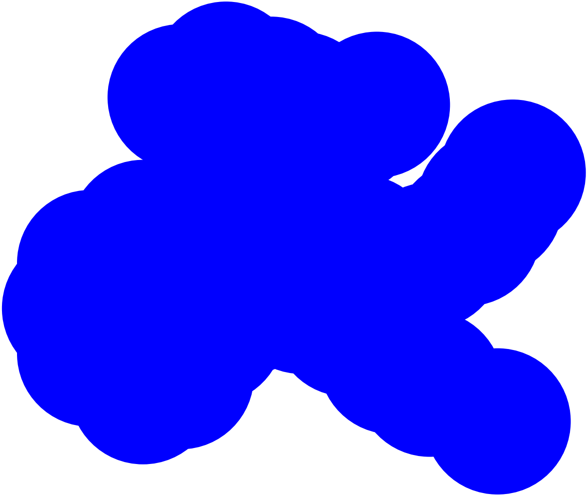
$\dim(H_0) = 1 \qquad \dim(H_1) = 1$

Example: expanding balls



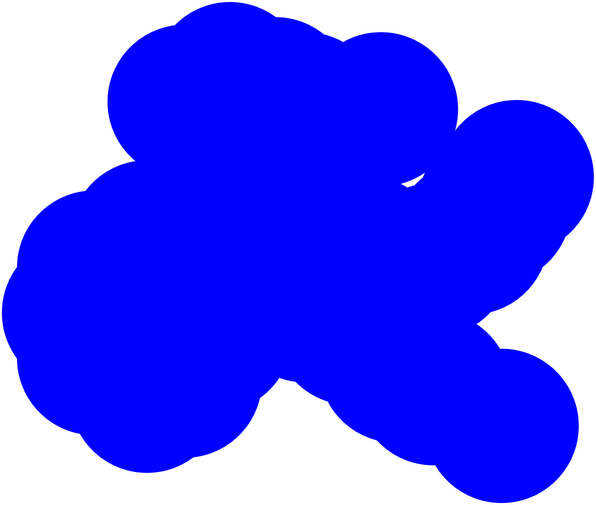
$\dim(H_0) = 1 \qquad \dim(H_1) = 1$

Example: expanding balls



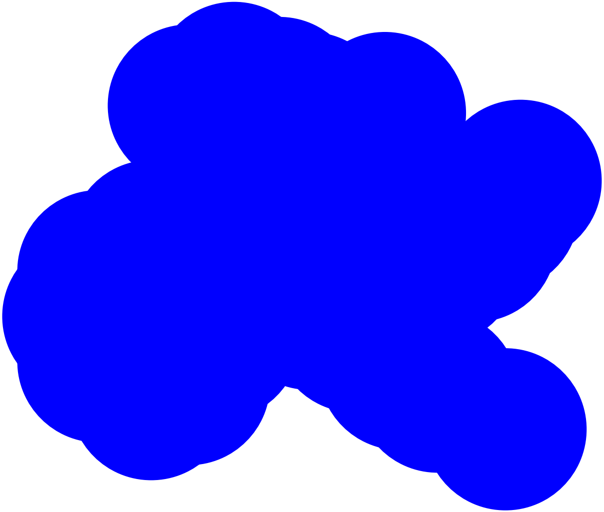
$\dim(H_0) = 1 \qquad \dim(H_1) = 0$

Example: expanding balls



$\dim(H_0) = 1 \qquad \dim(H_1) = 1$

Example: expanding balls



$\dim(H_0) = 1 \qquad \dim(H_1) = 0$

Persistent homology

Build X step by step

- measure evolving topology.

Def. Suppose X is a filtered space, meaning X is a union of an increasing sequence of subspaces: $\varnothing = X_0 \subset X_1 \subset \cdots \subset X_m = X$.

- The persistent homology of this filtration is $H_i X_1 \to H_i X_2 \to \cdots \to H_i X_m$, a sequence of vector space homomorphisms.
- A feature persists from j to k if it appears first in $H_i X_j$ and last in $H_i X_k$.

Examples:

1. Given a function $f : X \to \mathbb{R}$, let $X_t = \{x \in X \mid f(x) \leq t\}$. Good choice of $t_0, \ldots, t_m \in \mathbb{R}$: the values of t across which $H_i X_t$ changes.
2. Any simplicial complex: build it simplex by simplex in some order.

History. invented by [Frosini, Landi 1999], [Robins 1999];
[Edelsbrunner, Letscher, Zomorodian 2002]: includes efficient computation;
[Carlsson, Zomorodian 2009]: multiparameter persistence;
[Knudson and many others]: further theoretical developments, applications

Persistent homology

Build X step by step

- measure evolving topology.

Def. Suppose X is a filtered space, meaning X is a union of an increasing sequence of subspaces: $\varnothing = X_0 \subset X_1 \subset \cdots \subset X_m = X$.

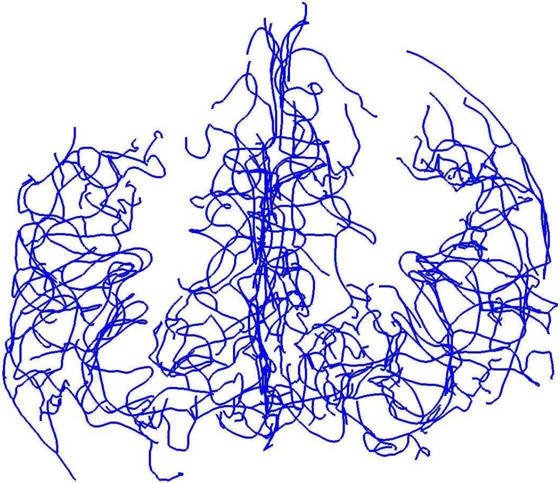
- The persistent homology of this filtration is $H_i X_1 \to H_i X_2 \to \cdots \to H_i X_m$, a sequence of vector space homomorphisms.
- A feature persists from j to k if it appears first in $H_i X_j$ and last in $H_i X_k$.

Examples:

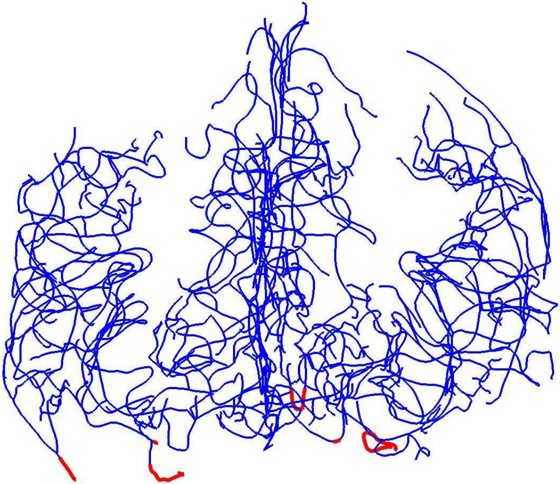
1. Given a function $f : X \to \mathbb{R}$, let $X_t = \{x \in X \mid f(x) \leq t\}$. Good choice of $t_0, \ldots, t_m \in \mathbb{R}$: the values of t across which $H_i X_t$ changes.
2. Any simplicial complex: build it simplex by simplex in some order.

History. invented by [Frosini, Landi 1999], [Robins 1999];
[Edelsbrunner, Letscher, Zomorodian 2002]: includes efficient computation;
[Carlsson, Zomorodian 2009]: multiparameter persistence;
[Knudson and many others]: further theoretical developments, applications

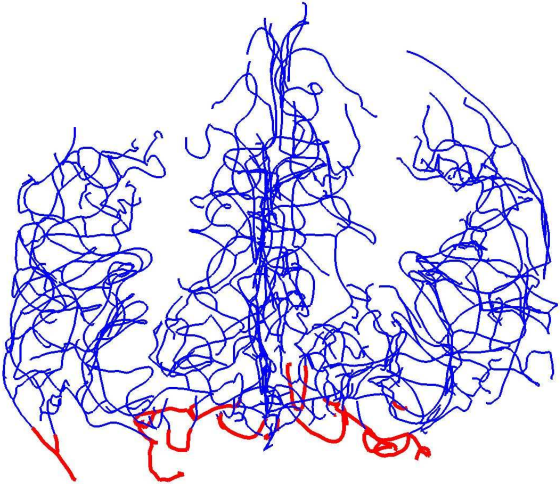
Example: filling brains



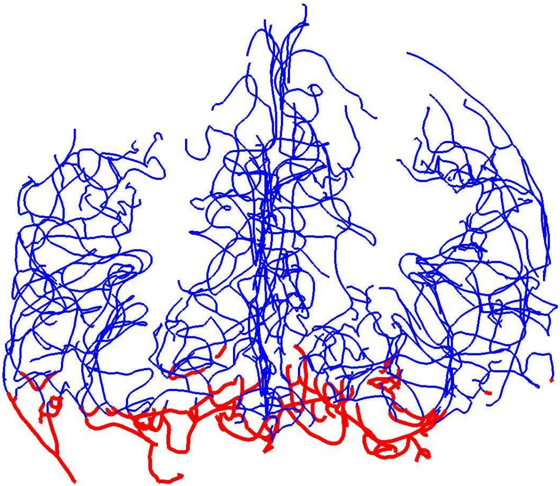
Example: filling brains



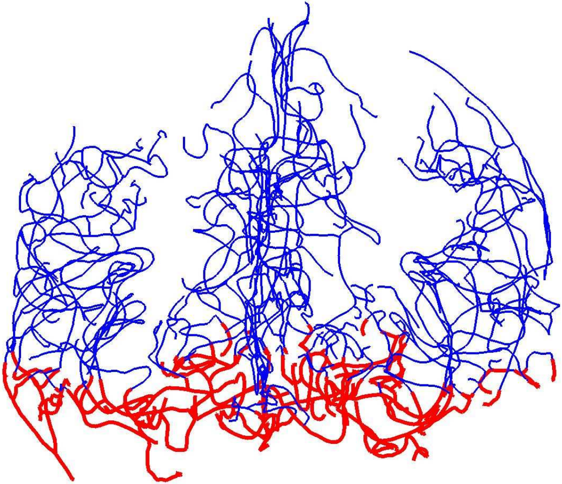
Example: filling brains



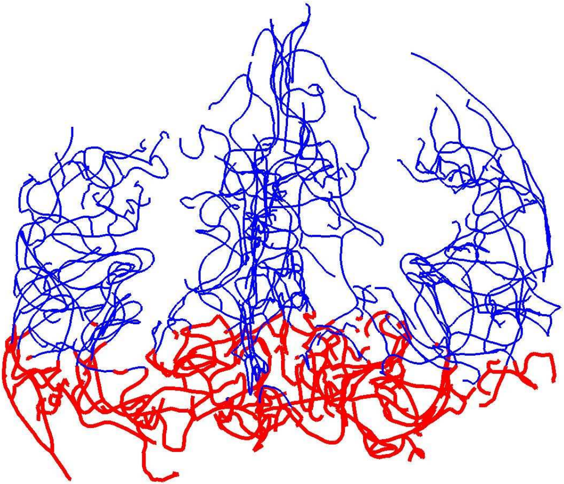
Example: filling brains



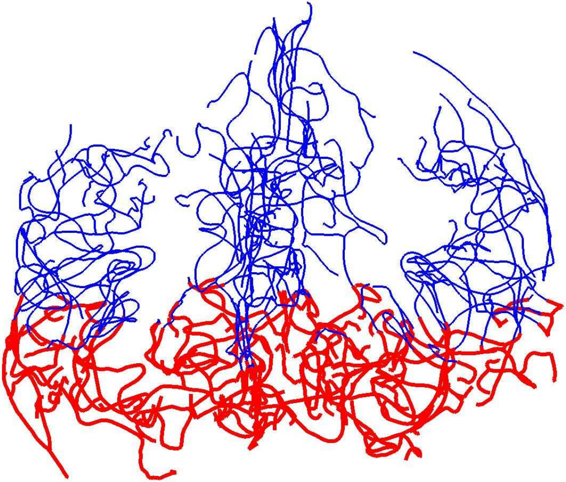
Example: filling brains



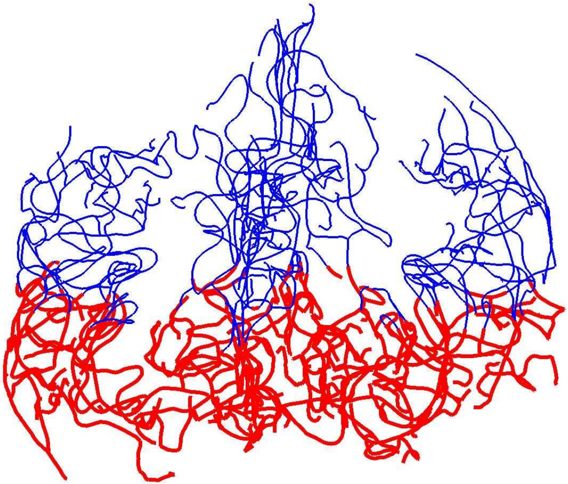
Example: filling brains



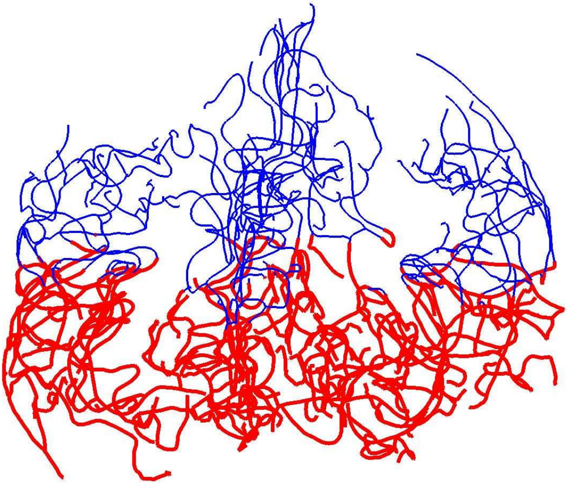
Example: filling brains



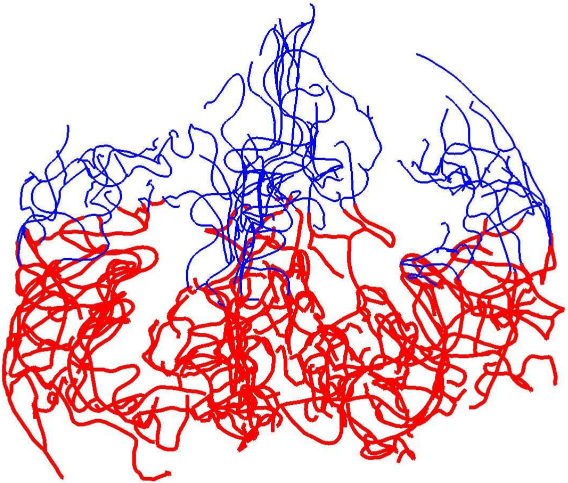
Example: filling brains



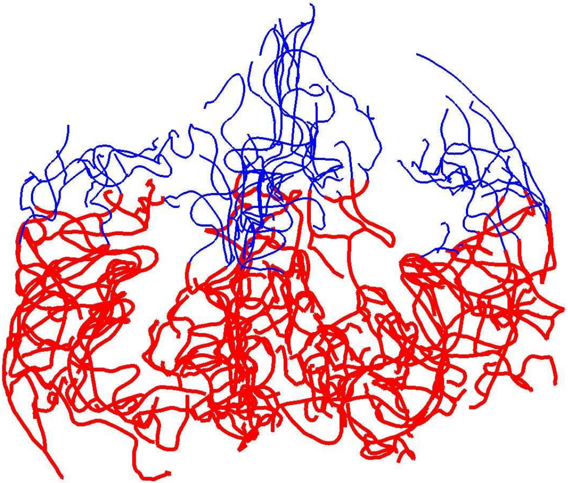
Example: filling brains



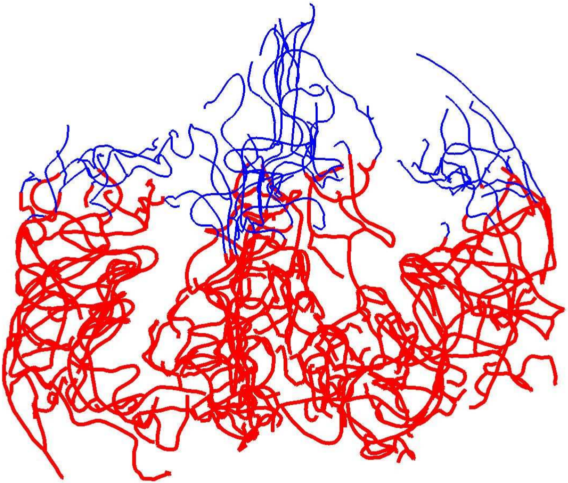
Example: filling brains



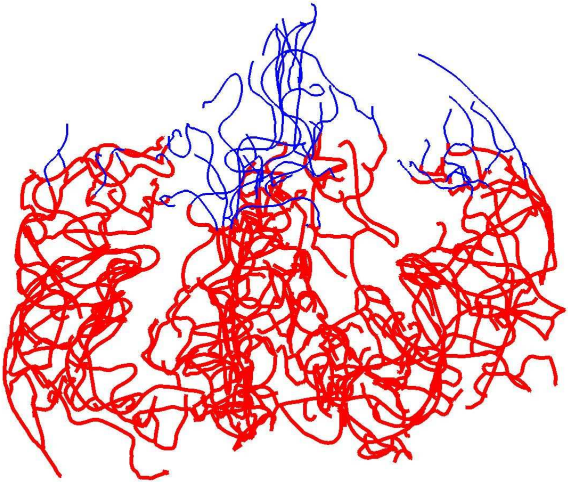
Example: filling brains



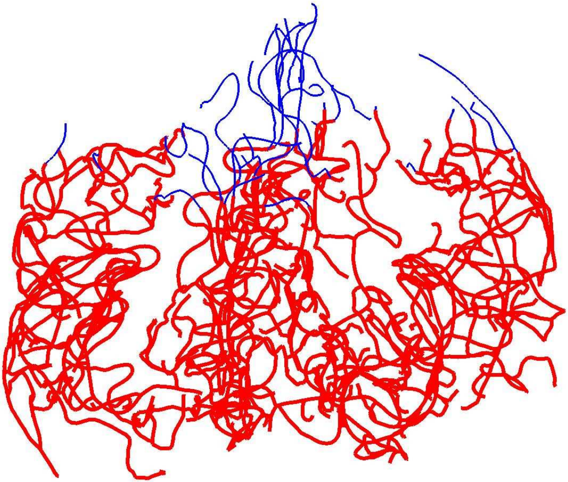
Example: filling brains



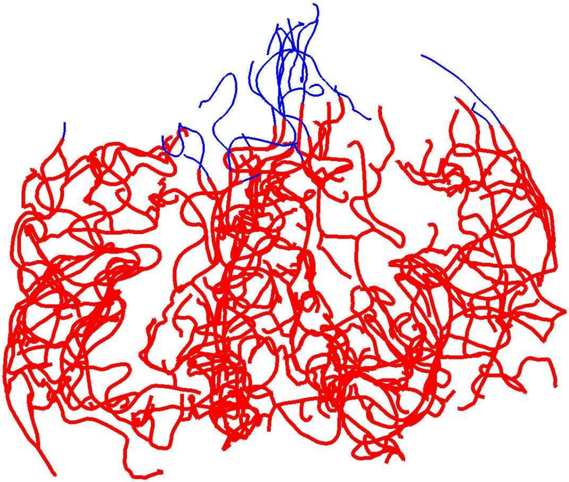
Example: filling brains



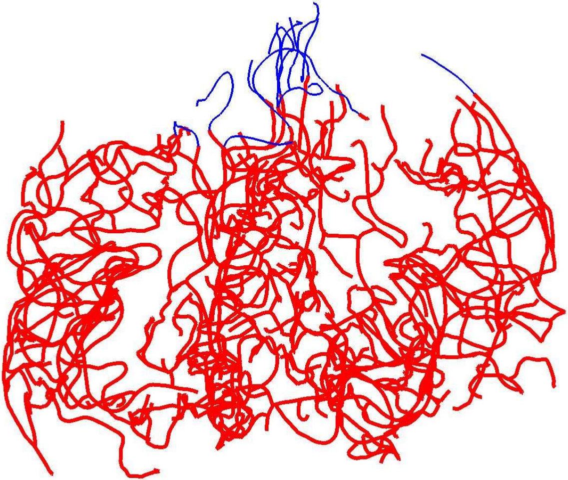
Example: filling brains



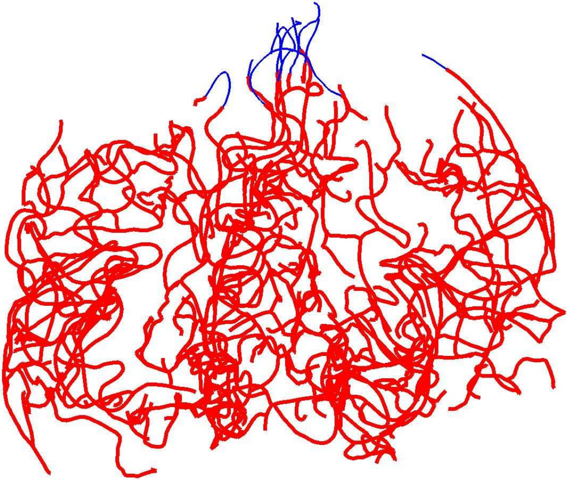
Example: filling brains



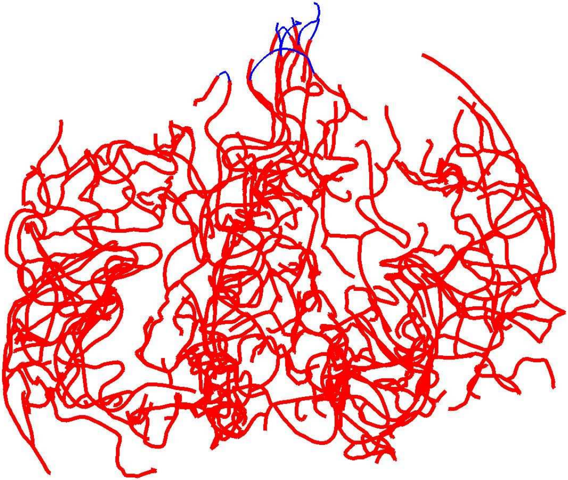
Example: filling brains



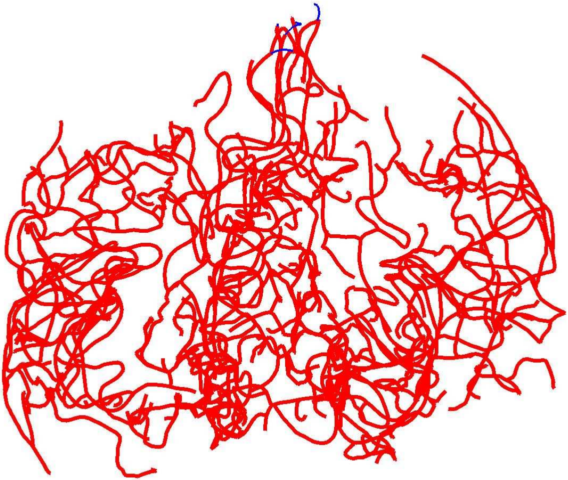
Example: filling brains



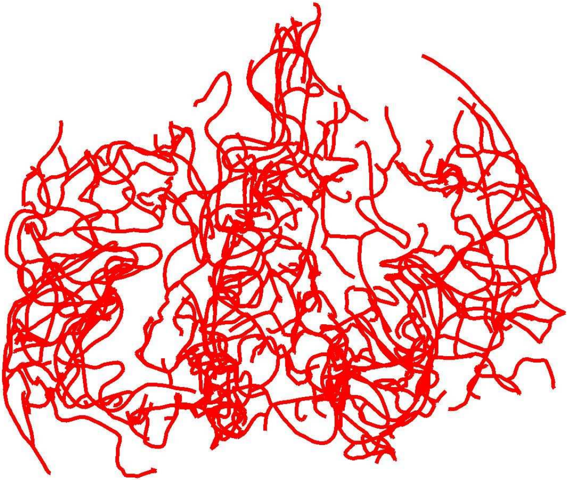
Example: filling brains



Example: filling brains



Example: filling brains



Persistent homology

Build X step by step

- measure evolving topology.

Def. Suppose X is a filtered space, meaning X is a union of an increasing sequence of subspaces: $\varnothing = X_0 \subset X_1 \subset \cdots \subset X_m = X$.

- The persistent homology of this filtration is $H_i X_1 \to H_i X_2 \to \cdots \to H_i X_m$, a sequence of vector space homomorphisms.
- A feature persists from j to k if it appears first in $H_i X_j$ and last in $H_i X_k$.

Examples:

1. Given a function $f : X \to \mathbb{R}$, let $X_t = \{x \in X \mid f(x) \leq t\}$. Good choice of $t_0, \ldots, t_m \in \mathbb{R}$: the values of t across which $H_i X_t$ changes.
2. Any simplicial complex: build it simplex by simplex in some order.

History. invented by [Frosini, Landi 1999], [Robins 1999];
[Edelsbrunner, Letscher, Zomorodian 2002]: includes efficient computation;
[Carlsson, Zomorodian 2009]: multiparameter persistence;
[Knudson and many others]: further theoretical developments, applications

Persistent homology

Build X step by step

- measure evolving topology.

Def. Suppose X is a filtered space, meaning X is a union of an increasing sequence of subspaces: $\varnothing = X_0 \subset X_1 \subset \cdots \subset X_m = X$.

- The persistent homology of this filtration is $H_i X_1 \to H_i X_2 \to \cdots \to H_i X_m$, a sequence of vector space homomorphisms.
- A feature persists from j to k if it appears first in $H_i X_j$ and last in $H_i X_k$.

Examples:

1. Given a function $f : X \to \mathbb{R}$, let $X_t = \{x \in X \mid f(x) \leq t\}$. Good choice of $t_0, \ldots, t_m \in \mathbb{R}$: the values of t across which $H_i X_t$ changes.
2. Any simplicial complex: build it simplex by simplex in some order.

History. invented by [Frosini, Landi 1999], [Robins 1999];
[Edelsbrunner, Letscher, Zomorodian 2002]: includes efficient computation;
[Carlsson, Zomorodian 2009]: multiparameter persistence;
[Knudson and many others]: further theoretical developments, applications

Persistent homology

Build X step by step

- measure evolving topology.

Def. Suppose X is a filtered space, meaning X is a union of an increasing sequence of subspaces: $\varnothing = X_0 \subset X_1 \subset \cdots \subset X_m = X$.

- The persistent homology of this filtration is $H_i X_1 \to H_i X_2 \to \cdots \to H_i X_m$, a sequence of vector space homomorphisms.
- A feature persists from j to k if it appears first in $H_i X_j$ and last in $H_i X_k$.

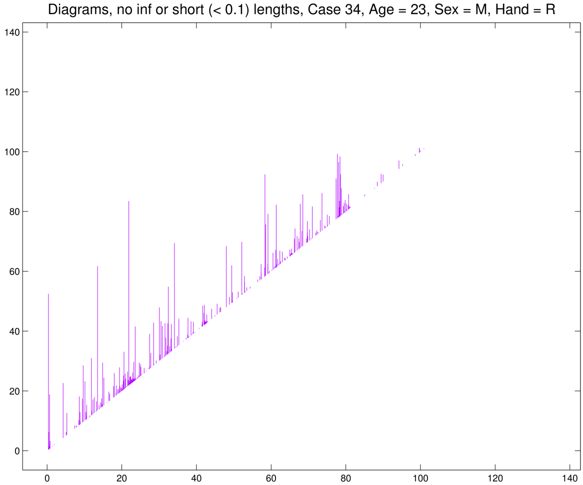
Examples:

1. Given a function $f : X \to \mathbb{R}$, let $X_t = \{x \in X \mid f(x) \leq t\}$. Good choice of $t_0, \ldots, t_m \in \mathbb{R}$: the values of t across which $H_i X_t$ changes.
2. Any simplicial complex: build it simplex by simplex in some order.

History. invented by [Frosini, Landi 1999], [Robins 1999];
[Edelsbrunner, Letscher, Zomorodian 2002]: includes efficient computation;
[Carlsson, Zomorodian 2009]: multiparameter persistence;
[Knudson and many others]: further theoretical developments, applications

Bar codes

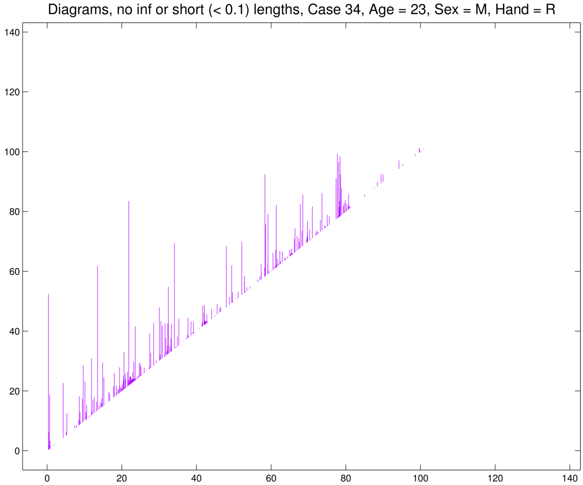
Data structure: 3D tree ⇝ bar code / lace array / persistence diagram:



- multiset of (vertical) line segments $[t, t']$ (plotted at x-coordinate t)
- one for each class with birth time t and death time t'.

Bar codes

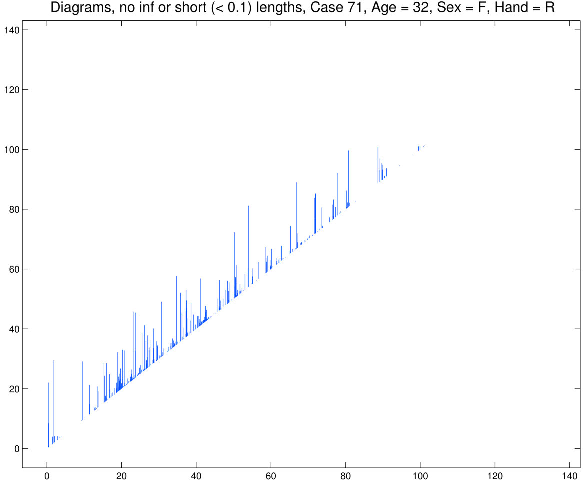
Data structure: 3D tree ⇝ bar code / lace array / persistence diagram:



- multiset of (vertical) line segments $[t, t']$ (plotted at x-coordinate t)
- one for each class with birth time t and death time t'.

Bar codes

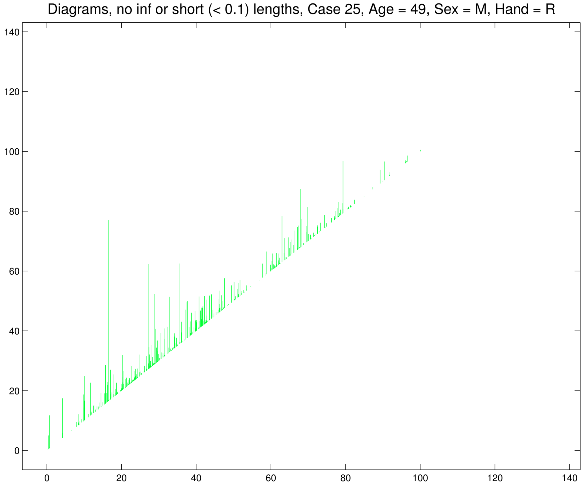
Data structure: 3D tree ⇝ bar code / lace array / persistence diagram:



- multiset of (vertical) line segments $[t, t']$ (plotted at x-coordinate t)
- one for each class with birth time t and death time t'.

Bar codes

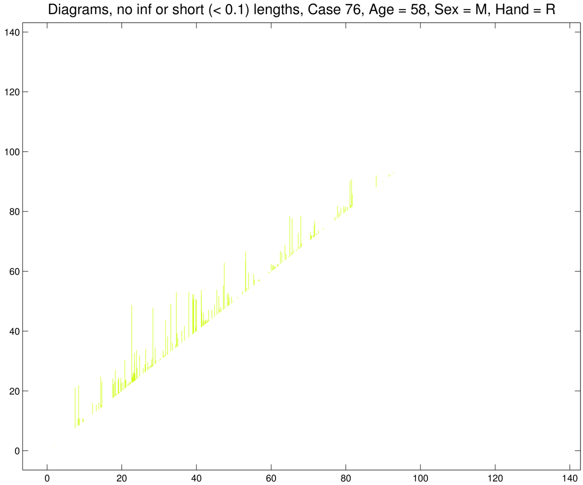
Data structure: 3D tree ⇝ bar code / lace array / persistence diagram:



- multiset of (vertical) line segments $[t, t']$ (plotted at x-coordinate t)
- one for each class with birth time t and death time t'.

Bar codes

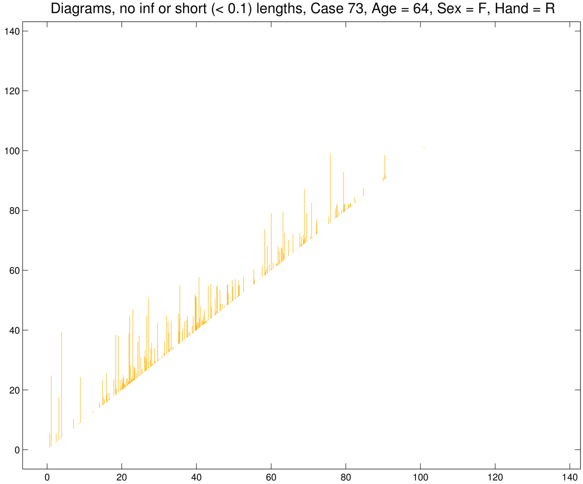
Data structure: 3D tree ⇝ bar code / lace array / persistence diagram:



- multiset of (vertical) line segments $[t, t']$ (plotted at x-coordinate t)
- one for each class with birth time t and death time t'.

Bar codes

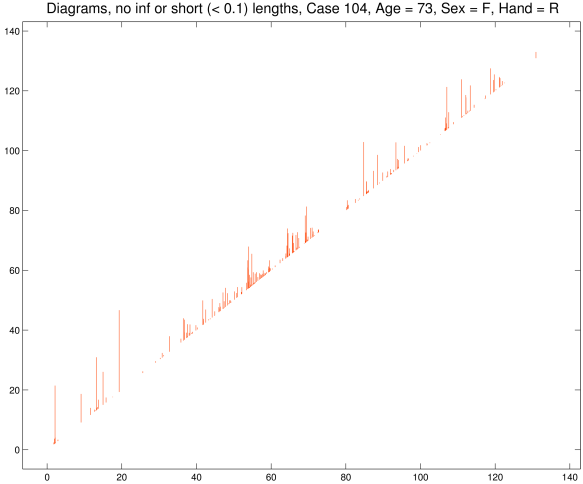
Data structure: 3D tree ⇝ bar code / lace array / persistence diagram:



- multiset of (vertical) line segments $[t, t']$ (plotted at x-coordinate t)
- one for each class with birth time t and death time t'.

Bar codes

Data structure: 3D tree ⇝ bar code / lace array / persistence diagram:



- multiset of (vertical) line segments $[t, t']$ (plotted at x-coordinate t)
- one for each class with birth time t and death time t'.

Sweep filtration

Goal: statistical analysis taking into account

- 3D structure, in particular
- “bendiness”, or “tortuosity”

Sweep filtration

Goal: statistical analysis taking into account

- 3D structure, in particular
- “bendiness”, or “tortuosity”

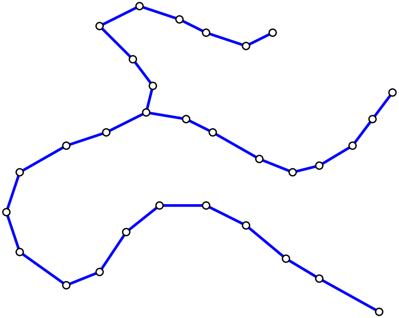
Filter by sweeping across with a plane:

Sweep filtration

Goal: statistical analysis taking into account

- 3D structure, in particular
- “bendiness”, or “tortuosity”

Filter by sweeping across with a plane:

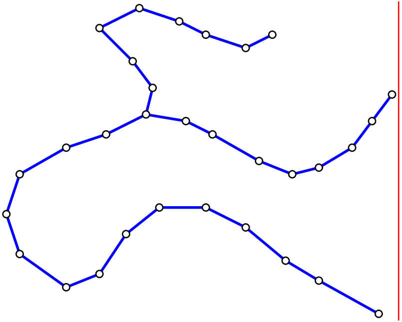


Sweep filtration

Goal: statistical analysis taking into account

- 3D structure, in particular
- “bendiness”, or “tortuosity”

Filter by sweeping across with a plane:

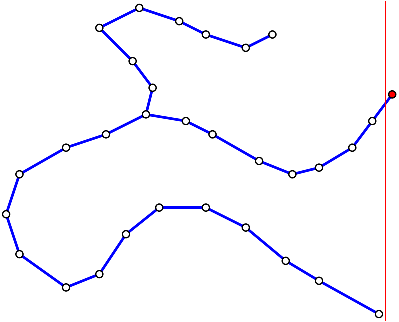


Sweep filtration

Goal: statistical analysis taking into account

- 3D structure, in particular
- “bendiness”, or “tortuosity”

Filter by sweeping across with a plane:

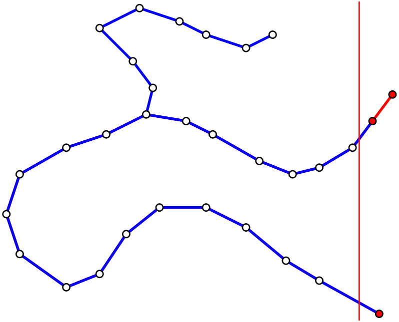


Sweep filtration

Goal: statistical analysis taking into account

- 3D structure, in particular
- “bendiness”, or “tortuosity”

Filter by sweeping across with a plane:

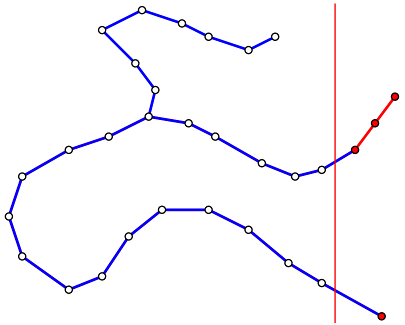


Sweep filtration

Goal: statistical analysis taking into account

- 3D structure, in particular
- “bendiness”, or “tortuosity”

Filter by sweeping across with a plane:

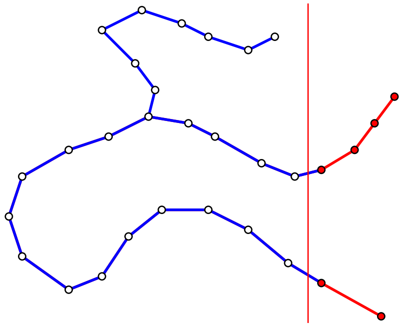


Sweep filtration

Goal: statistical analysis taking into account

- 3D structure, in particular
- “bendiness”, or “tortuosity”

Filter by sweeping across with a plane:

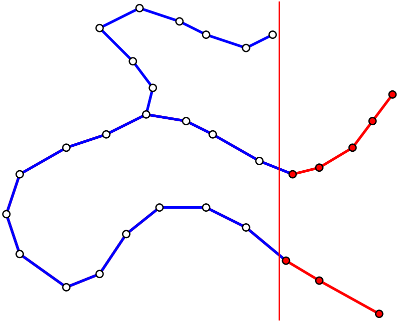


Sweep filtration

Goal: statistical analysis taking into account

- 3D structure, in particular
- “bendiness”, or “tortuosity”

Filter by sweeping across with a plane:

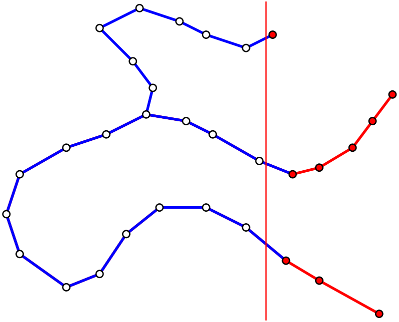


Sweep filtration

Goal: statistical analysis taking into account

- 3D structure, in particular
- “bendiness”, or “tortuosity”

Filter by sweeping across with a plane:

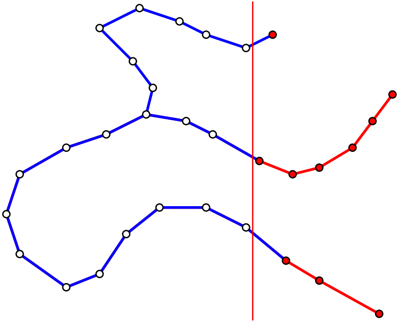


Sweep filtration

Goal: statistical analysis taking into account

- 3D structure, in particular
- “bendiness”, or “tortuosity”

Filter by sweeping across with a plane:

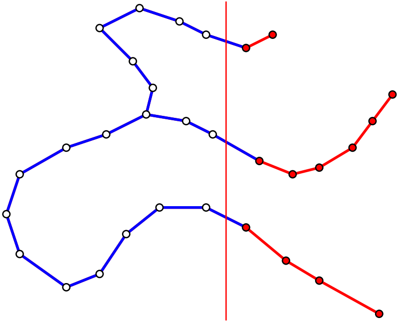


Sweep filtration

Goal: statistical analysis taking into account

- 3D structure, in particular
- “bendiness”, or “tortuosity”

Filter by sweeping across with a plane:

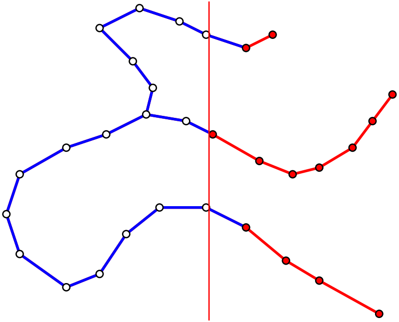


Sweep filtration

Goal: statistical analysis taking into account

- 3D structure, in particular
- “bendiness”, or “tortuosity”

Filter by sweeping across with a plane:

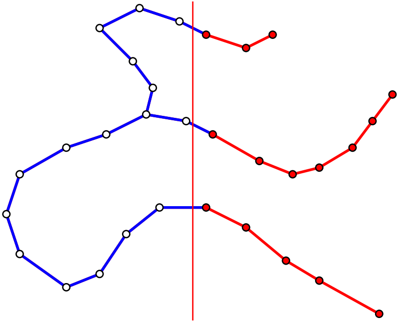


Sweep filtration

Goal: statistical analysis taking into account

- 3D structure, in particular
- “bendiness”, or “tortuosity”

Filter by sweeping across with a plane:

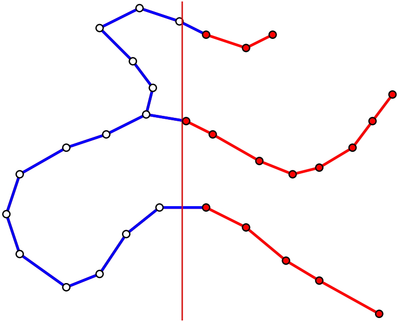


Sweep filtration

Goal: statistical analysis taking into account

- 3D structure, in particular
- “bendiness”, or “tortuosity”

Filter by sweeping across with a plane:

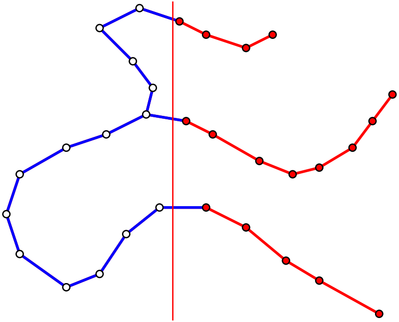


Sweep filtration

Goal: statistical analysis taking into account

- 3D structure, in particular
- "bendiness", or "tortuosity"

Filter by sweeping across with a plane:

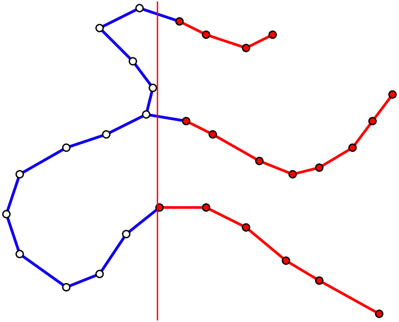


Sweep filtration

Goal: statistical analysis taking into account

- 3D structure, in particular
- "bendiness", or "tortuosity"

Filter by sweeping across with a plane:

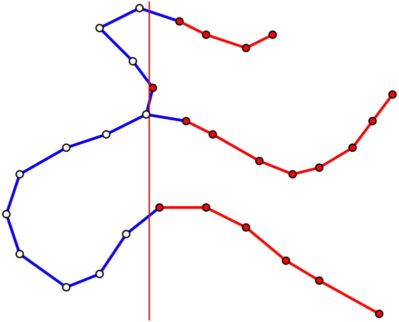


Sweep filtration

Goal: statistical analysis taking into account

- 3D structure, in particular
- “bendiness”, or “tortuosity”

Filter by sweeping across with a plane:

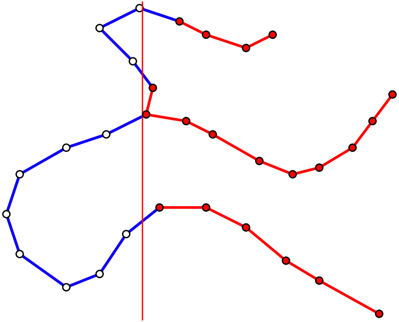


Sweep filtration

Goal: statistical analysis taking into account

- 3D structure, in particular
- “bendiness”, or “tortuosity”

Filter by sweeping across with a plane:

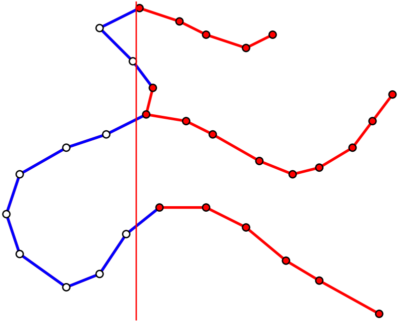


Sweep filtration

Goal: statistical analysis taking into account

- 3D structure, in particular
- “bendiness”, or “tortuosity”

Filter by sweeping across with a plane:

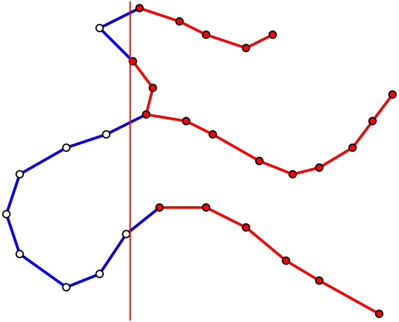


Sweep filtration

Goal: statistical analysis taking into account

- 3D structure, in particular
- “bendiness”, or “tortuosity”

Filter by sweeping across with a plane:

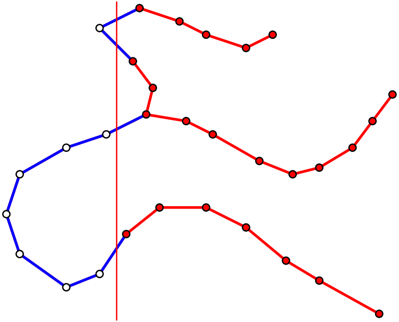


Sweep filtration

Goal: statistical analysis taking into account

- 3D structure, in particular
- “bendiness”, or “tortuosity”

Filter by sweeping across with a plane:

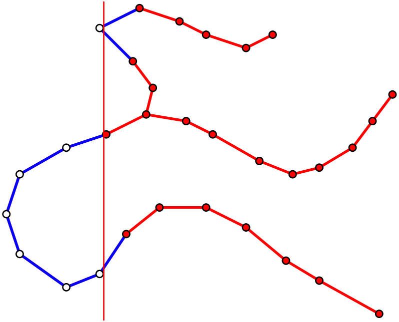


Sweep filtration

Goal: statistical analysis taking into account

- 3D structure, in particular
- “bendiness”, or “tortuosity”

Filter by sweeping across with a plane:

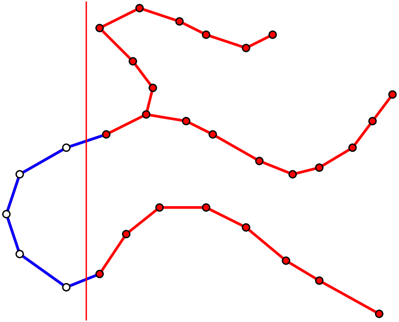


Sweep filtration

Goal: statistical analysis taking into account

- 3D structure, in particular
- “bendiness”, or “tortuosity”

Filter by sweeping across with a plane:

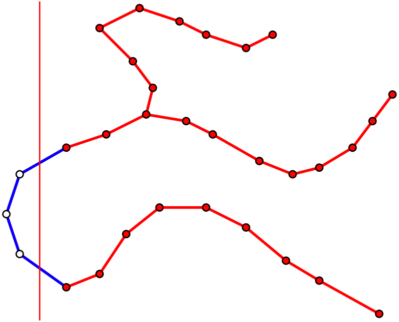


Sweep filtration

Goal: statistical analysis taking into account

- 3D structure, in particular
- “bendiness”, or “tortuosity”

Filter by sweeping across with a plane:

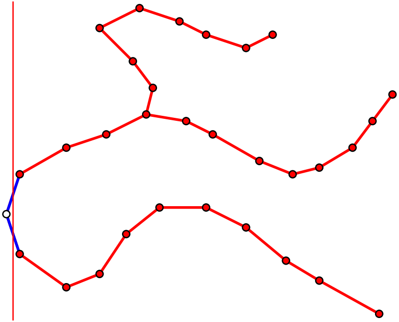


Sweep filtration

Goal: statistical analysis taking into account

- 3D structure, in particular
- “bendiness”, or “tortuosity”

Filter by sweeping across with a plane:

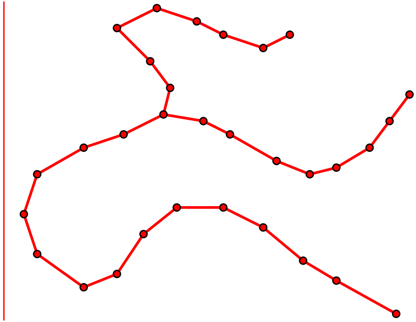


Sweep filtration

Goal: statistical analysis taking into account

- 3D structure, in particular
- “bendiness”, or “tortuosity”

Filter by sweeping across with a plane:

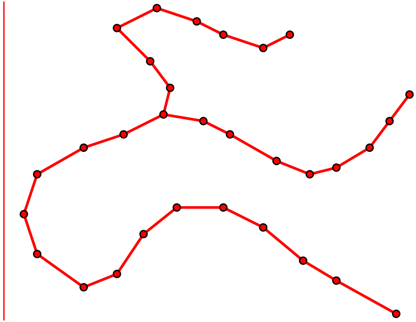


Sweep filtration

Goal: statistical analysis taking into account

- 3D structure, in particular
- “bendiness”, or “tortuosity”

Filter by sweeping across with a plane:



Record:

- birth time of each new component
- death of each component (when it joins to an older component)

Sweep filtration

Goal: statistical analysis taking into account

- 3D structure, in particular
- “bendiness”, or “tortuosity”

Filter by sweeping across with a plane:

Record:

- birth time of each new component
- death of each component (when it joins to an older component)

Sweep filtration

Goal: statistical analysis taking into account

- 3D structure, in particular
- “bendiness”, or “tortuosity”

Filter by sweeping across with a plane:

Record:

- birth time of each new component
- death of each component (when it joins to an older component)

Sweep filtration

Goal: statistical analysis taking into account

- 3D structure, in particular
- “bendiness”, or “tortuosity”

Filter by sweeping across with a plane:

Record:

- birth time of each new component
- death of each component (when it joins to an older component)

Sweep filtration

Goal: statistical analysis taking into account

- 3D structure, in particular
- “bendiness”, or “tortuosity”

Filter by sweeping across with a plane:

Record:

- birth time of each new component
- death of each component (when it joins to an older component)

Sweep filtration

Goal: statistical analysis taking into account

- 3D structure, in particular
- “bendiness”, or “tortuosity”

Filter by sweeping across with a plane:

Record:

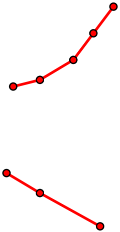
- birth time of each new component
- death of each component (when it joins to an older component)

Sweep filtration

Goal: statistical analysis taking into account

- 3D structure, in particular
- “bendiness”, or “tortuosity”

Filter by sweeping across with a plane:



Record:

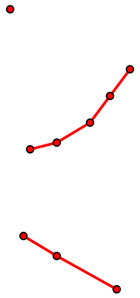
- birth time of each new component
- death of each component (when it joins to an older component)

Sweep filtration

Goal: statistical analysis taking into account

- 3D structure, in particular
- “bendiness”, or “tortuosity”

Filter by sweeping across with a plane:



Record:

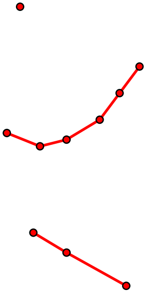
- birth time of each new component
- death of each component (when it joins to an older component)

Sweep filtration

Goal: statistical analysis taking into account

- 3D structure, in particular
- "bendiness", or "tortuosity"

Filter by sweeping across with a plane:



Record:

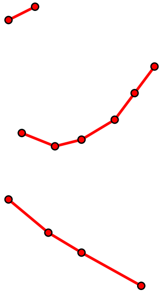
- birth time of each new component
- death of each component (when it joins to an older component)

Sweep filtration

Goal: statistical analysis taking into account

- 3D structure, in particular
- “bendiness”, or “tortuosity”

Filter by sweeping across with a plane:



Record:

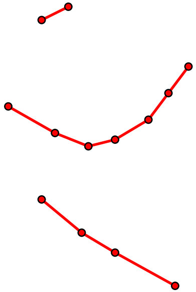
- birth time of each new component
- death of each component (when it joins to an older component)

Sweep filtration

Goal: statistical analysis taking into account

- 3D structure, in particular
- “bendiness”, or “tortuosity”

Filter by sweeping across with a plane:



Record:

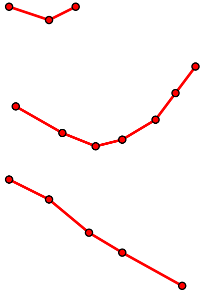
- birth time of each new component
- death of each component (when it joins to an older component)

Sweep filtration

Goal: statistical analysis taking into account

- 3D structure, in particular
- “bendiness”, or “tortuosity”

Filter by sweeping across with a plane:



Record:

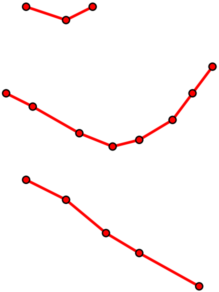
- birth time of each new component
- death of each component (when it joins to an older component)

Sweep filtration

Goal: statistical analysis taking into account

- 3D structure, in particular
- “bendiness”, or “tortuosity”

Filter by sweeping across with a plane:



Record:

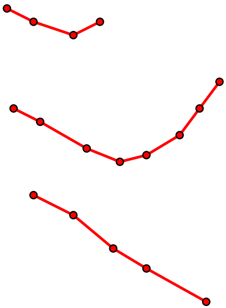
- birth time of each new component
- death of each component (when it joins to an older component)

Sweep filtration

Goal: statistical analysis taking into account

- 3D structure, in particular
- “bendiness”, or “tortuosity”

Filter by sweeping across with a plane:



Record:

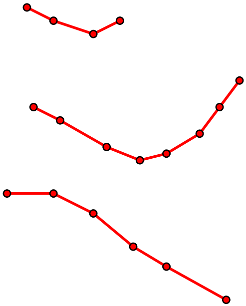
- birth time of each new component
- death of each component (when it joins to an older component)

Sweep filtration

Goal: statistical analysis taking into account

- 3D structure, in particular
- “bendiness”, or “tortuosity”

Filter by sweeping across with a plane:



Record:

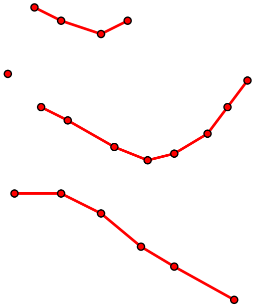
- birth time of each new component
- death of each component (when it joins to an older component)

Sweep filtration

Goal: statistical analysis taking into account

- 3D structure, in particular
- "bendiness", or "tortuosity"

Filter by sweeping across with a plane:



Record:

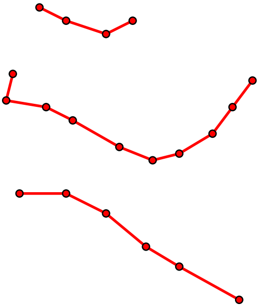
- birth time of each new component
- death of each component (when it joins to an older component)

Sweep filtration

Goal: statistical analysis taking into account

- 3D structure, in particular
- “bendiness”, or “tortuosity”

Filter by sweeping across with a plane:



Record:

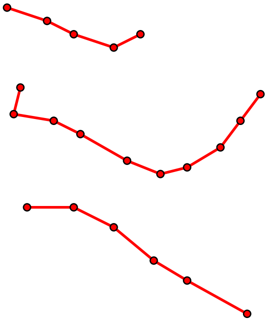
- birth time of each new component
- death of each component (when it joins to an older component)

Sweep filtration

Goal: statistical analysis taking into account

- 3D structure, in particular
- “bendiness”, or “tortuosity”

Filter by sweeping across with a plane:



Record:

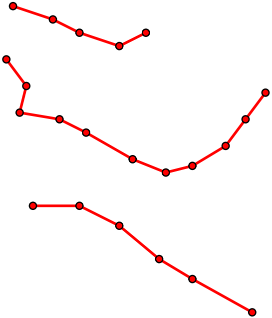
- birth time of each new component
- death of each component (when it joins to an older component)

Sweep filtration

Goal: statistical analysis taking into account

- 3D structure, in particular
- “bendiness”, or “tortuosity”

Filter by sweeping across with a plane:



Record:

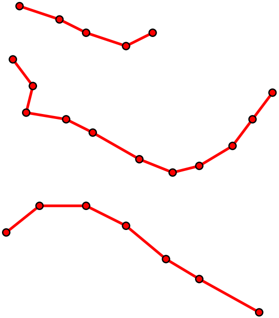
- birth time of each new component
- death of each component (when it joins to an older component)

Sweep filtration

Goal: statistical analysis taking into account

- 3D structure, in particular
- “bendiness”, or “tortuosity”

Filter by sweeping across with a plane:



Record:

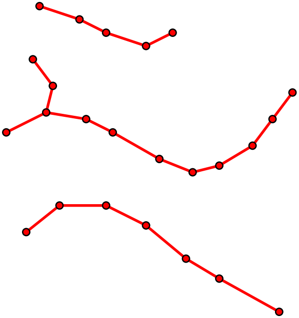
- birth time of each new component
- death of each component (when it joins to an older component)

Sweep filtration

Goal: statistical analysis taking into account

- 3D structure, in particular
- “bendiness”, or “tortuosity”

Filter by sweeping across with a plane:



Record:

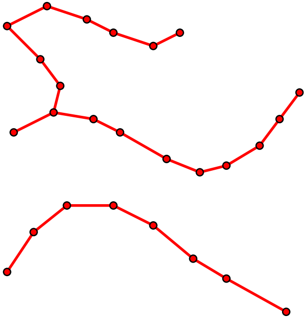
- birth time of each new component
- death of each component (when it joins to an older component)

Sweep filtration

Goal: statistical analysis taking into account

- 3D structure, in particular
- “bendiness”, or “tortuosity”

Filter by sweeping across with a plane:



Record:

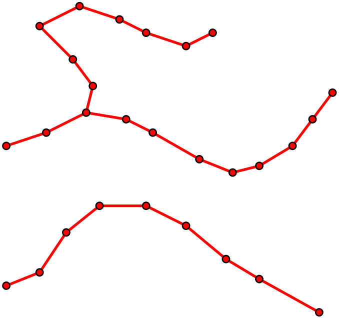
- birth time of each new component
- death of each component (when it joins to an older component)

Sweep filtration

Goal: statistical analysis taking into account

- 3D structure, in particular
- “bendiness”, or “tortuosity”

Filter by sweeping across with a plane:



Record:

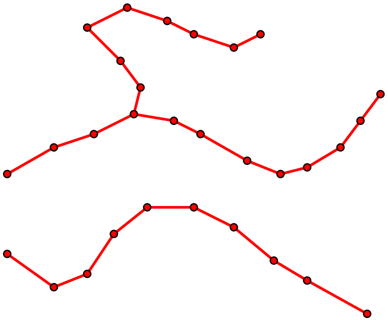
- birth time of each new component
- death of each component (when it joins to an older component)

Sweep filtration

Goal: statistical analysis taking into account

- 3D structure, in particular
- "bendiness", or "tortuosity"

Filter by sweeping across with a plane:



Record:

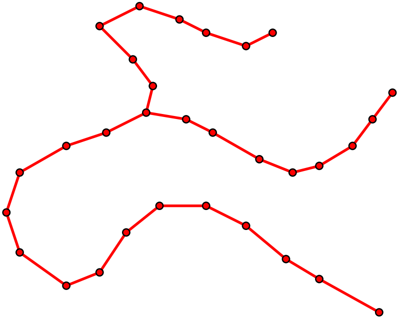
- birth time of each new component
- death of each component (when it joins to an older component)

Sweep filtration

Goal: statistical analysis taking into account

- 3D structure, in particular
- “bendiness”, or “tortuosity”

Filter by sweeping across with a plane:



Record:

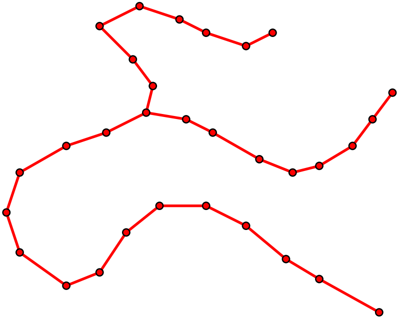
- birth time of each new component
- death of each component (when it joins to an older component)

Sweep filtration

Goal: statistical analysis taking into account

- 3D structure, in particular
- “bendiness”, or “tortuosity”

Filter by sweeping across with a plane:



Record:

- birth time of each new component
- death of each component (when it joins to an older component)

Easily computable (if dim X is low; ambient space dim irrelevant).

Statistical analysis

Reduce to linear methods. 3D tree $\rightsquigarrow$ bar code $\rightsquigarrow$ vector in $\mathbb{R}^{100}$:

- top 100 bar lengths, in decreasing order, log scale
- correlate first principal component score vs. age

Conclusions. [Bendich, Marron, M.—, Pieloch, Skwerer 2014]
Longest bars in older brains tend to be shorter and later.

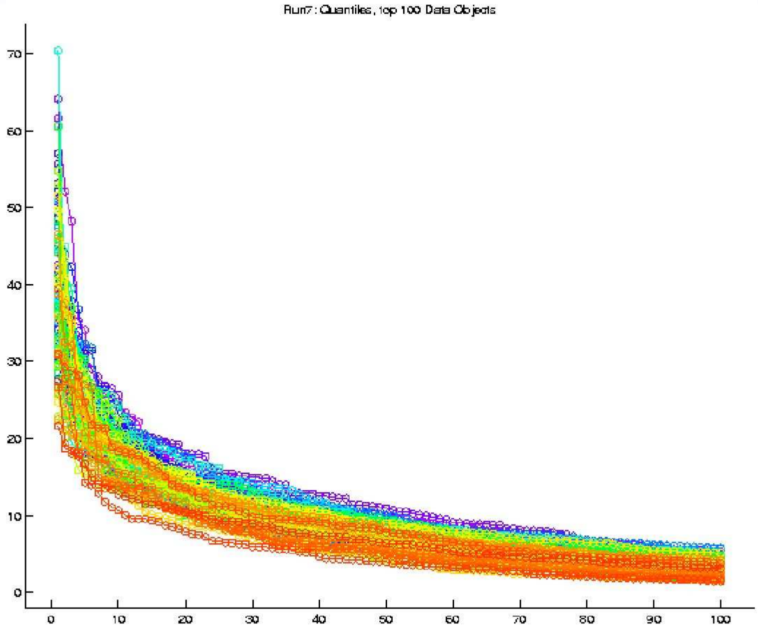
- Pearson correlation 0.52663
- p-value 3.0127×10^{-8} strongly significant

Remarks. Results essentially unchanged after

- rescaling to account for natural variation in overall brain size (force standard deviation of the set of bar lengths to equal 1)
- rescaling to account for known correlation of age vs. total vessel length L [Bullitt, et al. 2005] (divide by L, $\sqrt{L}$, or $\sqrt[3]{L}$)
- repeating the analysis with residuals from regression between feature vector and total length.

Moral. Persistent homology can topologically detect statistically significant geometric motifs.

Top 100 bars


Run7 : Quantiles, top 100 Data Objects
70
60
50
40
30
20
10
0
0
10
20
30
40
50
60
70
80
90
100

Statistical analysis

Reduce to linear methods. 3D tree $\rightsquigarrow$ bar code $\rightsquigarrow$ vector in $\mathbb{R}^{100}$:

- top 100 bar lengths, in decreasing order, log scale
- correlate first principal component score vs. age

Conclusions. [Bendich, Marron, M.—, Pieloch, Skwerer 2016]
Longest bars in older brains tend to be shorter and later.

- Pearson correlation 0.52663
- p-value 3.0127×10^{-8} strongly significant

Remarks. Results essentially unchanged after

- rescaling to account for natural variation in overall brain size (force standard deviation of the set of bar lengths to equal 1)
- rescaling to account for known correlation of age vs. total vessel length L [Bullitt, et al. 2005] (divide by L, $\sqrt{L}$, or $\sqrt[3]{L}$)
- repeating the analysis with residuals from regression between feature vector and total length.

Moral. Persistent homology can topologically detect statistically significant geometric motifs.

Statistical analysis

Reduce to linear methods. 3D tree $\rightsquigarrow$ bar code $\rightsquigarrow$ vector in $\mathbb{R}^{100}$:

- top 100 bar lengths, in decreasing order, log scale
- correlate first principal component score vs. age

Conclusions. [Bendich, Marron, M.—, Pieloch, Skwerer 2016]
Longest bars in older brains tend to be shorter and later.

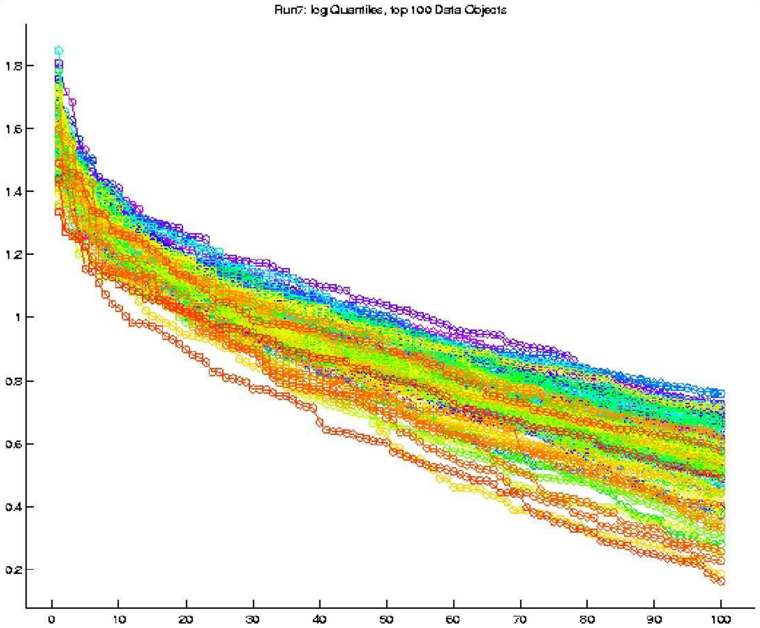
- Pearson correlation 0.52663
- p-value 3.0127×10^{-8} strongly significant

Remarks. Results essentially unchanged after

- rescaling to account for natural variation in overall brain size (force standard deviation of the set of bar lengths to equal 1)
- rescaling to account for known correlation of age vs. total vessel length L [Bullitt, et al. 2005] (divide by L, $\sqrt{L}$, or $\sqrt[3]{L}$)
- repeating the analysis with residuals from regression between feature vector and total length.

Moral. Persistent homology can topologically detect statistically significant geometric motifs.

Top 100 bars: log scale



Statistical analysis

Reduce to linear methods. 3D tree $\rightsquigarrow$ bar code $\rightsquigarrow$ vector in $\mathbb{R}^{100}$:

- top 100 bar lengths, in decreasing order, log scale
- correlate first principal component score vs. age

Conclusions. [Bendich, Marron, M.—, Pieloch, Skwerer 2014]
Longest bars in older brains tend to be shorter and later.

- Pearson correlation 0.52663
- p-value 3.0127×10^{-8} strongly significant

Remarks. Results essentially unchanged after

- rescaling to account for natural variation in overall brain size (force standard deviation of the set of bar lengths to equal 1)
- rescaling to account for known correlation of age vs. total vessel length L [Bullitt, et al. 2005] (divide by L, $\sqrt{L}$, or $\sqrt[3]{L}$)
- repeating the analysis with residuals from regression between feature vector and total length.

Moral. Persistent homology can topologically detect statistically significant geometric motifs.

Statistical analysis

Reduce to linear methods. 3D tree $\rightsquigarrow$ bar code $\rightsquigarrow$ vector in $\mathbb{R}^{100}$:

- top 100 bar lengths, in decreasing order, log scale
- correlate first principal component score vs. age

Conclusions. [Bendich, Marron, M.—, Pieloch, Skwerer 2014]
Longest bars in older brains tend to be shorter and later.

- Pearson correlation 0.52663
- p-value 3.0127×10^{-8} strongly significant

Remarks. Results essentially unchanged after

- rescaling to account for natural variation in overall brain size (force standard deviation of the set of bar lengths to equal 1)
- rescaling to account for known correlation of age vs. total vessel length L [Bullitt, et al. 2005] (divide by L, $\sqrt{L}$, or $\sqrt[3]{L}$)
- repeating the analysis with residuals from regression between feature vector and total length.

Moral. Persistent homology can topologically detect statistically significant geometric motifs.

Statistical analysis

Reduce to linear methods. 3D tree $\rightsquigarrow$ bar code $\rightsquigarrow$ vector in $\mathbb{R}^{100}$:

- top 100 bar lengths, in decreasing order, log scale
- correlate first principal component score vs. age

Conclusions. [Bendich, Marron, M.—, Pieloch, Skwerer 2014]
Longest bars in older brains tend to be shorter and later.

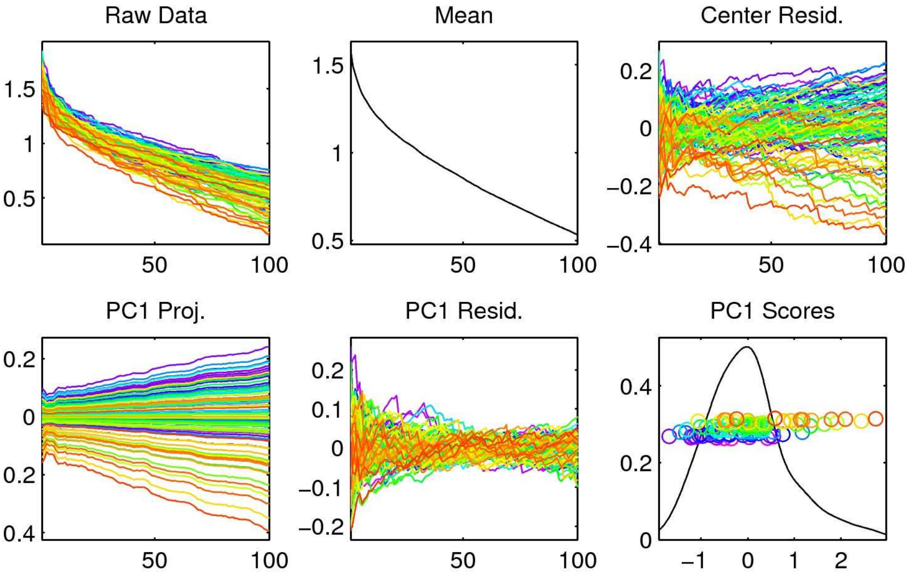
- Pearson correlation 0.52663
- p-value 3.0127×10^{-8} strongly significant

Remarks. Results essentially unchanged after

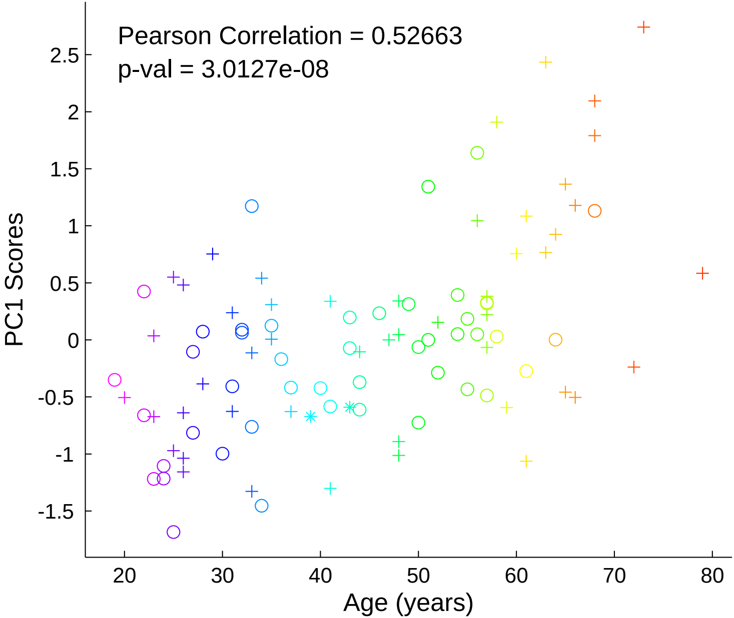
- rescaling to account for natural variation in overall brain size (force standard deviation of the set of bar lengths to equal 1)
- rescaling to account for known correlation of age vs. total vessel length L [Bullitt, et al. 2005] (divide by L, $\sqrt{L}$, or $\sqrt[3]{L}$)
- repeating the analysis with residuals from regression between feature vector and total length.

Moral. Persistent homology can topologically detect statistically significant geometric motifs.

Age vs. PC1


Raw Data
1.5
1
0.5
50
100
Mean
1.5
1
0.5
50
100
Center Resid.
0.2
0
−0.2
−0.4
50
100
PC1 Proj.
0.2
0
0.2
0.4
50
100
PC1 Resid.
0.2
0.1
0
−0.1
−0.2
50
100
PC1 Scores
0.4
0.2
0
−1
0
1
2

Age vs. PC1


Pearson Correlation = 0.52663
p-val = 3.0127e-08
PC1 Scores
Age (years)
2.5
2
1.5
1
0.5
0
-0.5
-1
-1.5
20
30
40
50
60
70
80

Statistical analysis

Reduce to linear methods. 3D tree $\rightsquigarrow$ bar code $\rightsquigarrow$ vector in $\mathbb{R}^{100}$:

- top 100 bar lengths, in decreasing order, log scale
- correlate first principal component score vs. age

Conclusions. [Bendich, Marron, M.—, Pieloch, Skwerer 2014]
Longest bars in older brains tend to be shorter and later.

- Pearson correlation 0.52663
- p-value 3.0127×10^{-8} strongly significant

Remarks. Results essentially unchanged after

- rescaling to account for natural variation in overall brain size (force standard deviation of the set of bar lengths to equal 1)
- rescaling to account for known correlation of age vs. total vessel length L [Bullitt, et al. 2005] (divide by L, $\sqrt{L}$, or $\sqrt[3]{L}$)
- repeating the analysis with residuals from regression between feature vector and total length.

Moral. Persistent homology can topologically detect statistically significant geometric motifs.

Statistical analysis

Reduce to linear methods. 3D tree $\rightsquigarrow$ bar code $\rightsquigarrow$ vector in $\mathbb{R}^{100}$:

- top 100 bar lengths, in decreasing order, log scale
- correlate first principal component score vs. age

Conclusions. [Bendich, Marron, M.—, Pieloch, Skwerer 2014]
Longest bars in older brains tend to be shorter and later.

- Pearson correlation 0.52663
- p-value 3.0127×10^{-8} strongly significant

Remarks. Results essentially unchanged after

- rescaling to account for natural variation in overall brain size (force standard deviation of the set of bar lengths to equal 1)
- rescaling to account for known correlation of age vs. total vessel length L [Bullitt, et al. 2005] (divide by L, $\sqrt{L}$, or $\sqrt[3]{L}$)
- repeating the analysis with residuals from regression between feature vector and total length.

Moral. Persistent homology can topologically detect statistically significant geometric motifs.

Statistical analysis

Reduce to linear methods. 3D tree $\rightsquigarrow$ bar code $\rightsquigarrow$ vector in $\mathbb{R}^{100}$:

- top 100 bar lengths, in decreasing order, log scale
- correlate first principal component score vs. age

Conclusions. [Bendich, Marron, M.—, Pieloch, Skwerer 2014]
Longest bars in older brains tend to be shorter and later.

- Pearson correlation 0.52663
- p-value 3.0127×10^{-8} strongly significant

Remarks. Results essentially unchanged after

- rescaling to account for natural variation in overall brain size (force standard deviation of the set of bar lengths to equal 1)
- rescaling to account for known correlation of age vs. total vessel length L [Bullitt, et al. 2005] (divide by L, $\sqrt{L}$, or $\sqrt[3]{L}$)
- repeating the analysis with residuals from regression between feature vector and total length.

Moral. Persistent homology can topologically detect statistically significant geometric motifs.

Statistical analysis

Reduce to linear methods. 3D tree $\rightsquigarrow$ bar code $\rightsquigarrow$ vector in $\mathbb{R}^{100}$:

- top 100 bar lengths, in decreasing order, log scale
- correlate first principal component score vs. age

Conclusions. [Bendich, Marron, M.—, Pieloch, Skwerer 2014]
Longest bars in older brains tend to be shorter and later.

- Pearson correlation 0.52663
- p-value 3.0127×10^{-8} strongly significant

Remarks. Results essentially unchanged after

- rescaling to account for natural variation in overall brain size (force standard deviation of the set of bar lengths to equal 1)
- rescaling to account for known correlation of age vs. total vessel length L [Bullitt, et al. 2005] (divide by L, $\sqrt{L}$, or $\sqrt[3]{L}$)
- repeating the analysis with residuals from regression between feature vector and total length.

Moral. Persistent homology can topologically detect statistically significant geometric motifs.

Statistical analysis

Reduce to linear methods. 3D tree $\rightsquigarrow$ bar code $\rightsquigarrow$ vector in $\mathbb{R}^{100}$:

- top 100 bar lengths, in decreasing order, log scale
- correlate first principal component score vs. age

Conclusions. [Bendich, Marron, M.—, Pieloch, Skwerer 2014]
Longest bars in older brains tend to be shorter and later.

- Pearson correlation 0.52663
- p-value 3.0127×10^{-8} strongly significant

Remarks. Results essentially unchanged after

- rescaling to account for natural variation in overall brain size (force standard deviation of the set of bar lengths to equal 1)
- rescaling to account for known correlation of age vs. total vessel length L [Bullitt, et al. 2005] (divide by L, $\sqrt{L}$, or $\sqrt[3]{L}$)
- repeating the analysis with residuals from regression between feature vector and total length.

Moral. Persistent homology can topologically detect statistically significant geometric motifs.

Statistical analysis

Reduce to linear methods. 3D tree $\rightsquigarrow$ bar code $\rightsquigarrow$ vector in $\mathbb{R}^{100}$:

- top 100 bar lengths, in decreasing order, log scale
- correlate first principal component score vs. age

Conclusions. [Bendich, Marron, M.—, Pieloch, Skwerer 2014]
Longest bars in older brains tend to be shorter and later.

- Pearson correlation 0.52663
- p-value 3.0127×10^{-8} strongly significant

Remarks. Results essentially unchanged after

- rescaling to account for natural variation in overall brain size (force standard deviation of the set of bar lengths to equal 1)
- rescaling to account for known correlation of age vs. total vessel length L [Bullitt, et al. 2005] (divide by L, $\sqrt{L}$, or $\sqrt[3]{L}$)
- repeating the analysis with residuals from regression between feature vector and total length.

Moral. Persistent homology can topologically detect statistically significant geometric motifs.

Reflections on persistent homology

Where did the best correlation occur?

- How did we choose top 100 bar lengths?
- What choices yield the best correlation? Why?

Persistent homology mantra: most significant features

- are “biggest”
- live “far from the diagonal” in bar codes.

For brain artery trees.

- Not surprising that very short bars $\leftrightarrow$ noise, although in future studies they might not.
- While biggest features are important,
- they hinder strength of correlation.

Morals.

- Importance $\not\Rightarrow$ significance for geometric features.
- Persistent homology can detect significant features lying between important and noise.

Reflections on persistent homology

Where did the best correlation occur?

- How did we choose top 100 bar lengths?
- What choices yield the best correlation? Why?

Persistent homology mantra: most significant features

- are “biggest”
- live “far from the diagonal” in bar codes.

For brain artery trees.

- Not surprising that very short bars $\leftrightarrow$ noise, although in future studies they might not.
- While biggest features are important,
- they hinder strength of correlation.

Morals.

- Importance $\not\Rightarrow$ significance for geometric features.
- Persistent homology can detect significant features lying between important and noise.

Reflections on persistent homology

Where did the best correlation occur?

- How did we choose top 100 bar lengths?
- What choices yield the best correlation? Why?

Persistent homology mantra: most significant features

- are "biggest"
- live "far from the diagonal" in bar codes.

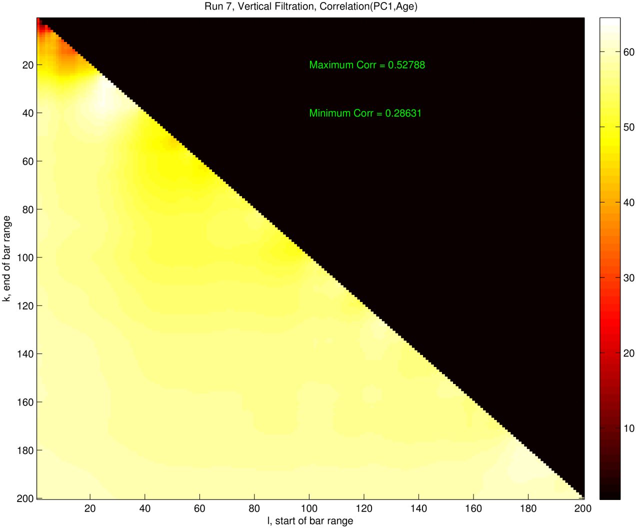
For brain artery trees.

- Not surprising that very short bars $\leftrightarrow$ noise, although in future studies they might not.
- While biggest features are important,
- they hinder strength of correlation.

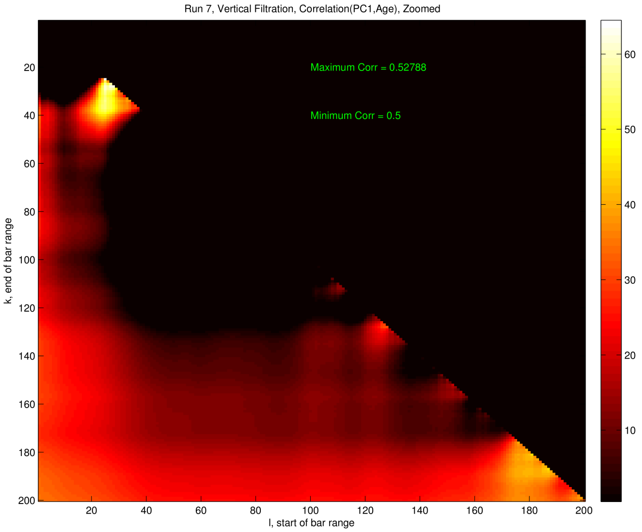
Morals.

- Importance $\not\Rightarrow$ significance for geometric features.
- Persistent homology can detect significant features lying between important and noise.

Top 100 bars


Run 7, Vertical Filtration, Correlation(PC1,Age)
Maximum Corr = 0.52788
Minimum Corr = 0.28631
k, end of bar range
l, start of bar range
20
40
60
80
100
120
140
160
180
200
10
30
50

Top 100 bars


Run 7, Vertical Filtration, Correlation(PC1,Age), Zoomed
Maximum Corr = 0.52788
Minimum Corr = 0.5
k, end of bar range
l, start of bar range
20
40
60
80
100
120
140
160
180
200
10
20
30
40
50
60

Reflections on persistent homology

Where did the best correlation occur?

- How did we choose top 100 bar lengths?
- What choices yield the best correlation? Why?

Persistent homology mantra: most significant features

- are “biggest”
- live “far from the diagonal” in bar codes.

For brain artery trees.

- Not surprising that very short bars $\leftrightarrow$ noise, although in future studies they might not.
- While biggest features are important,
- they hinder strength of correlation.

Morals.

- Importance $\not\Rightarrow$ significance for geometric features.
- Persistent homology can detect significant features lying between important and noise.

Reflections on persistent homology

Where did the best correlation occur?

- How did we choose top 100 bar lengths?
- What choices yield the best correlation? Why?

Persistent homology mantra: most significant features

- are “biggest”
- live “far from the diagonal” in bar codes.

For brain artery trees.

- Not surprising that very short bars $\leftrightarrow$ noise, although in future studies they might not.
- While biggest features are important,
- they hinder strength of correlation.

Morals.

- Importance $\not\Rightarrow$ significance for geometric features.
- Persistent homology can detect significant features lying between important and noise.

Reflections on persistent homology

Where did the best correlation occur?

- How did we choose top 100 bar lengths?
- What choices yield the best correlation? Why?

Persistent homology mantra: most significant features

- are “biggest”
- live “far from the diagonal” in bar codes.

For brain artery trees.

- Not surprising that very short bars $\leftrightarrow$ noise, although in future studies they might not.
- While biggest features are important,
- they hinder strength of correlation.

Morals.

- Importance $\not\Rightarrow$ significance for geometric features.
- Persistent homology can detect significant features lying between important and noise.

Lesson for students. Integration of biology, math, stat, and computation in research and application.

Future directions

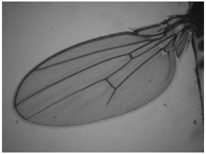
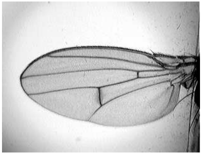
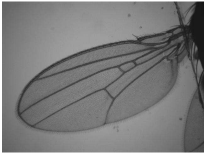
- fruit fly wings (with Houle, Thomas, Curry, Beriwal)

- lung arteries (with McLean et al., Marron)

- fMRI (with Lazar et al.)

Future directions

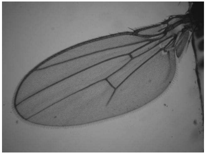
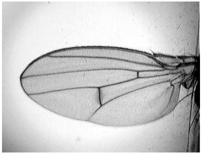
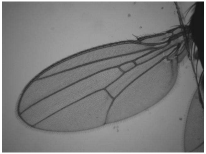
- fruit fly wings (with Houle, Thomas, Curry, Beriwal)

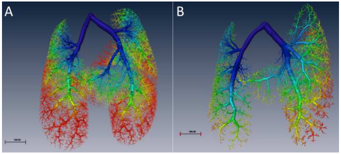
- lung arteries (with McLean et al., Marron)
- fMRI (with Lazar et al.)

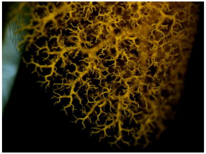
Future directions

- fruit fly wings (with Houle, Thomas, Curry, Beriwal)

- lung arteries (with McLean et al., Marron)

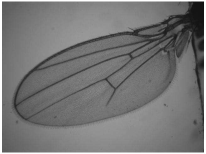
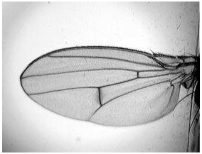
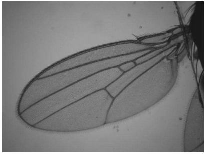




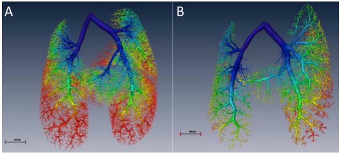
- fMRI (with Lazar et al.)

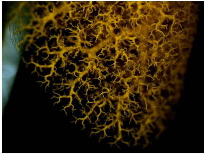
Future directions

- fruit fly wings (with Houle, Thomas, Curry, Beriwal)

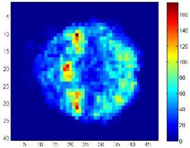
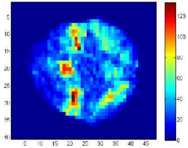
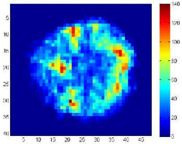




- lung arteries (with McLean et al., Marron)



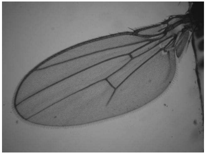


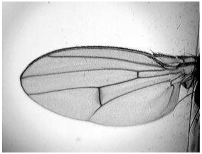
- fMRI (with Lazar et al.)

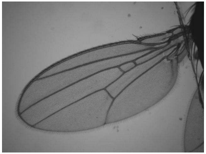




Future directions

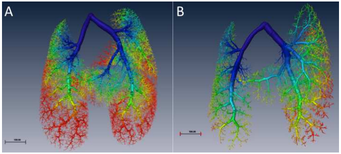
- fruit fly wings (with Houle, Thomas, Curry, Beriwal)

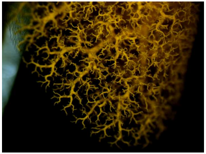




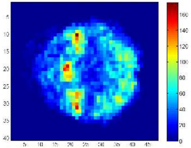


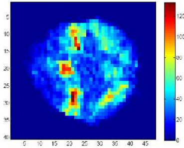
- lung arteries (with McLean et al., Marron)

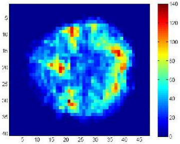




- fMRI (with Lazar et al.)







Thank You