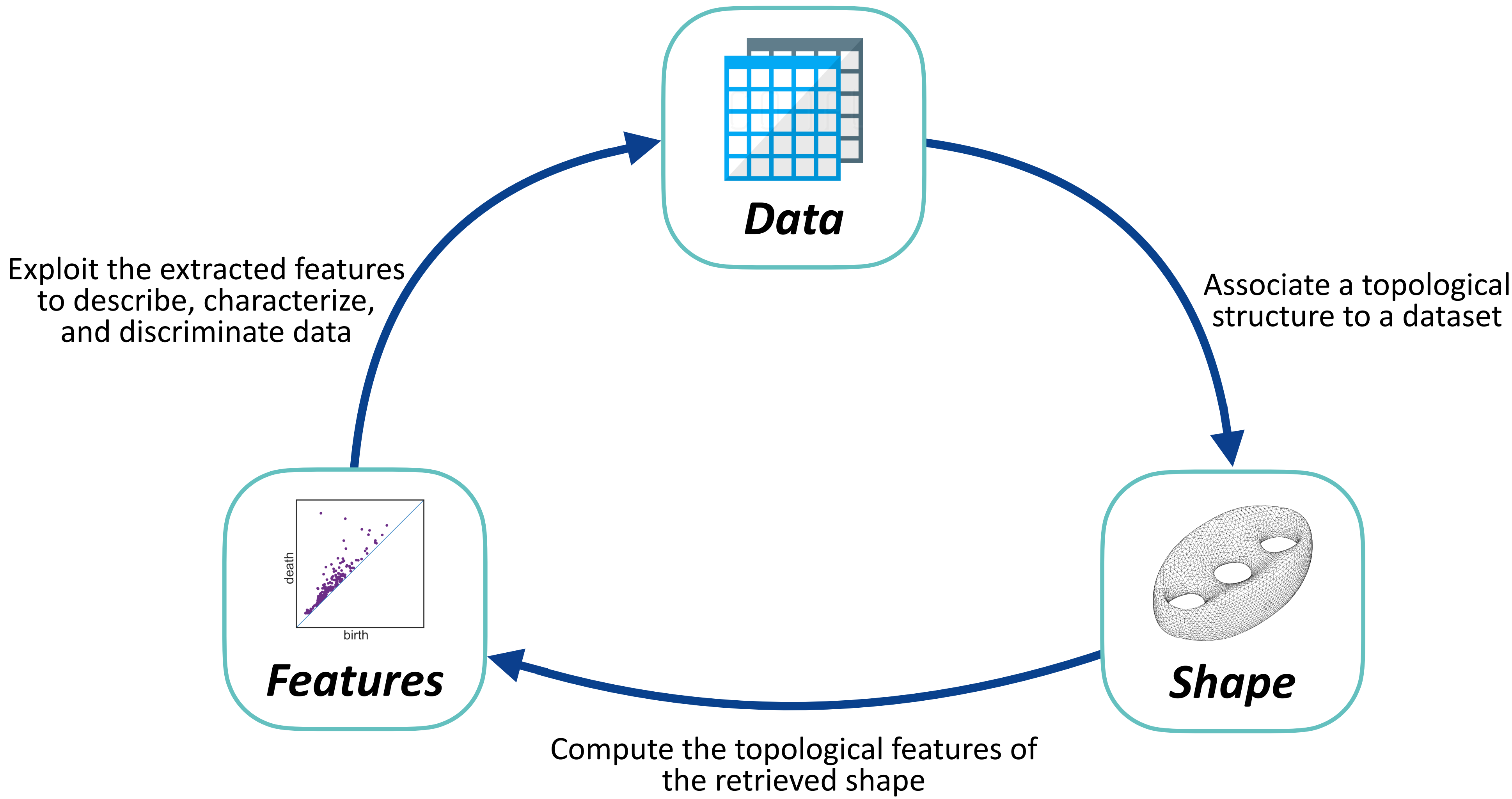
Topological Data Analysis

Persistence-Based Kernels

Ulderico Fugacci

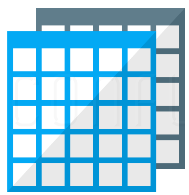
CNR - IMATI

Topological Data Analysis

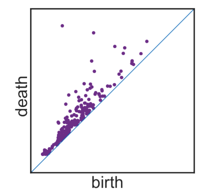


Kernels for Persistent Homology

Topological Data Analysis allows for assigning to (almost) ***any dataset*** a collection of features representing a ***topological summary*** of the input data



Data



Features

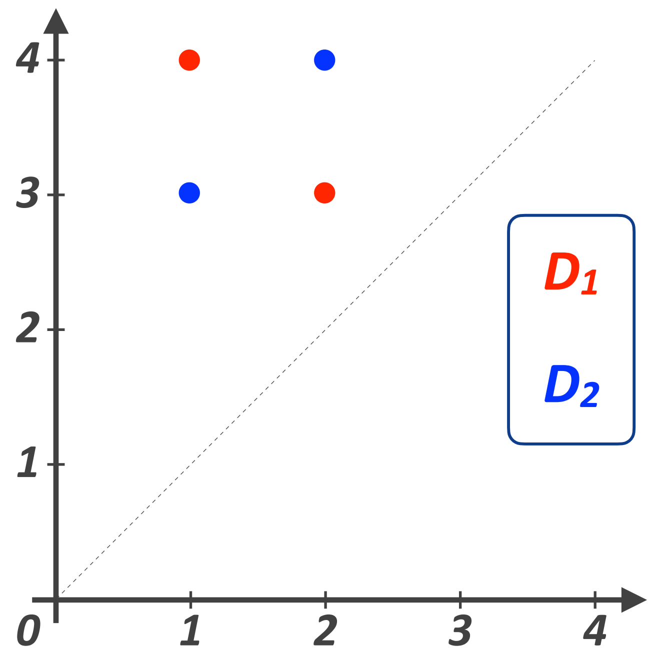
Goal:

Today, we address one main question:

- ***Is this information immediately suitable for statistics and machine learning?***

Kernels for Persistent Homology

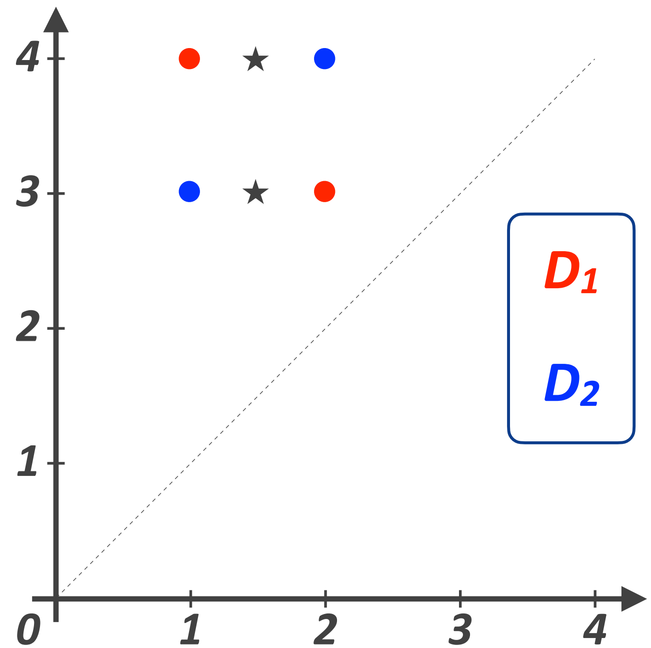
A Naive Example:



Mean *of persistence diagrams is* **not unique**

Kernels for Persistent Homology

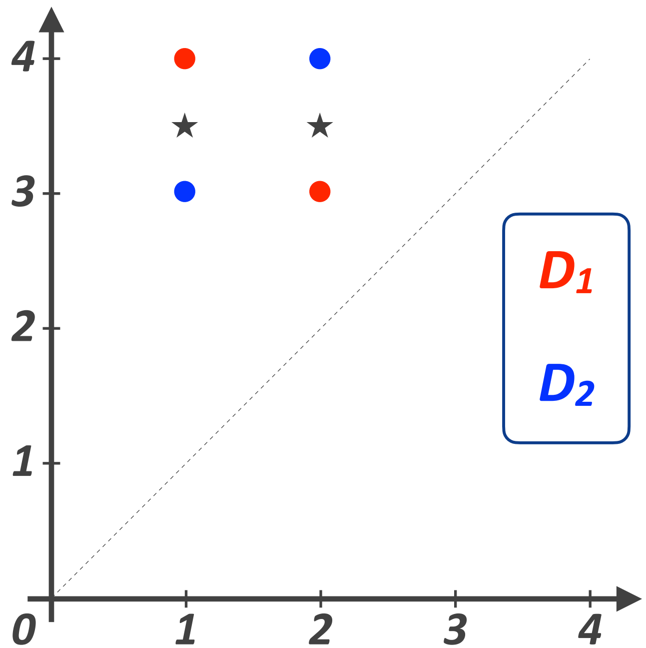
A Naive Example:



Mean** of persistence diagrams is **not unique

Kernels for Persistent Homology

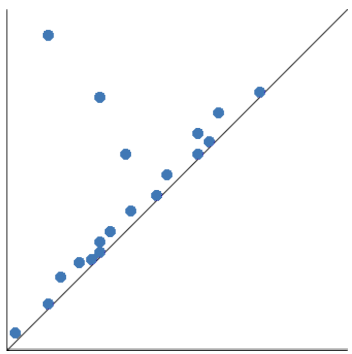
A Naive Example:

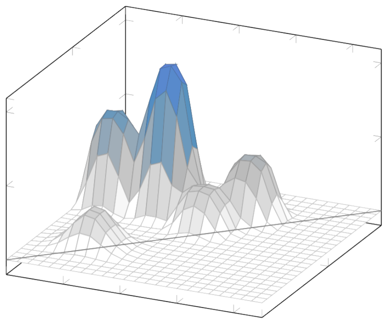


Mean** of persistence diagrams is **not unique

Kernels for Persistent Homology

Adopted Strategy:





*Represent persistence diagrams as elements of a **Hilbert space***

Kernels for Persistent Homology

Definitions:

A ***Hilbert space H*** *is*

a ***real or complex vector space*** *endowed with an* ***inner product***

$\langle \cdot, \cdot \rangle : H \times H \longrightarrow \mathbb{R}$ *such that, with respect to the distance induced by* $\langle \cdot, \cdot \rangle$,

H is a ***complete metric space***

Recall that, a metric space H is called ***complete*** *if*

every Cauchy sequence in H converges in H

Kernels for Persistent Homology

Example:

The space ***L^2*** of ***square-integrable functions*** on $\mathbb{R}^2$ is a Hilbert space

- $$\|f\|_{L^2} := \left(\int_{\mathbb{R}^2} |f|^2 d\mu \right)^{\frac{1}{2}} < +\infty$$
- $$\langle f, g \rangle_{L^2} := \int_{\mathbb{R}^2} f \cdot g \; d\mu$$

Kernels for Persistent Homology

Kernel Trick:

$$\begin{tikzcd} \Omega \times \Omega \arrow[r, "k"] \arrow[d, "\phi \times \phi"'] & \mathbb{R} \\ H \times H \arrow[ur, "\langle \cdot , \cdot \rangle"'] & \end{tikzcd}$$

Definition:

*A **kernel k** for an input space Ω is a map $k: \Omega \times \Omega \longrightarrow \mathbb{R}$ such that there exist a **Hilbert space H** and a **feature map** $\phi: \Omega \longrightarrow H$ for which*

$$k(X, Y) = \langle \phi(X), \phi(Y) \rangle$$

Kernels for Persistent Homology

Pseudo-Distance:

A kernel k: Ω × Ω ⟶ ℝ implicitly induces on Ω

a ***pseudo-distance*** ***d_k: Ω × Ω*** ⟶ ℝ defined, for each X, Y ∈ Ω, as

$$d_k(X,Y) := \|\phi(X) - \phi(Y)\|_H = \Big(k(X,X) + k(Y,Y) - 2k(X,Y)\Big)^{1/2}$$

Stability:

*A kernel k: Ω × Ω ⟶ ℝ is **stable** w.r.t a distance d in Ω*

if there is a constant C > 0 such that, for all X, Y ∈ Ω,

$$d_k(X,Y) \leq C \cdot d(X,Y)$$

Kernels for Persistent Homology

Our Goal:

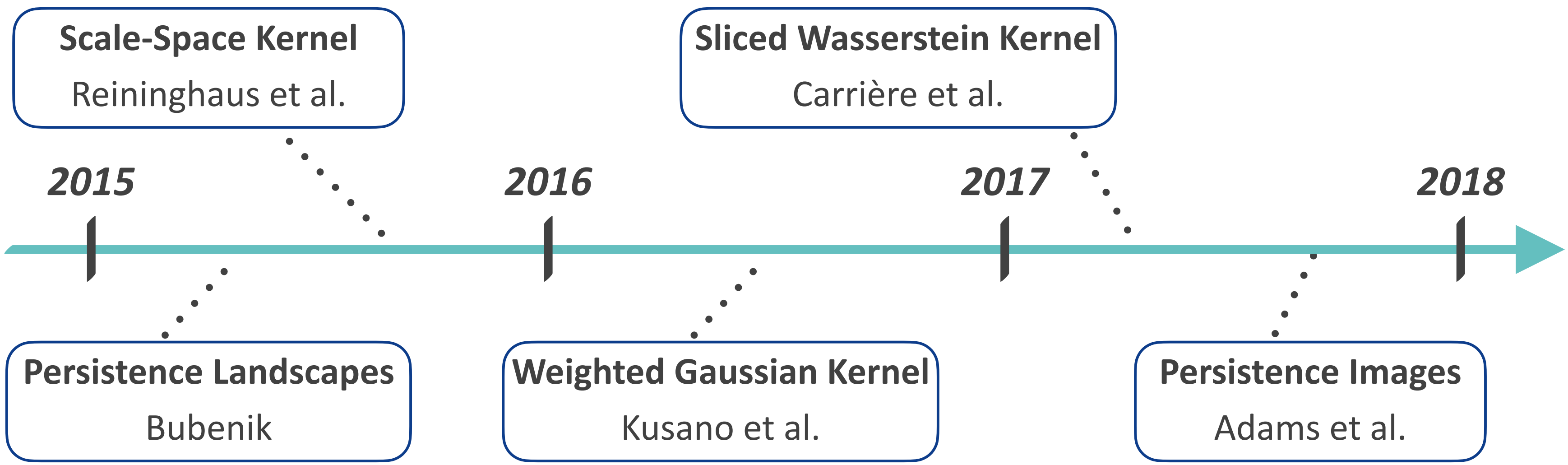
Defining a ***kernel*** for the set Ω of finite ***persistence diagrams***:

- ***Stable***
- ***Easy to be computed***
- Possible endowed with an ***explicit feature map*** ϕ***: Ω*** $\longrightarrow$ ***H***

The idea of a kernel for persistence diagrams has

- *Originally* ***born in the '90s*** *(see [Donatini et al. 1998; Ferri et al. 1998])*
- *Spread in the literature and* ***widely adopted in applications just recently***

Kernels for Persistent Homology

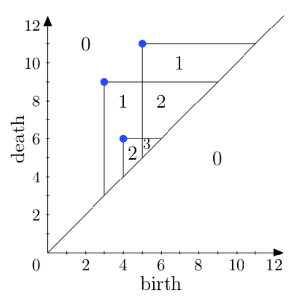


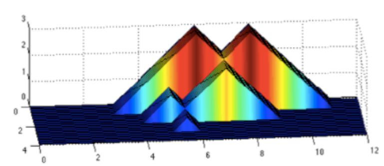
Kernels for Persistent Homology

2015 **2016** **2017** **2018**

Persistence Landscapes

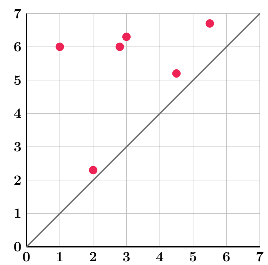
Bubenik

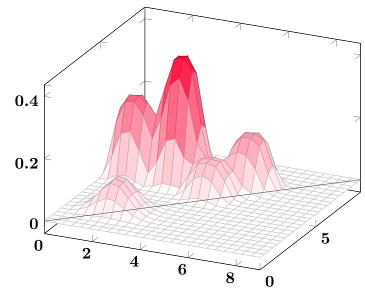




Kernels for Persistent Homology

Scale-Space Kernel
Reininghaus et al.





2015
2016
2017
2018

Persistence Landscapes
Bubenik

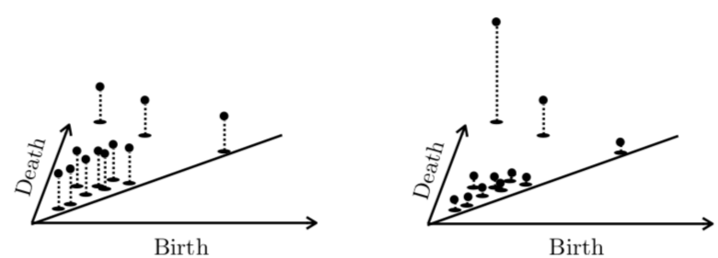
Kernels for Persistent Homology

Scale-Space Kernel
Reininghaus et al.

2015 ***2016*** ***2017*** ***2018***

Persisten

Weighted Gaussian Kernel
Kusano et al.



Kernels for Persistent Homology

$$k_{SW}(D, E) := \exp\left(-\frac{d_{SW}(D, E)}{2\sigma^2}\right)$$

Sliced Wasserstein Kernel

Carrière et al.

Scale

Reini

2015 2016 2017 2018

Persistence Landscapes

Bubenik

Weighted Gaussian Kernel

Kusano et al.

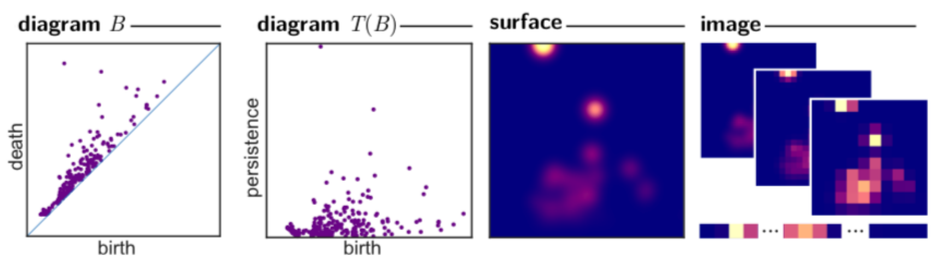
Kernels for Persistent Homology

Scale-Space Kernel
Reininghaus et al.

Sliced Wasserstein Kernel
Carrière et al.

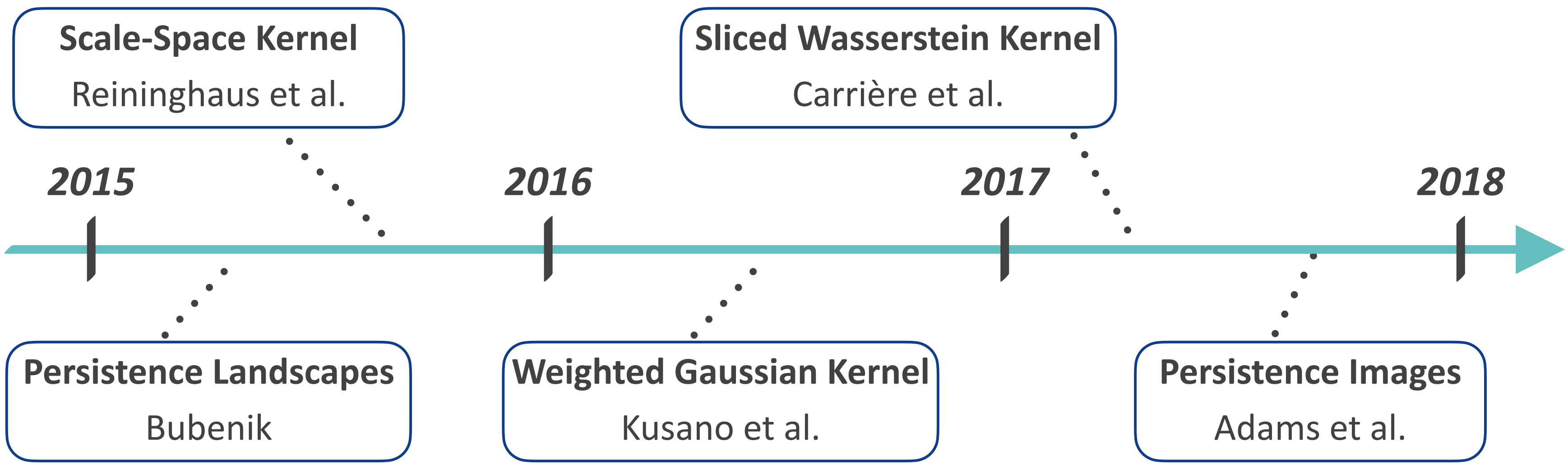
2015 ***2016*** ***2017*** ***2018***

Persi



Persistence Images
Adams et al.

Kernels for Persistent Homology



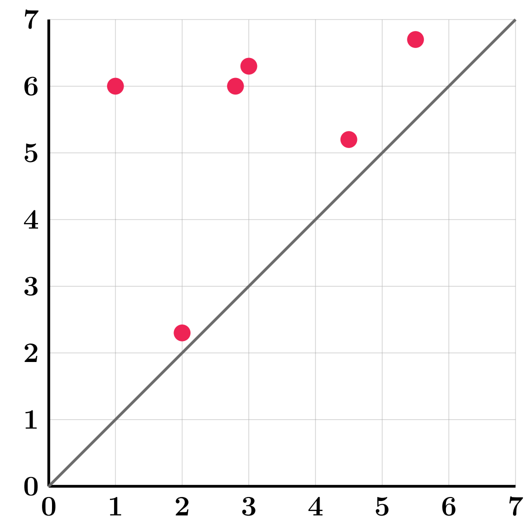
Kernels for Persistent Homology

Scale-Space Kernel:

Given a finite persistence diagram D, we consider the solution $\phi : \Delta^+ \times \mathbb{R}_{\geq 0} \longrightarrow \mathbb{R}$ of the following ***heat diffusion problem***:

- having a Dirichlet ***boundary condition*** on the diagonal
- setting as an ***initial condition*** a sum of Dirac deltas

$$\begin{cases} \Delta_p \phi = \partial_\sigma \phi & \text{in } \Delta^+ \times \mathbb{R}_{>0} \\ \phi = 0 & \text{on } \partial\Delta^+ \times \mathbb{R}_{\geq 0} \\ \phi = \sum_{q \in D} \boldsymbol{\delta}_q & \text{on } \Delta^+ \times \{0\} \end{cases}$$



Kernels for Persistent Homology

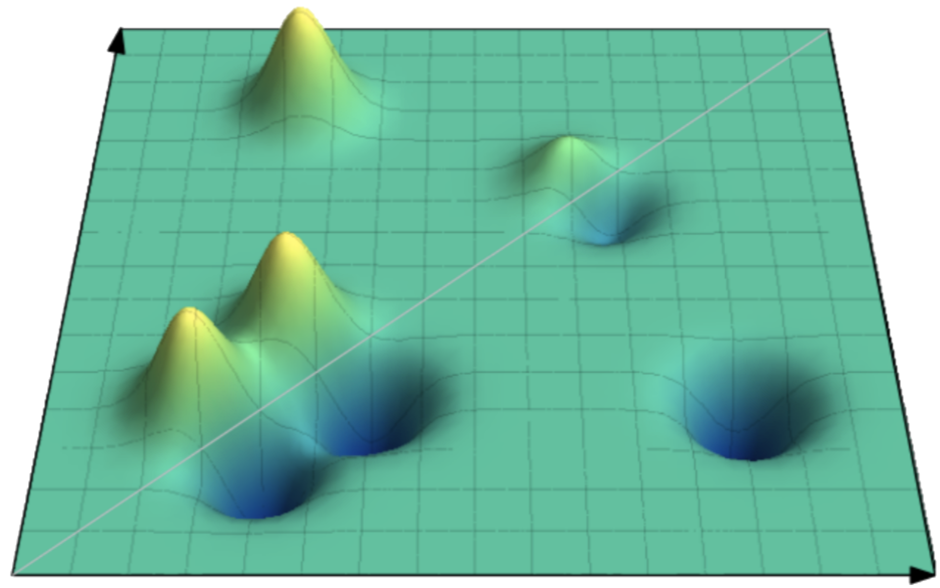
Scale-Space Kernel:

A ***solution*** is found by:

- ***extending*** Δ^+ to $\mathbb{R}^2$
- ***replacing*** the initial condition with

$$\phi = \Sigma_{q \in D} \boldsymbol{\delta}_q - \boldsymbol{\delta}_{q'} \qquad \text{on } \mathbb{R}^2 \times \{0\}$$

where, if $q=(a,b)$, then $q'=(b,a)$



Solution:

$$\phi_\sigma(p) = \frac{1}{4\pi\sigma} \sum_{q \in D} \Big(\exp\big(-\frac{\|p-q\|^2}{4\sigma}\big) - \exp\big(-\frac{\|p-q'\|^2}{4\sigma}\big)\Big)$$

Kernels for Persistent Homology

Scale-Space Kernel:

Stability Theorem:

Given two finite persistence diagrams D, E, we have that

$$\|\phi_\sigma(D) - \phi_\sigma(E)\|_{L^2} \leq \frac{1}{2\sqrt{\pi}\sigma} d_{W,1}(D, E)$$

*where, for $r \geq 1$, the **r-Wasserstein distance** is defined as*

$$d_{W,r}(D, E) := \left(\inf_{\gamma} \sum_{p \in D} \|p - \gamma(p)\|_\infty^r \right)^{1/r}$$

with γ running over all bijections from D to E

Kernels for Persistent Homology

Definitions:

A kernel k for the set Ω of finite persistence diagrams is called:

- ***additive*** if, for all D, E, F ∈ Ω, $k(D \cup E, F) = k(D, F) + k(E, F)$
- ***trivial*** if, for all D, E ∈ Ω, $k(D, E) = 0$

Theorem:

*Any **non-trivial additive** kernel k for the set Ω*

*is **not stable** with respect to $d_{W,r}$ for any $1 < r \leq \infty$*

(*Notice that* $d_{W,\infty} = d_B$)

Kernels for Persistent Homology

Sliced Wasserstein Kernel:

A standard way to construct a kernel is to exponentiate the negative of an Euclidean distance

$$k_\sigma(X,Y) := \exp\left(-\frac{\|X-Y\|^2}{2\sigma^2}\right)$$

Theorem:

$$k_\sigma(X,Y) := \exp\left(-\frac{f(X,Y)}{2\sigma^2}\right)$$

*defines a valid kernel for all $\sigma > 0$ **if and only if** f is conditionally negative definite function*

I.e., for any $n \in \mathbb{N}$, for any $X_1, \ldots, X_n \in \Omega$, and for any $a_1, \ldots, a_n \in \mathbb{R}$ such that $\Sigma_i\, a_i = 0$,

one has $\Sigma_{i,j}\, a_i\, a_j\, f(X_i, X_j) \leq 0$

Kernels for Persistent Homology

Sliced Wasserstein Kernel:

Issue:

*None of the already introduced distances (and neither their squares) between persistence diagrams **is conditionally negative definite***

Solution:

*The **Sliced Wasserstein distance** $\mathbf{d_{sw}}$ is specifically designed to be **conditionally negative definite***

*Based on it, one can define the **Sliced Wasserstein kernel** $\mathbf{k_{sw}}$ as*

$$k_{SW}(D,E) := \exp\left(-\frac{d_{SW}(D,E)}{2\sigma^2}\right)$$

Bibliography

General References:

- ***Books on TDA:***
  - A. J. Zomorodian. ***Topology for computing.*** Cambridge University Press, 2005.
  - H. Edelsbrunner, J. Harer. ***Computational topology: an introduction***. American Mathematical Society, 2010.
  - R. W. Ghrist. ***Elementary applied topology***. Seattle: Createspace, 2014.
- ***Papers on TDA:***
  - G. Carlsson. ***Topology and data***. Bulletin of the American Mathematical Society 46.2, pages 255-308, 2009.

Today's References:

- ***(Proto-)Kernels for Persistent Homology:***
  - P. Donatini, P. Frosini, A. Lovato. ***Size functions for signature recognition.*** Proc. of SPIE, 3454, pages 178-183, 1998.
  - M. Ferri, P. Frosini, A. Lovato, C. Zambelli. ***Point selection: a new comparison scheme for size functions.*** Proc. 3rd Asian Conference on Computer Vision, vol. I, pages 329-337, 1998.

Bibliography

Today's References:

- ***Kernels for Persistent Homology:***

  - P. Bubenik. ***Statistical topological data analysis using persistence landscapes.*** Journal of Machine Learning Research, 16.1, pages 77-102, 2015.
  - J. Reininghaus, S. Huber, U. Bauer, R. Kwitt. ***A stable multi-scale kernel for topological machine learning***. Proc. of IEEE Conference on Computer Vision and Pattern Recognition, pages 4741-4748, 2015.
  - G. Kusano, Y. Hiraoka, K. Fukumizu. ***Persistence weighted Gaussian kernel for topological data analysis.*** International Conference on Machine Learning, pages 2004-2013, 2016.
  - M. Carrière, M. Cuturi, S. Oudot. ***Sliced Wasserstein kernel for persistence diagrams.*** Proc. of the 34th International Conference on Machine Learning, 70, pages 664-673, 2017.
  - H. Adams et al. ***Persistence images: a stable vector representation of persistent homology.*** Journal of Machine Learning Research, 18.1, pages 218-252, 2017.