Topology for statistical analysis of brain artery images

Ezra Miller

Duke University, Department of Mathematics
ezra@math.duke.edu

joint with

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

Summer Workshop in Math (SWiM)
Duke University

24 June 2020
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:
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:
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:
Topological space $X \leadsto$ homology $H_iX$ for each dimension $i$.

- set of “$i$-dimensional holes” in $X$
Homology

Topological space $X \rightsquigarrow$ homology $H_iX$ for each dimension $i$.
- set of “$i$-dimensional holes” in $X$
Topological space $X \leadsto$ homology $H_iX$ for each dimension $i$.

- set of “$i$-dimensional holes” in $X$
Homology

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

- set of “$i$-dimensional holes” in $X$

\[ \# H_1 = 1 \]
Topological space $X \rightsquigarrow$ homology $H_i X$ for each dimension $i$.

- set of “$i$-dimensional holes” in $X$

$\# H_1 = 1$
Topological space $X \leadsto$ homology $H_i X$ for each dimension $i$.

- set of “$i$-dimensional holes” in $X$

\[
\begin{align*}
\# H_1 &= 1 \\
\# H_1 &= 0
\end{align*}
\]
Topological space $X \rightsquigarrow$ homology $H_i X$ for each dimension $i$.

- set of “$i$-dimensional holes” in $X$

\[
\begin{align*}
#H_1 &= 1 \\
#H_2 &= 1
\end{align*}
\]
Topological space $X \leadsto$ homology $H_iX$ for each dimension $i$.

- set of “$i$-dimensional holes” in $X$

\[
\begin{align*}
\# H_1 &= 1 \\
\# H_1 &= 0 \\
\# H_2 &= 1
\end{align*}
\]
Homology

Topological space $X \leadsto$ homology $H_i X$ for each dimension $i$.
- set of “$i$-dimensional holes” in $X$

$\# H_1 = 1$
$\# H_1 = 0$
$\# H_2 = 1$
Homology

Topological space $X \rightsquigarrow$ homology $H_i X$ for each dimension $i$.
- set of “$i$-dimensional holes” in $X$

$\# H_1 = 1$
$\# H_1 = 0$
$\# H_2 = 1$
Topological space $X \rightsquigarrow$ homology $H_i X$ for each dimension $i$.

- set of “$i$-dimensional holes” in $X$

\[
\begin{align*}
\#H_1 &= 1 \\
\#H_1 &= 0 \\
\#H_2 &= 1 \\
\#H_1 &= 2
\end{align*}
\]
Homology

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

- set of “$i$-dimensional holes” in $X$

\[
\#H_1 = \begin{cases} 1 & \text{for } i = 1 \\ 0 & \text{for } i = 2 \end{cases} \quad \#H_2 = \begin{cases} 0 & \text{for } i = 1 \\ 1 & \text{for } i = 2 \end{cases}
\]
Homology

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

- set of “$i$-dimensional holes” in $X$

$\# H_1 = 1$  \hspace{1cm}  $\# H_1 = 0$  \hspace{1cm}  $\# H_1 = 2$

$\# H_2 = 1$  \hspace{1cm}  $\# H_2 = 1$

- $i = 0$ case: $H_0$ is the set of connected components of $X$
Topological space $X \rightsquigarrow$ homology $H_i X$ for each dimension $i$.

- set of “$i$-dimensional holes” in $X$

\[-\begin{array}{l}
\#H_0 = 1 \\
\#H_1 = 1 \\
\#H_2 = 1 \\
\#H_0 = 1 \\
\#H_1 = 0 \\
\#H_2 = 1 \\
\#H_0 = 1 \\
\#H_1 = 2 \\
\#H_2 = 1
\end{array}\]

- $i = 0$ case: $H_0$ is the set of connected components of $X$
Homology

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

- set of “$i$-dimensional holes” in $X$

![Diagram showing $H_0$, $H_1$, and $H_2$ for different topological spaces.]

$\# H_0 = 1$
$\# H_1 = 1$
$\# H_2 = 1$
$\# H_0 = 1$
$\# H_1 = 0$
$\# H_2 = 1$
$\# H_0 = 1$
$\# H_1 = 2$
$\# H_2 = 1$

- $i = 0$ case: $H_0$ is the set of connected components of $X$

Formally: homology $H_i X$ is a vector space for each dimension $i$. 
Persistent homology

Build $X$ step by step

- measure evolving topology.

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

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

**Example:** Given a function $f : X \to \mathbb{R}$, let $X_t = \{ x \in X \mid f(x) \leq t \}$. Good choice of $t_0, \ldots, t_m \in \mathbb{R}$: the values of $t$ across which $H_i X_t$ changes.

**History.** invented by [Frosini, Landi 1999], [Robins 1999], [Edelsbrunner, Letscher, Zomorodian 2002]: includes efficient computation; [Carlsson, Scolamiero, Turner, many others]: additional theory, applications
Persistent homology

Build X step by step

- measure evolving topology.

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

- The persistent homology of this filtration is $H_i X_1 \to H_i X_2 \to \cdots \to H_i X_m$, a sequence of sets.

- A feature persists from $j$ to $k$ if it appears first in $H_i X_j$ and last in $H_i X_k$.

Example: Given a function $f : X \to \mathbb{R}$, let $X_t = \{ x \in X \mid f(x) \leq t \}$. Good choice of $t_0, \ldots, t_m \in \mathbb{R}$: the values of $t$ across which $H_i X_t$ changes.

History. invented by [Frosini, Landi 1999], [Robins 1999], [Edelsbrunner, Letscher, Zomorodian 2002]: includes efficient computation; [Carlsson, Scolamiero, Turner, many others]: additional theory, applications
Example: expanding balls
Example: expanding balls
Example: expanding balls
Example: expanding balls
Example: expanding balls
Example: expanding balls
Example: expanding balls
Example: expanding balls
Example: expanding balls
Example: expanding balls
Example: expanding balls
Example: expanding balls
Example: expanding balls
Example: expanding balls
Example: expanding balls
Example: expanding balls
Example: expanding balls
Example: expanding balls
Example: expanding balls
Example: expanding balls
Example: expanding balls
Example: expanding balls
Persistent homology

Build $X$ step by step

- measure evolving topology.

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

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

**Example:** Given a function $f : X \rightarrow \mathbb{R}$, let $X_t = \{ x \in X \mid f(x) \leq t \}$. Good choice of $t_0, \ldots, t_m \in \mathbb{R}$: the values of $t$ across which $H_i X_t$ changes.

**History.** invented by [Frosini, Landi 1999], [Robins 1999], [Edelsbrunner, Letscher, Zomorodian 2002]: includes efficient computation; [Carlsson, Scolamiero, Turner, many others]: additional theory, applications
Persistent homology

Build $X$ step by step

- measure evolving topology.

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

- The **persistent homology** of this filtration is $H_iX_1 \to H_iX_2 \to \cdots \to H_iX_m$, a sequence of sets.
- A feature **persists** from $j$ to $k$ if it appears first in $H_iX_j$ and last in $H_iX_k$.

**Example:** Given a function $f: X \to \mathbb{R}$, let $X_t = \{x \in X \mid f(x) \leq t\}$. Good choice of $t_0, \ldots, t_m \in \mathbb{R}$: the values of $t$ across which $H_iX_t$ changes.

**History.** invented by [Frosini, Landi 1999], [Robins 1999], [Edelsbrunner, Letscher, Zomorodian 2002]: includes efficient computation; [Carlsson, Scolamiero, Turner, many others]: additional theory, applications
Persistent homology

Build $X$ step by step

- measure evolving topology.

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

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

Example: Given a function $f : X \rightarrow \mathbb{R}$, let $X_t = \{ x \in X \mid f(x) \leq t \}$. Good choice of $t_0, \ldots, t_m \in \mathbb{R}$: the values of $t$ across which $H_i X_t$ changes.

History. invented by [Frosini, Landi 1999], [Robins 1999], [Edelsbrunner, Letscher, Zomorodian 2002]: includes efficient computation; [Carlsson, Scolamiero, Turner, many others]: additional theory, applications
Example: expanding balls
Example: expanding balls

$\#H_0 = 31$
Example: expanding balls

\[ \# H_0 = 31 \]
Example: expanding balls

\[ \#H_0 = 31 \]
Example: expanding balls

\[ \#H_0 = 26 \]
Example: expanding balls

$\#H_0 = 21$
Example: expanding balls

\[ \#H_0 = 12 \]
Example: expanding balls

\[ \# H_0 = 6 \]
Example: expanding balls

\[ \#H_0 = 2 \]
Example: expanding balls

\[ \#H_0 = 2 \]
Example: expanding balls

\[ \#H_0 = 1 \quad \#H_1 = 2 \]
Example: expanding balls

$\#H_0 = 1 \quad \#H_1 = 1$
Example: expanding balls

\[ \#H_0 = 1 \quad \#H_1 = 1 \]
Example: expanding balls

\[ \#H_0 = 1 \quad \#H_1 = 3 \]
Example: expanding balls

\[ \#H_0 = 1 \quad \#H_1 = 1 \]
Example: expanding balls

\[ \#H_0 = 1 \quad \#H_1 = 1 \]
Example: expanding balls

\[ \#H_0 = 1 \quad \#H_1 = 1 \]
Example: expanding balls

\#H_0 = 1 \quad \#H_1 = 1
Example: expanding balls

\#H_0 = 1 \quad \#H_1 = 0
Example: expanding balls

\[ \#H_0 = 1 \quad \#H_1 = 1 \]
Example: expanding balls

\[ \#H_0 = 1 \quad \#H_1 = 0 \]
Persistent homology

Build $X$ step by step

• measure evolving topology.

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

• The persistent homology of this filtration is $H_i X_1 \to H_i X_2 \to \cdots \to H_i X_m$, a sequence of sets.

• A feature persists from $j$ to $k$ if it appears first in $H_i X_j$ and last in $H_i X_k$.

Example: Given a function $f : X \to \mathbb{R}$, let $X_t = \{ x \in X \mid f(x) \leq t \}$. Good choice of $t_0, \ldots, t_m \in \mathbb{R}$: the values of $t$ across which $H_i X_t$ changes.

History. invented by [Frosini, Landi 1999], [Robins 1999], [Edelsbrunner, Letscher, Zomorodian 2002]: includes efficient computation; [Carlsson, Scolamiero, Turner, many others]: additional theory, applications
Persistent homology

Build X step by step

- measure evolving topology.

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

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

**Example:** Given a function $f : X \to \mathbb{R}$, let $X_t = \{ x \in X \mid f(x) \leq t \}$. Good choice of $t_0, \ldots, t_m \in \mathbb{R}$: the values of $t$ across which $H_i X_t$ changes.

**History.** invented by [Frosini, Landi 1999], [Robins 1999], [Edelsbrunner, Letscher, Zomorodian 2002]: includes efficient computation; [Carlsson, Scolamiero, Turner, many others]: additional theory, applications.
Example: filling brains
Example: filling brains
Example: filling brains
Example: filling brains
Example: filling brains
Example: filling brains
Example: filling brains
Example: filling brains
Example: filling brains
Example: filling brains
Example: filling brains
Example: filling brains
Example: filling brains
Example: filling brains
Example: filling brains
Example: filling brains
Example: filling brains
Example: filling brains
Example: filling brains
Example: filling brains
Persistent homology

Build $X$ step by step

- measure evolving topology.

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

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

**Example**: Given a function $f : X \to \mathbb{R}$, let $X_t = \{ x \in X \mid f(x) \leq t \}$. Good choice of $t_0, \ldots, t_m \in \mathbb{R}$: the values of $t$ across which $H_i X_t$ changes.

**History.** invented by [Frosini, Landi 1999], [Robins 1999], [Edelsbrunner, Letscher, Zomorodian 2002]: includes efficient computation; [Carlsson, Scolamiero, Turner, many others]: additional theory, applications
Persistent homology

Build $X$ step by step

- measure evolving topology.

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

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

**Example:** Given a function $f : X \to \mathbb{R}$, let $X_t = \{ x \in X \mid f(x) \leq t \}$. Good choice of $t_0, \ldots, t_m \in \mathbb{R}$: the values of $t$ across which $H_i X_t$ changes.

**History.** invented by [Frosini, Landi 1999], [Robins 1999], [Edelsbrunner, Letscher, Zomorodian 2002]: includes efficient computation; [Carlsson, Scolamiero, Turner, many others]: additional theory, applications
Bar codes

Data structure: 3D tree $\leadsto$ bar code

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

Data structure: 3D tree $\mapsto$ bar code

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

Data structure: 3D tree \(\leadsto\) bar code

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

Data structure: 3D tree $\leadsto$ bar code

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

Diagrams, no inf or short ($< 0.1$) lengths, Case 25, Age = 49, Sex = M, Hand = R
Bar codes

Data structure: 3D tree $\leadsto$ bar code

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

Data structure: 3D tree $\leadsto$ bar code

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

Data structure: 3D tree $\rightsquigarrow$ bar code

- multiset of (vertical) line segments $[t, t']$ (plotted at $x$-coordinate $t$)
- one for each feature 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:
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:
**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 2016]
Longest bars in older brains tend to be shorter and later.
- Pearson correlation $0.52663$: linear correlation markedly present
- $p$-value $3.0127 \times 10^{-8}$ strongly significant

Similar results after accounting for
- natural variation in overall brain size or
- known correlation of age vs. total vessel length [Bullitt, et al. 2005].

Moral. Persistent homology can topologically detect statistically significant geometric motifs.
Top 100 bars

Run1: Quantiles, top 100 Data Objects
**Statistical analysis**

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

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

**Conclusions [Bendich, Marron, M.—, Pieloch, Skwerer 2016]**

Longest bars in older brains tend to be shorter and later.

- Pearson correlation 0.52663: linear correlation markedly present
- $p$-value $3.0127 \times 10^{-8}$ strongly significant

Similar results after accounting for

- natural variation in overall brain size or
- known correlation of age vs. total vessel length [Bullitt, et al. 2005].

**Moral.** Persistent homology can topologically detect statistically significant geometric motifs.
Statistical analysis

Reduce to linear methods. 3D tree $\leadsto$ bar code $\leadsto$ vector in $\mathbb{R}^{100}$:
- top 100 bar lengths, in decreasing order, log scale
- correlate first principal component score vs. age

Conclusions [Bendich, Marron, M.—, Pieloch, Skwerer 2016]
Longest bars in older brains tend to be shorter and later.
- Pearson correlation 0.52663: linear correlation markedly present
- $p$-value $3.0127 \times 10^{-8}$ strongly significant

Similar results after accounting for
- natural variation in overall brain size or
- known correlation of age vs. total vessel length [Bullitt, et al. 2005].

Moral. Persistent homology can topologically detect statistically significant geometric motifs.
Top 100 bars: log scale
Statistical analysis

Reduce to linear methods. 3D tree $\leadsto$ bar code $\leadsto$ vector in $\mathbb{R}^{100}$:
- top 100 bar lengths, in decreasing order, log scale
- correlate first principal component score vs. age

Conclusions [Bendich, Marron, M.—, Pieloch, Skwerer 2016]
Longest bars in older brains tend to be shorter and later.
- Pearson correlation $0.52663$: linear correlation markedly present
  - $p$-value $3.0127 \times 10^{-8}$ strongly significant

Similar results after accounting for
- natural variation in overall brain size or
- known correlation of age vs. total vessel length [Bullitt, et al. 2005].

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

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

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

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

- Pearson correlation 0.52663: linear correlation markedly present
- $p$-value $3.0127 \times 10^{-8}$ strongly significant

Similar results after accounting for
- natural variation in overall brain size or
- known correlation of age vs. total vessel length [Bullitt, et al. 2005].

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

Reduce to linear methods. 3D tree $\rightsquigarrow$ bar code $\rightsquigarrow$ vector in $\mathbb{R}^{100}$:
- top 100 bar lengths, in decreasing order, log scale
- correlate first principal component score vs. age

Conclusions [Bendich, Marron, M.—, Pieloch, Skwerer 2016]
Longest bars in older brains tend to be shorter and later.
- Pearson correlation 0.52663: linear correlation markedly present
- $p$-value $3.0127 \times 10^{-8}$ strongly significant

Similar results after accounting for
- natural variation in overall brain size or
- known correlation of age vs. total vessel length [Bullitt, et al. 2005].

Moral. Persistent homology can topologically detect statistically significant geometric motifs.
Age vs. PC1

Raw Data

Mean

Center Resid.

PC1 Proj.

PC1 Resid.

PC1 Scores

Artery trees
Homology
Persistence
Bar codes
Statistical analysis
Age vs. PC1

Pearson Correlation = 0.52663
p-val = 3.0127e−08
Statistical analysis

Reduce to linear methods. 3D tree $\rightsquigarrow$ bar code $\rightsquigarrow$ vector in $\mathbb{R}^{100}$:
- top 100 bar lengths, in decreasing order, log scale
- correlate first principal component score vs. age

Conclusions [Bendich, Marron, M.—, Pieloch, Skwerer 2016]
Longest bars in older brains tend to be shorter and later.
- Pearson correlation 0.52663: linear correlation markedly present
- $p$-value $3.0127 \times 10^{-8}$ strongly significant

Similar results after accounting for
- natural variation in overall brain size or
- known correlation of age vs. total vessel length [Bullitt, et al. 2005].

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

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

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

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

- Pearson correlation 0.52663: linear correlation markedly present
- $p$-value $3.0127 \times 10^{-8}$ strongly significant

Similar results after accounting for
- natural variation in overall brain size or
- known correlation of age vs. total vessel length [Bullitt, et al. 2005].

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

Reduce to linear methods. 3D tree $\Rightarrow$ bar code $\Rightarrow$ vector in $\mathbb{R}^{100}$:
- top 100 bar lengths, in decreasing order, log scale
- correlate first principal component score vs. age

Conclusions [Bendich, Marron, M.—, Pieloch, Skwerer 2016]
Longest bars in older brains tend to be shorter and later.
- Pearson correlation $0.52663$: linear correlation markedly present
- $p$-value $3.0127 \times 10^{-8}$ strongly significant

Similar results after accounting for
- natural variation in overall brain size or
- known correlation of age vs. total vessel length [Bullitt, et al. 2005].

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

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

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

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

- **Pearson correlation 0.52663**: linear correlation markedly present
- **$p$-value $3.0127 \times 10^{-8}$ strongly significant**

Similar results after accounting for
- natural variation in overall brain size or
- known correlation of age vs. total vessel length [Bullitt, et al. 2005].

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

Reduce to linear methods. 3D tree $\rightsquigarrow$ bar code $\rightsquigarrow$ vector in $\mathbb{R}^{100}$:
- top 100 bar lengths, in decreasing order, log scale
- correlate first principal component score vs. age

Conclusions [Bendich, Marron, M.—, Pieloch, Skwerer 2016]
Longest bars in older brains tend to be shorter and later.
- Pearson correlation $0.52663$: linear correlation markedly present
- $p$-value $3.0127 \times 10^{-8}$ strongly significant

Similar results after accounting for
- natural variation in overall brain size or
- known correlation of age vs. total vessel length [Bullitt, et al. 2005].

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

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

Reduce to linear methods. 3D tree $\rightsquigarrow$ bar code $\rightsquigarrow$ vector in $\mathbb{R}^{100}$:
- top 100 bar lengths, in decreasing order, log scale
- correlate first principal component score vs. age

Conclusions [Bendich, Marron, M.—, Pieloch, Skwerer 2016]
Longest bars in older brains tend to be shorter and later.
- Pearson correlation 0.52663: linear correlation markedly present
- $p$-value $3.0127 \times 10^{-8}$ strongly significant

Similar results after accounting for
- natural variation in overall brain size or
- known correlation of age vs. total vessel length [Bullitt, et al. 2005].

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

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

Thank You