On Inverse Problems in TDA

— joint work with E. Solomon (Brown Math.) — arXiv: 1712.03630
The preimage problem in the data Sciences

Data

Features

(feature design or learning)

bag of words, word2vec
shape contexts, heat kernels
node2vec, Laplacian fact., rand. walks
dim. reduction, auto-encoders, etc.
The preimage problem in the data Sciences

Data

Features

Can the feature map be inverted?
- Right inverse (∃ preimage): interpretable AI
- Left inverse (∃! preimage): reliable interpretation

Scenarios: dictionaries, deep layers, stats, etc.

dim. reduction, auto-encoders, etc.
Can the feature map be inverted?

- Right inverse (\( \exists \) preimage): interpretable AI
- Left inverse (\( \exists! \) preimage): reliable interpretation

Scenarios: dictionaries, deep layers, stats, etc.
The preimage problem in the data Sciences

Can the feature map be inverted?
- Right inverse (∃ preimage): interpretable AI
- Left inverse (∃! preimage): reliable interpretation

Scenarios: dictionaries, deep layers, stats, etc.
TDA and the preimage problem
TDA and the preimage problem

Model \xrightarrow{\text{(inference)}} \text{Descriptor} \xrightarrow{\text{(decoding)}} \text{Data} \xrightarrow{\text{(sampling)}} \text{Model}

Mathematical abstraction:
- compact metric spaces modulo isometries
- Gromov-Hausdorff distance

Mathematical abstraction:
- compact metric spaces modulo isometries
- Gromov-Hausdorff distance

- compact metric space

Lipschitz operator

right inverse: realize barcode as the PH of some isom. class
left inverse: characterize isom. class uniquely
TDA and the preimage problem

Mathematical abstraction:
- compact metric spaces modulo isometries
- Gromov-Hausdorff distance

Model \[\xrightarrow{(inference)}\] Model

(left inverse: characterize isom. class uniquely)

Data \[\xrightarrow{(decoding)}\] Descriptor (TDA)

(right inverse: realize barcode as the PH of some isom. class)

Lipschitz operator

compact metric space
Fact: [Folklore] Any (graded) persistence module $\mathbb{R} \to \text{vect}_k$ can be realized as the (graded) $\tilde{P}H$ of a piecewise-constant function on a bouquet of spheres.
Right inverses for TDA

**Fact:** [Folklore] Any (graded) persistence module $\mathbb{R} \rightarrow \text{vect}_k$ can be realized as the (graded) $\tilde{P}H$ of a piecewise-constant function on a bouquet of spheres.

**Thm:** [Curry, Reiss] [Botnan, Fluhr]
Any (graded) barcode can be realized as the level-set $\tilde{P}H$ of some stratified map on some stratified space.
Right inverses (local) for TDA

\[ u \in \mathbb{R}^{nd} \rightarrow \mathbb{R}^{2^n-1} \]

**Thm:** [Gameiro, Hiraoka, Obayashi]

(i) *Generic* point cloud $\rightarrow \exists \Omega \ni u \in \mathbb{R}^{nd}$ over which the correspondence $u \mapsto v$ can be extended to a map $f : \Omega \rightarrow \mathbb{R}^{2^n-1}$ computing persistence barcodes.

(ii) For $\Omega$ small enough, $f$ is of class $C^\infty$.

Observation: pairing given by order of distances is constant in small enough $O$. 
Right inverses (local) for TDA

Thm: [Gameiro, Hiraoka, Obayashi]

(i) Generic point cloud $\Rightarrow \exists \Omega \ni u$ in $\mathbb{R}^{nd}$ over which the correspondence $u \mapsto v$ can be extended to a map $f : \Omega \rightarrow \mathbb{R}^{2^n - 1}$ computing persistence barcodes.

(ii) For $\Omega$ small enough, $f$ is of class $C^\infty$.

→ adapt Newton-Raphson continuation method to build right inverse of $f$ in $f(\Omega)$ (Jacobian matrix of $f$ can be singular $\Rightarrow$ use pseudo-inverse)
Left inverses?

- Unions of (open) balls — Čech/Rips/Delaunay filtrations

\[
\text{dgm} \mathcal{C}(P, \ell_2) = \{(0, +\infty)\} \sqcup \{(0, \frac{1}{2})\} \sqcup \{(0, \frac{1}{2})\}
\]

\[
\text{dgm} \mathcal{R}(P, \ell_2) = \{(0, +\infty)\} \sqcup \{(0, 1)\} \sqcup \{(0, 1)\}
\]

⇒ diagrams for different values of \(\alpha\) are indistinguishable
Left inverses?

- Unions of (open) balls — Čech/Rips/Delaunay filtrations

**Prop:** [Folklore]
For any metric tree $(X, d_X)$:

$$dgm \mathcal{R}(X, d_X) = dgm \mathcal{C}(X, d_X) = \{(0, +\infty)\}$$

$\Rightarrow$ no information on the metric

$X$ is 0-hyperbolic

$\Rightarrow$ metric balls are convex

$\Rightarrow$ geodesic triangles are tripods
Left inverses?

- Unions of (open) balls — Čech/Rips/Delaunay filtrations
- Reeb graphs

⇒ Reeb graphs are indistinguishable from their diagrams
Left inverses?

- Unions of (open) balls — Čech/Rips/Delaunay filtrations
- Reeb graphs
- Real-valued functions

**Prop:** [Folklore]
Given \( f : X \to \mathbb{R} \) and \( h : Y \to X \) homeomorphism,

\[
dgm f \circ h = dgm f
\]

Too large a group of transformations...
Left inverses?

- Unions of (open) balls — Čech/Rips/Delaunay filtrations
- Reeb graphs
- Real-valued functions

possible solutions:

- richer topological invariants (e.g. persistent homotopy)
- use multiple filter functions (aggregation vs multipersistence)
Persistent Homology Transform (PHT)

\[ (X, d_X) \text{ (compact)} \]

\[ \mathcal{F} = \{ f_w \}_{w \in W} \]

\[ \mathbb{R} \]

\[ \text{PHT}(X) = \{ \text{dgm } f_w \mid w \in W \} \]

(diagrams, \( d_b \))
**Thm:** [Boyer, Curry, Mukherjee, Turner 2014, 2018]  
[Ghrist, Levanger, Mai 2018]  
Let $\mathcal{F} = \{\langle \cdot, w \rangle \}_{w \in \mathbb{S}^{d-1}}$, where $d$ is fixed. Then, PHT is injective on the class of semialgebraic sets in $\mathbb{R}^d$.  

**Still true** for a fixed finite set of directions (of size exponential in $d$). [Curry, Mukherjee, Turner]
**Thm:** [Boyer, Curry, Mukherjee, Turner 2014, 2018]  
[Ghrist, Levanger, Mai 2018]

Let $\mathcal{F} = \{\langle \cdot, w \rangle \}_{w \in \mathbb{S}^{d-1}}$, where $d$ is fixed. Then, PHT is injective on the class of semialgebraic sets in $\mathbb{R}^d$.

**Corollary:** PHT is a **sufficient statistic** for such sets  
$\Rightarrow$ parametric inference
Given a compact length space \((X, d_X)\), take \(\mathcal{F} = \{d_X(\cdot, x)\}_{x \in X}\)
Given a compact length space \((X, d_X)\), take \(\mathcal{F} = \{d_X(\cdot, x)\}_{x \in X}\)

**Thm (local stability):** [Carrière, O., Ovsjanikov 2015]

Let \((X, d_X)\) and \((Y, d_Y)\) be compact length spaces with positive convexity radius \((\rho(X), \rho(Y) > 0)\). Let \(x \in X\) and \(y \in Y\). If \(d_{GH}((X, x), (Y, y)) \leq \frac{1}{20} \min\{\rho(X), \rho(Y)\}\), then

\[
d_b(dgm d_X(\cdot, x), dgm d_Y(\cdot, y)) \leq 20 d_{GH}((X, x), (Y, y)).
\]
Given a compact length space \((X, d_X)\), take \(\mathcal{F} = \{d_X(\cdot, x)\}_{x \in X}\)

**Corollary (local stability of PHT):**
Let \((X, d_X)\) and \((Y, d_Y)\) be compact length spaces with positive convexity radius \((\varrho(X), \varrho(Y) > 0)\).
If \(d_{GH}(X, Y) \leq \frac{1}{20} \min\{\varrho(X), \varrho(Y)\}\), then
\[
d_H(\text{PHT}(X), \text{PHT}(Y)) \leq 20 d_{GH}((X, x), (Y, y)).
\]

**Thm (local stability):** [Carrière, O., Ovsjanikov 2015]
Let \((X, d_X)\) and \((Y, d_Y)\) be compact length spaces with positive convexity radius \((\varrho(X), \varrho(Y) > 0)\). Let \(x \in X\) and \(y \in Y\).
If \(d_{GH}((X, x), (Y, y)) \leq \frac{1}{20} \min\{\varrho(X), \varrho(Y)\}\), then
\[
d_b(\text{dgm } d_X(\cdot, x), \text{dgm } d_Y(\cdot, y)) \leq 20 d_{GH}((X, x), (Y, y)).
\]
Given a compact length space \((X, d_X)\), take \(\mathcal{F} = \{d_X(\cdot, x)\}_{x \in X}\).

**Corollary (local stability of PHT):**

Let \((X, d_X)\) and \((Y, d_Y)\) be compact length spaces with positive convexity radius \((\varrho(X), \varrho(Y) > 0)\).

If \(d_{GH}(X, Y) \leq \frac{1}{20} \min\{\varrho(X), \varrho(Y)\}\), then

\[
d_H(\text{PHT}(X), \text{PHT}(Y)) \leq 20 d_{GH}((X, x), (Y, y)).
\]
Focus: compact metric graphs (1-dimensional stratified length spaces)

PHT: \( \mathcal{F} = \{d_X(\cdot, x)\}_{x \in X} \), \( \text{dgm} = \text{extended persistence diagram} \)

**Thm (global stability):** [Dey, Shi, Wang 2015]
For any compact metric graphs \( X, Y \),
\[
d_H(\text{PHT}(X), \text{PHT}(Y)) \leq 18 d_{GH}(X, Y).
\]

**Thm (density):** [Gromov]
Compact metric graphs are GH-dense among the compact length spaces.
PHT for metric graphs

Focus: compact metric graphs (1-dimensional stratified length spaces)

PHT: $\mathcal{F} = \{d_X(\cdot, x)\}_{x \in X}$, dgm = extended persistence diagram

Thm (global stability): [Dey, Shi, Wang 2015]
For any compact metric graphs $X, Y$,

$$d_H(\text{PHT}(X), \text{PHT}(Y)) \leq 18 d_{\text{GH}}(X, Y).$$

Thm (density): [Gromov]
Compact metric graphs are GH-dense among the compact length spaces.

Q: injectivity of PHT on metric graphs?
Bad news: PHT is not injective on all compact metric graphs.

\[ \text{PHT}(X) = \text{PHT}(Y) \text{ while } X \not\simeq Y \]
PHT for metric graphs

**Bad news:** PHT is not injective on all compact metric graphs

$$\text{PHT}(X) = \text{PHT}(Y) \text{ while } X \not\simeq Y$$

**Note:** $\text{Aut}(X)$ is non-trivial, hence $\Psi_X : x \mapsto \text{dgm} \text{ d}_X(\cdot, x)$ is not injective
Let \( \text{Inj}_\Psi = \{ X \text{ compact metric graph s.t. } \Psi_X \text{ is injective} \} \)

**Thm 1:**
PHT is injective on \( \text{Inj}_\Psi \).

**Thm 2:**
\( \text{Inj}_\Psi \) is GH-dense among the compact metric graphs.

Note: \( \Psi_X \) injective \( \iff \) \( \text{Aut}(X) \) trivial
Let $\text{Inj}_\Psi = \{ X \text{ compact metric graph s.t. } \Psi_X \text{ is injective} \}$

**Thm 1:**
PHT is injective on $\text{Inj}_\Psi$.

**Thm 2:**
$\text{Inj}_\Psi$ is GH-dense among the compact metric graphs.

**Corollary:**
There is a GH-dense subset of the compact length spaces on which PHT is injective.

Note: $\Psi_X$ injective $\iff$ $\text{Aut}(X)$ trivial
PHT for metric graphs

Let $\text{Inj}_\Psi = \{ X \text{ compact metric graph s.t. } \Psi_X \text{ is injective}\}$

**Thm 1:**
PHT is injective on $\text{Inj}_\Psi$.

**Thm 2:**
$\text{Inj}_\Psi$ is GH-dense among the compact metric graphs.

**Corollary:**
There is a GH-dense subset of the compact length spaces on which PHT is injective.

**Thm 3:**
PHT is GH-locally injective on compact metric graphs.
Generative model:

metric graph $\equiv$ combinatorial graph $(V, E) +$ edge weights $E \to \mathbb{R}_+$

mixture (proba. mass function, proba. measure with density on $\mathbb{R}_+^{\{|E|\}}$)
Generic injectivity

Generative model:

\[
\text{metric graph } \equiv \text{combinatorial graph } (V, E) + \text{edge weights } E \to \mathbb{R}_+
\]

mixture (proba. mass function, proba. measure with density on \(\mathbb{R}_+^{|E|}\))

Thm 4:
Under this model, there is a full-measure subset of the metric graphs on which PHT is injective.
Generic injectivity

Generative model:

metric graph \equiv \text{combinatorial graph } (V, E) + \text{ edge weights } E \rightarrow \mathbb{R}_+

mixture (\text{ proba. mass function }, \text{ proba. measure with density on } \mathbb{R}_+^{\mid E\mid})

Thm 4:
Under this model, there is a full-measure subset of the metric graphs on which PHT is injective.

Questions:
- is PHT a sufficient statistic for metric graphs?
- are finitely many basepoints enough? algorithm?
- what about higher-dimensional stratified spaces?
Proof outline for Thm 1

Let \( \text{Inj}_\Psi = \{X \text{ compact metric graph s.t. } \Psi_X \text{ is injective}\} \)

**Thm 1:**
PHT is injective on \( \text{Inj}_\Psi \).

**Thm 2:**
\( \text{Inj}_\Psi \) is GH-dense among the compact metric graphs.

**Corollary:**
There is a GH-dense subset of the compact length spaces on which PHT is injective.

**Thm 3:**
PHT is GH-locally injective on compact metric graphs.
Prop:
If $X$ is not a circle, then $\Psi_X$ is a local isometry:

$$\forall x \exists U_x \forall y \in U_x \ d_X(x, y) = d_b(\Psi_X(x), \Psi_X(y))$$

Proof outline for Thm 1
Proof outline for Thm 1

Prop:
If \( X \) is not a circle, then \( \Psi_X \) is a local isometry:

\[
\forall x \exists U_x \ \forall y \in U_x \ d_X(x, y) = d_b(\Psi_X(x), \Psi_X(y))
\]

Corollary:
If \( \Psi_X \) is injective, then \( \Psi_X \) is a (global) isometry from \( (X, d_X) \) to \( (\text{PHT}(X), \hat{d}_b) \).
Proof outline for Thm 2

Let $\text{Inj}_\Psi = \{ X \text{ compact metric graph s.t. } \Psi_X \text{ is injective} \}$

**Thm 1:**
PHT is injective on $\text{Inj}_\Psi$.

**Thm 2:**
$\text{Inj}_\Psi$ is GH-dense among the compact metric graphs.

**Corollary:**
There is a GH-dense subset of the compact length spaces on which PHT is injective.

**Thm 3:**
PHT is GH-locally injective on compact metric graphs.
Proof outline for Thm 2

Given \((X, d_X)\), for any \(\varepsilon > 0\) build an \(\varepsilon\)-approximation \((X_\varepsilon, d_{X_\varepsilon})\) in \(d_{GH}\)

Break symmetries by **cactification**:  

- subdivide edges  
- add hanging branches (**thorns**) with distinct lengths
Proof outline for Thm 2

Given \((X, d_X)\), for any \(\varepsilon > 0\) build an \(\varepsilon\)-approximation \((X_\varepsilon, d_{X_\varepsilon})\) in \(d_{\text{GH}}\)

Break symmetries by **cactification**:

- subdivide edges
- add hanging branches (\textit{thorns}) with distinct lengths

\[ \rightarrow (X_\varepsilon, d_{X_\varepsilon}) \] parametrized by distances to thorn bases and tips

\[ \rightarrow \] these distances appear in the persistence diagrams
Proof outline for Thm 3

Let $\text{Inj}_\Psi = \{X \text{ compact metric graph s.t. } \Psi_X \text{ is injective}\}$

**Thm 1:**
PHT is injective on $\text{Inj}_\Psi$.

**Thm 2:**
$\text{Inj}_\Psi$ is GH-dense among the compact metric graphs.

**Corollary:**
There is a GH-dense subset of the compact length spaces on which PHT is injective.

**Thm 3:**
PHT is GH-locally injective on compact metric graphs.
Prop: The map \((X, d_X, x) \mapsto R_{d_X}(\cdot, x)\) is injective.
Proof outline for Thm 3

**Prop:**
The map \((X, d_X, x) \mapsto R_{d_X}(\cdot, x)\) is injective.

**Thm:** [Carrière, O. 2017]
The map \(R_f \mapsto \text{dgm } f\) is GH-locally injective.
Generic injectivity

Generative model:

metric graph \equiv \text{combinatorial graph } (V, E) \ + \ \text{edge weights } E \rightarrow \mathbb{R}_+

\text{mixture (proba. mass function, proba. measure with density on } \mathbb{R}_+^{\mid E \mid})

**Thm 4:**
Under this model, there is a full-measure subset of the metric graphs on which PHT is injective.

Proof outline:
- for (almost) any fixed combinatorial graph \( G \), \( \Psi_G \) is \textit{generically} injective.
- deal with exceptions (e.g. linear graphs) explicitly
The preimage problem in the data Sciences

**Data**

( feature design or learning )

[Image of chairs]

---

**Features**

\( \in \mathbb{R}^n \)

- bag of words, word2vec
- shape contexts, heat kernels
- node2vec, Laplacian fact., rand. walks
- dim. reduction, auto-encoders, etc.