Persistent homology analysis of brain artery trees

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.)

[arXiv:stat.AP/1411.6652]

Joint Mathematics Meetings, San Antonio

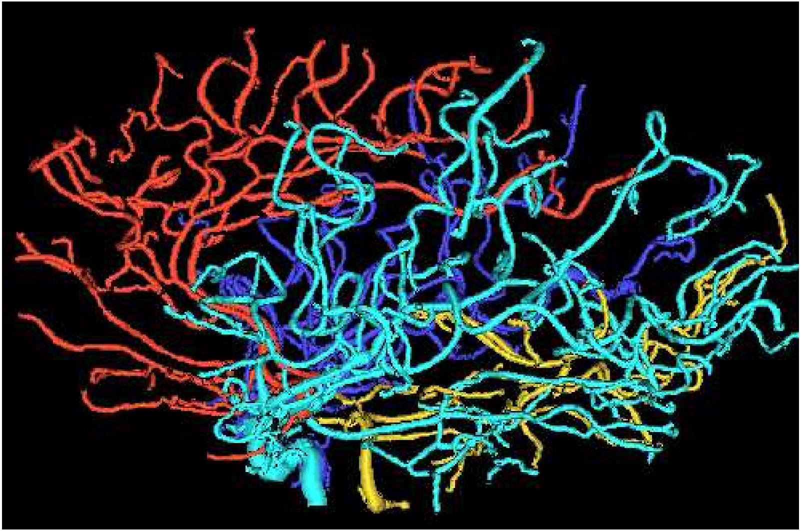
11 January 2015

Brain artery trees

Goal: Statistical analysis taking 3D geometry into account

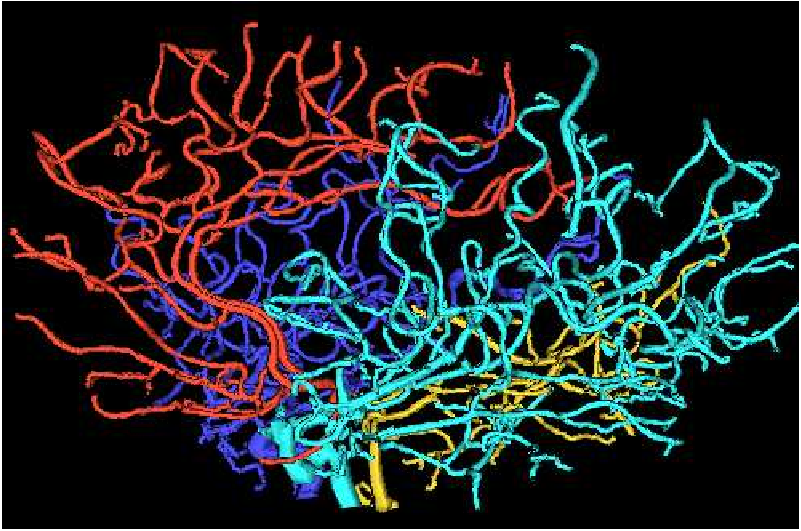
- predict stroke tendency
- screen for loci of pathology, such as tumors
- explore how age affects vascularization

Tube tracking



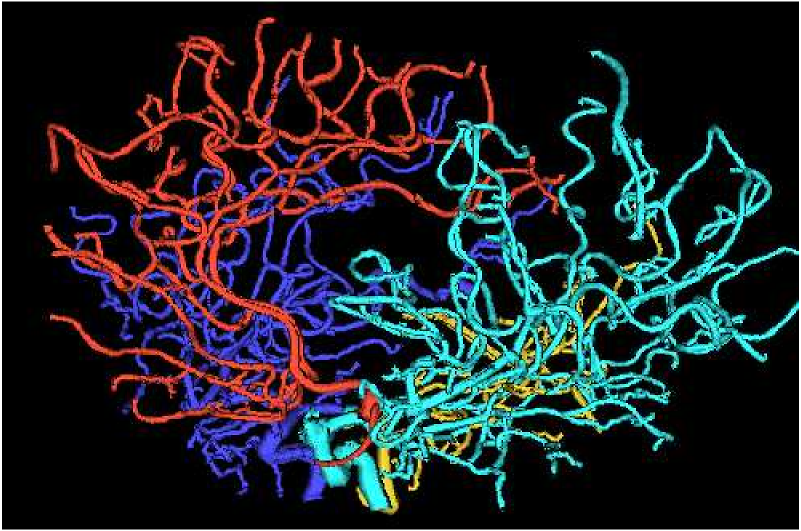
[Bullitt and Aylward, 2002]

Tube tracking



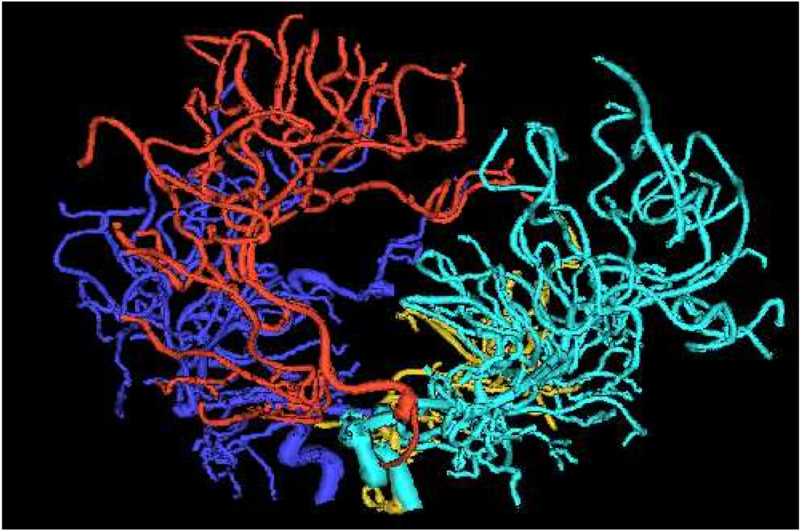
[Bullitt and Aylward, 2002]

Tube tracking



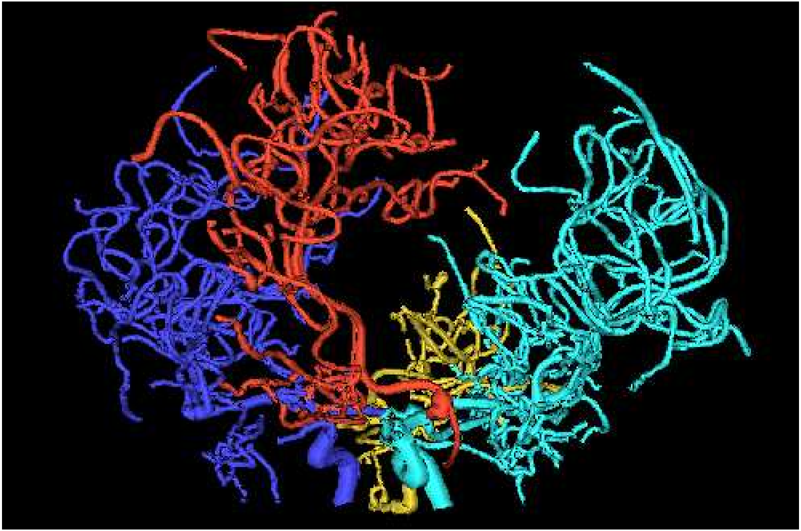
[Bullitt and Aylward, 2002]

Tube tracking



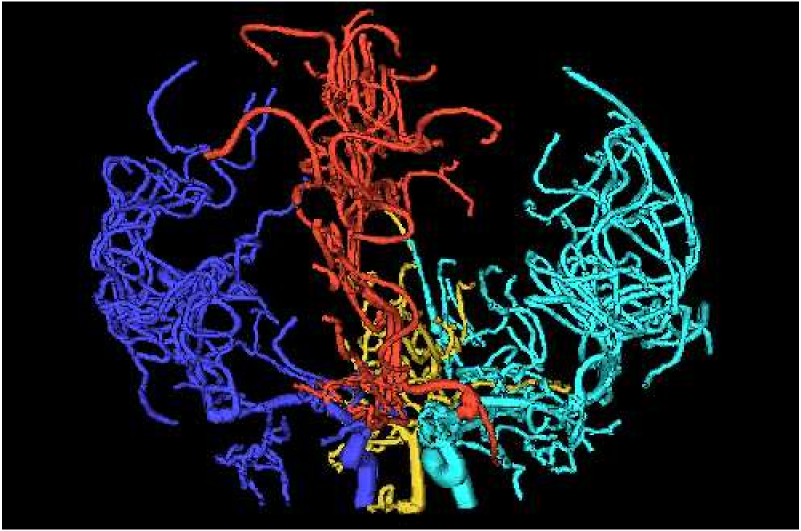
[Bullitt and Aylward, 2002]

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

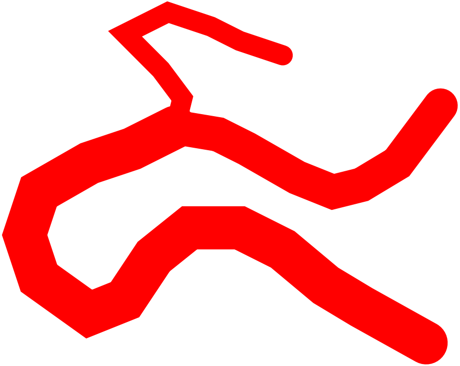
The images:

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:

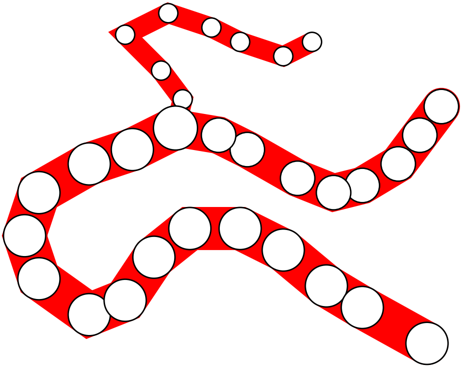


Brain artery trees

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:

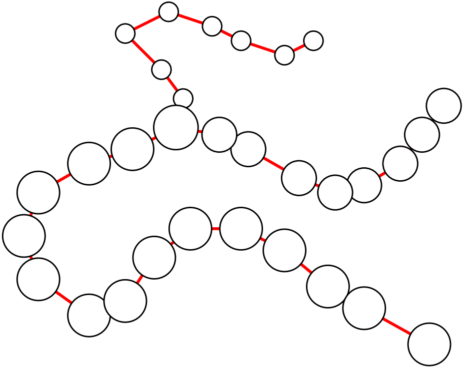


Brain artery trees

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

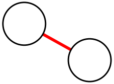
The data structure:

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:

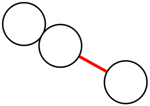


Brain artery trees

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:

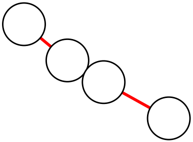


Brain artery trees

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:

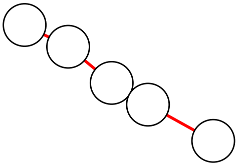


Brain artery trees

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:

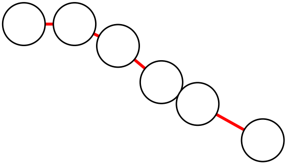


Brain artery trees

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:

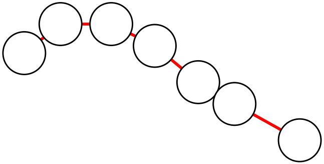


Brain artery trees

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:

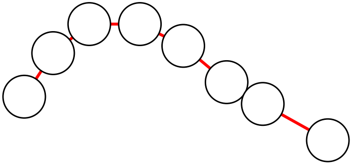


Brain artery trees

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:

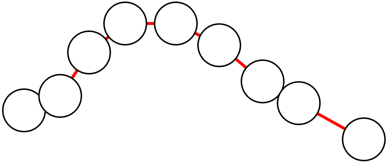


Brain artery trees

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:

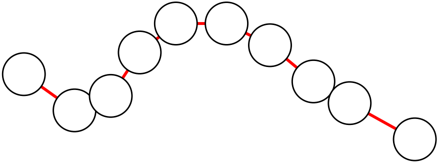


Brain artery trees

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:

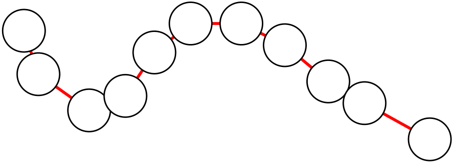


Brain artery trees

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:

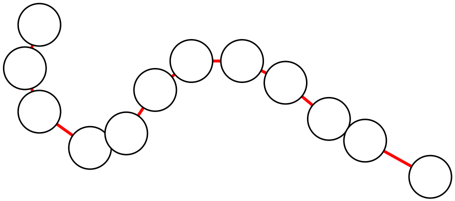


Brain artery trees

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:

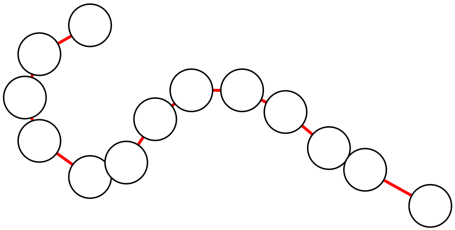


Brain artery trees

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:

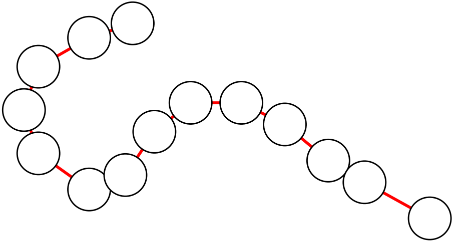


Brain artery trees

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:

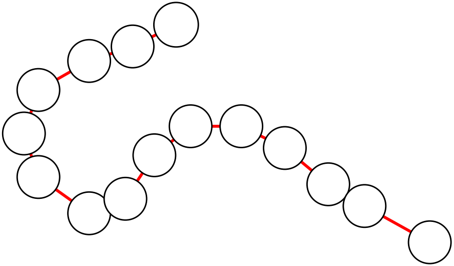


Brain artery trees

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:

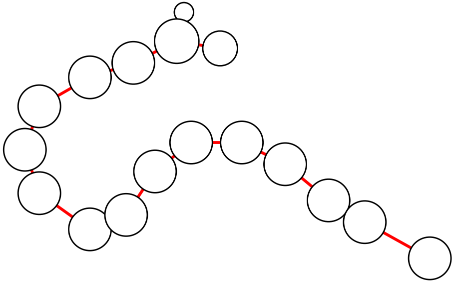


Brain artery trees

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:

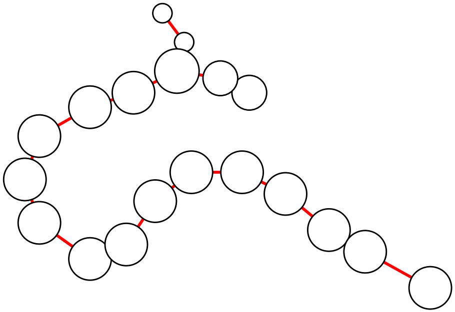


Brain artery trees

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:

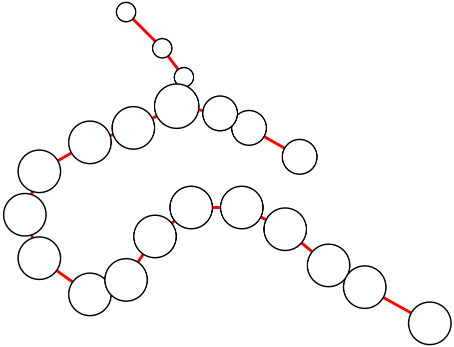


Brain artery trees

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:

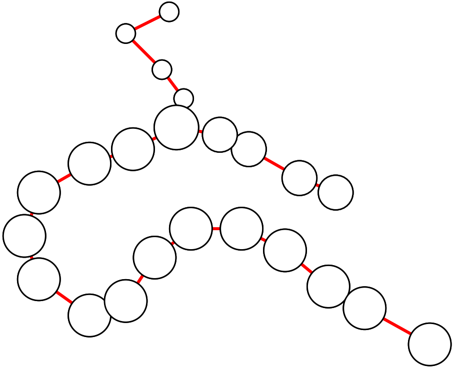


Brain artery trees

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:

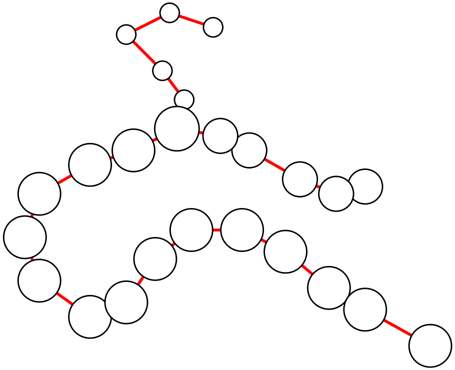


Brain artery trees

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:

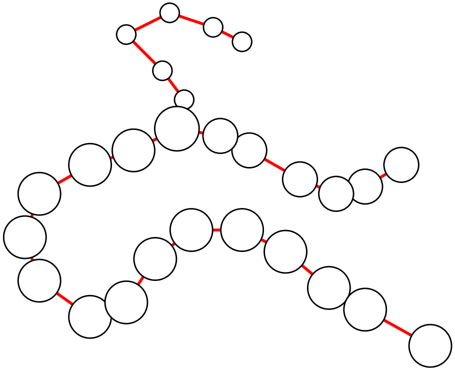


Brain artery trees

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:

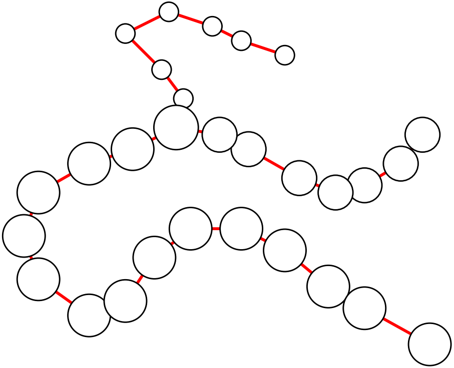


Brain artery trees

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:

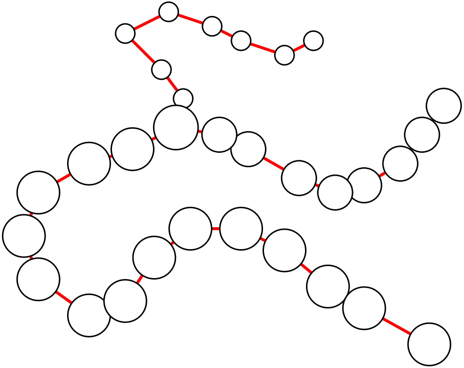


Brain artery trees

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:

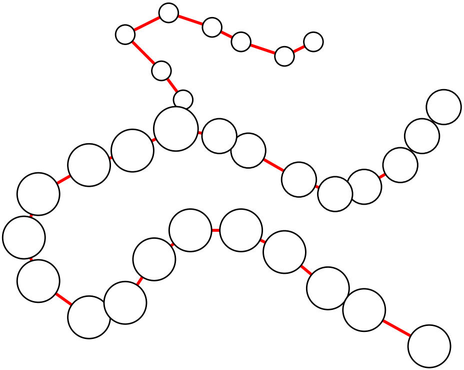


Brain artery trees

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:

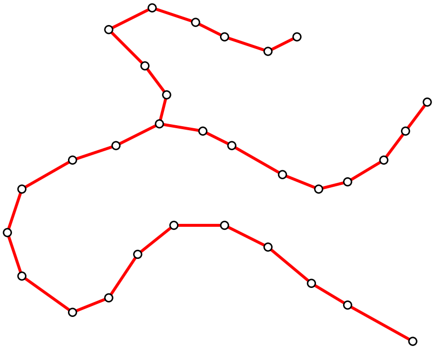


Brain artery trees

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:

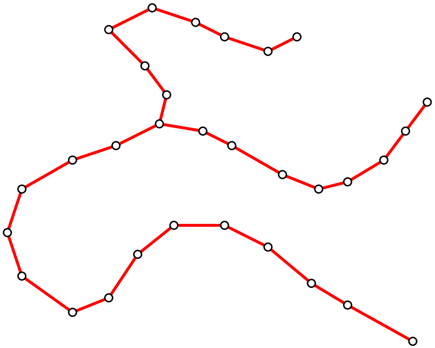


Brain artery trees

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:



Persistent homology

Fix a topological space X

- 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

Fix a topological space X

- 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

Fix a topological space X

- build X step by step
- measure evolving topology

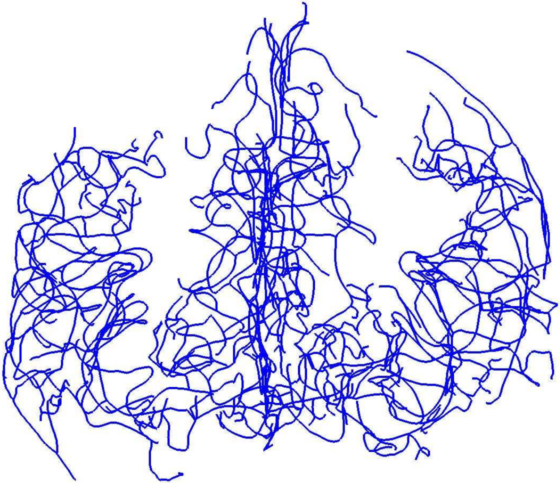
Def. Let $X_\bullet$ be a filtered space, meaning $\varnothing = X_0 \subset X_1 \subset \cdots \subset X_m = X$. The persistent homology $H_i X_\bullet$ is $H_i X_1 \to H_i X_2 \to \cdots \to H_i X_m$, a sequence of vector space homomorphisms.

Examples:

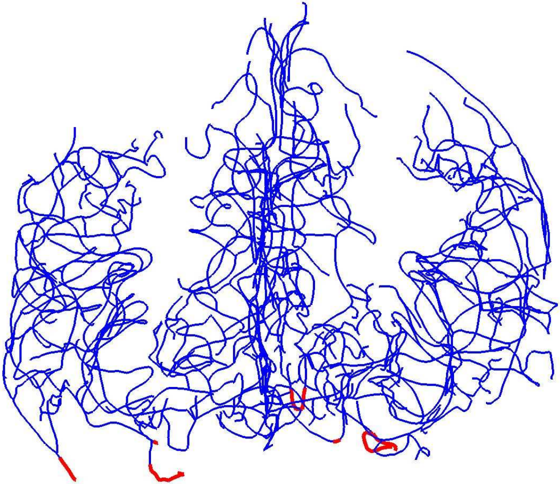
1. Given a function $f : X \to \mathbb{R}$, let $X_k = f^{-1}\big((-\infty, t_k]\big)$. Good choice of $t_0, \ldots, t_m \in \mathbb{R}$: the values of t across which $H_i X_t$ changes.
2. Any simplicial complex: build it simplex by simplex in some order.

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

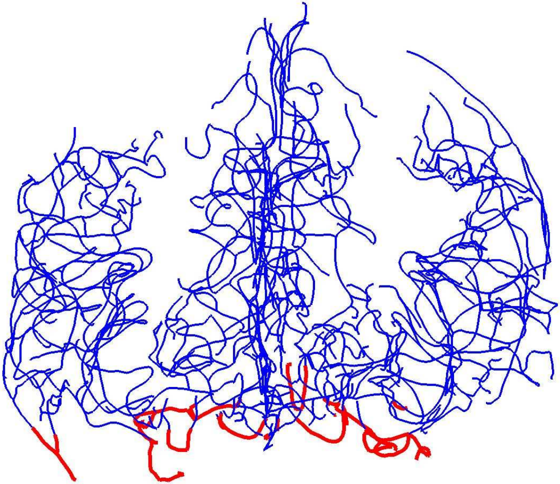
Example: filling brains



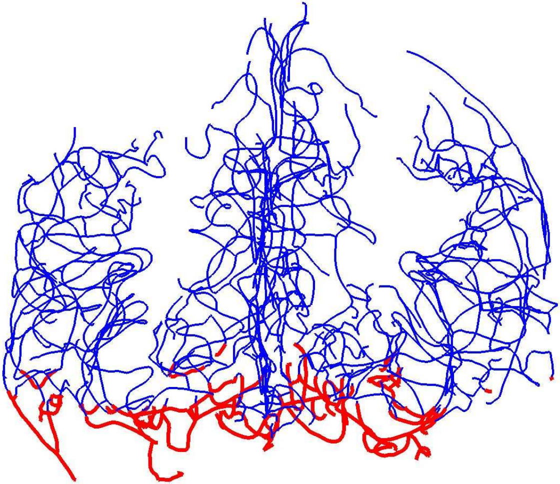
Example: filling brains



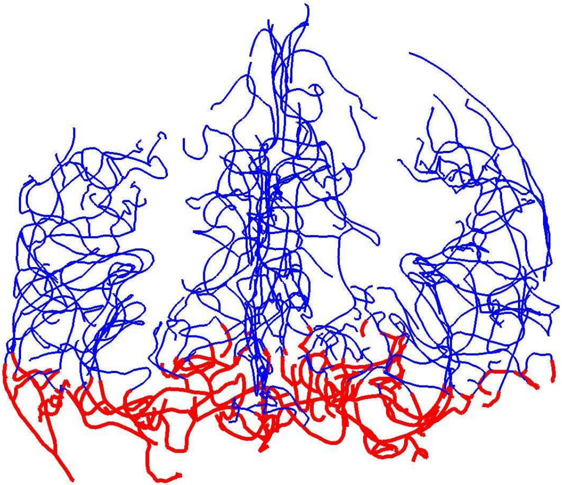
Example: filling brains



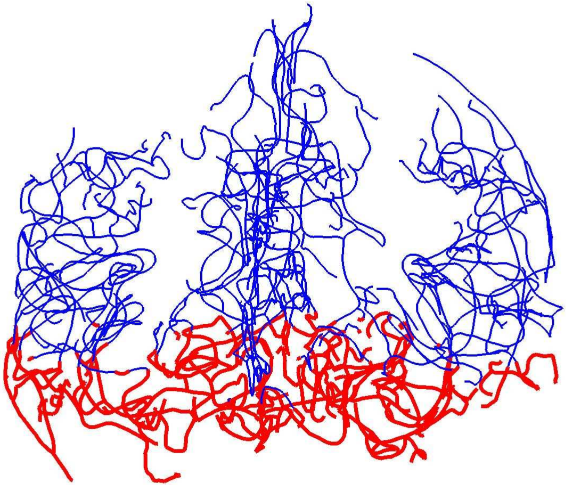
Example: filling brains



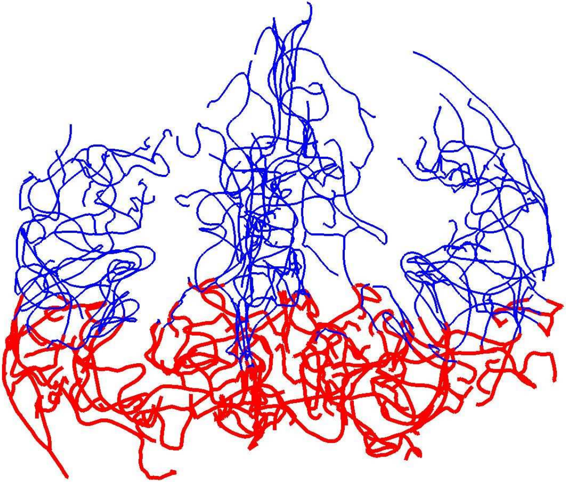
Example: filling brains



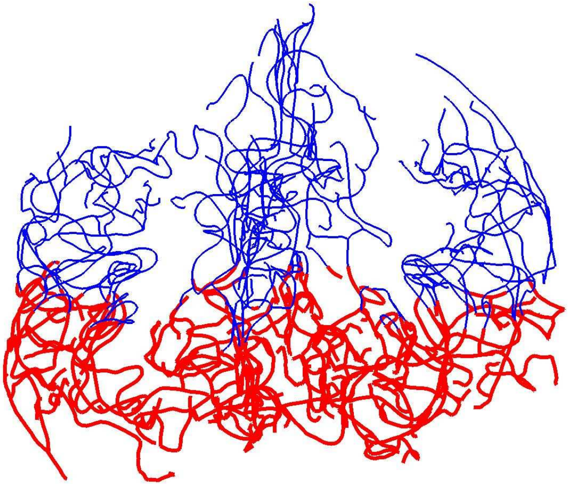
Example: filling brains



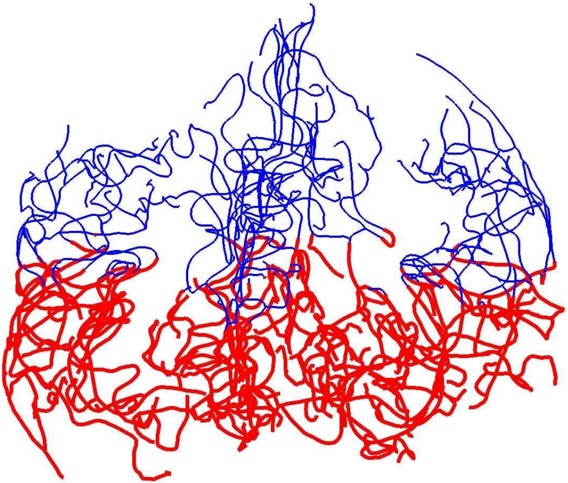
Example: filling brains



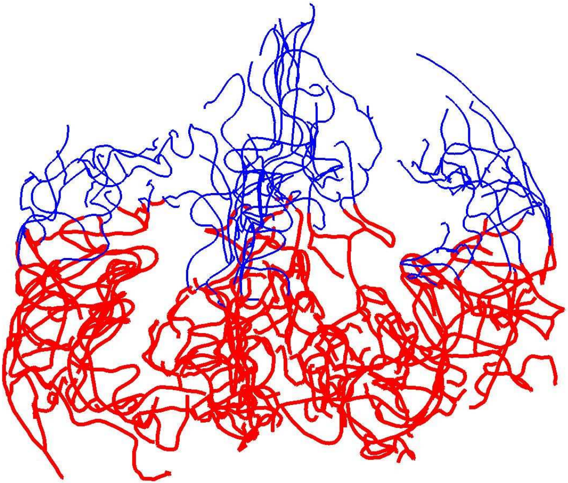
Example: filling brains



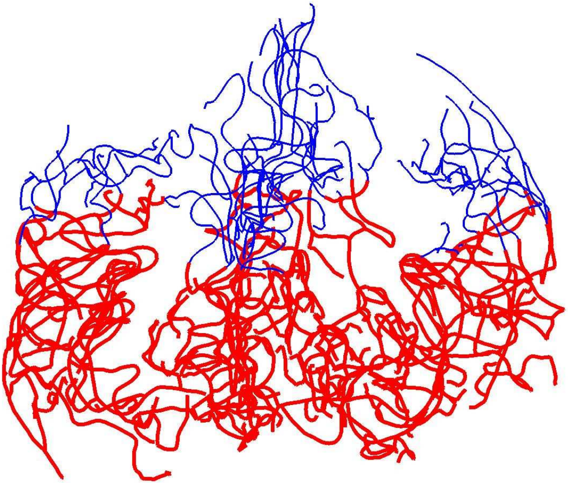
Example: filling brains



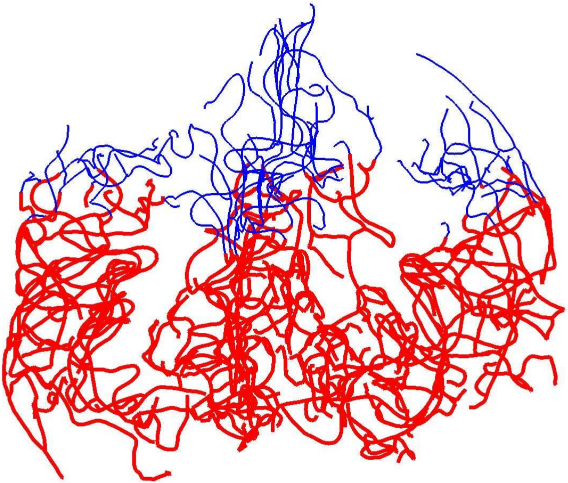
Example: filling brains



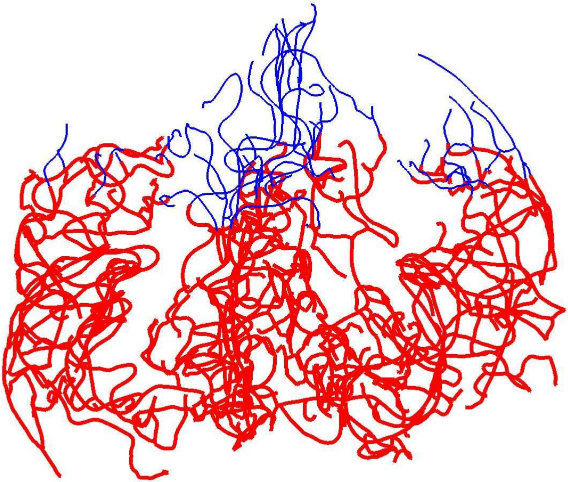
Example: filling brains



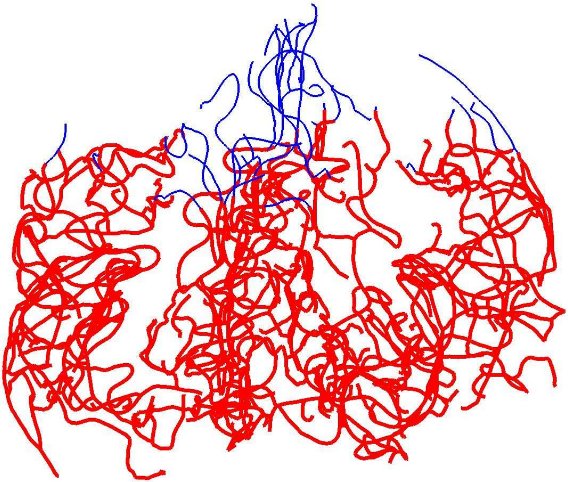
Example: filling brains



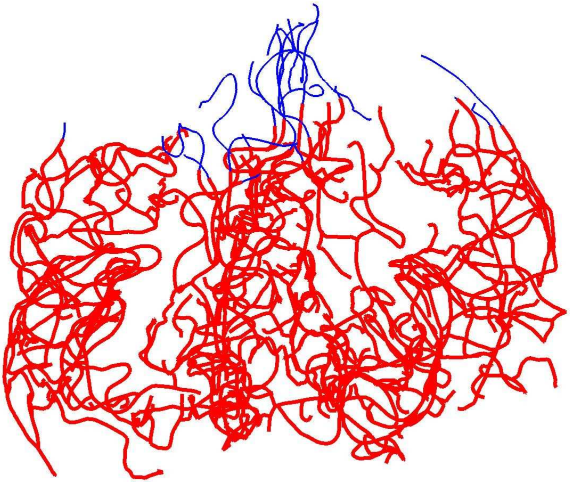
Example: filling brains



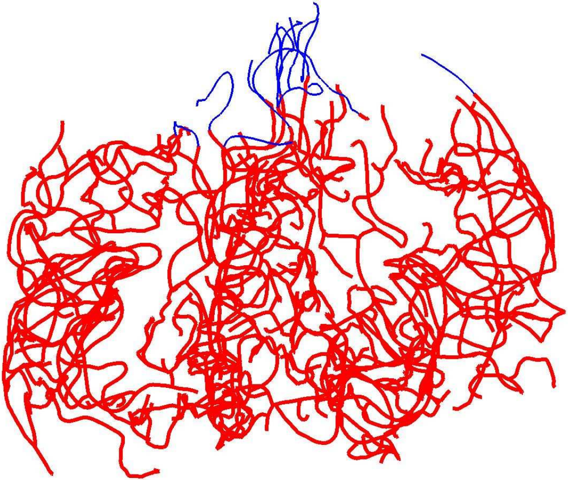
Example: filling brains



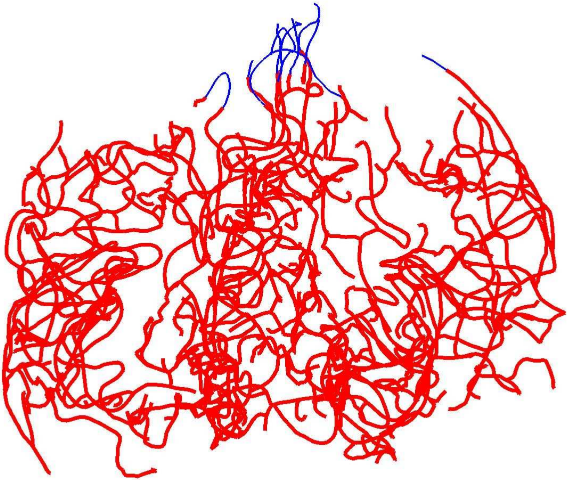
Example: filling brains



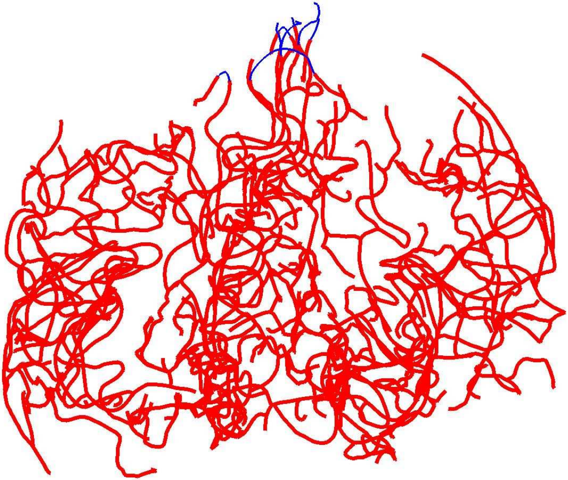
Example: filling brains



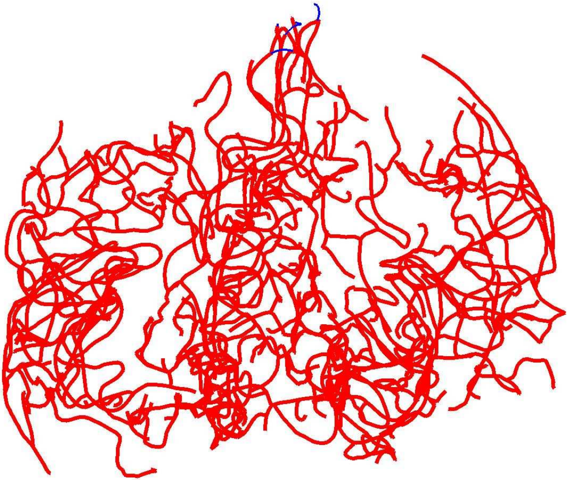
Example: filling brains



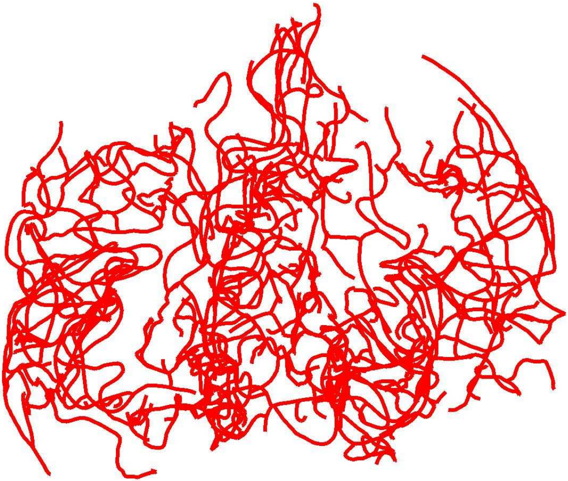
Example: filling brains



Example: filling brains



Example: filling brains



Persistent homology

Fix a topological space X

- 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

Fix a topological space X

- 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

Fix a topological space X

- 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

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”

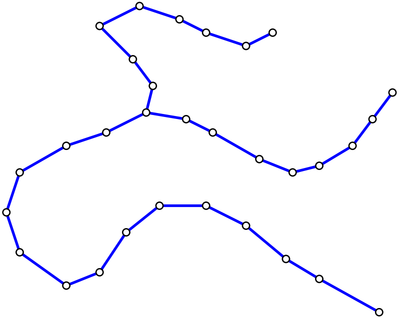
Filter by sweeping across with a plane:

Sweep filtration

Goal: statistical analysis taking into account

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

Filter by sweeping across with a plane:

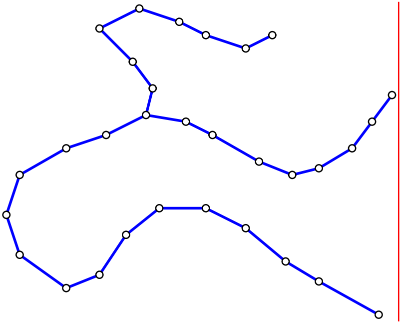


Sweep filtration

Goal: statistical analysis taking into account

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

Filter by sweeping across with a plane:

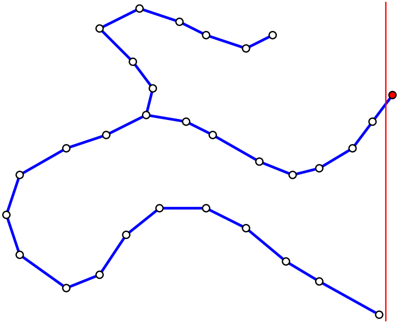


Sweep filtration

Goal: statistical analysis taking into account

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

Filter by sweeping across with a plane:

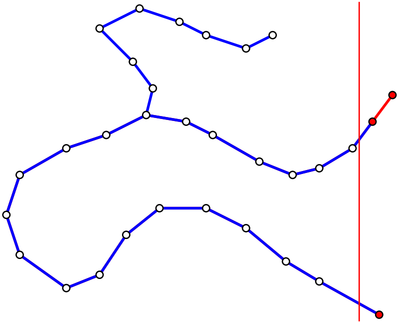


Sweep filtration

Goal: statistical analysis taking into account

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

Filter by sweeping across with a plane:

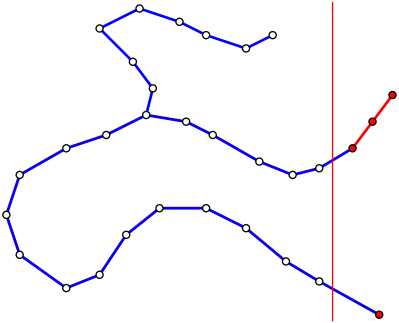


Sweep filtration

Goal: statistical analysis taking into account

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

Filter by sweeping across with a plane:

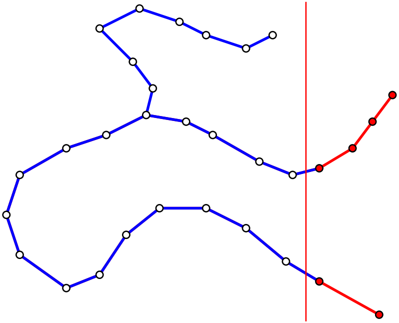


Sweep filtration

Goal: statistical analysis taking into account

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

Filter by sweeping across with a plane:

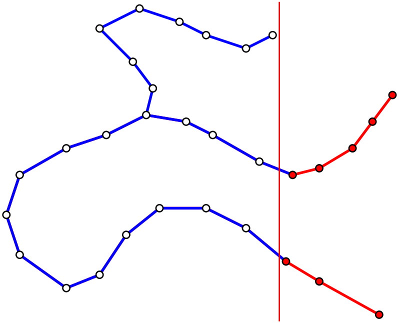


Sweep filtration

Goal: statistical analysis taking into account

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

Filter by sweeping across with a plane:

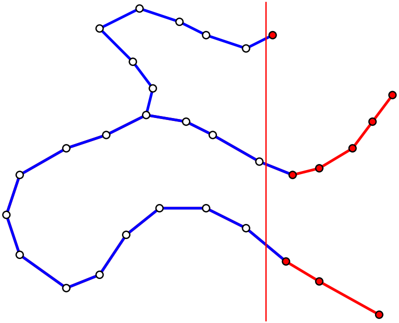


Sweep filtration

Goal: statistical analysis taking into account

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

Filter by sweeping across with a plane:

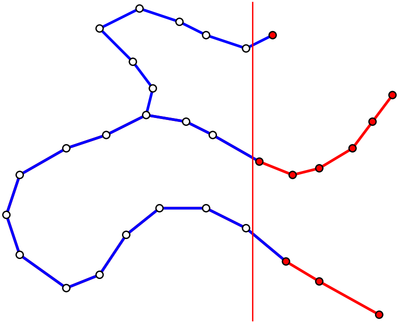


Sweep filtration

Goal: statistical analysis taking into account

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

Filter by sweeping across with a plane:

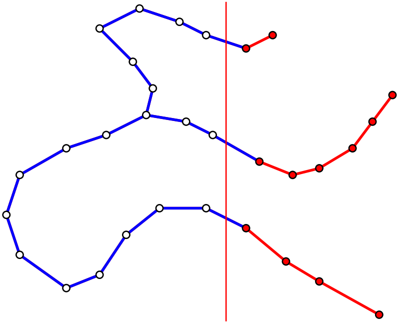


Sweep filtration

Goal: statistical analysis taking into account

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

Filter by sweeping across with a plane:

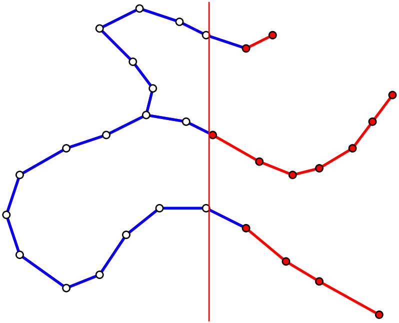


Sweep filtration

Goal: statistical analysis taking into account

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

Filter by sweeping across with a plane:

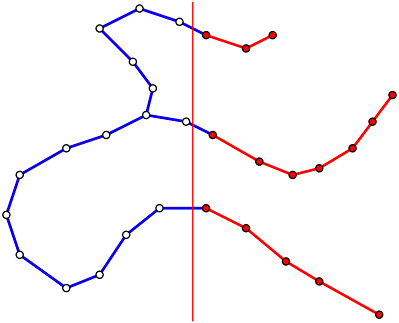


Sweep filtration

Goal: statistical analysis taking into account

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

Filter by sweeping across with a plane:

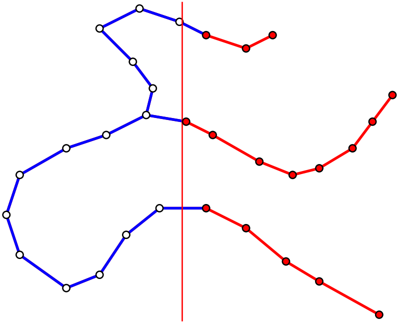


Sweep filtration

Goal: statistical analysis taking into account

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

Filter by sweeping across with a plane:

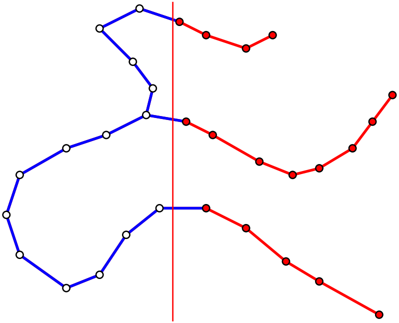


Sweep filtration

Goal: statistical analysis taking into account

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

Filter by sweeping across with a plane:

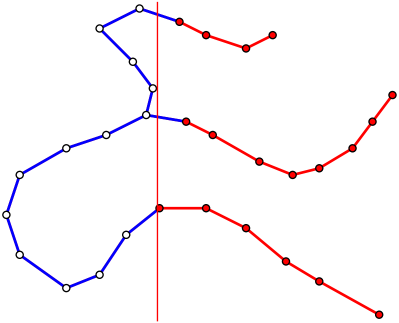


Sweep filtration

Goal: statistical analysis taking into account

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

Filter by sweeping across with a plane:

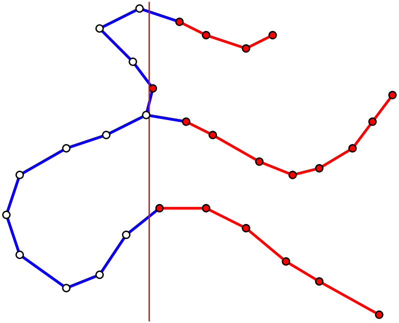


Sweep filtration

Goal: statistical analysis taking into account

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

Filter by sweeping across with a plane:

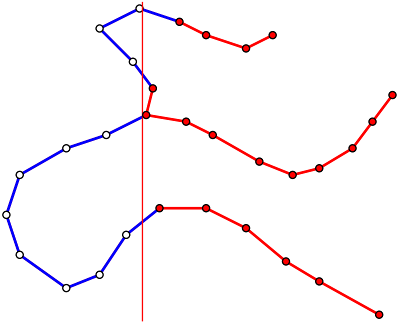


Sweep filtration

Goal: statistical analysis taking into account

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

Filter by sweeping across with a plane:

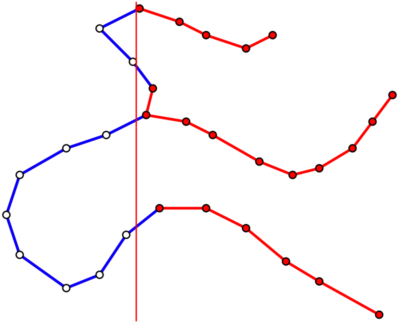


Sweep filtration

Goal: statistical analysis taking into account

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

Filter by sweeping across with a plane:

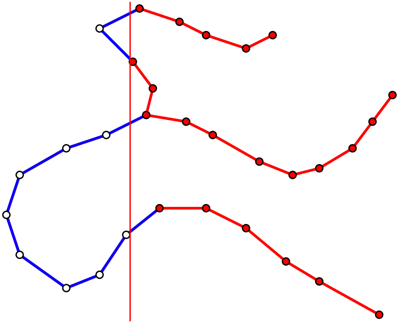


Sweep filtration

Goal: statistical analysis taking into account

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

Filter by sweeping across with a plane:

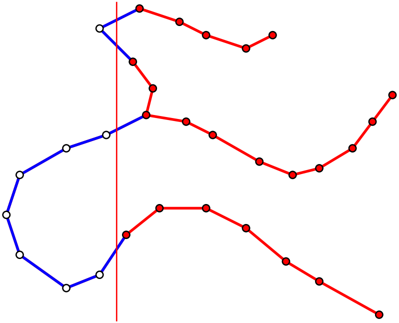


Sweep filtration

Goal: statistical analysis taking into account

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

Filter by sweeping across with a plane:

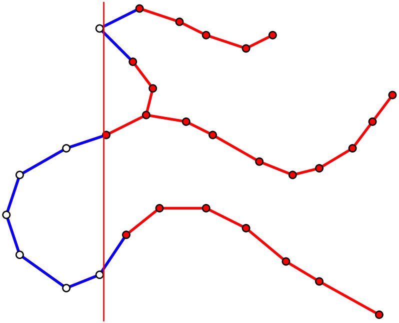


Sweep filtration

Goal: statistical analysis taking into account

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

Filter by sweeping across with a plane:

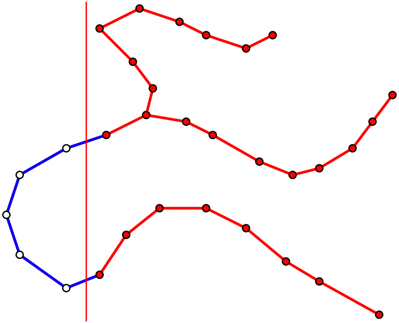


Sweep filtration

Goal: statistical analysis taking into account

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

Filter by sweeping across with a plane:

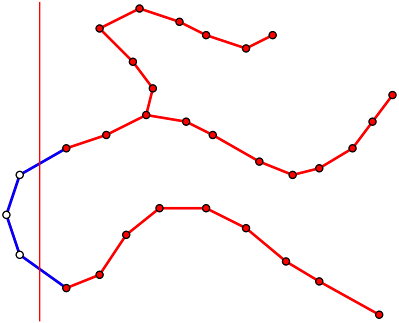


Sweep filtration

Goal: statistical analysis taking into account

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

Filter by sweeping across with a plane:

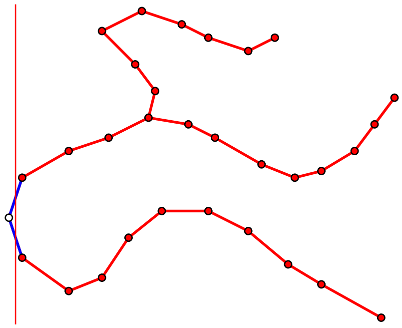


Sweep filtration

Goal: statistical analysis taking into account

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

Filter by sweeping across with a plane:

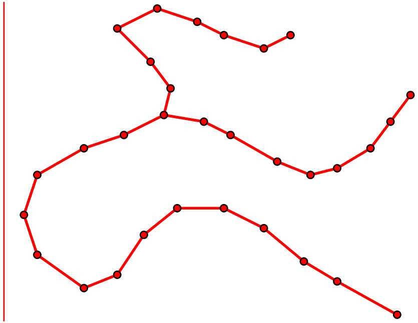


Sweep filtration

Goal: statistical analysis taking into account

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

Filter by sweeping across with a plane:

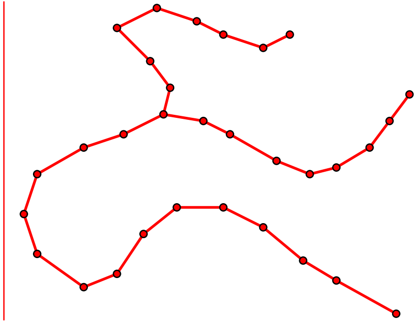


Sweep filtration

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:

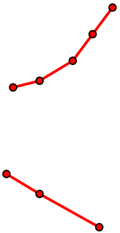
- birth time of each new component
- death of each component (when it joins to an older component)

Sweep filtration

Goal: statistical analysis taking into account

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

Filter by sweeping across with a plane:



Record:

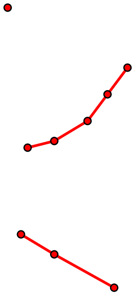
- birth time of each new component
- death of each component (when it joins to an older component)

Sweep filtration

Goal: statistical analysis taking into account

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

Filter by sweeping across with a plane:



Record:

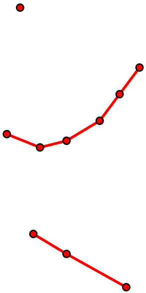
- birth time of each new component
- death of each component (when it joins to an older component)

Sweep filtration

Goal: statistical analysis taking into account

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

Filter by sweeping across with a plane:



Record:

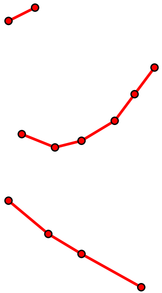
- birth time of each new component
- death of each component (when it joins to an older component)

Sweep filtration

Goal: statistical analysis taking into account

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

Filter by sweeping across with a plane:



Record:

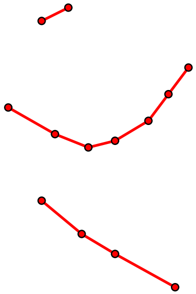
- birth time of each new component
- death of each component (when it joins to an older component)

Sweep filtration

Goal: statistical analysis taking into account

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

Filter by sweeping across with a plane:



Record:

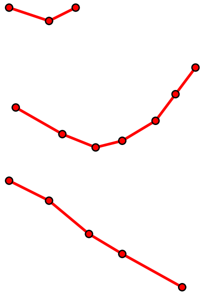
- birth time of each new component
- death of each component (when it joins to an older component)

Sweep filtration

Goal: statistical analysis taking into account

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

Filter by sweeping across with a plane:



Record:

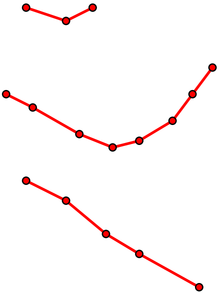
- birth time of each new component
- death of each component (when it joins to an older component)

Sweep filtration

Goal: statistical analysis taking into account

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

Filter by sweeping across with a plane:



Record:

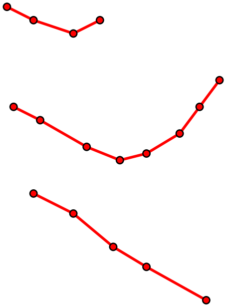
- birth time of each new component
- death of each component (when it joins to an older component)

Sweep filtration

Goal: statistical analysis taking into account

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

Filter by sweeping across with a plane:



Record:

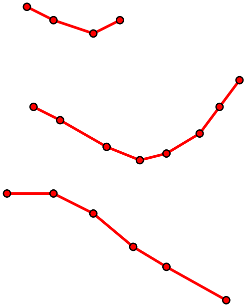
- birth time of each new component
- death of each component (when it joins to an older component)

Sweep filtration

Goal: statistical analysis taking into account

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

Filter by sweeping across with a plane:



Record:

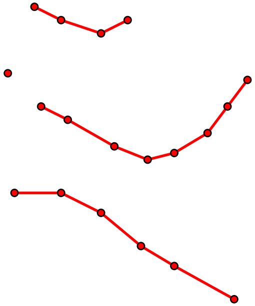
- birth time of each new component
- death of each component (when it joins to an older component)

Sweep filtration

Goal: statistical analysis taking into account

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

Filter by sweeping across with a plane:



Record:

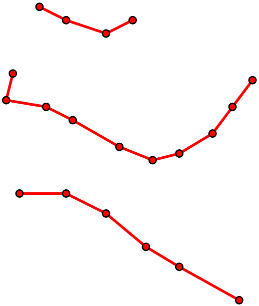
- birth time of each new component
- death of each component (when it joins to an older component)

Sweep filtration

Goal: statistical analysis taking into account

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

Filter by sweeping across with a plane:



Record:

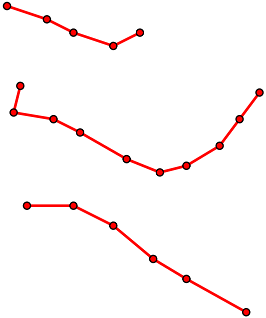
- birth time of each new component
- death of each component (when it joins to an older component)

Sweep filtration

Goal: statistical analysis taking into account

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

Filter by sweeping across with a plane:



Record:

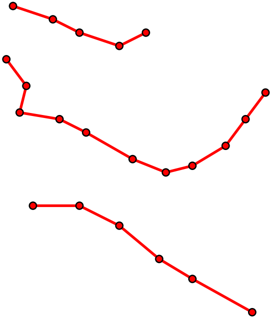
- birth time of each new component
- death of each component (when it joins to an older component)

Sweep filtration

Goal: statistical analysis taking into account

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

Filter by sweeping across with a plane:



Record:

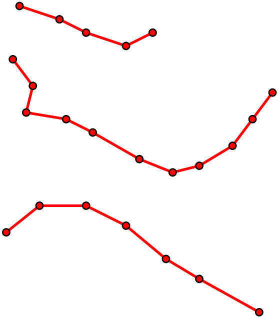
- birth time of each new component
- death of each component (when it joins to an older component)

Sweep filtration

Goal: statistical analysis taking into account

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

Filter by sweeping across with a plane:



Record:

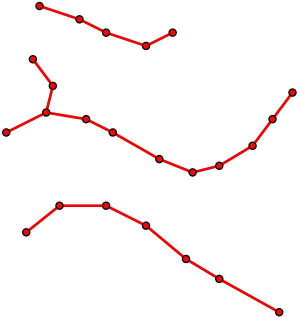
- birth time of each new component
- death of each component (when it joins to an older component)

Sweep filtration

Goal: statistical analysis taking into account

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

Filter by sweeping across with a plane:



Record:

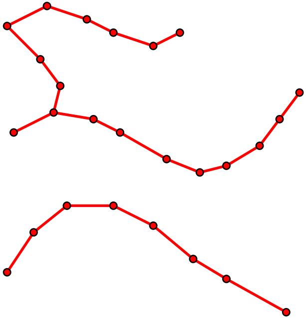
- birth time of each new component
- death of each component (when it joins to an older component)

Sweep filtration

Goal: statistical analysis taking into account

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

Filter by sweeping across with a plane:



Record:

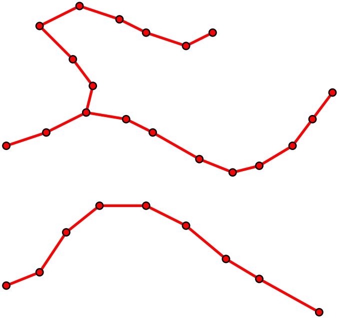
- birth time of each new component
- death of each component (when it joins to an older component)

Sweep filtration

Goal: statistical analysis taking into account

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

Filter by sweeping across with a plane:



Record:

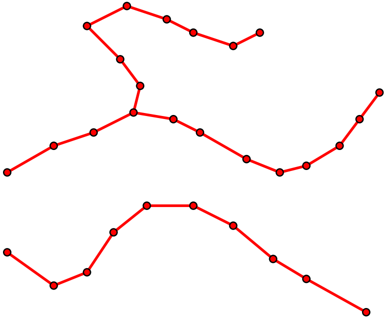
- birth time of each new component
- death of each component (when it joins to an older component)

Sweep filtration

Goal: statistical analysis taking into account

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

Filter by sweeping across with a plane:



Record:

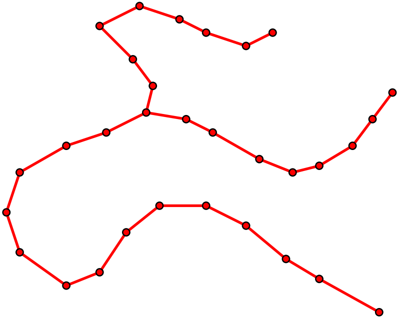
- birth time of each new component
- death of each component (when it joins to an older component)

Sweep filtration

Goal: statistical analysis taking into account

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

Filter by sweeping across with a plane:



Record:

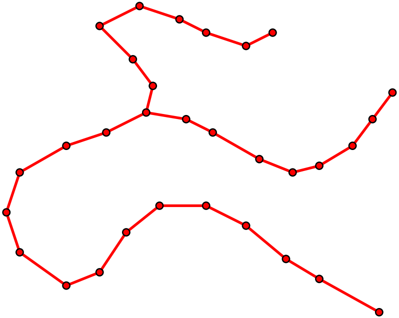
- birth time of each new component
- death of each component (when it joins to an older component)

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; dim of ambient space is irrelevant.

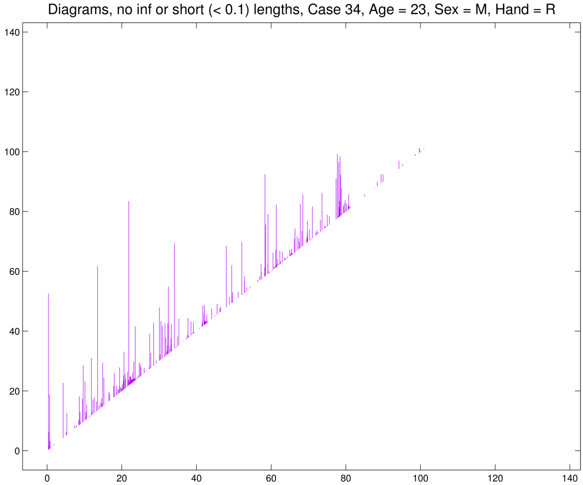
Bar codes

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

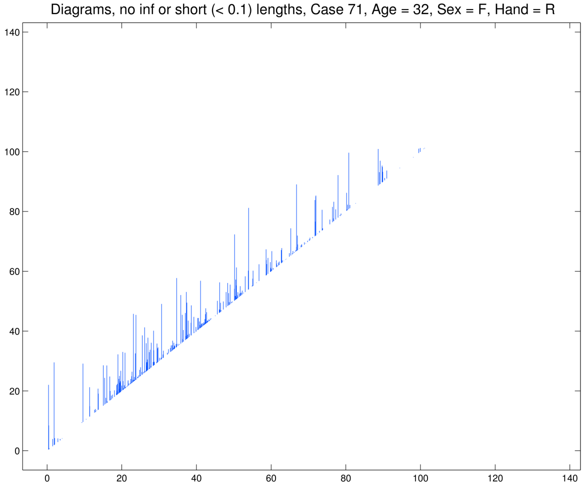
Data structure: 3D tree ⇝ bar code / lace array / persistence diagram:



- 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

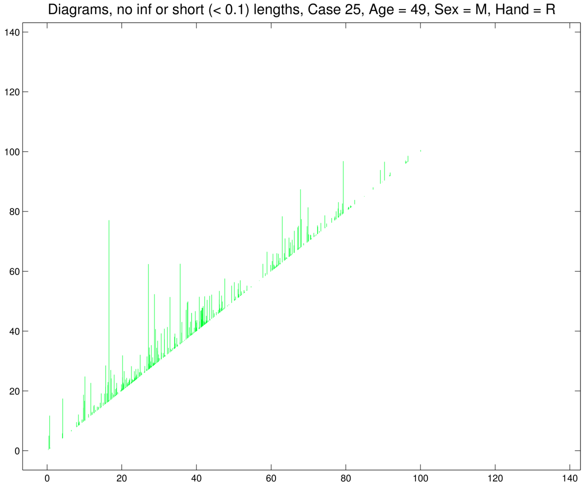
Data structure: 3D tree ⇝ bar code / lace array / persistence diagram:



- 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

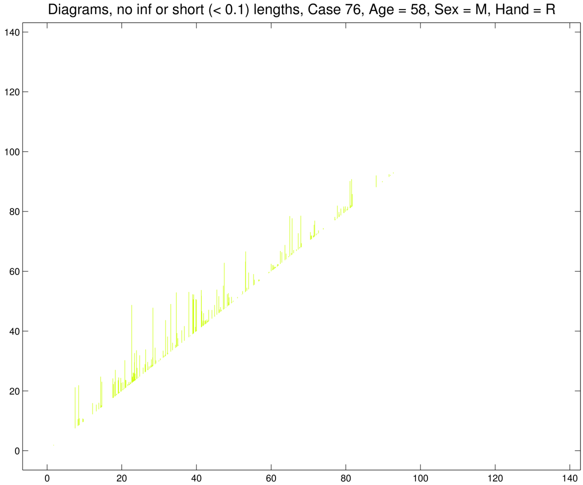
Data structure: 3D tree ⇝ bar code / lace array / persistence diagram:



- 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'.

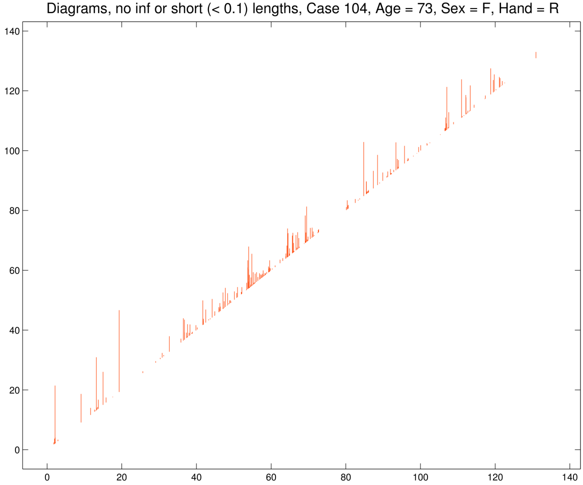
Bar codes

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'.

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.

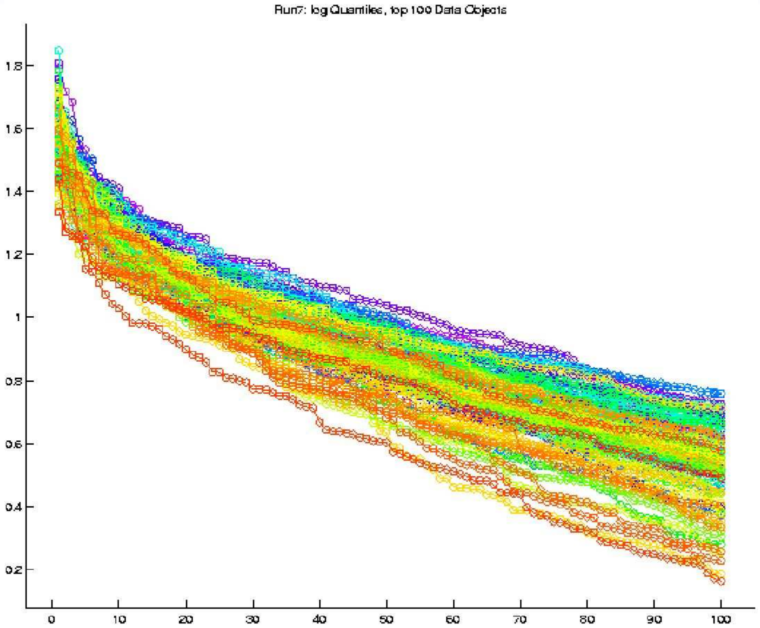
- Pearson correlation 0.52663
- p-value 3.0127×10^{-8} strongly significant

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.

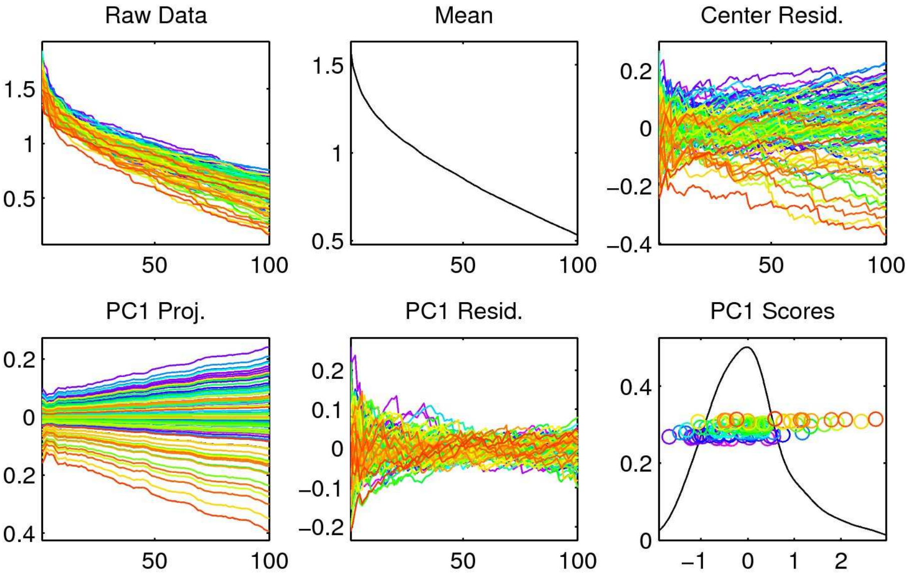
- Pearson correlation 0.52663
- p-value 3.0127×10^{-8} strongly significant

Remarks. Results essentially unchanged after

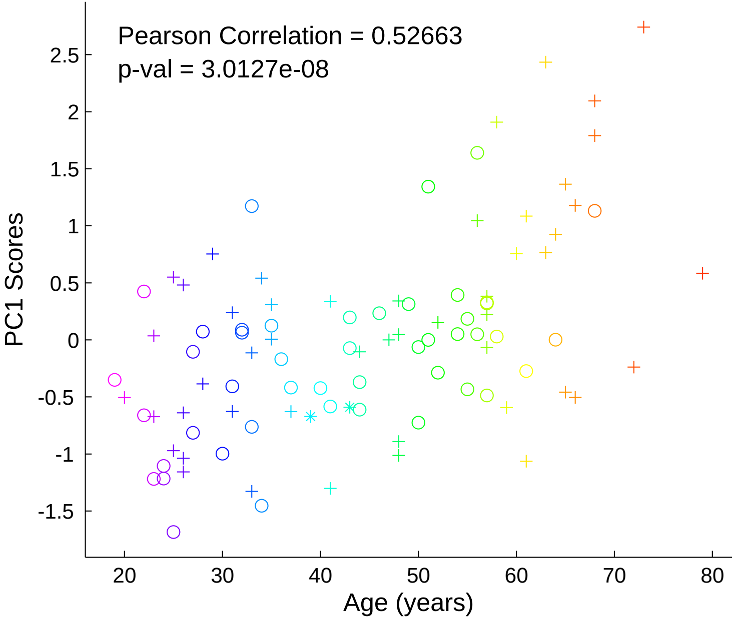
- rescaling to account for natural variation in overall brain size (force standard deviation of the set of bar lengths to equal 1)
- rescaling to account for known correlation of age vs. total vessel length L [Bullitt, et al. 2005] (divide by L, $\sqrt{L}$, or $\sqrt[3]{L}$)
- repeating the analysis 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.
  (Challenge problem: detect meaningful minute features.)
- 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. (Challenge problem: detect meaningful minute features.)
- 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.

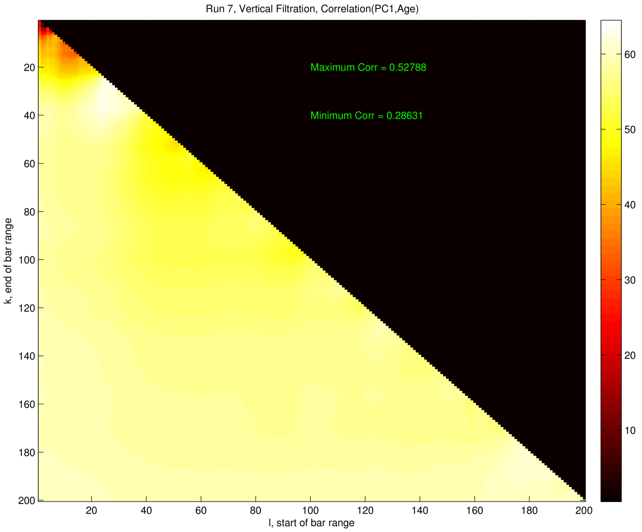
For brain artery trees.

- Not surprising that very short bars $\leftrightarrow$ noise,
  although in future studies they might not.
  (Challenge problem: detect meaningful minute features.)
- While biggest features are important,
- they hinder strength of correlation.

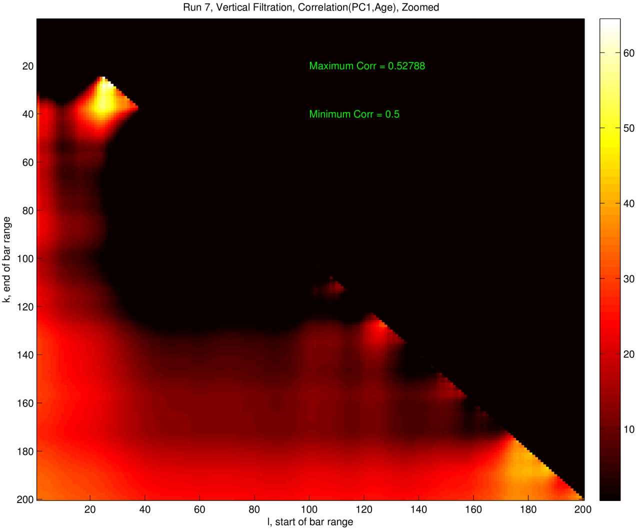
Lessons.

- 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.

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 $\sim .25$

Conclusion.

- 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 $\sim .25$

Conclusion.

- 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 $\sim .25$

Conclusion.

- 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 $\sim .25$

Conclusion.

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

Future directions

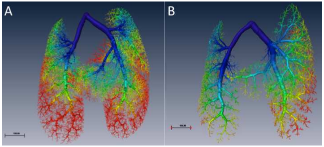
Lung vasculature. (with McLean et al., Bendich, Marron)

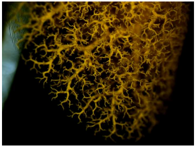
fMRI. (with Lazar et al.): classification using persistent homology

Fruit fly wings. (with Houle, Bendich, Cruz, A. Thomas):

Future directions

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



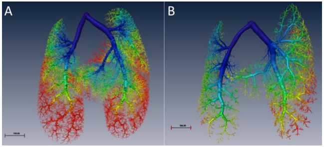


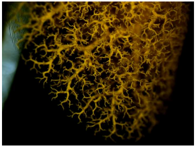
fMRI. (with Lazar et al.): classification using persistent homology

Fruit fly wings. (with Houle, Bendich, Cruz, A. Thomas):

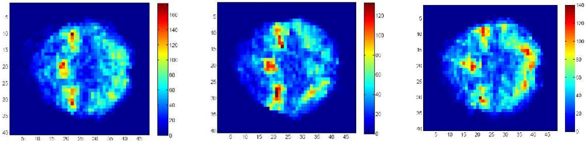
Future directions

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





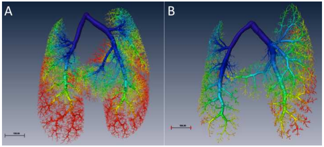
fMRI. (with Lazar et al.): classification using persistent homology

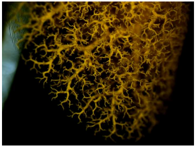


Fruit fly wings. (with Houle, Bendich, Cruz, A. Thomas):

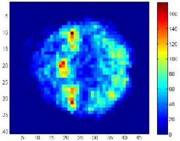
Future directions

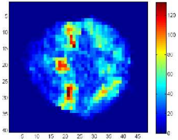
Lung vasculature. (with McLean et al., Bendich, Marron)

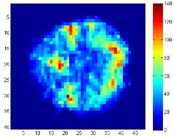




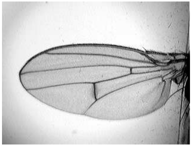
fMRI. (with Lazar et al.): classification using persistent homology

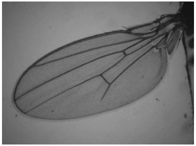


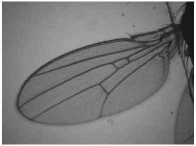




Fruit fly wings. (with Houle, Bendich, Cruz, A. Thomas):

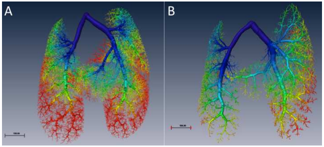


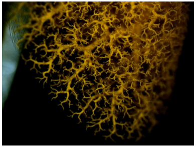




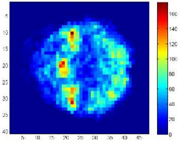
Future directions

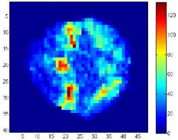
Lung vasculature. (with McLean et al., Bendich, Marron)

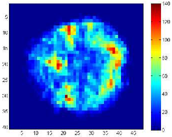




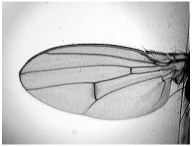
fMRI. (with Lazar et al.): classification using persistent homology

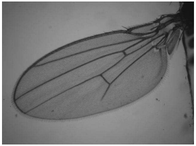


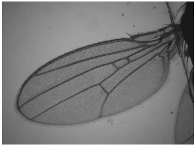




Fruit fly wings. (with Houle, Bendich, Cruz, A. Thomas):







Thank You