Applying persistent homology to brain artery and vein imaging

Ezra Miller

Duke University, Department of Mathematics

ezra@math.duke.edu

joint with

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

University of Georgia

12 November 2014

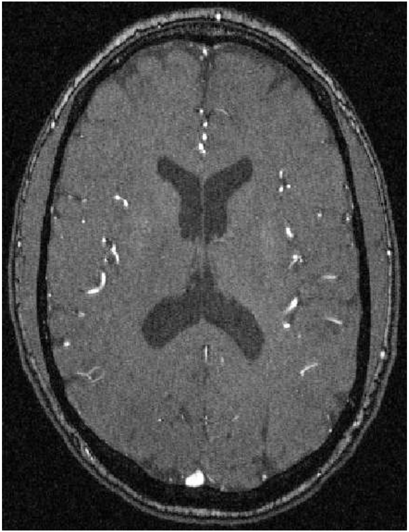
Outline

Brain artery trees

Goal: Statistical analysis taking 3D geometry into account

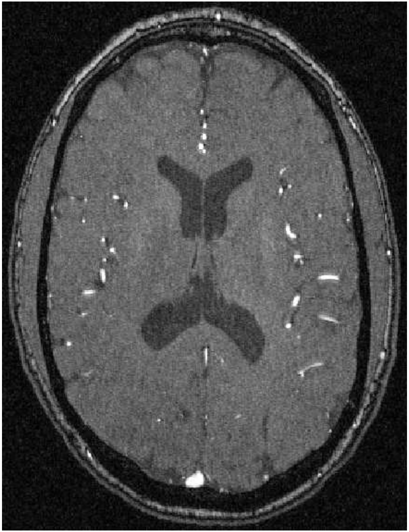
- predict stroke tendency
- screen for loci of pathology, such as tumors
- explore how age affects vascularization

Magnetic Resonance Angiography (MRA)



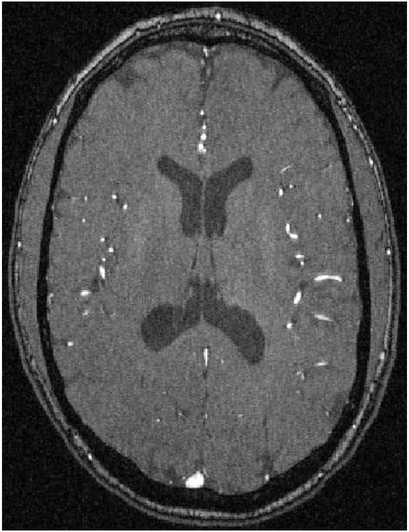
from Elizabeth Bullitt, Dept. of Neurosurgery, UNC-CH

Magnetic Resonance Angiography (MRA)



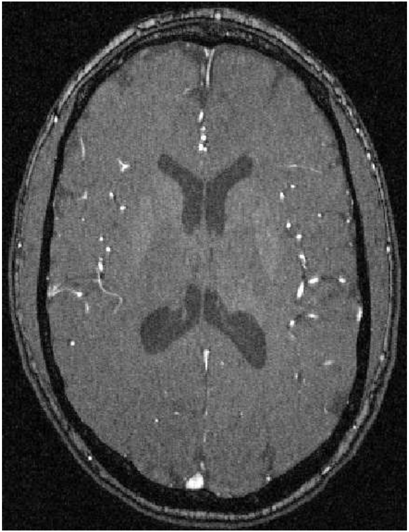
from Elizabeth Bullitt, Dept. of Neurosurgery, UNC-CH

Magnetic Resonance Angiography (MRA)



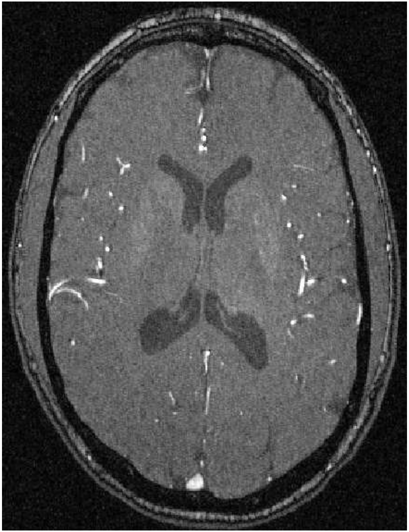
from Elizabeth Bullitt, Dept. of Neurosurgery, UNC-CH

Magnetic Resonance Angiography (MRA)



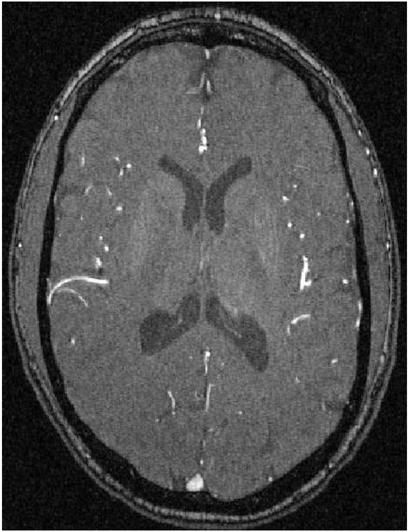
from Elizabeth Bullitt, Dept. of Neurosurgery, UNC-CH

Magnetic Resonance Angiography (MRA)



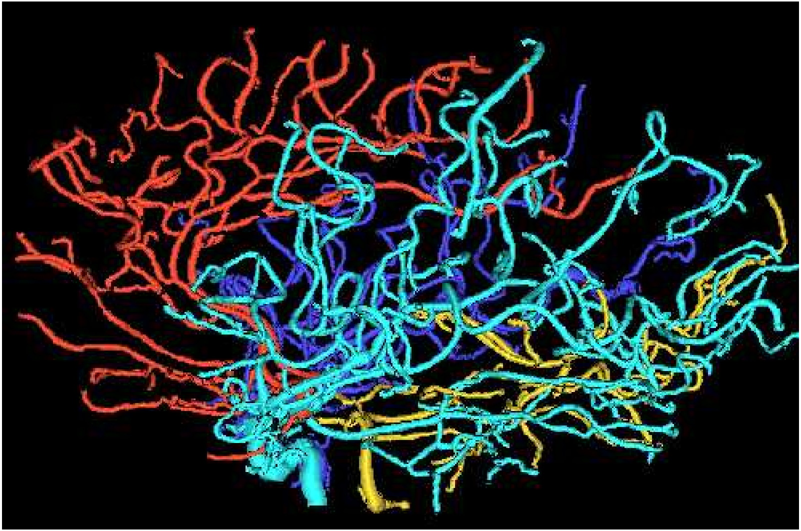
from Elizabeth Bullitt, Dept. of Neurosurgery, UNC-CH

Magnetic Resonance Angiography (MRA)



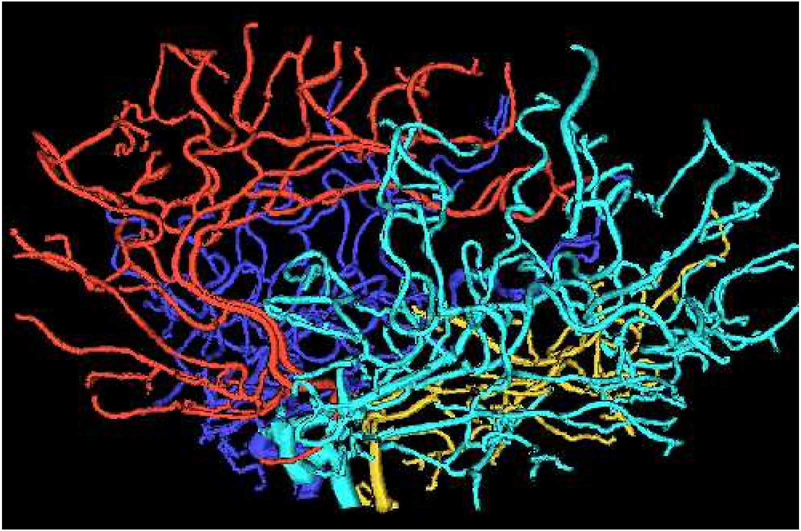
from Elizabeth Bullitt, Dept. of Neurosurgery, UNC-CH

Tube tracking



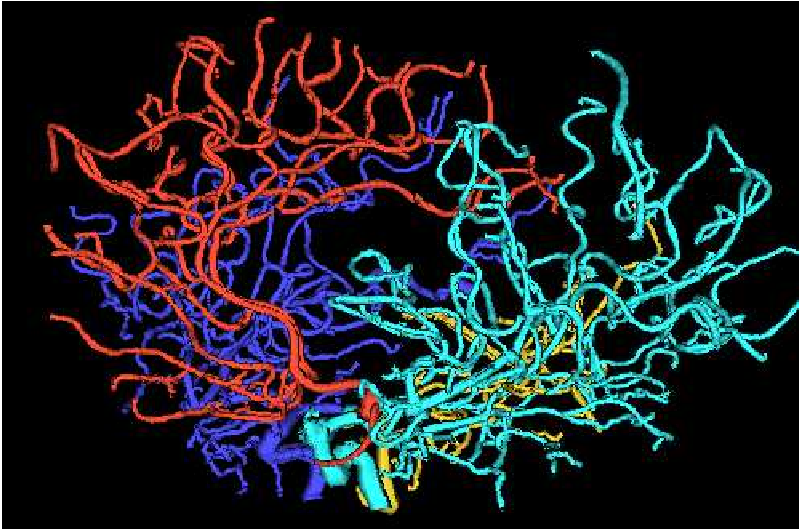
[Bullitt and Aylward, 2002]

Tube tracking



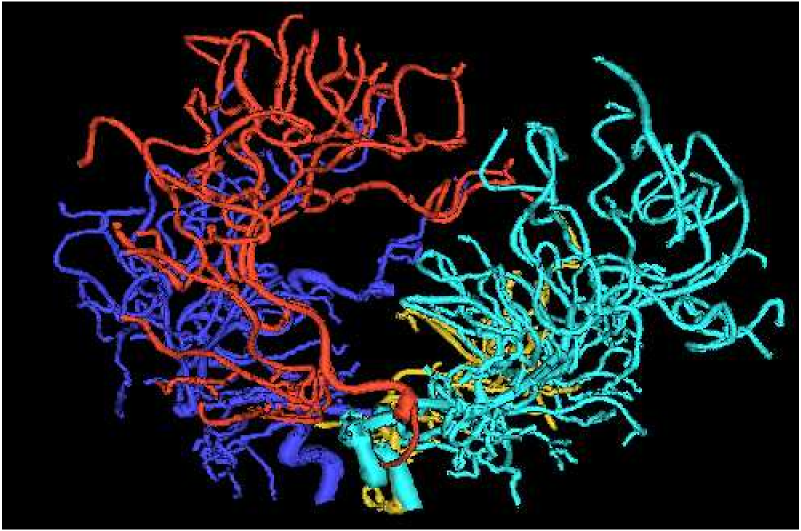
[Bullitt and Aylward, 2002]

Tube tracking



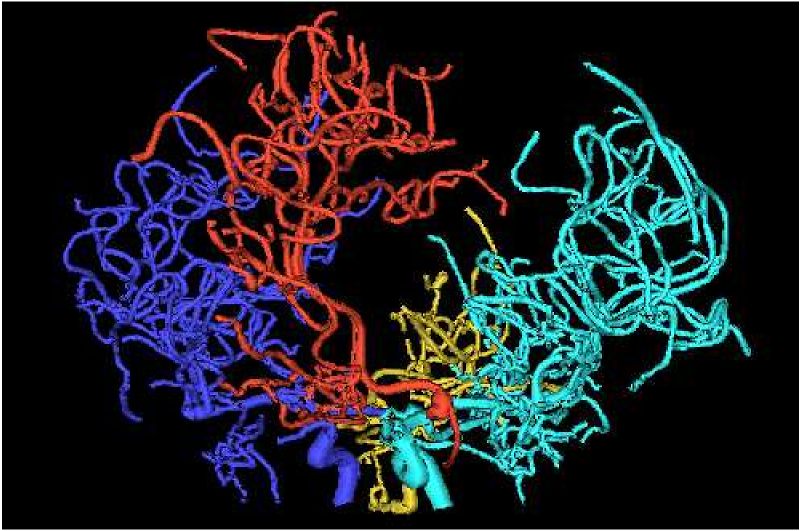
[Bullitt and Aylward, 2002]

Tube tracking



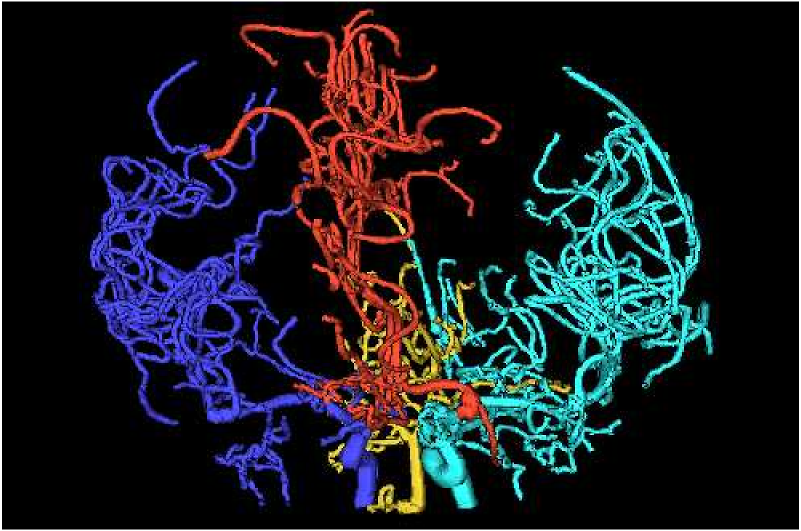
[Bullitt and Aylward, 2002]

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

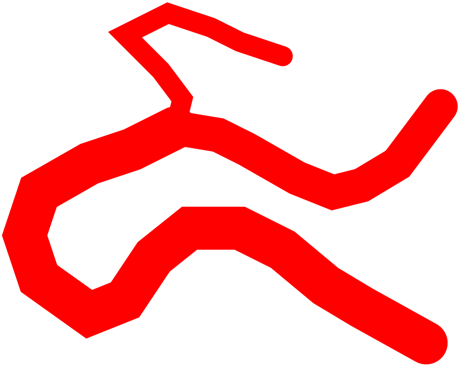
The images:

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:

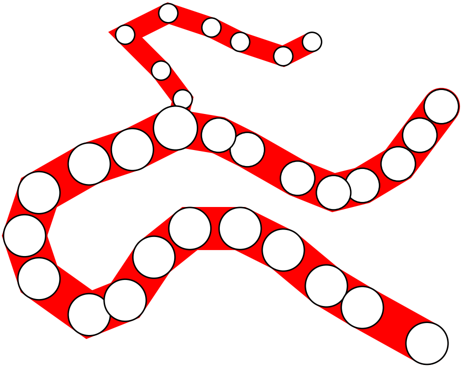


Brain artery trees

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:

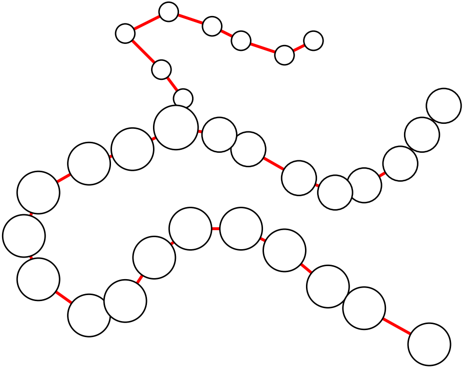


Brain artery trees

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

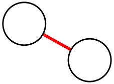
The data structure:

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:

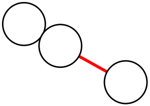


Brain artery trees

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:

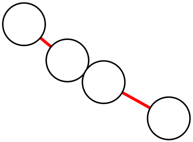


Brain artery trees

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:

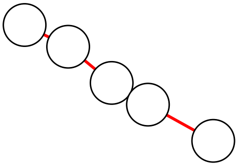


Brain artery trees

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:

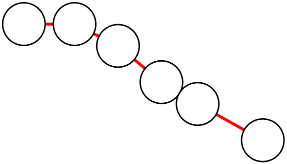


Brain artery trees

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:

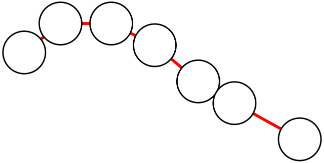


Brain artery trees

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:

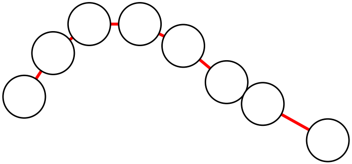


Brain artery trees

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:

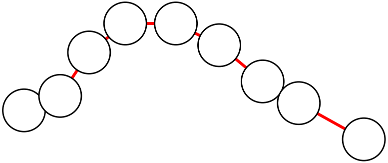


Brain artery trees

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:

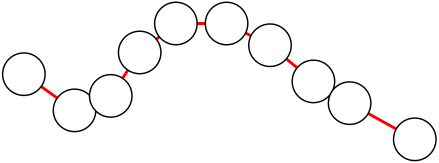


Brain artery trees

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:

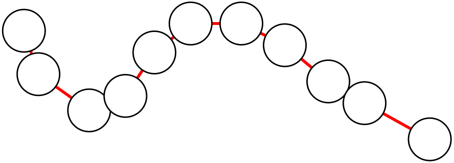


Brain artery trees

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:

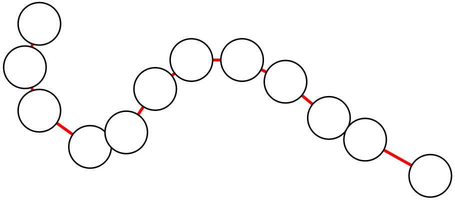


Brain artery trees

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:

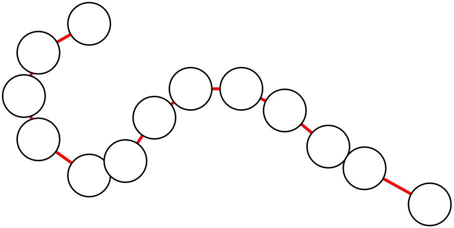


Brain artery trees

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:

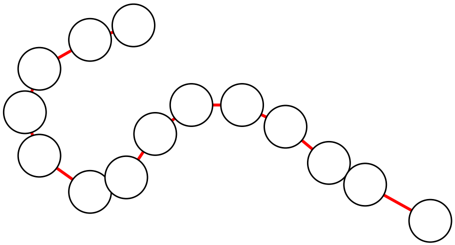


Brain artery trees

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:

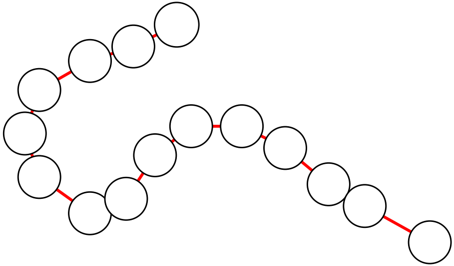


Brain artery trees

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:

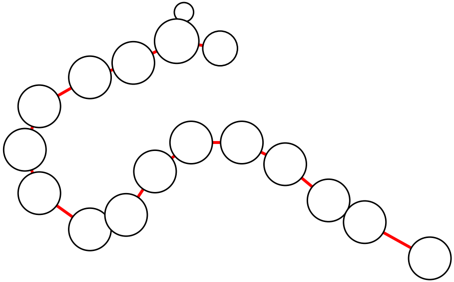


Brain artery trees

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:

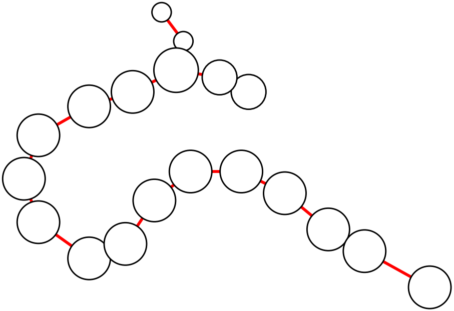


Brain artery trees

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:

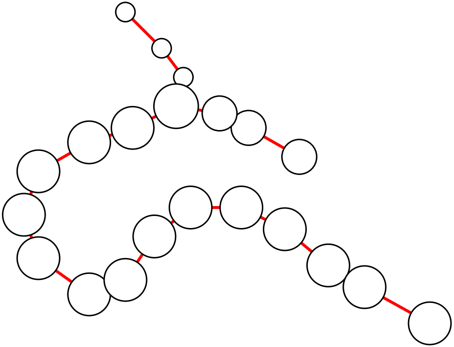


Brain artery trees

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:

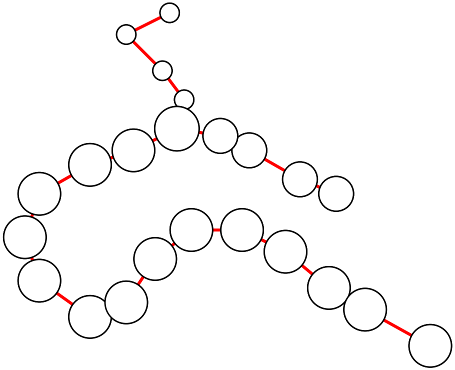


Brain artery trees

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:

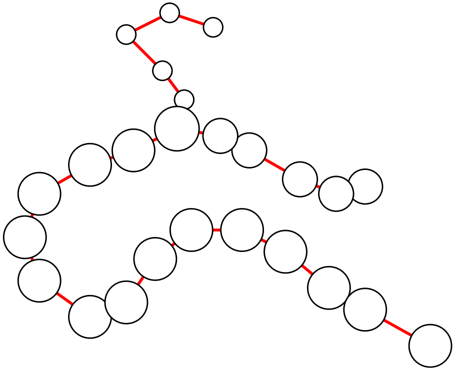


Brain artery trees

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:

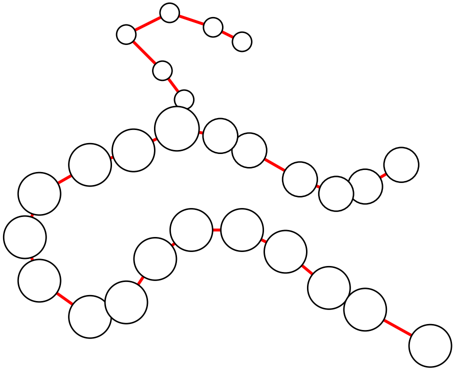


Brain artery trees

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:

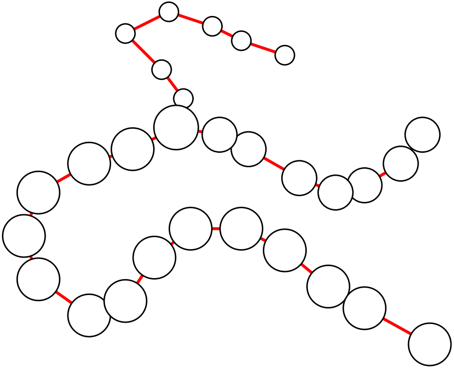


Brain artery trees

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:

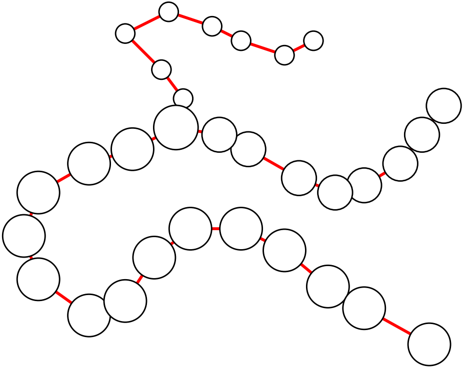


Brain artery trees

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:

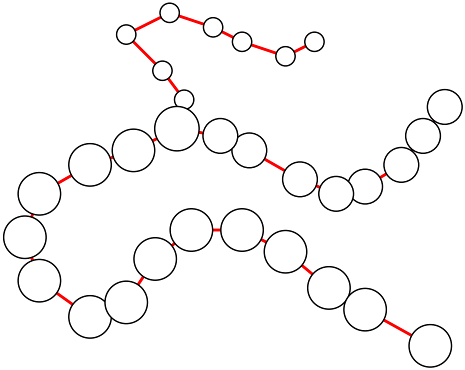


Brain artery trees

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:

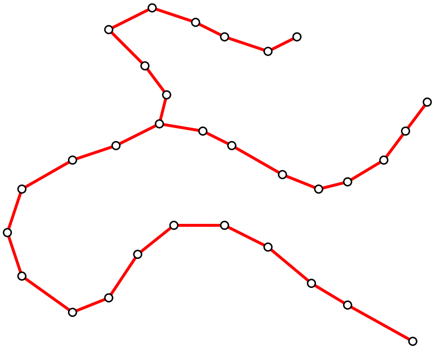


Brain artery trees

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:

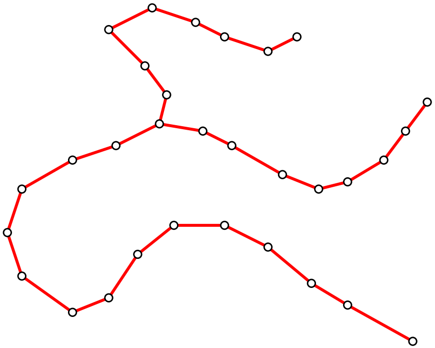


Brain artery trees

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:



Prior analyses

Discrete methods [Aydin, et al. 2009]

- disregard metric and embedding
- compare combinatorial structures
- no correlations detected

Phylogenetic trees [SAMSI WG 2013]

- connect cortical surface landmarks to nearest leaves
- apply averaging algorithm [M.—, Owen, Provan; Bačák 2012] in tree space [Billera, Holmes, Vogtmann 2001]
- too combinatorial again: found nothing but sticky mean at origin

Dyck paths [Dan Shen and J.S. Marron, et al. 2014]

- pay attention to edge lengths but disregard 3D embedding
- complicated tree pruning
- Pearson correlation $\sim .25$

Premise.

- combinatorics and branch length not enough;
- location and twist are crucial.

Prior analyses

Discrete methods [Aydin, et al. 2009]

- disregard metric and embedding
- compare combinatorial structures
- no correlations detected

Phylogenetic trees [SAMSI WG 2013]

- connect cortical surface landmarks to nearest leaves
- apply averaging algorithm [M.—, Owen, Provan; Bačák 2012] in tree space [Billera, Holmes, Vogtmann 2001]
- too combinatorial again: found nothing but sticky mean at origin

Dyck paths [Dan Shen and J.S. Marron, et al. 2014]

- pay attention to edge lengths but disregard 3D embedding
- complicated tree pruning
- Pearson correlation $\sim .25$

Premise.

- combinatorics and branch length not enough;
- location and twist are crucial.

Prior analyses

Discrete methods [Aydin, et al. 2009]

- disregard metric and embedding
- compare combinatorial structures
- no correlations detected

Phylogenetic trees [SAMSI WG 2013]

- connect cortical surface landmarks to nearest leaves
- apply averaging algorithm [M.—, Owen, Provan; Bačák 2012] in tree space [Billera, Holmes, Vogtmann 2001]
- too combinatorial again: found nothing but sticky mean at origin

Dyck paths [Dan Shen and J.S. Marron, et al. 2014]

- pay attention to edge lengths but disregard 3D embedding
- complicated tree pruning
- Pearson correlation $\sim .25$

Premise.

- combinatorics and branch length not enough;
- location and twist are crucial.

Prior analyses

Discrete methods [Aydin, et al. 2009]

- disregard metric and embedding
- compare combinatorial structures
- no correlations detected

Phylogenetic trees [SAMSI WG 2013]

- connect cortical surface landmarks to nearest leaves
- apply averaging algorithm [M.—, Owen, Provan; Bačák 2012] in tree space [Billera, Holmes, Vogtmann 2001]
- too combinatorial again: found nothing but sticky mean at origin

Dyck paths [Dan Shen and J.S. Marron, et al. 2014]

- pay attention to edge lengths but disregard 3D embedding
- complicated tree pruning
- Pearson correlation $\sim .25$

Premise.

- combinatorics and branch length not enough;
- location and twist are crucial.

Homology

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

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

Homology

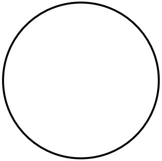
Topological space $X \rightsquigarrow$ homology $H_i X$ for each dimension i.

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

Homology

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

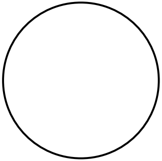
- vector space that measures "i-dimensional holes" in X



Homology

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

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

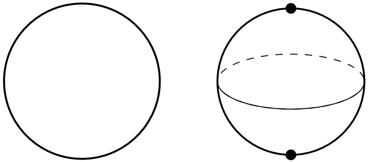


$\dim(H_1) = 1$

Homology

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

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

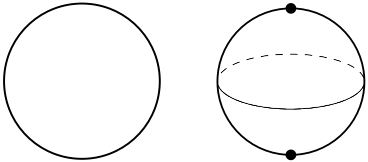


$\dim(H_1) = 1$

Homology

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

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

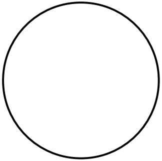


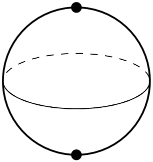
$\dim(H_1) = 1$ $\dim(H_1) = 0$

Homology

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

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





$\dim(H_1) = 1$

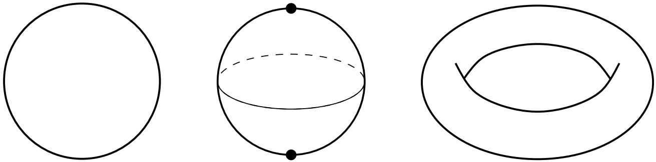
$\dim(H_1) = 0$

$\dim(H_2) = 1$

Homology

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

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



$\dim(H_1) = 1$

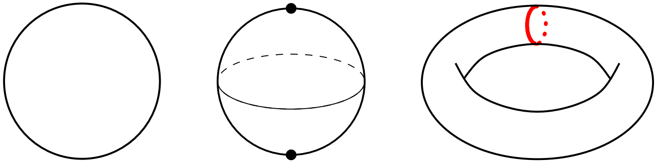
$\dim(H_1) = 0$

$\dim(H_2) = 1$

Homology

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

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



$\dim(H_1) = 1$

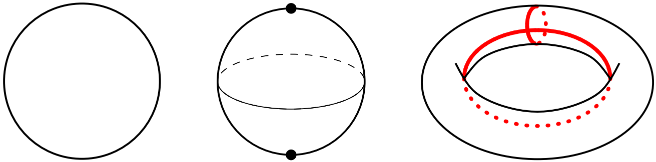
$\dim(H_1) = 0$

$\dim(H_2) = 1$

Homology

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

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



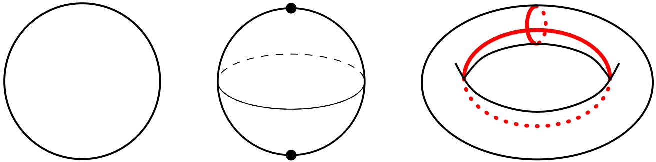
$\dim(H_1) = 1$ $\dim(H_1) = 0$

$\dim(H_2) = 1$

Homology

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

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



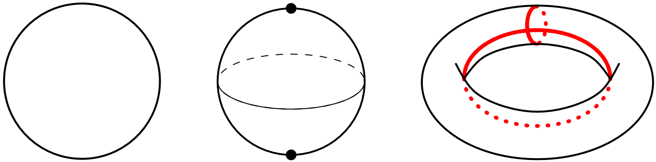
$\dim(H_1) = 1$ $\dim(H_1) = 0$ $\dim(H_1) = 2$

$\dim(H_2) = 1$

Homology

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

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



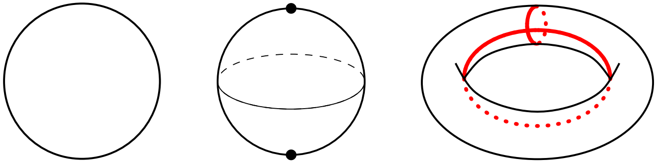
$\dim(H_1) = 1$ $\dim(H_1) = 0$ $\dim(H_1) = 2$

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

Homology

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

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



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

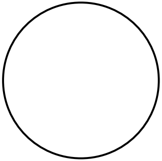
$\dim(H_2) = 1$ $\quad$ $\dim(H_2) = 1$

- $i = 0$ case: H_i counts connected components of X

Homology

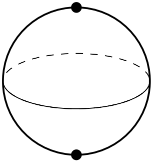
Topological space $X \rightsquigarrow$ homology $H_i X$ for each dimension i.

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



$\dim(H_0) = 1$

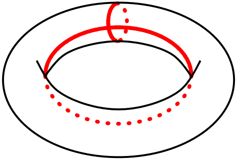
$\dim(H_1) = 1$



$\dim(H_0) = 1$

$\dim(H_1) = 0$

$\dim(H_2) = 1$



$\dim(H_0) = 1$

$\dim(H_1) = 2$

$\dim(H_2) = 1$

- $i = 0$ case: H_i counts connected components of X

Persistent homology

Build X step by step

- measure evolving topology.

Def. Let $X_\bullet$ be a filtered space, meaning $\varnothing = X_0 \subset X_1 \subset \cdots \subset X_m = X$. The persistent homology $H_i X_\bullet$ is $H_i X_1 \to H_i X_2 \to \cdots \to H_i X_m$, a sequence of vector space homomorphisms.

Examples:

1. Given a function $f : X \to \mathbb{R}$, let $X_k = f^{-1}\big((-\infty, t_k]\big)$. 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; [many others, including Carlsson]: further developments

Persistent homology

Build X step by step

- measure evolving topology.

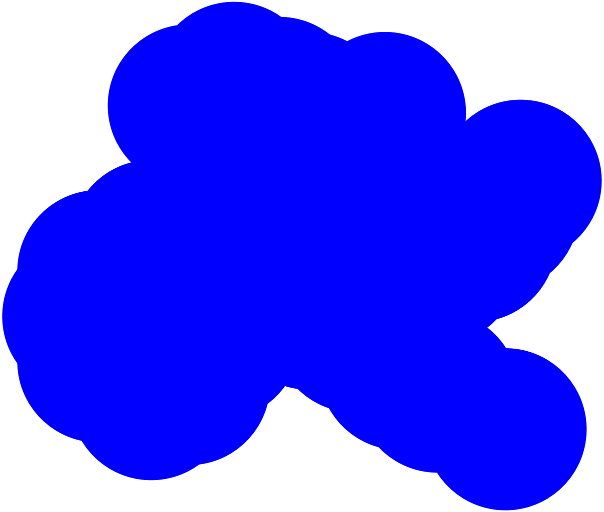
Def. Let $X_\bullet$ be a filtered space, meaning $\varnothing = X_0 \subset X_1 \subset \cdots \subset X_m = X$. The persistent homology $H_i X_\bullet$ is $H_i X_1 \to H_i X_2 \to \cdots \to H_i X_m$, a sequence of vector space homomorphisms.

Examples:

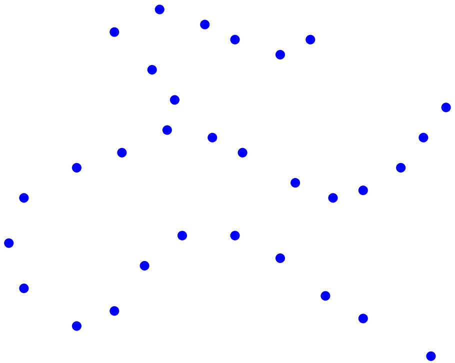
1. Given a function $f : X \to \mathbb{R}$, let $X_k = f^{-1}\big((-\infty, t_k]\big)$. Good choice of $t_0, \ldots, t_m \in \mathbb{R}$: the values of t across which $H_i X_t$ changes.
2. Any simplicial complex: build it simplex by simplex in some order.

History. invented by [Frosini, Landi 1999], [Robins 1999], [Edelsbrunner, Letscher, Zomorodian 2002]: includes efficient computation; [many others, including Carlsson]: further developments

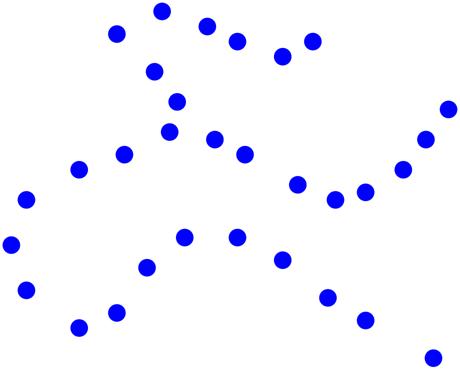
Example: expanding balls



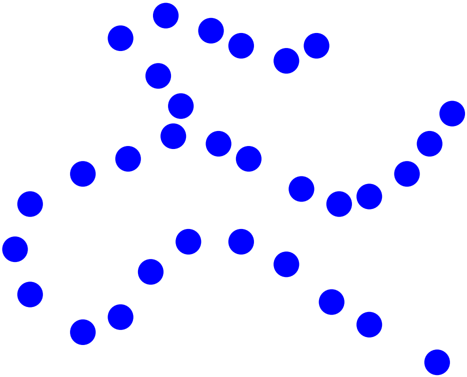
Example: expanding balls



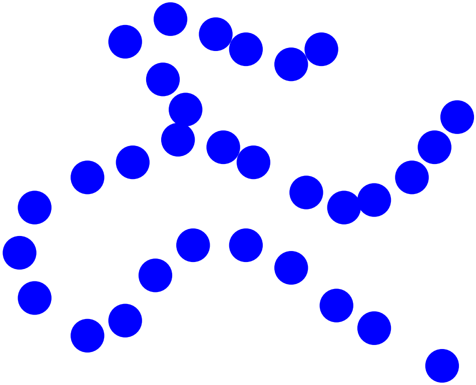
Example: expanding balls



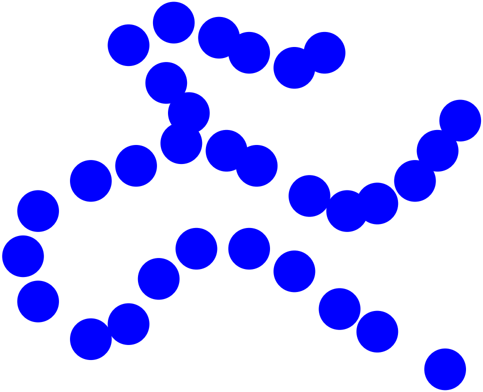
Example: expanding balls



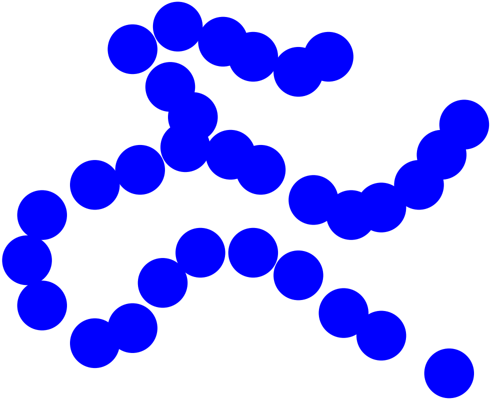
Example: expanding balls



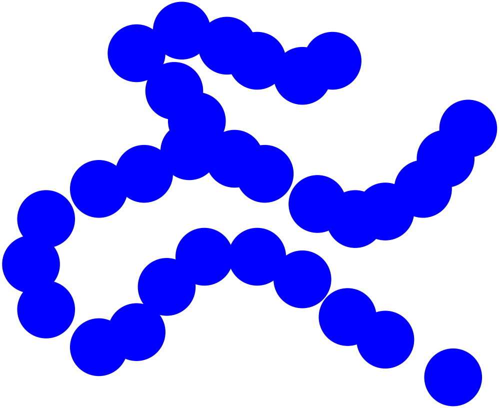
Example: expanding balls



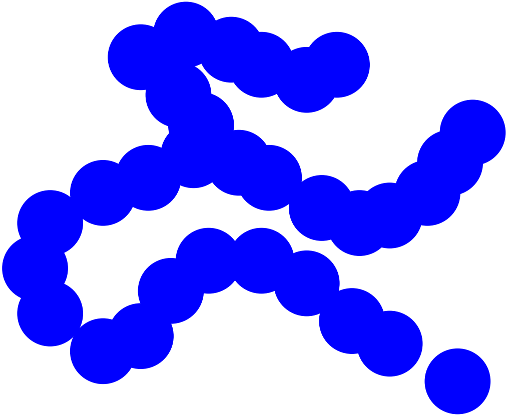
Example: expanding balls



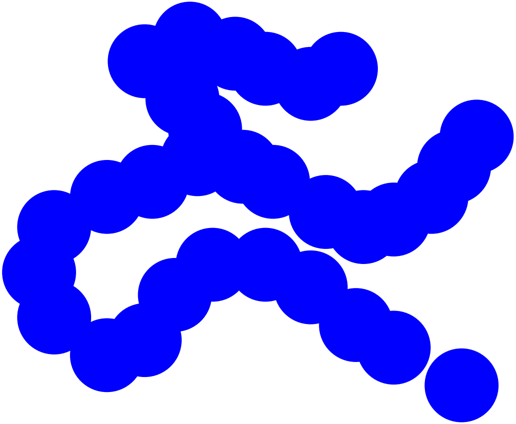
Example: expanding balls



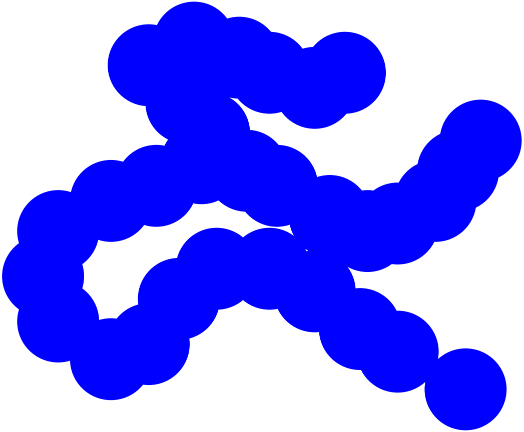
Example: expanding balls



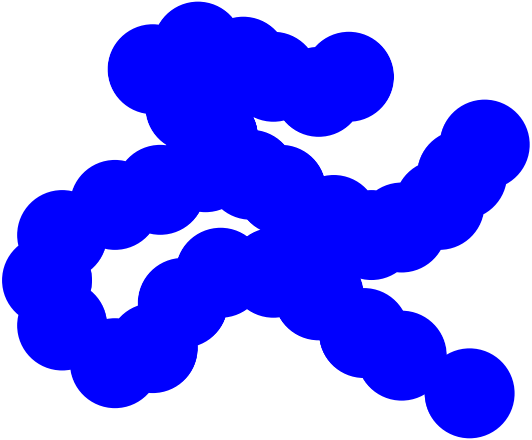
Example: expanding balls



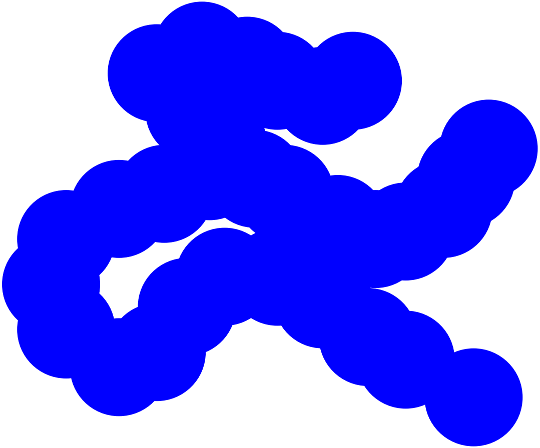
Example: expanding balls



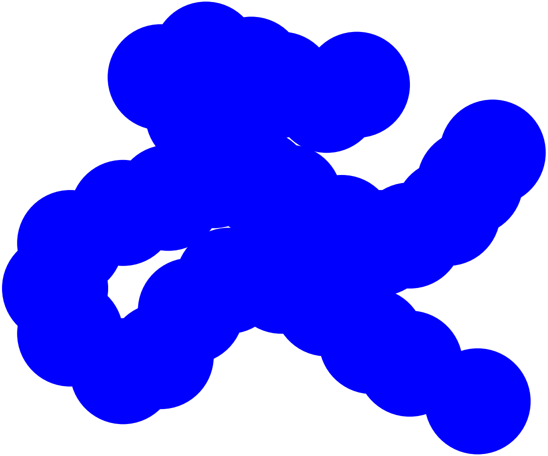
Example: expanding balls



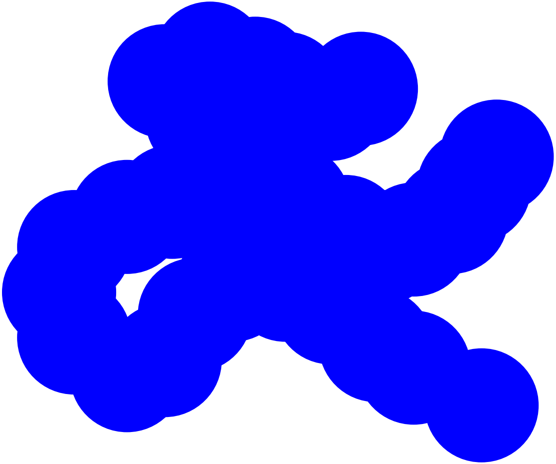
Example: expanding balls



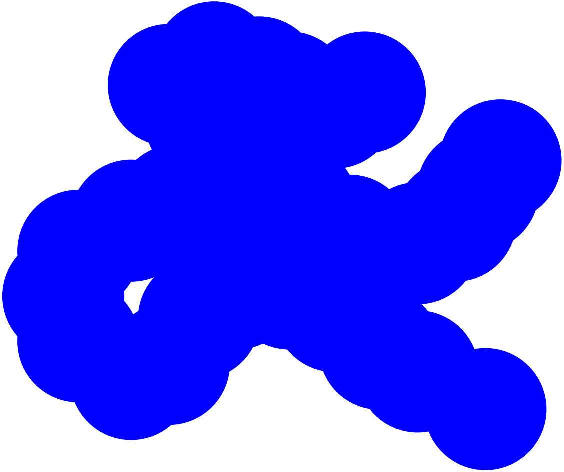
Example: expanding balls



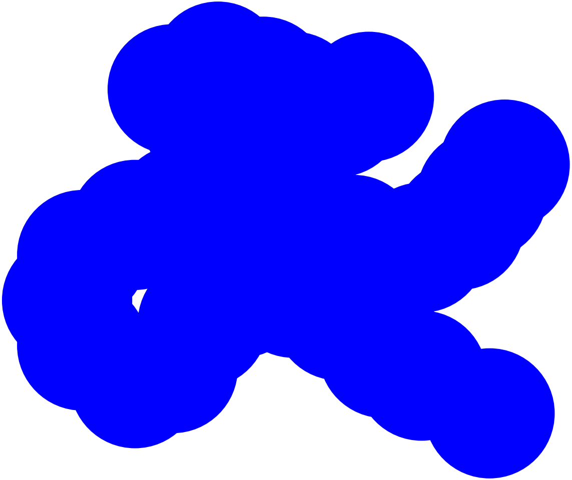
Example: expanding balls



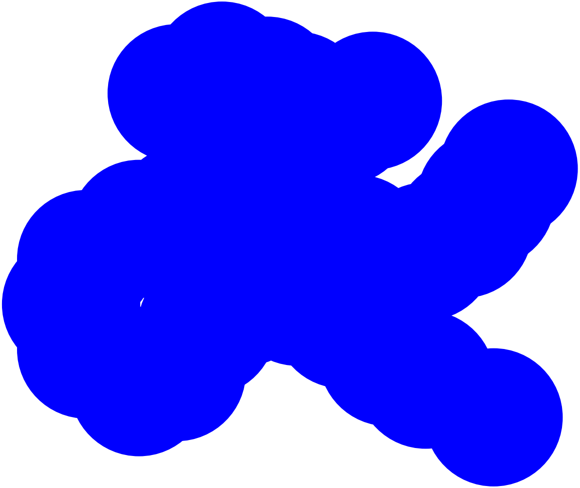
Example: expanding balls



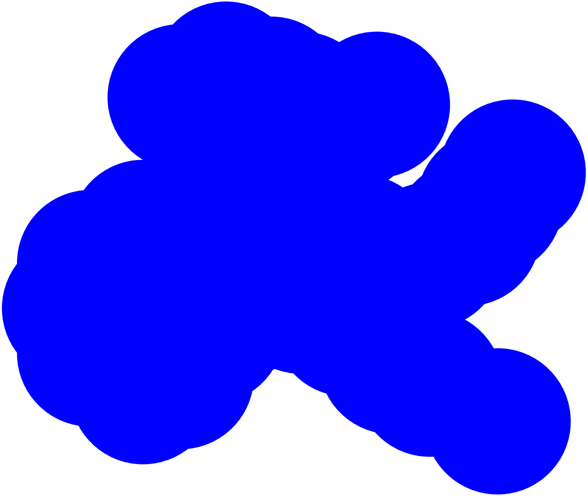
Example: expanding balls



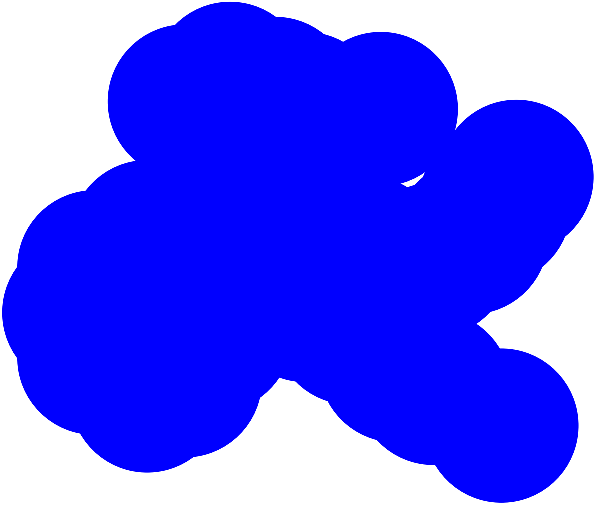
Example: expanding balls



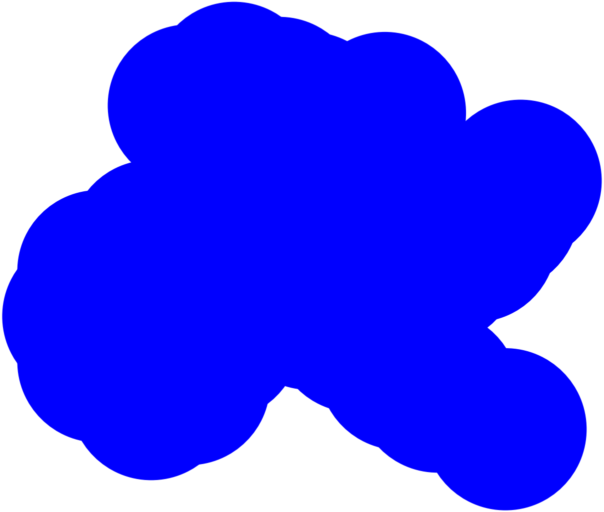
Example: expanding balls



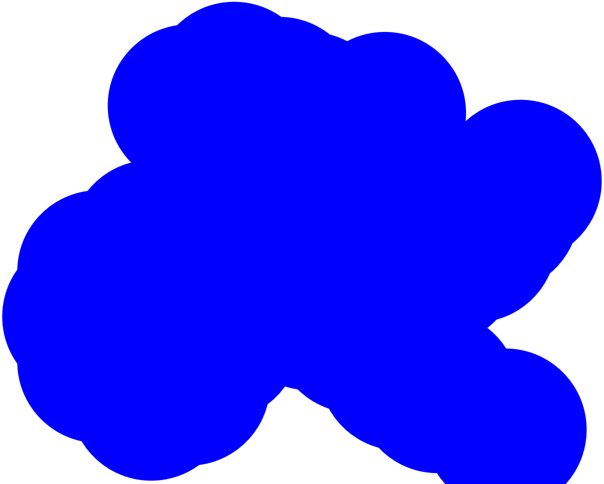
Example: expanding balls



Example: expanding balls



Example: expanding balls



Persistent homology

Build X step by step

- measure evolving topology.

Def. Let $X_\bullet$ be a filtered space, meaning $\varnothing = X_0 \subset X_1 \subset \cdots \subset X_m = X$. The persistent homology $H_i X_\bullet$ is $H_i X_1 \to H_i X_2 \to \cdots \to H_i X_m$, a sequence of vector space homomorphisms.

Examples:

1. Given a function $f : X \to \mathbb{R}$, let $X_k = f^{-1}\big((-\infty, t_k]\big)$. 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; [many others, including Carlsson]: further developments

Persistent homology

Build X step by step

- measure evolving topology.

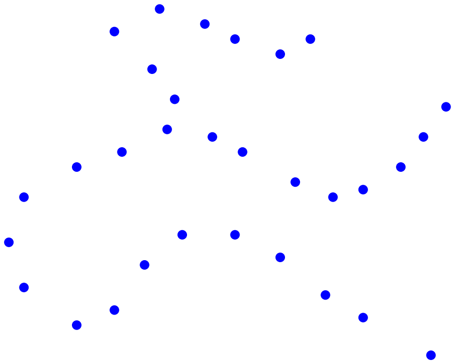
Def. Let $X_\bullet$ be a filtered space, meaning $\varnothing = X_0 \subset X_1 \subset \cdots \subset X_m = X$. The persistent homology $H_i X_\bullet$ is $H_i X_1 \to H_i X_2 \to \cdots \to H_i X_m$, a sequence of vector space homomorphisms.

Examples:

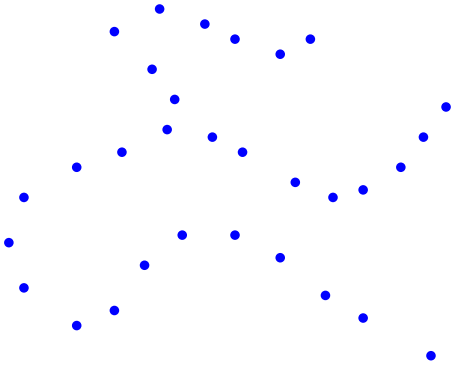
1. Given a function $f : X \to \mathbb{R}$, let $X_k = f^{-1}\big((-\infty, t_k]\big)$. Good choice of $t_0, \ldots, t_m \in \mathbb{R}$: the values of t across which $H_i X_t$ changes.
2. Any simplicial complex: build it simplex by simplex in some order.

History. invented by [Frosini, Landi 1999], [Robins 1999], [Edelsbrunner, Letscher, Zomorodian 2002]: includes efficient computation; [many others, including Carlsson]: further developments

Example: expanding balls

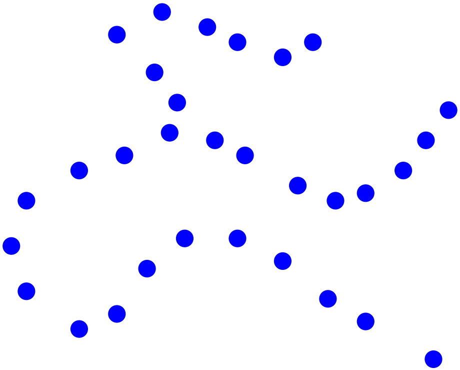


Example: expanding balls



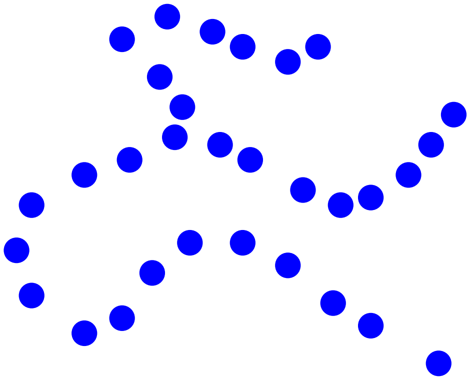
$\dim(H_0) = 31$

Example: expanding balls



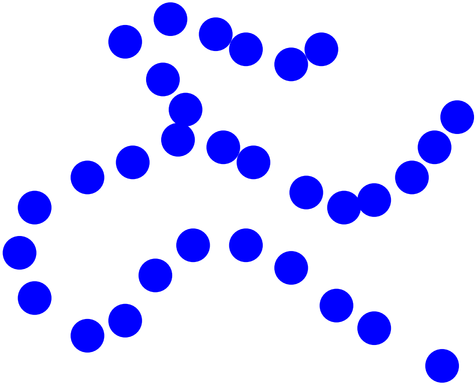
$\dim(H_0) = 31$

Example: expanding balls



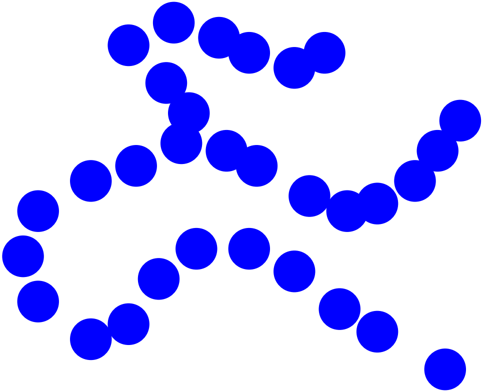
$\dim(H_0) = 31$

Example: expanding balls



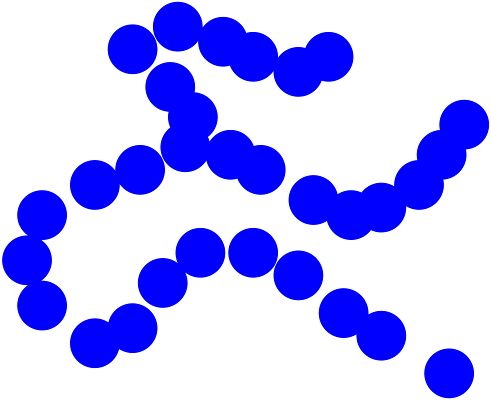
$\dim(H_0) = 26$

Example: expanding balls



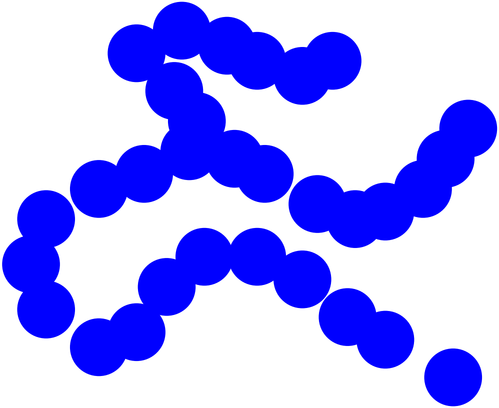
$\dim(H_0) = 21$

Example: expanding balls



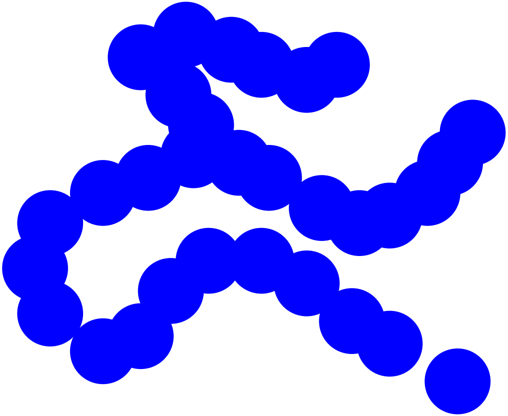
$\dim(H_0) = 12$

Example: expanding balls



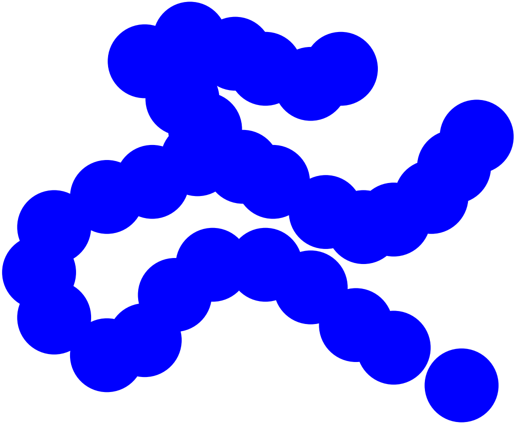
$\dim(H_0) = 6$

Example: expanding balls



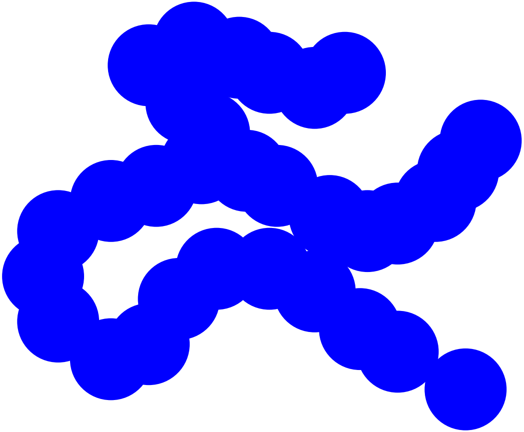
$\dim(H_0) = 2$

Example: expanding balls



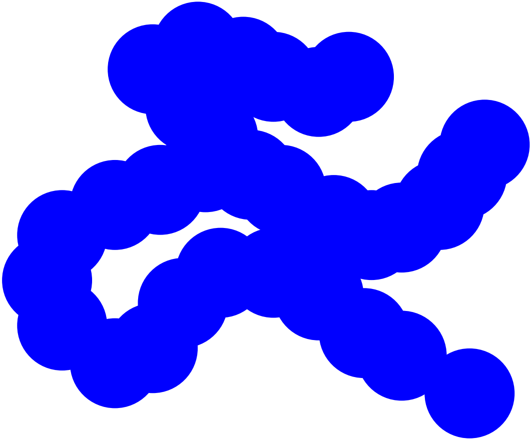
$\dim(H_0) = 2$

Example: expanding balls



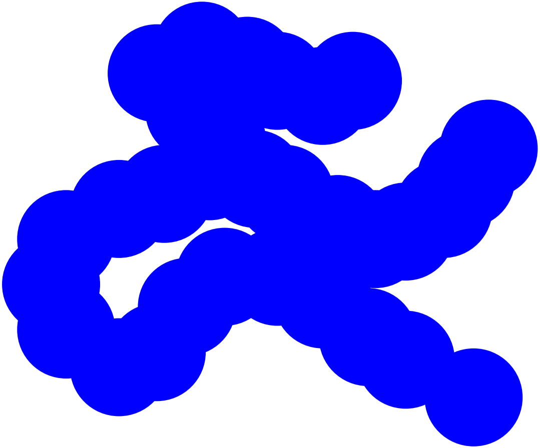
$\dim(H_0) = 1$ $\dim(H_1) = 2$

Example: expanding balls



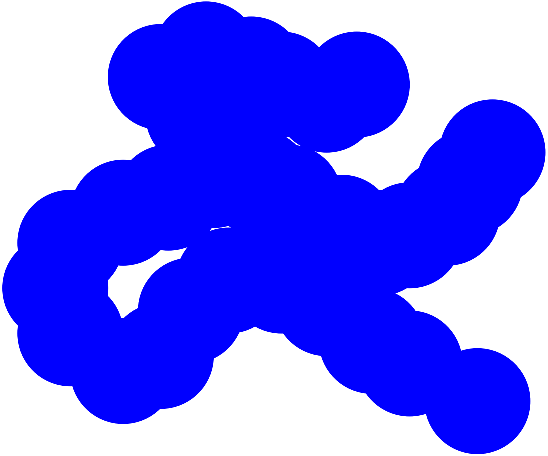
$\dim(H_0) = 1 \qquad \dim(H_1) = 1$

Example: expanding balls



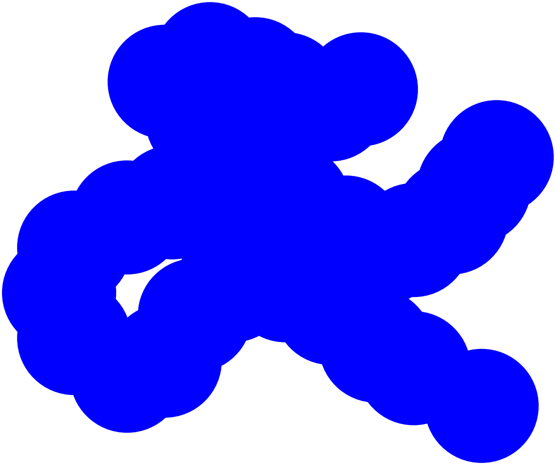
$\dim(H_0) = 1 \qquad \dim(H_1) = 1$

Example: expanding balls



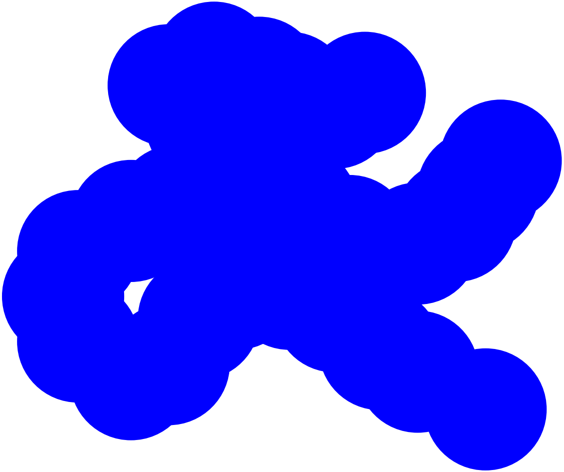
$\dim(H_0) = 1 \qquad \dim(H_1) = 3$

Example: expanding balls



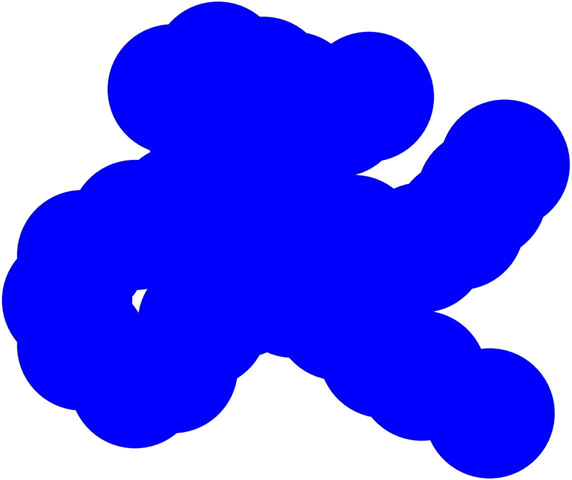
$\dim(H_0) = 1 \qquad \dim(H_1) = 1$

Example: expanding balls



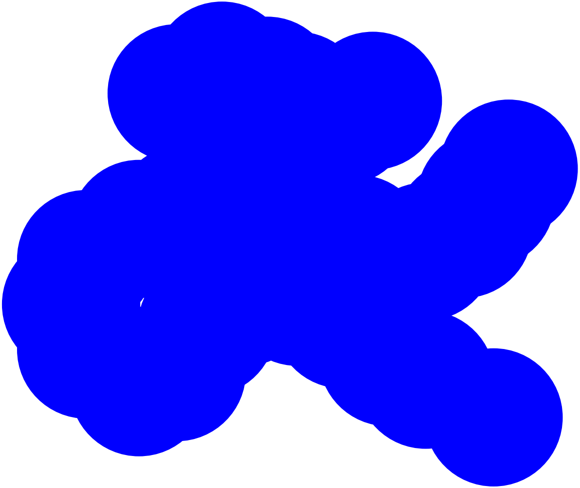
$\dim(H_0) = 1 \qquad \dim(H_1) = 1$

Example: expanding balls



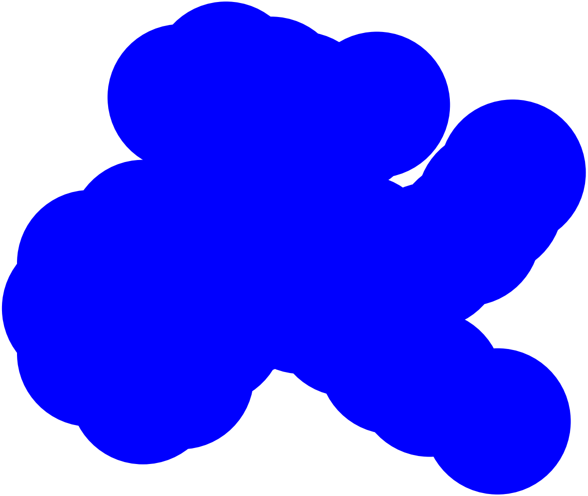
$\dim(H_0) = 1 \qquad \dim(H_1) = 1$

Example: expanding balls



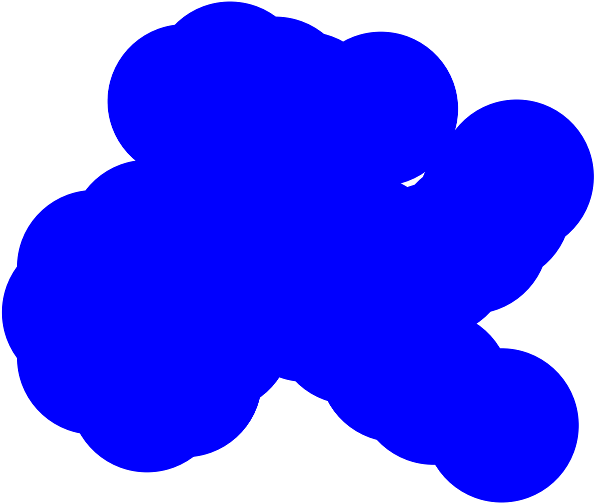
$\dim(H_0) = 1 \qquad \dim(H_1) = 1$

Example: expanding balls



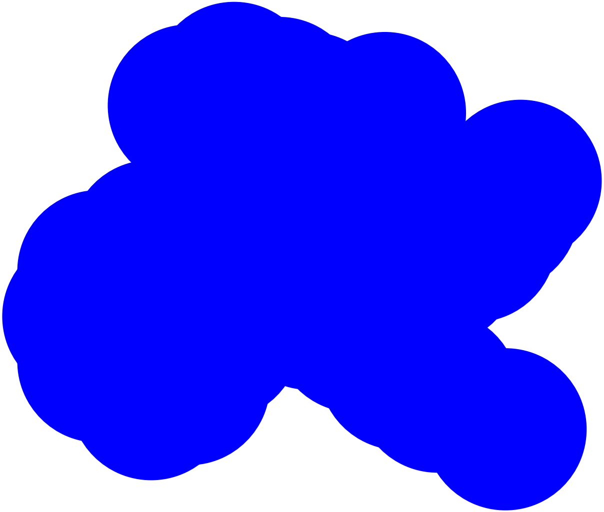
$\dim(H_0) = 1$ $\dim(H_1) = 0$

Example: expanding balls



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

Example: expanding balls



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

Persistent homology

Build X step by step

- measure evolving topology.

Def. Let $X_\bullet$ be a filtered space, meaning $\varnothing = X_0 \subset X_1 \subset \cdots \subset X_m = X$. The persistent homology $H_i X_\bullet$ is $H_i X_1 \to H_i X_2 \to \cdots \to H_i X_m$, a sequence of vector space homomorphisms.

Examples:

1. Given a function $f : X \to \mathbb{R}$, let $X_k = f^{-1}\big((-\infty, t_k]\big)$. 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; [many others, including Carlsson]: further developments

Persistent homology

Build X step by step

- measure evolving topology.

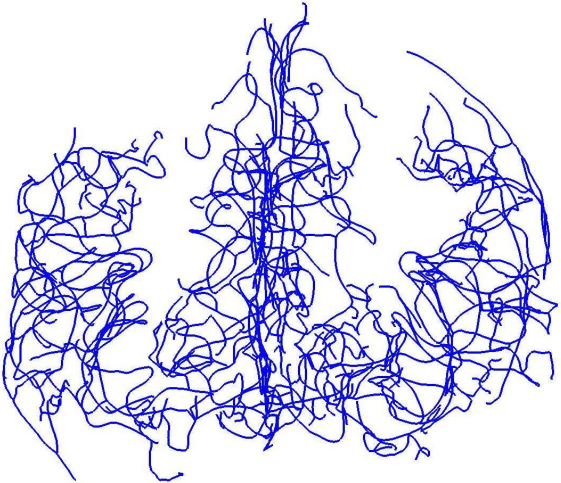
Def. Let $X_\bullet$ be a filtered space, meaning $\varnothing = X_0 \subset X_1 \subset \cdots \subset X_m = X$. The persistent homology $H_i X_\bullet$ is $H_i X_1 \to H_i X_2 \to \cdots \to H_i X_m$, a sequence of vector space homomorphisms.

Examples:

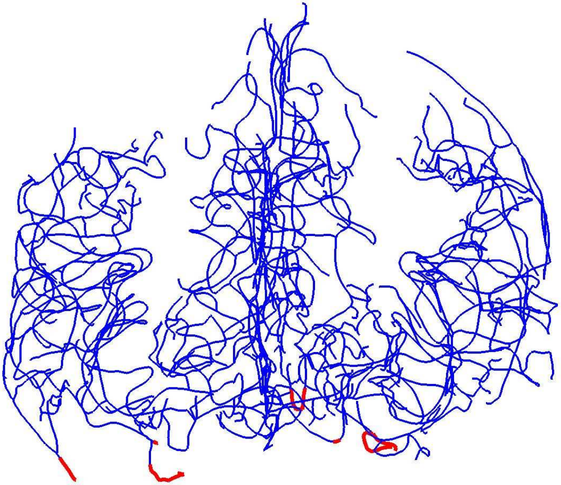
1. Given a function $f : X \to \mathbb{R}$, let $X_k = f^{-1}\big((-\infty, t_k]\big)$. Good choice of $t_0, \ldots, t_m \in \mathbb{R}$: the values of t across which $H_i X_t$ changes.
2. Any simplicial complex: build it simplex by simplex in some order.

History. invented by [Frosini, Landi 1999], [Robins 1999], [Edelsbrunner, Letscher, Zomorodian 2002]: includes efficient computation; [many others, including Carlsson]: further developments

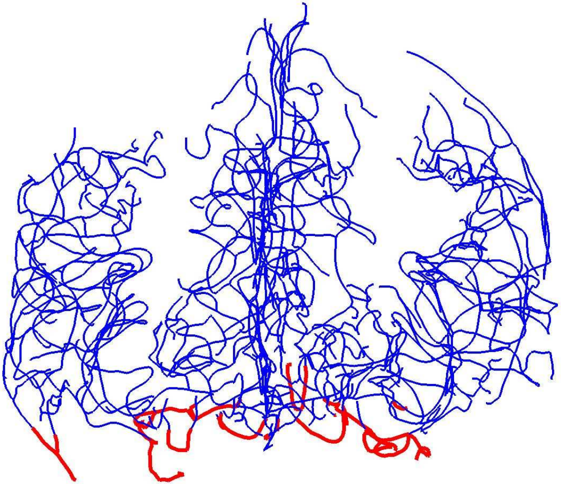
Example: filling brains



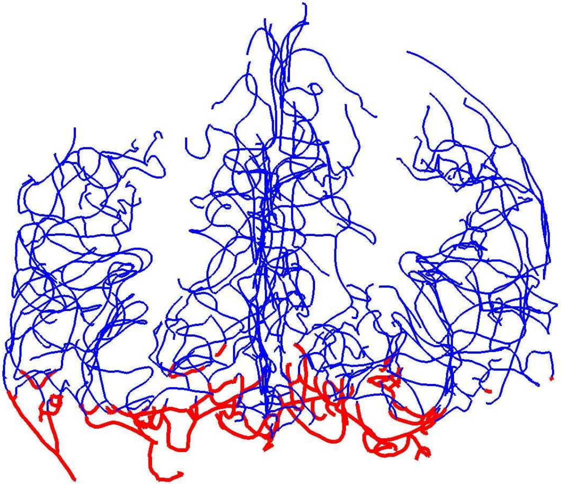
Example: filling brains



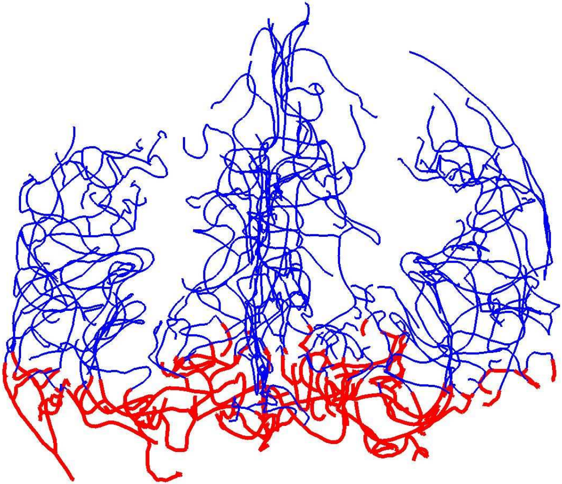
Example: filling brains



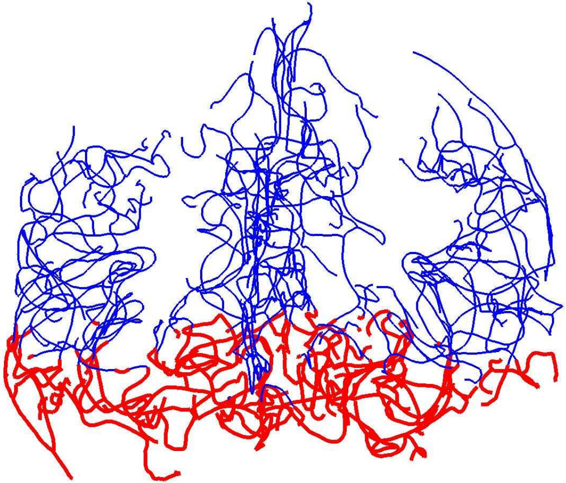
Example: filling brains



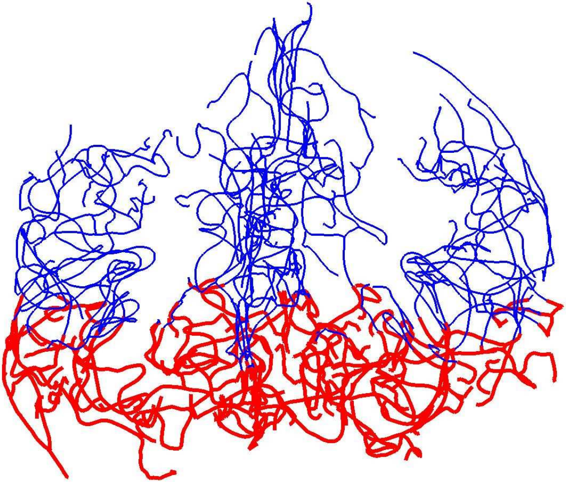
Example: filling brains



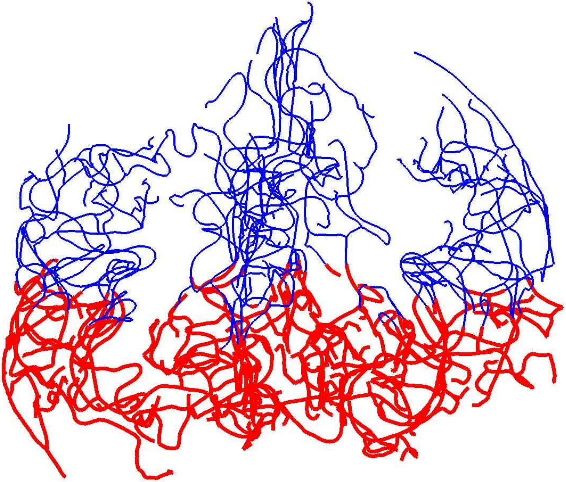
Example: filling brains



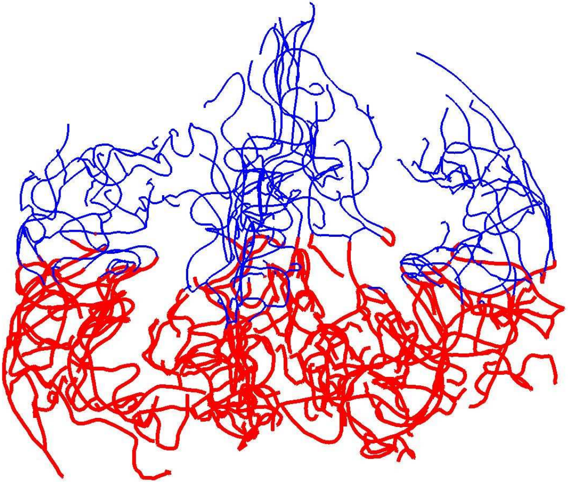
Example: filling brains



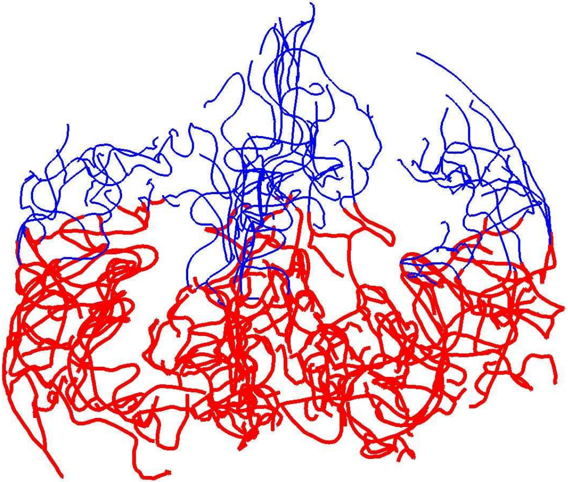
Example: filling brains



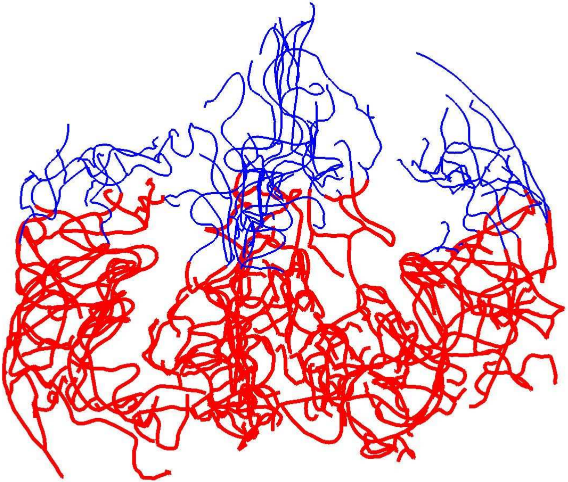
Example: filling brains



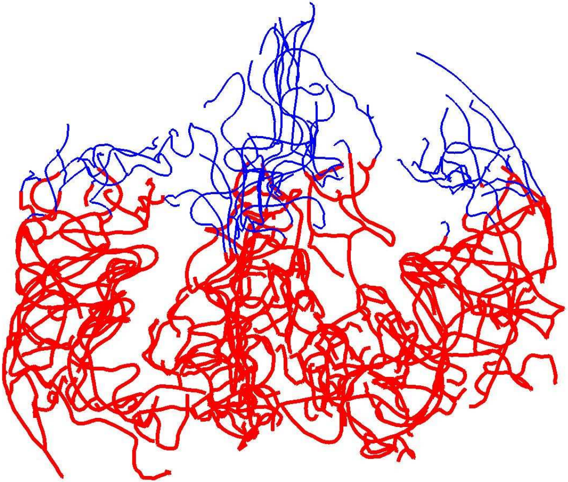
Example: filling brains



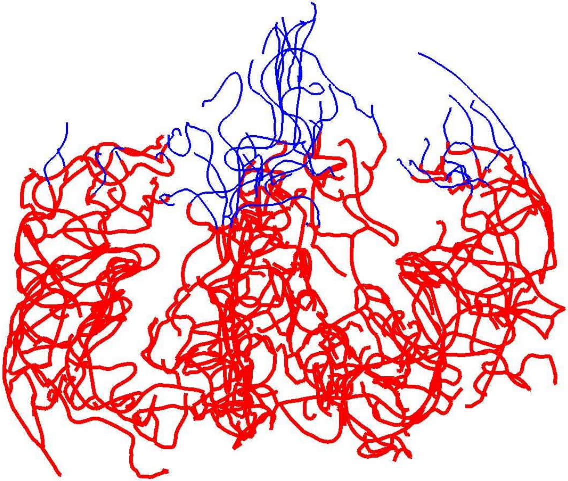
Example: filling brains



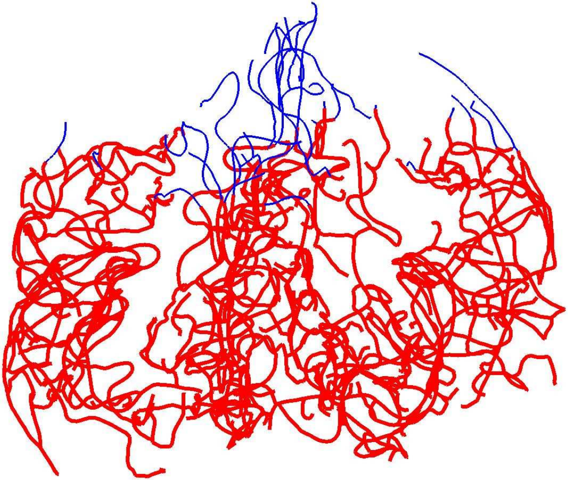
Example: filling brains



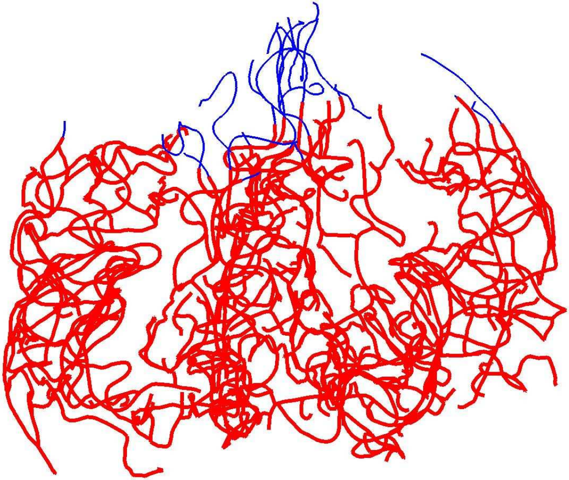
Example: filling brains



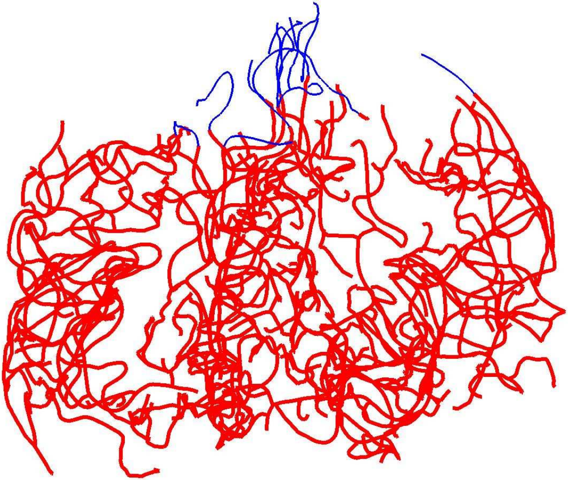
Example: filling brains



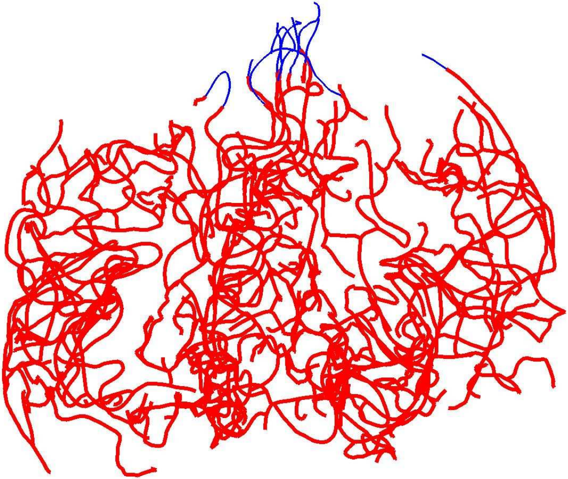
Example: filling brains



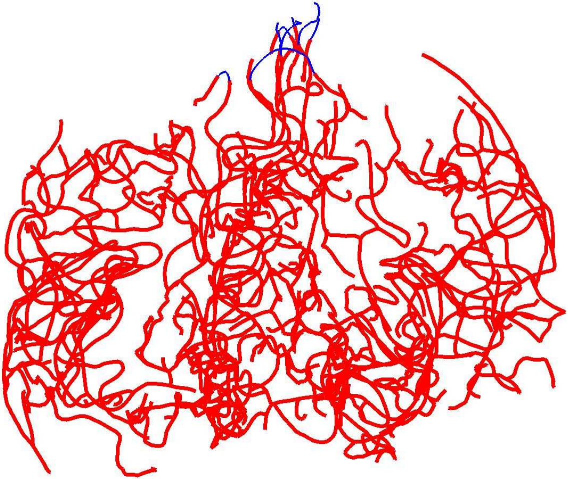
Example: filling brains



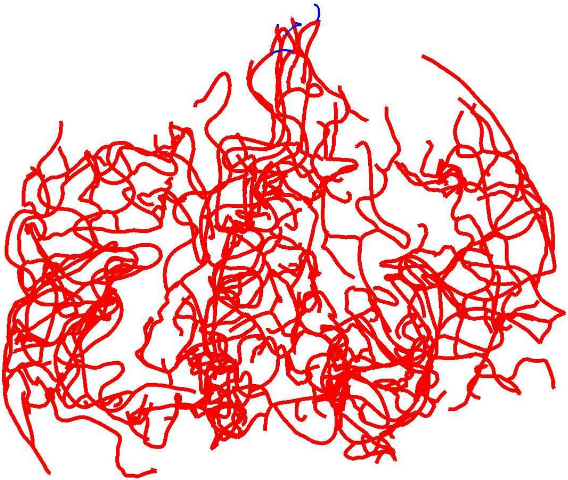
Example: filling brains



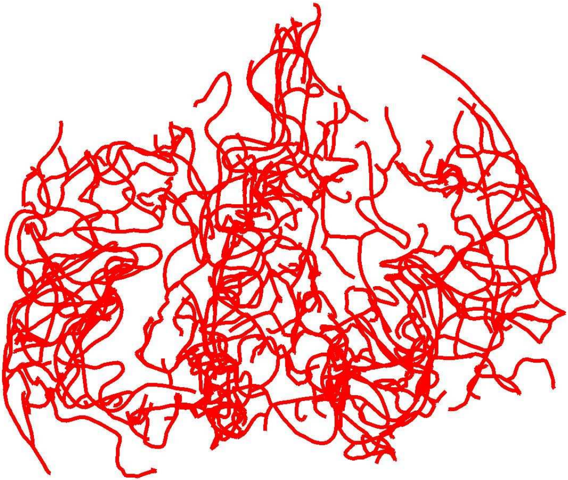
Example: filling brains



Example: filling brains



Example: filling brains



Persistent homology

Build X step by step

- measure evolving topology.

Def. Let $X_\bullet$ be a filtered space, meaning $\varnothing = X_0 \subset X_1 \subset \cdots \subset X_m = X$. The persistent homology $H_i X_\bullet$ is $H_i X_1 \to H_i X_2 \to \cdots \to H_i X_m$, a sequence of vector space homomorphisms.

Examples:

1. Given a function $f : X \to \mathbb{R}$, let $X_k = f^{-1}\big((-\infty, t_k]\big)$. 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; [many others, including Carlsson]: further developments

Persistent homology

Build X step by step

- measure evolving topology.

Def. Let $X_\bullet$ be a filtered space, meaning $\varnothing = X_0 \subset X_1 \subset \cdots \subset X_m = X$. The persistent homology $H_i X_\bullet$ is $H_i X_1 \to H_i X_2 \to \cdots \to H_i X_m$, a sequence of vector space homomorphisms.

Examples:

1. Given a function $f : X \to \mathbb{R}$, let $X_k = f^{-1}\big((-\infty, t_k]\big)$. 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; [many others, including Carlsson]: further developments

Persistent homology

Build X step by step

- measure evolving topology.

Def. Let $X_\bullet$ be a filtered space, meaning $\varnothing = X_0 \subset X_1 \subset \cdots \subset X_m = X$. The persistent homology $H_i X_\bullet$ is $H_i X_1 \to H_i X_2 \to \cdots \to H_i X_m$, a sequence of vector space homomorphisms.

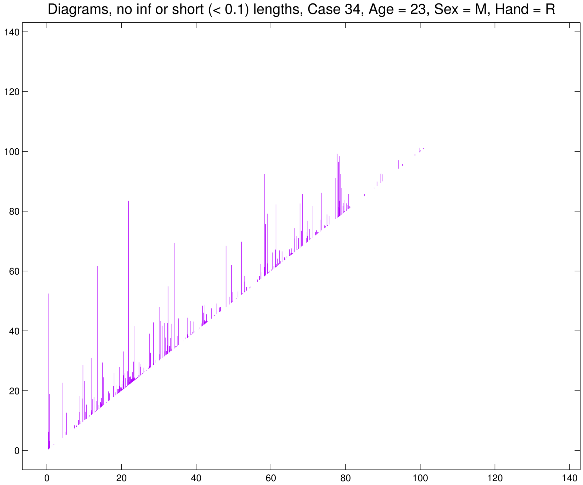
Examples:

1. Given a function $f : X \to \mathbb{R}$, let $X_k = f^{-1}\big((-\infty, t_k]\big)$. 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; [many others, including Carlsson]: further developments

Bar codes

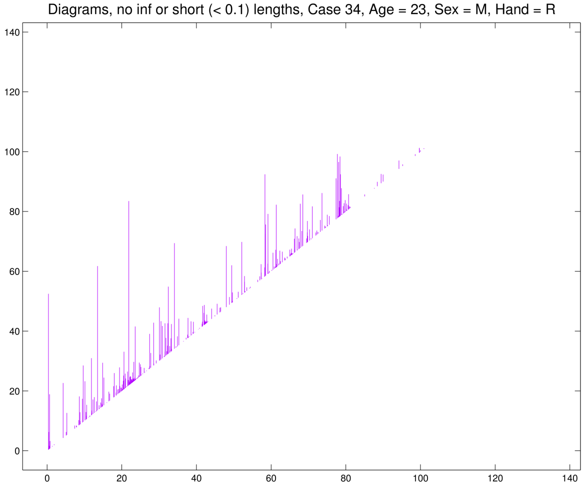
Data structure: 3D tree ⇝ bar code / lace array / persistence diagram:



- 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

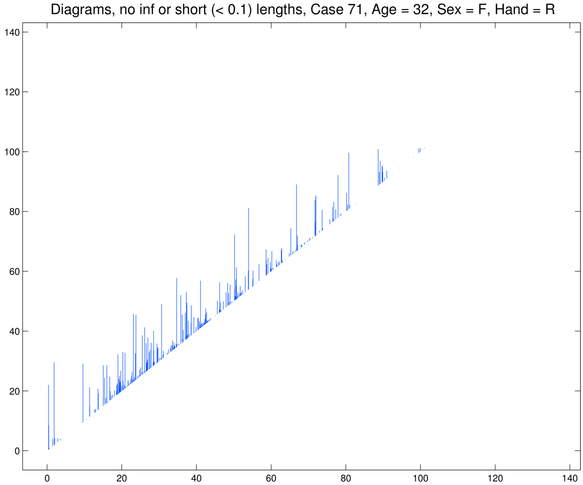
Data structure: 3D tree ⇝ bar code / lace array / persistence diagram:



- 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

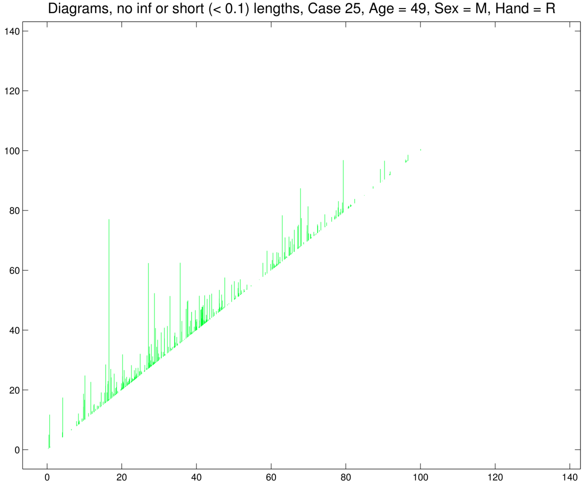
Data structure: 3D tree ⇝ bar code / lace array / persistence diagram:



- 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

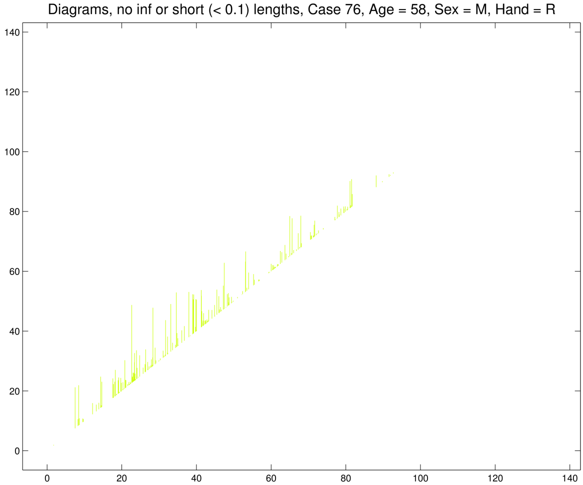
Data structure: 3D tree ⇝ bar code / lace array / persistence diagram:



- 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

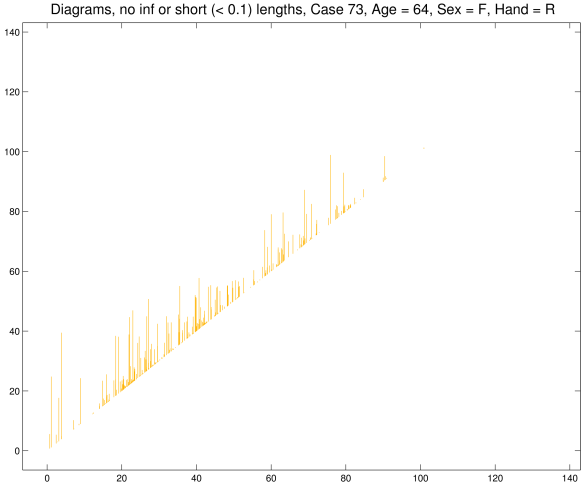
Data structure: 3D tree ⇝ bar code / lace array / persistence diagram:



- 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

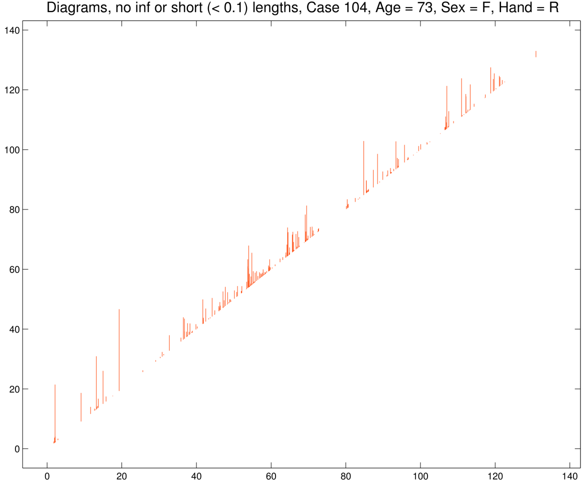
Data structure: 3D tree ⇝ bar code / lace array / persistence diagram:



- 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”

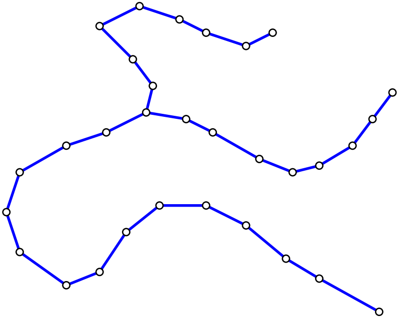
Filter by sweeping across with a plane:

Sweep filtration

Goal: statistical analysis taking into account

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

Filter by sweeping across with a plane:

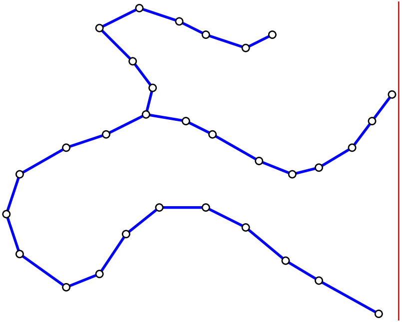


Sweep filtration

Goal: statistical analysis taking into account

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

Filter by sweeping across with a plane:

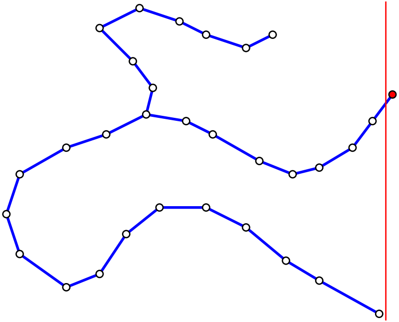


Sweep filtration

Goal: statistical analysis taking into account

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

Filter by sweeping across with a plane:

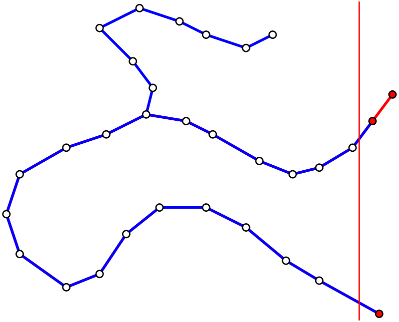


Sweep filtration

Goal: statistical analysis taking into account

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

Filter by sweeping across with a plane:

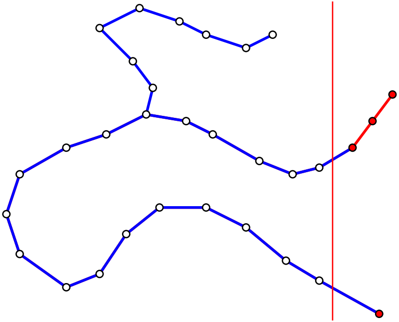


Sweep filtration

Goal: statistical analysis taking into account

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

Filter by sweeping across with a plane:

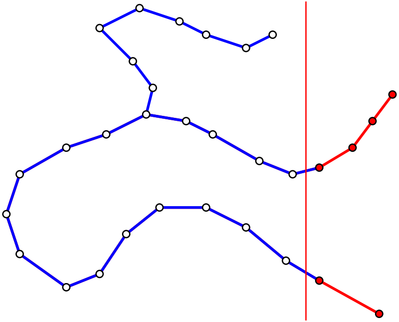


Sweep filtration

Goal: statistical analysis taking into account

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

Filter by sweeping across with a plane:

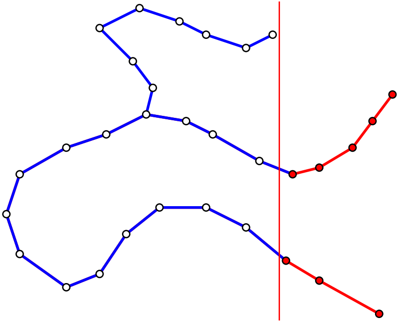


Sweep filtration

Goal: statistical analysis taking into account

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

Filter by sweeping across with a plane:

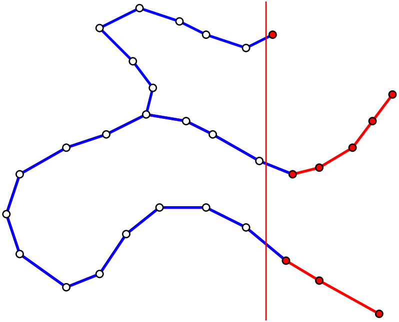


Sweep filtration

Goal: statistical analysis taking into account

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

Filter by sweeping across with a plane:

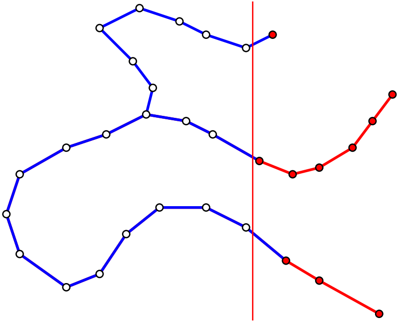


Sweep filtration

Goal: statistical analysis taking into account

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

Filter by sweeping across with a plane:

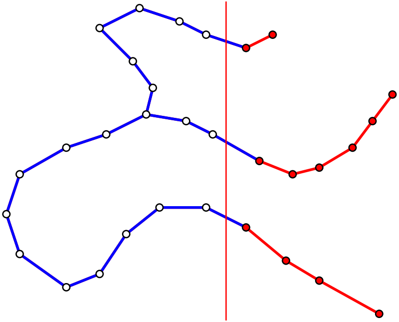


Sweep filtration

Goal: statistical analysis taking into account

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

Filter by sweeping across with a plane:

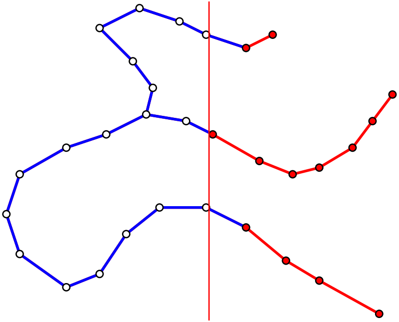


Sweep filtration

Goal: statistical analysis taking into account

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

Filter by sweeping across with a plane:

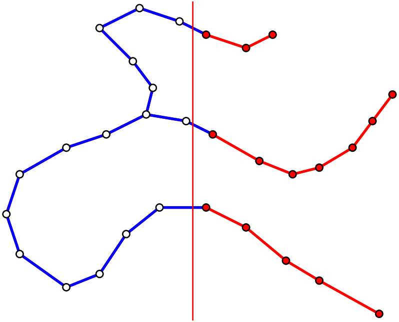


Sweep filtration

Goal: statistical analysis taking into account

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

Filter by sweeping across with a plane:

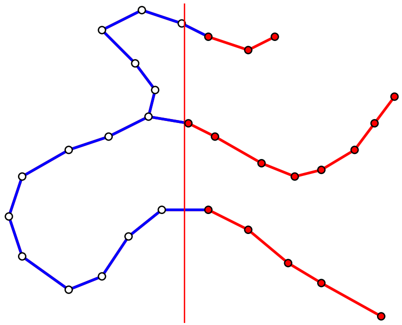


Sweep filtration

Goal: statistical analysis taking into account

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

Filter by sweeping across with a plane:

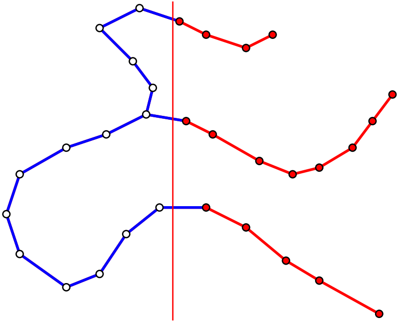


Sweep filtration

Goal: statistical analysis taking into account

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

Filter by sweeping across with a plane:

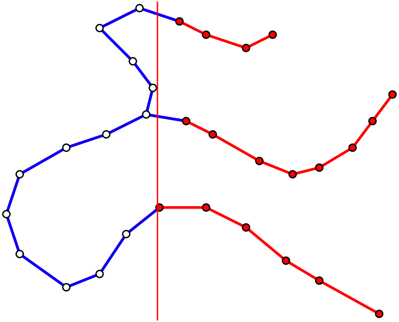


Sweep filtration

Goal: statistical analysis taking into account

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

Filter by sweeping across with a plane:

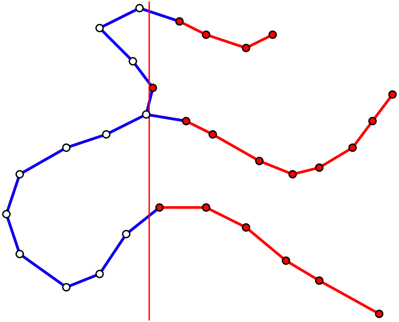


Sweep filtration

Goal: statistical analysis taking into account

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

Filter by sweeping across with a plane:

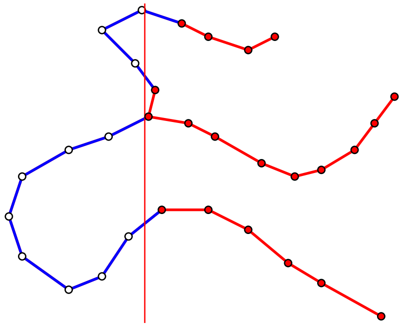


Sweep filtration

Goal: statistical analysis taking into account

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

Filter by sweeping across with a plane:

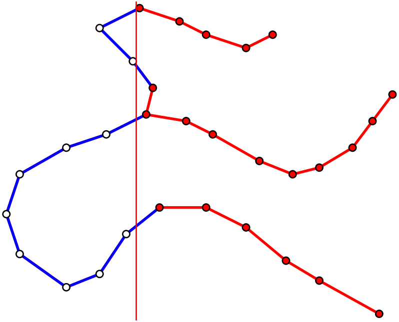


Sweep filtration

Goal: statistical analysis taking into account

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

Filter by sweeping across with a plane:

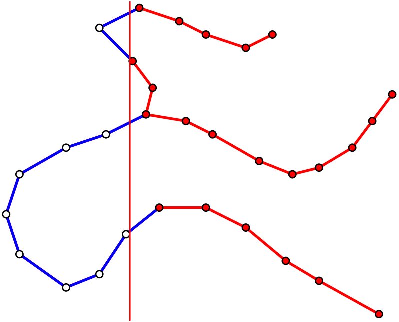


Sweep filtration

Goal: statistical analysis taking into account

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

Filter by sweeping across with a plane:

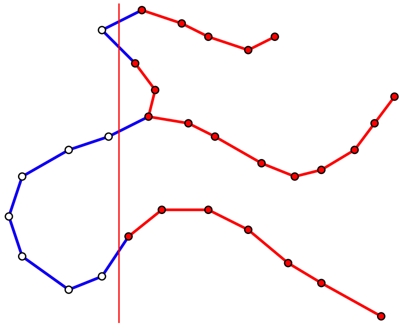


Sweep filtration

Goal: statistical analysis taking into account

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

Filter by sweeping across with a plane:

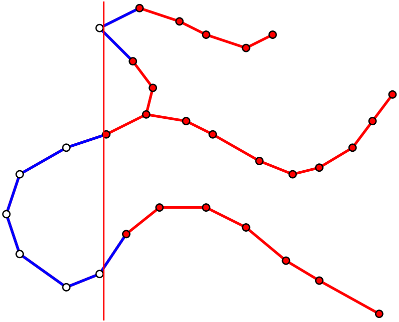


Sweep filtration

Goal: statistical analysis taking into account

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

Filter by sweeping across with a plane:

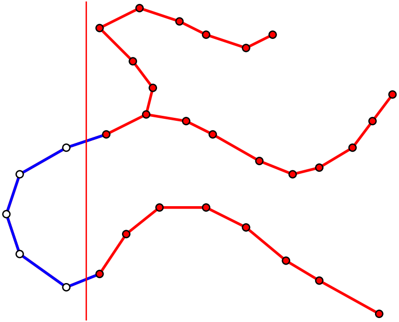


Sweep filtration

Goal: statistical analysis taking into account

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

Filter by sweeping across with a plane:

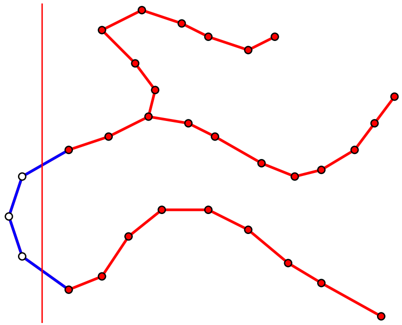


Sweep filtration

Goal: statistical analysis taking into account

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

Filter by sweeping across with a plane:

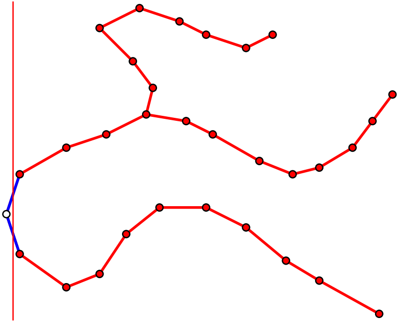


Sweep filtration

Goal: statistical analysis taking into account

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

Filter by sweeping across with a plane:

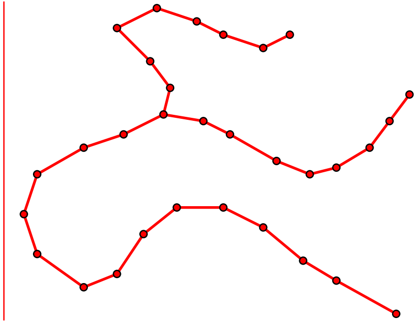


Sweep filtration

Goal: statistical analysis taking into account

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

Filter by sweeping across with a plane:

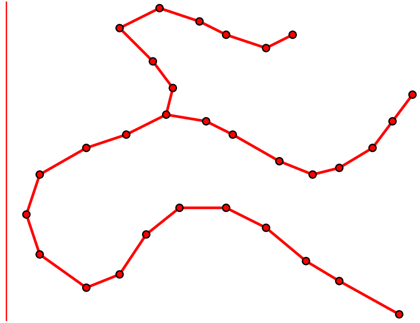


Sweep filtration

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:

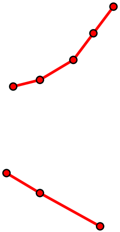
- birth time of each new component
- death of each component (when it joins to an older component)

Sweep filtration

Goal: statistical analysis taking into account

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

Filter by sweeping across with a plane:



Record:

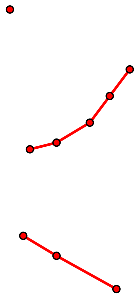
- birth time of each new component
- death of each component (when it joins to an older component)

Sweep filtration

Goal: statistical analysis taking into account

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

Filter by sweeping across with a plane:



Record:

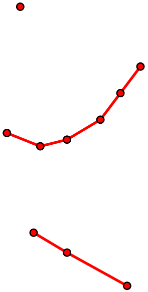
- birth time of each new component
- death of each component (when it joins to an older component)

Sweep filtration

Goal: statistical analysis taking into account

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

Filter by sweeping across with a plane:



Record:

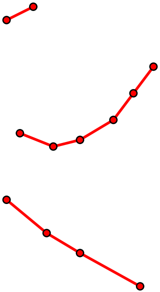
- birth time of each new component
- death of each component (when it joins to an older component)

Sweep filtration

Goal: statistical analysis taking into account

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

Filter by sweeping across with a plane:



Record:

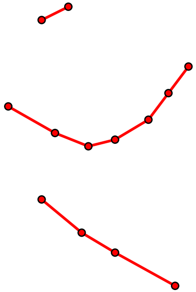
- birth time of each new component
- death of each component (when it joins to an older component)

Sweep filtration

Goal: statistical analysis taking into account

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

Filter by sweeping across with a plane:



Record:

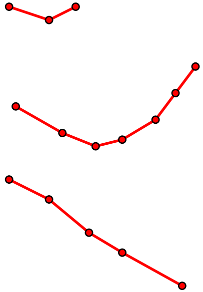
- birth time of each new component
- death of each component (when it joins to an older component)

Sweep filtration

Goal: statistical analysis taking into account

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

Filter by sweeping across with a plane:



Record:

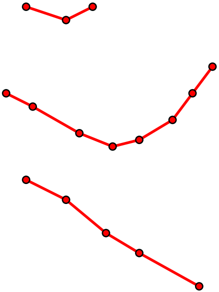
- birth time of each new component
- death of each component (when it joins to an older component)

Sweep filtration

Goal: statistical analysis taking into account

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

Filter by sweeping across with a plane:



Record:

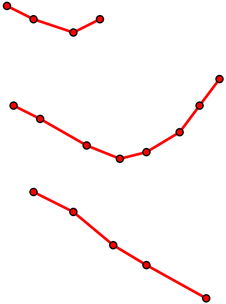
- birth time of each new component
- death of each component (when it joins to an older component)

Sweep filtration

Goal: statistical analysis taking into account

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

Filter by sweeping across with a plane:



Record:

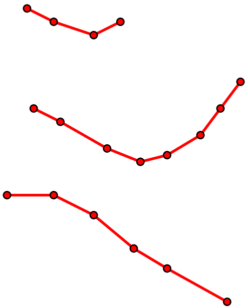
- birth time of each new component
- death of each component (when it joins to an older component)

Sweep filtration

Goal: statistical analysis taking into account

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

Filter by sweeping across with a plane:



Record:

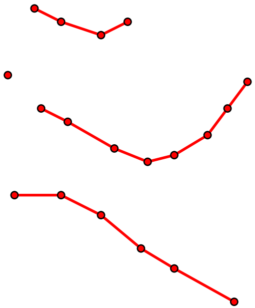
- birth time of each new component
- death of each component (when it joins to an older component)

Sweep filtration

Goal: statistical analysis taking into account

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

Filter by sweeping across with a plane:



Record:

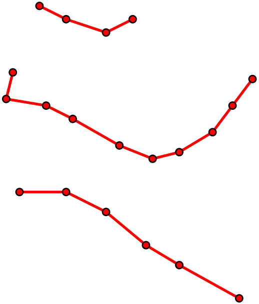
- birth time of each new component
- death of each component (when it joins to an older component)

Sweep filtration

Goal: statistical analysis taking into account

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

Filter by sweeping across with a plane:



Record:

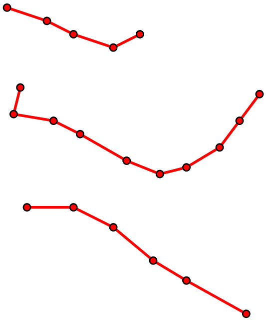
- birth time of each new component
- death of each component (when it joins to an older component)

Sweep filtration

Goal: statistical analysis taking into account

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

Filter by sweeping across with a plane:



Record:

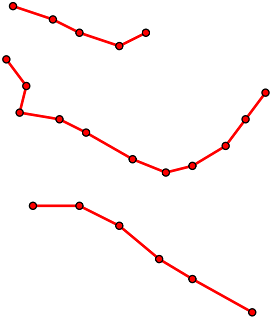
- birth time of each new component
- death of each component (when it joins to an older component)

Sweep filtration

Goal: statistical analysis taking into account

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

Filter by sweeping across with a plane:



Record:

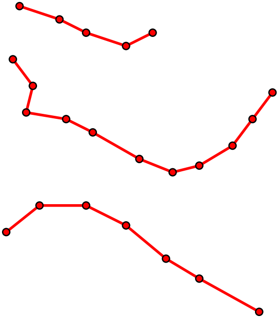
- birth time of each new component
- death of each component (when it joins to an older component)

Sweep filtration

Goal: statistical analysis taking into account

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

Filter by sweeping across with a plane:



Record:

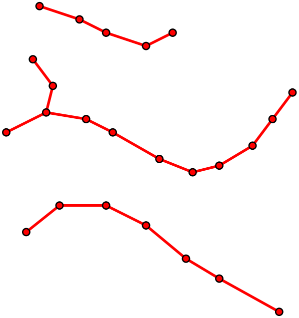
- birth time of each new component
- death of each component (when it joins to an older component)

Sweep filtration

Goal: statistical analysis taking into account

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

Filter by sweeping across with a plane:



Record:

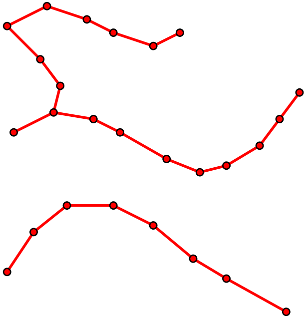
- birth time of each new component
- death of each component (when it joins to an older component)

Sweep filtration

Goal: statistical analysis taking into account

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

Filter by sweeping across with a plane:



Record:

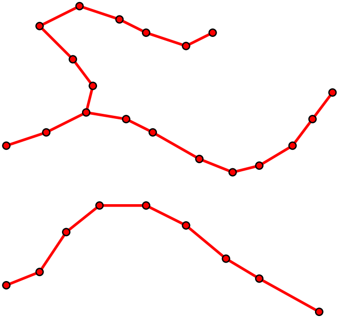
- birth time of each new component
- death of each component (when it joins to an older component)

Sweep filtration

Goal: statistical analysis taking into account

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

Filter by sweeping across with a plane:



Record:

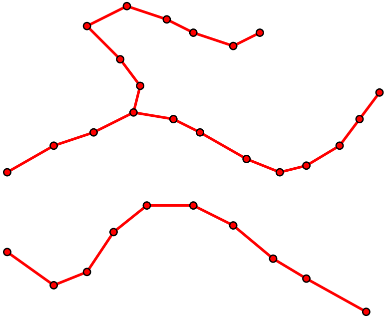
- birth time of each new component
- death of each component (when it joins to an older component)

Sweep filtration

Goal: statistical analysis taking into account

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

Filter by sweeping across with a plane:



Record:

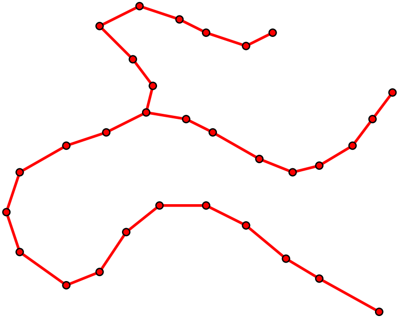
- birth time of each new component
- death of each component (when it joins to an older component)

Sweep filtration

Goal: statistical analysis taking into account

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

Filter by sweeping across with a plane:



Record:

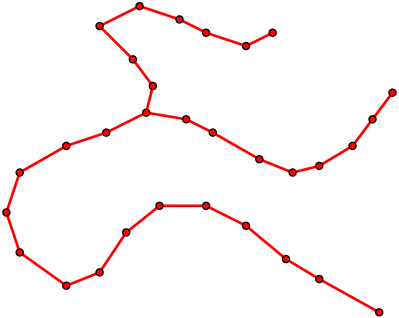
- birth time of each new component
- death of each component (when it joins to an older component)

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.

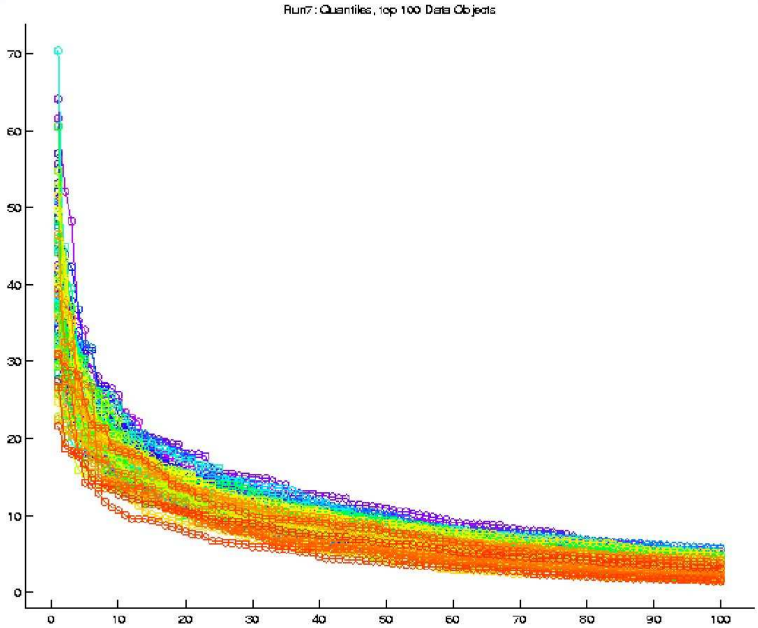
- Pearson correlation 0.52663
- p-value 3.0127×10^{-8} strongly significant

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 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.

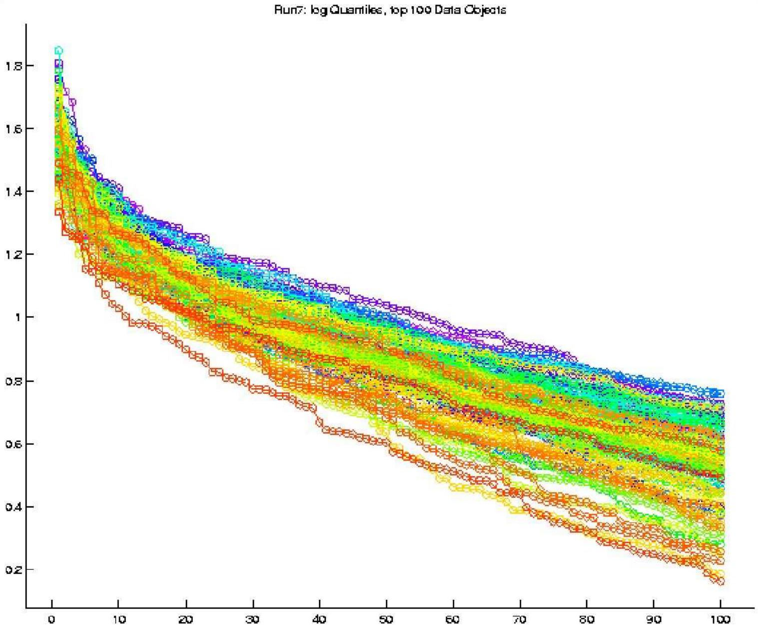
- Pearson correlation 0.52663
- p-value 3.0127×10^{-8} strongly significant

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.

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.

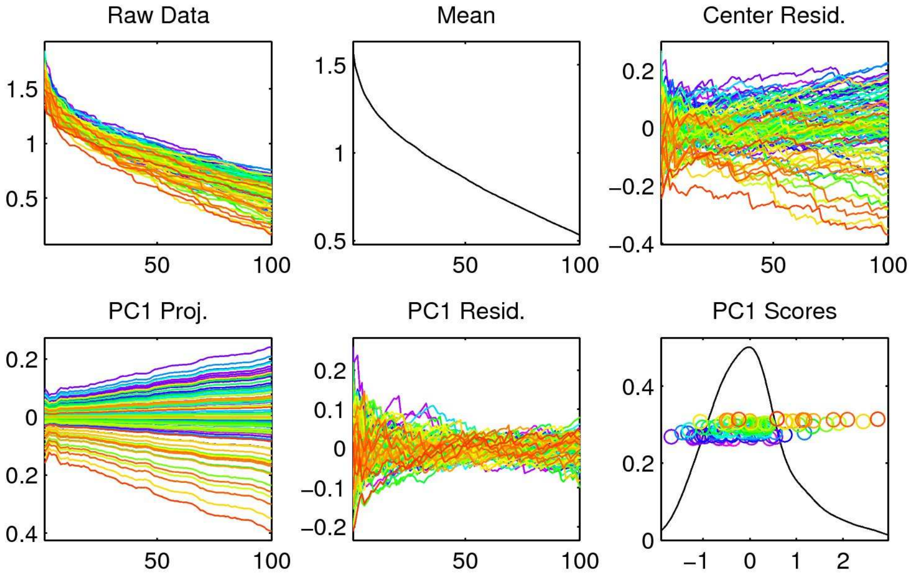
- Pearson correlation 0.52663
- p-value 3.0127×10^{-8} strongly significant

Remarks. Results essentially unchanged after

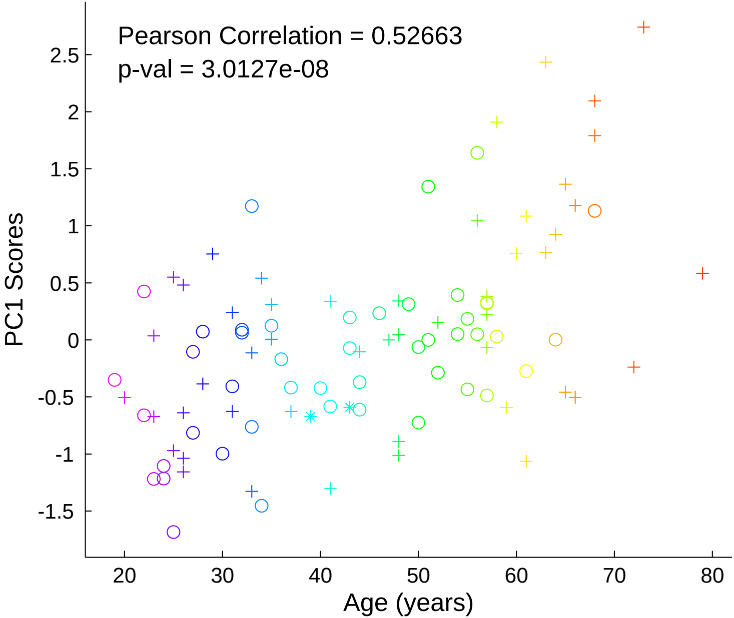
- rescaling to account for natural variation in overall brain size (force standard deviation of the set of bar lengths to equal 1)
- rescaling to account for known correlation of age vs. total vessel length L [Bullitt, et al. 2005] (divide by L, $\sqrt{L}$, or $\sqrt[3]{L}$)
- repeating the analysis 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
2.5
2
1.5
1
0.5
0
-0.5
-1
-1.5
20
30
40
50
60
70
80
Age (years)

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.

Lessons.

- 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.

Lessons.

- 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.

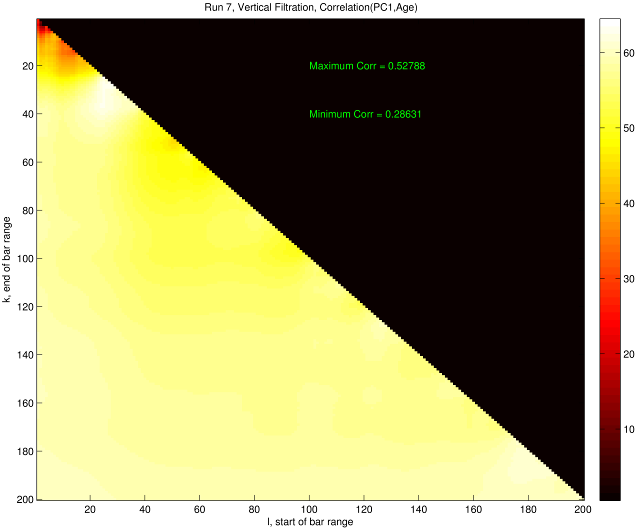
For brain artery trees.

- 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.

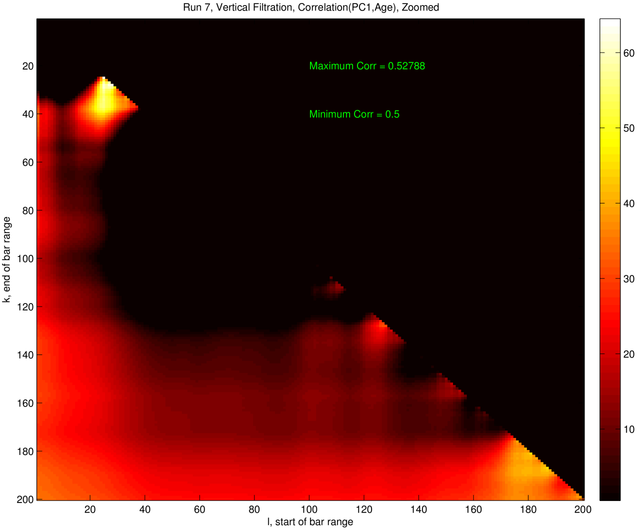
Lessons.

- 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
20
30
40
50
60

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.

Lessons.

- 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.

Lessons.

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

Next steps

Pathology.

- Can persistence (on brain artery trees, or lungs, . . .) detect or measure pathology?
- Compare "tortuosity" [Bullitt, et al.].
- Filter by (radius of) curvature to highlight high-frequencey bends.

Additional analyses.

- Explain residual strength of persistent homology age correlation by independent geometric measures; interpret anatomically.
- Check for overfitting: subsample.
- Other persistence methods, such as landscapes [Bubenik 2012].

Additional datasets.

- fruit fly wings (with Houle, Bendich, Cruz)
- lung vasculature (with McLean et al., Bendich, Marron)
- fMRI (with Lazar et al.)

Bar code statistics. Mean, variance, confidence: (metric) geometry of the space of persistence diagrams [Turner, Mileyko, Mukherjee, Harer 2012]

Next steps

Pathology.

- Can persistence (on brain artery trees, or lungs, . . .) detect or measure pathology?
- Compare "tortuosity" [Bullitt, et al.].
- Filter by (radius of) curvature to highlight high-frequencey bends.

Additional analyses.

- Explain residual strength of persistent homology age correlation by independent geometric measures; interpret anatomically.
- Check for overfitting: subsample.
- Other persistence methods, such as landscapes [Bubenik 2012].

Additional datasets.

- fruit fly wings (with Houle, Bendich, Cruz)
- lung vasculature (with McLean et al., Bendich, Marron)
- fMRI (with Lazar et al.)

Bar code statistics. Mean, variance, confidence: (metric) geometry of the space of persistence diagrams [Turner, Mileyko, Mukherjee, Harer 2012]

Next steps

Pathology.

- Can persistence (on brain artery trees, or lungs, . . .) detect or measure pathology?
- Compare "tortuosity" [Bullitt, et al.].
- Filter by (radius of) curvature to highlight high-frequencey bends.

Additional analyses.

- Explain residual strength of persistent homology age correlation by independent geometric measures; interpret anatomically.
- Check for overfitting: subsample.
- Other persistence methods, such as landscapes [Bubenik 2012].

Additional datasets.

- fruit fly wings (with Houle, Bendich, Cruz)
- lung vasculature (with McLean et al., Bendich, Marron)
- fMRI (with Lazar et al.)

Bar code statistics. Mean, variance, confidence: (metric) geometry of the space of persistence diagrams [Turner, Mileyko, Mukherjee, Harer 2012]

Next steps

Pathology.

- Can persistence (on brain artery trees, or lungs, . . .) detect or measure pathology?
- Compare “tortuosity” [Bullitt, et al.].
- Filter by (radius of) curvature to highlight high-frequencey bends.

Additional analyses.

- Explain residual strength of persistent homology age correlation by independent geometric measures; interpret anatomically.
- Check for overfitting: subsample.
- Other persistence methods, such as landscapes [Bubenik 2012].

Additional datasets.

- fruit fly wings (with Houle, Bendich, Cruz)
- lung vasculature (with McLean et al., Bendich, Marron)
- fMRI (with Lazar et al.)

Bar code statistics. Mean, variance, confidence: (metric) geometry of the space of persistence diagrams [Turner, Mileyko, Mukherjee, Harer 2012]

Next steps

Pathology.

- Can persistence (on brain artery trees, or lungs, . . .) detect or measure pathology?
- Compare “tortuosity” [Bullitt, et al.].
- Filter by (radius of) curvature to highlight high-frequencey bends.

Additional analyses.

- Explain residual strength of persistent homology age correlation by independent geometric measures; interpret anatomically.
- Check for overfitting: subsample.
- Other persistence methods, such as landscapes [Bubenik 2012].

Additional datasets.

- fruit fly wings (with Houle, Bendich, Cruz)
- lung vasculature (with McLean et al., Bendich, Marron)
- fMRI (with Lazar et al.)

Bar code statistics. Mean, variance, confidence: (metric) geometry of the space of persistence diagrams [Turner, Mileyko, Mukherjee, Harer 2012]

Next steps

Pathology.

- Can persistence (on brain artery trees, or lungs, . . .) detect or measure pathology?
- Compare "tortuosity" [Bullitt, et al.].
- Filter by (radius of) curvature to highlight high-frequencey bends.

Additional analyses.

- Explain residual strength of persistent homology age correlation by independent geometric measures; interpret anatomically.
- Check for overfitting: subsample.
- Other persistence methods, such as landscapes [Bubenik 2012].

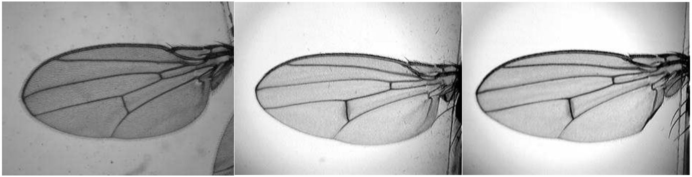
Additional datasets.

- fruit fly wings (with Houle, Bendich, Cruz)
- lung vasculature (with McLean et al., Bendich, Marron)
- fMRI (with Lazar et al.)

Bar code statistics. Mean, variance, confidence: (metric) geometry of the space of persistence diagrams [Turner, Mileyko, Mukherjee, Harer 2012]

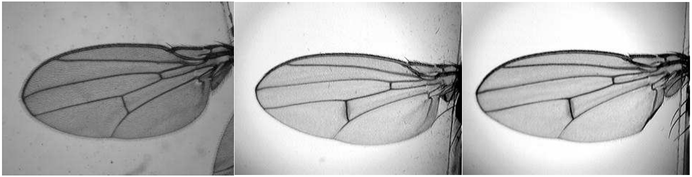
Fruit fly wings

Normal fly wings [photos from David Houle's lab]:

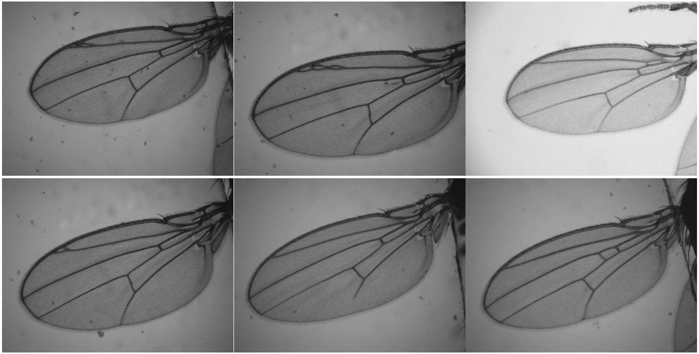


Fruit fly wings

Normal fly wings [photos from David Houle's lab]:



Topologically abnormal veins:



Biological background

What generates topological novelty?

[Houle, et al.]: selecting for certain continuous wing vein deformations yields

- skew toward more oddly shaped wings, but also
- much higher rate of topological novelty

Hypothesis

Topological novelty arises at the extreme of selection for continuous shape characteristics

Test the hypothesis

- "plot" wings in "form space"
- determine whether topological variants lie "in the direction of" continuous shape selected for. . .
- . . . at the extreme in that direction

Other questions

- predict which genes are involved given observed phenotype
- biologically determine which genes are involved, and correlate

Biological background

What generates topological novelty?

[Houle, et al.]: selecting for certain continuous wing vein deformations yields

- skew toward more oddly shaped wings, but also
- much higher rate of topological novelty

Hypothesis

Topological novelty arises at the extreme of selection for continuous shape characteristics

Test the hypothesis

- "plot" wings in "form space"
- determine whether topological variants lie "in the direction of" continuous shape selected for. . .
- . . . at the extreme in that direction

Other questions

- predict which genes are involved given observed phenotype
- biologically determine which genes are involved, and correlate

Biological background

What generates topological novelty?

[Houle, et al.]: selecting for certain continuous wing vein deformations yields

- skew toward more oddly shaped wings, but also
- much higher rate of topological novelty

Hypothesis

Topological novelty arises at the extreme of selection for continuous shape characteristics

Test the hypothesis

- "plot" wings in "form space"
- determine whether topological variants lie "in the direction of" continuous shape selected for. . .
- . . . at the extreme in that direction

Other questions

- predict which genes are involved given observed phenotype
- biologically determine which genes are involved, and correlate

Biological background

What generates topological novelty?

[Houle, et al.]: selecting for certain continuous wing vein deformations yields

- skew toward more oddly shaped wings, but also
- much higher rate of topological novelty

Hypothesis

Topological novelty arises at the extreme of selection for continuous shape characteristics

Test the hypothesis

- "plot" wings in "form space"
- determine whether topological variants lie "in the direction of" continuous shape selected for. . .
- . . . at the extreme in that direction

Other questions

- predict which genes are involved given observed phenotype
- biologically determine which genes are involved, and correlate

Biological background

What generates topological novelty?

[Houle, et al.]: selecting for certain continuous wing vein deformations yields

- skew toward more oddly shaped wings, but also
- much higher rate of topological novelty

Hypothesis

Topological novelty arises at the extreme of selection for continuous shape characteristics

Test the hypothesis

- "plot" wings in "form space"
- determine whether topological variants lie "in the direction of" continuous shape selected for...
- ...at the extreme in that direction

Other questions

- predict which genes are involved given observed phenotype
- biologically determine which genes are involved, and correlate

Stratified persistence

Goal. Statistical analysis encompassing topological vein variation, giving appropriate weight to new singular points in addition to varying shape

- compare phenotypic distance to genotypic distance; needs
- metric specifying distance between topologically distinct wings

Obstacles

- shape spaces need constant numbers of landmarks
- stability of persistent homology [Cohen-Steiner–Edelsbrunner–Harer 2007] ⇒ dense sampling understates short new features

Plan. Encode wing as 2-parameter persistence diagram

- 1st parameter: usual distance (expanding balls)
- 2nd parameter: immunity (intersection homology [Bendich, Harer 2011]): disallow interaction of larger strata with smaller ones

Progress. (with Houle, Bendich, Cruz)

- algorithm(!)
- with low complexity(!)

Stratified persistence

Goal. Statistical analysis encompassing topological vein variation, giving appropriate weight to new singular points in addition to varying shape

- compare phenotypic distance to genotypic distance; needs
- metric specifying distance between topologically distinct wings

Obstacles

- shape spaces need constant numbers of landmarks
- stability of persistent homology [Cohen-Steiner–Edelsbrunner–Harer 2007] ⇒ dense sampling understates short new features

Plan. Encode wing as 2-parameter persistence diagram

- 1st parameter: usual distance (expanding balls)
- 2nd parameter: immunity (intersection homology [Bendich, Harer 2011]): disallow interaction of larger strata with smaller ones

Progress. (with Houle, Bendich, Cruz)

- algorithm(!)
- with low complexity(!)

Stratified persistence

Goal. Statistical analysis encompassing topological vein variation, giving appropriate weight to new singular points in addition to varying shape

- compare phenotypic distance to genotypic distance; needs
- metric specifying distance between topologically distinct wings

Obstacles

- shape spaces need constant numbers of landmarks
- stability of persistent homology [Cohen-Steiner–Edelsbrunner–Harer 2007] $\Rightarrow$ dense sampling understates short new features

Plan. Encode wing as 2-parameter persistence diagram

- 1st parameter: usual distance (expanding balls)
- 2nd parameter: immunity (intersection homology [Bendich, Harer 2011]): disallow interaction of larger strata with smaller ones

Progress. (with Houle, Bendich, Cruz)

- algorithm(!)
- with low complexity(!)

Stratified persistence

Goal. Statistical analysis encompassing topological vein variation, giving appropriate weight to new singular points in addition to varying shape

- compare phenotypic distance to genotypic distance; needs
- metric specifying distance between topologically distinct wings

Obstacles

- shape spaces need constant numbers of landmarks
- stability of persistent homology [Cohen-Steiner–Edelsbrunner–Harer 2007] ⇒ dense sampling understates short new features

Plan. Encode wing as 2-parameter persistence diagram

- 1st parameter: usual distance (expanding balls)
- 2nd parameter: immunity (intersection homology [Bendich, Harer 2011]): disallow interaction of larger strata with smaller ones

Progress. (with Houle, Bendich, Cruz)

- algorithm(!)
- with low complexity(!)

Stratified persistence

Goal. Statistical analysis encompassing topological vein variation, giving appropriate weight to new singular points in addition to varying shape

- compare phenotypic distance to genotypic distance; needs
- metric specifying distance between topologically distinct wings

Obstacles

- shape spaces need constant numbers of landmarks
- stability of persistent homology [Cohen-Steiner–Edelsbrunner–Harer 2007] $\Rightarrow$ dense sampling understates short new features

Plan. Encode wing as 2-parameter persistence diagram

- 1st parameter: usual distance (expanding balls)
- 2nd parameter: immunity (intersection homology [Bendich, Harer 2011]): disallow interaction of larger strata with smaller ones

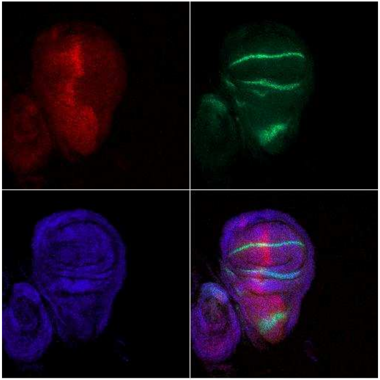
Progress. (with Houle, Bendich, Cruz)

- algorithm(!)
- with low complexity(!)

Developmental dream

Capture morphological or expression time series at cellular resolution to apply (stratified or ordinary) persistence in higher dimension.

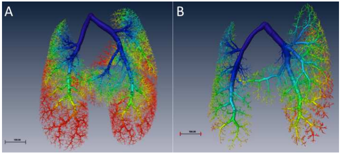
- take development into account: time series for expression levels or vein formation $\rightsquigarrow$ 3D (or higher-dim) geometric structures

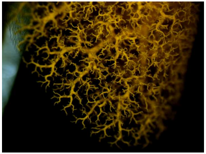


- compare genotypic and phenotypic distance
- reconstruct phylogeny from morphological measurements

Future directions

Lung vasculature. (with McLean et al., Bendich, Marron)





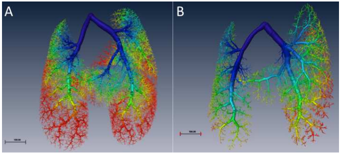
Options for application of (stratified) persistent homology:

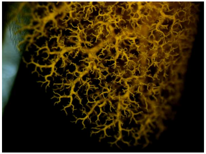
- expand blood vessel tree to fill 3D lung
- filter blood vessel tree by height
  vessel diameter
  curvature

fMRI. (with Lazar et al.): classification using persistent homology

Future directions

Lung vasculature. (with McLean et al., Bendich, Marron)





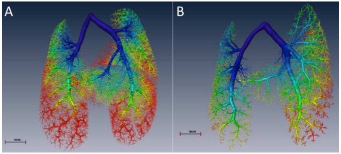
Options for application of (stratified) persistent homology:

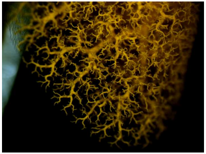
- expand blood vessel tree to fill 3D lung
- filter blood vessel tree by height
  vessel diameter
  curvature

fMRI. (with Lazar et al.): classification using persistent homology

Future directions

Lung vasculature. (with McLean et al., Bendich, Marron)





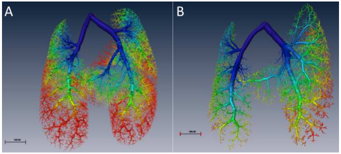
Options for application of (stratified) persistent homology:

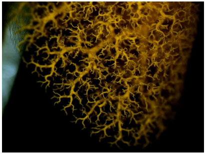
- expand blood vessel tree to fill 3D lung
- filter blood vessel tree by height
vessel diameter
curvature

fMRI. (with Lazar et al.): classification using persistent homology

Future directions

Lung vasculature. (with McLean et al., Bendich, Marron)

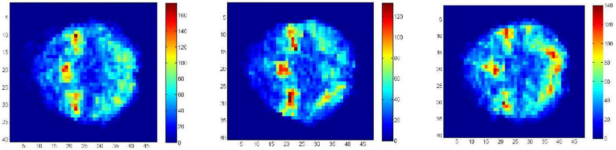




Options for application of (stratified) persistent homology:

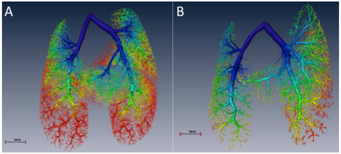
- expand blood vessel tree to fill 3D lung
- filter blood vessel tree by height
  vessel diameter
  curvature

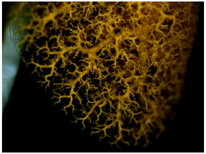
fMRI. (with Lazar et al.): classification using persistent homology



Future directions

Lung vasculature. (with McLean et al., Bendich, Marron)

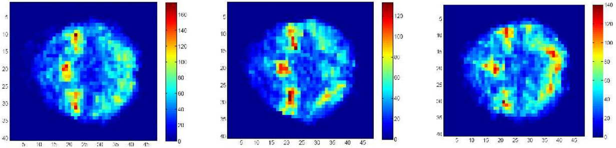




Options for application of (stratified) persistent homology:

- expand blood vessel tree to fill 3D lung
- filter blood vessel tree by height
  vessel diameter
  curvature

fMRI. (with Lazar et al.): classification using persistent homology



Thank You