

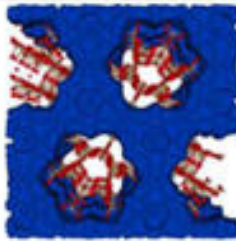

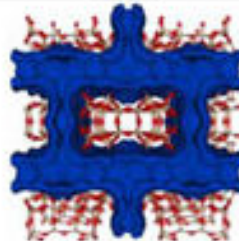
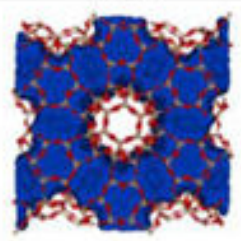
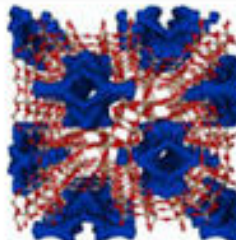
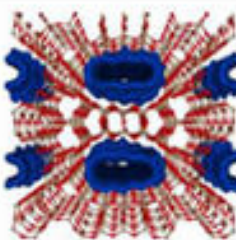
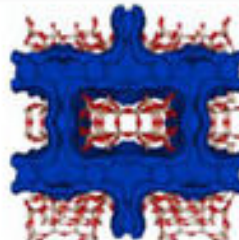

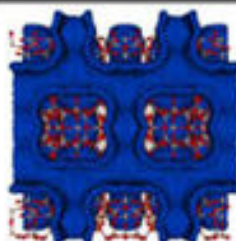
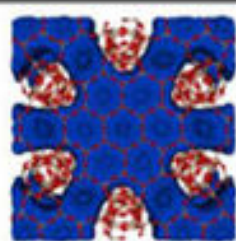
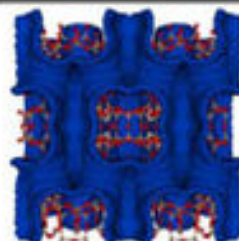
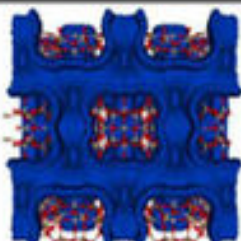
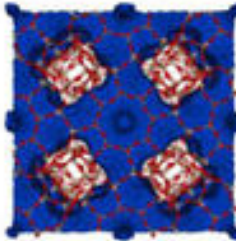
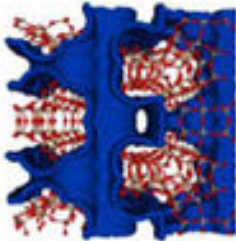
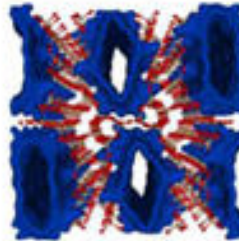
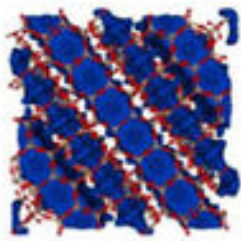
INF556

Topological Data Analysis (TDA)

Steve Oudot

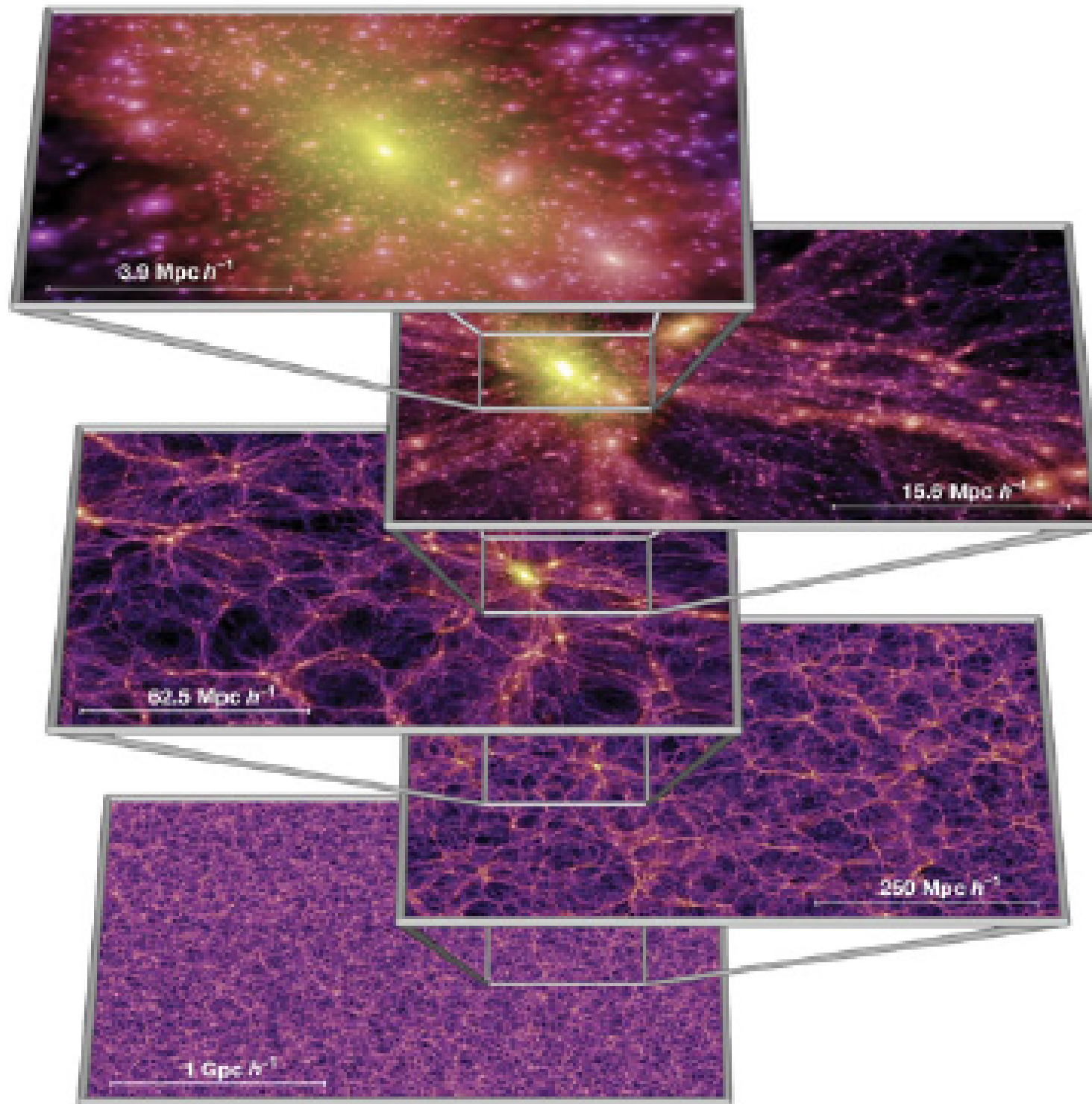
steve.oudot@inria.fr

Data has Shape, Shape carries Meaning

Seed	Descriptor	Selected Nth similar structure			
		1 st	2 nd	3 rd	4 th
SSF	PerH				
	ConD				
IWV	PerH				
	ConD				

Networks of cavities in granular materials determine their physical properties

Data has Shape, Shape carries Meaning



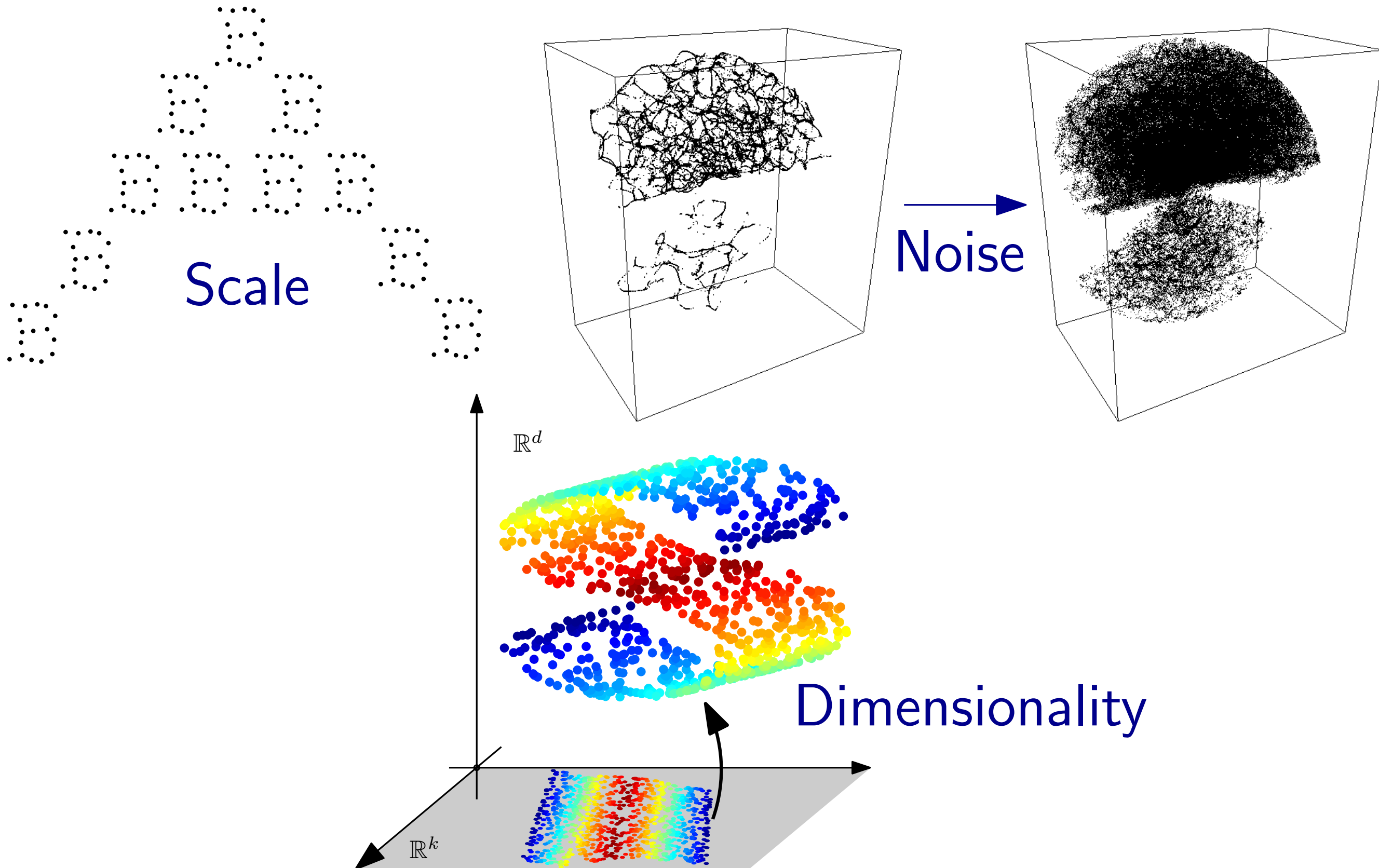
The *cosmic web* characterizes the structure of the Universe and its history

Data has Shape, Shape carries Meaning

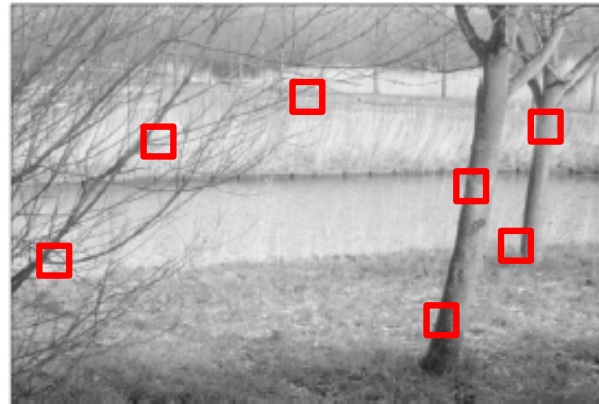
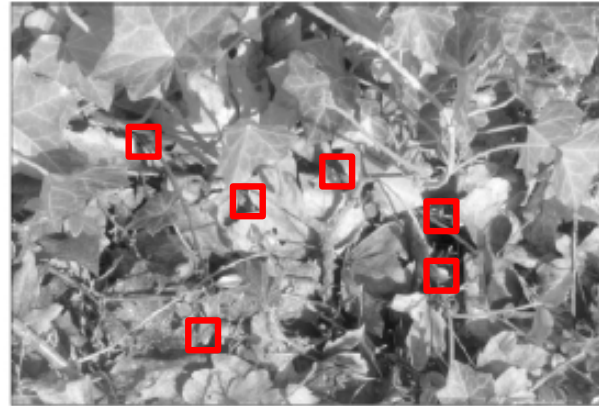
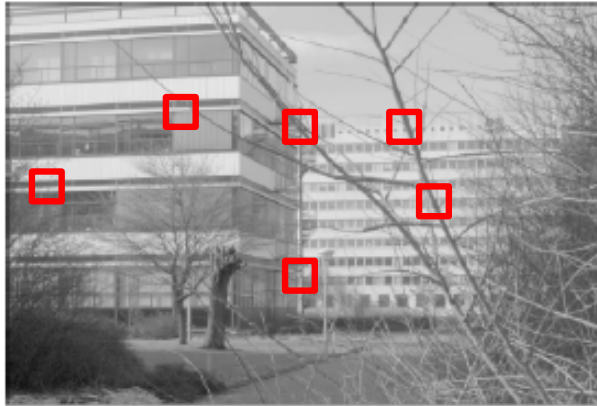
objectives: capture the *shape* of data, in order to:

- ▶ provide higher-level understanding (data mining)
- ▶ improve learning performances (ML)
- ▶ understand prediction models' behavior (DL)

Challenges in uncovering the *Shape* of Data



Challenges in uncovering the *Shape* of Data

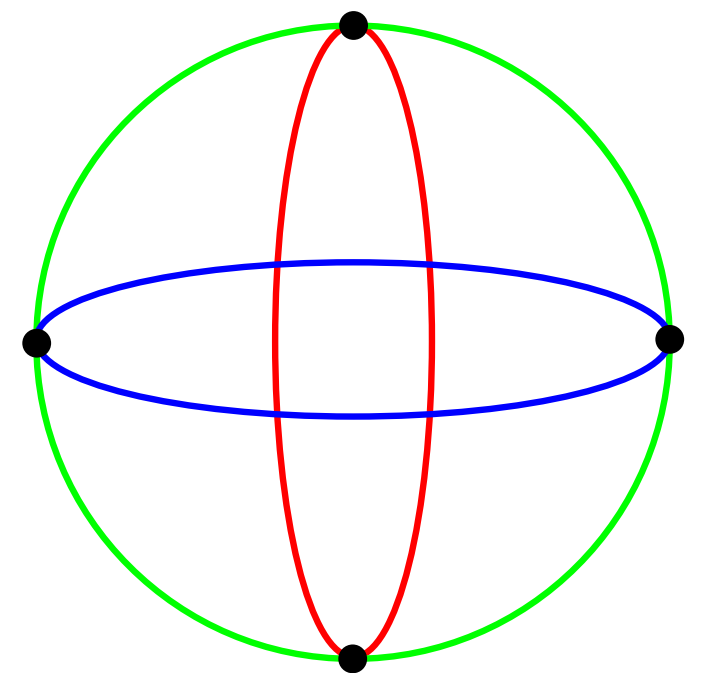
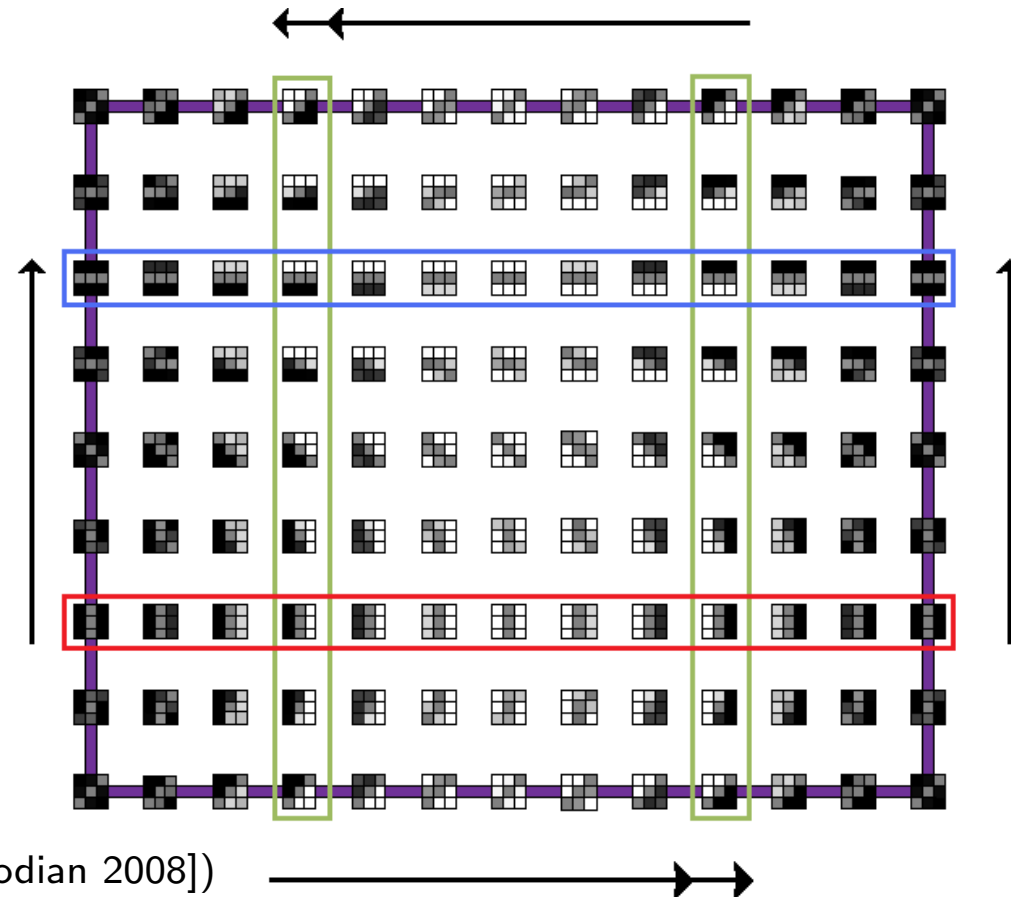


4 million data points in \mathbb{R}^9

(source: [Lee, Pederson, Mumford 2003])

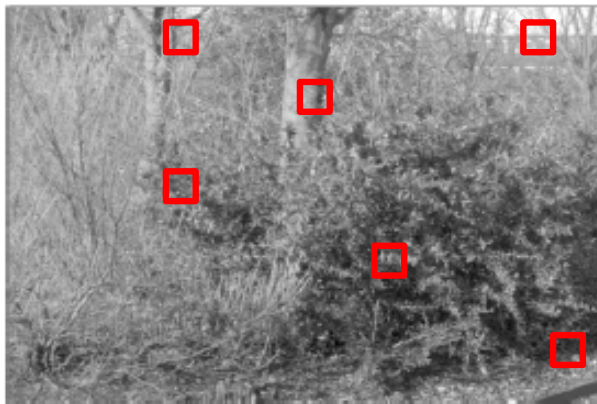
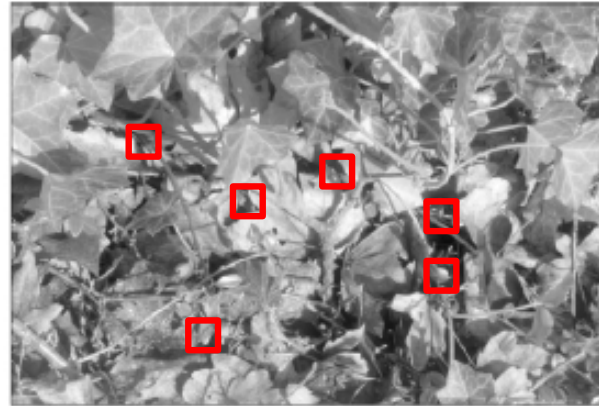
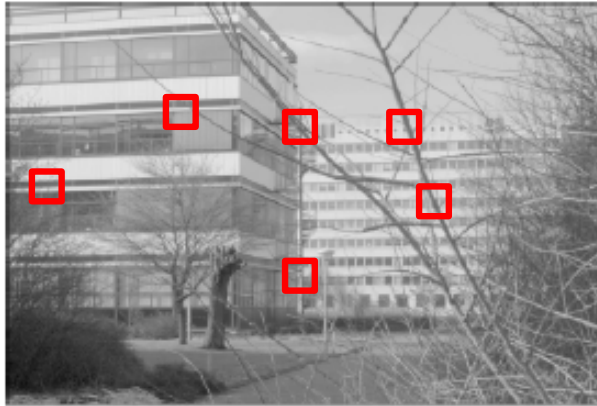
Motivation: study cognitive representation of space of images

Topology



(source: [Carlsson, Ishkhanov, de Silva, Zomorodian 2008])

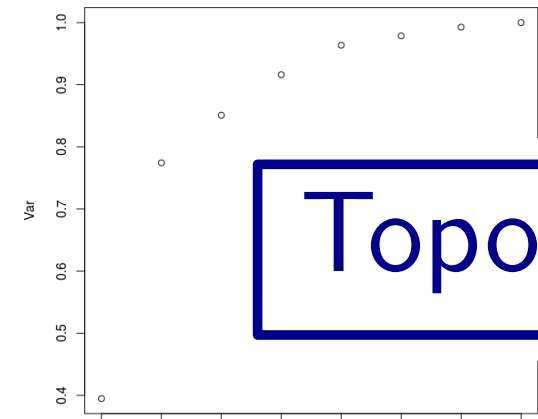
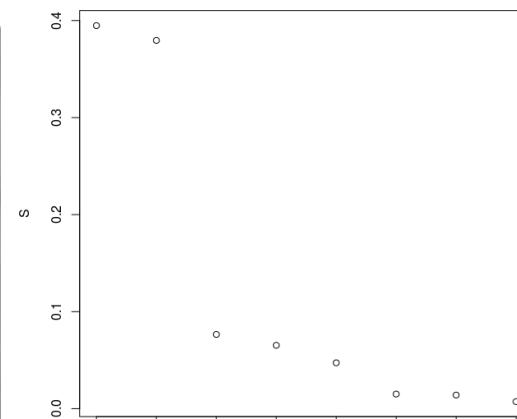
Challenges in uncovering the *Shape* of Data



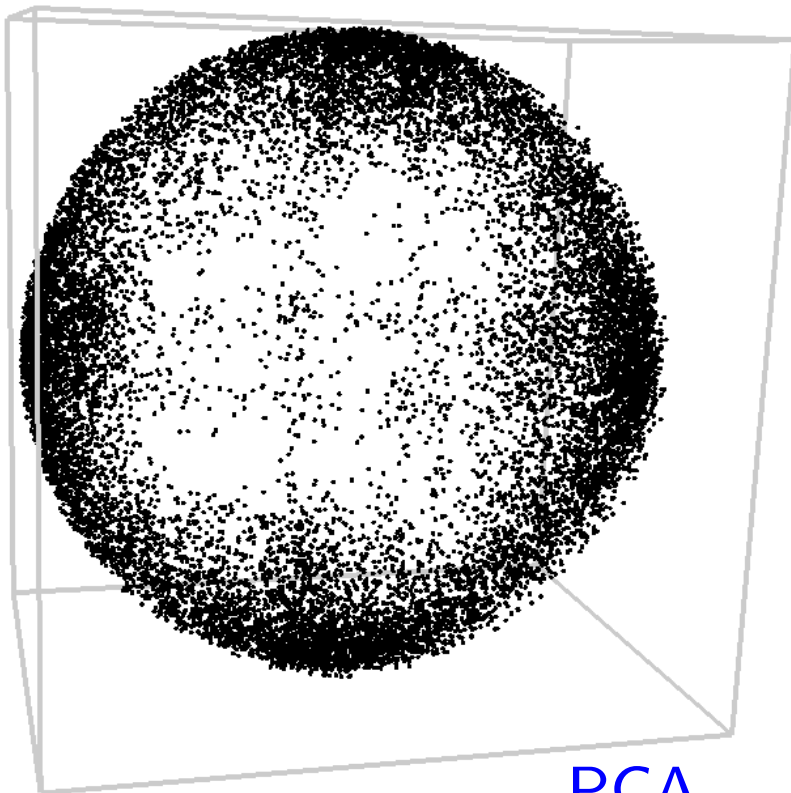
4 million data points in \mathbb{R}^9

(source: [Lee, Pederson, Mumford 2003])

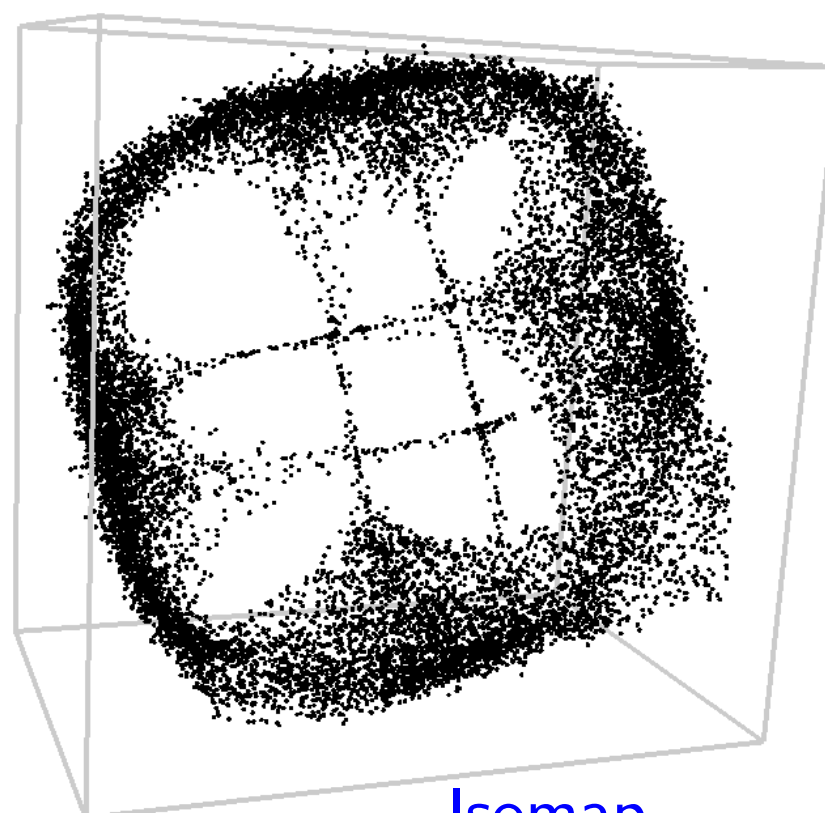
Motivation: study cognitive representation of space of images



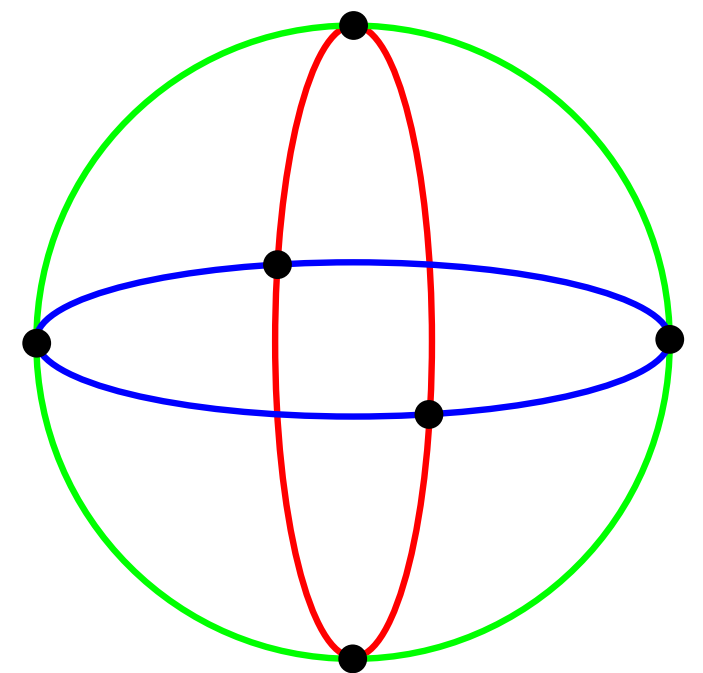
Topology



PCA



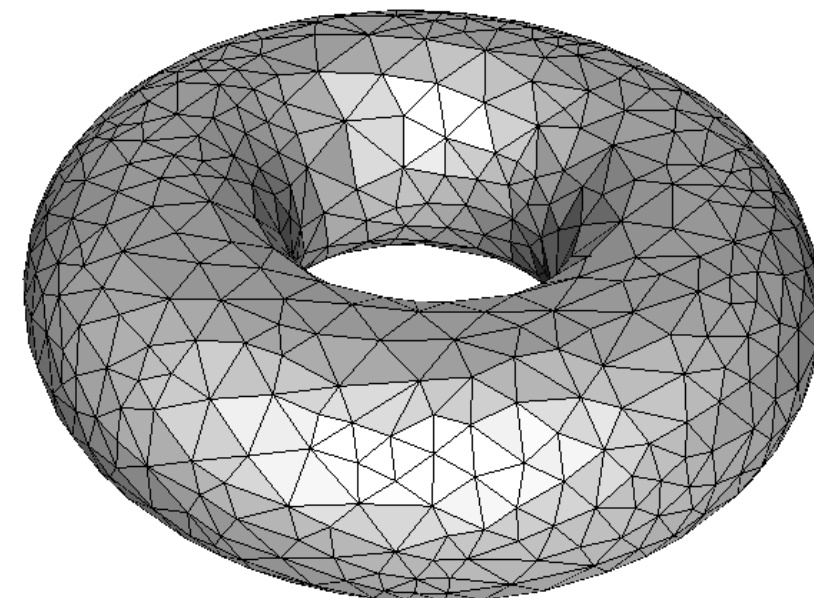
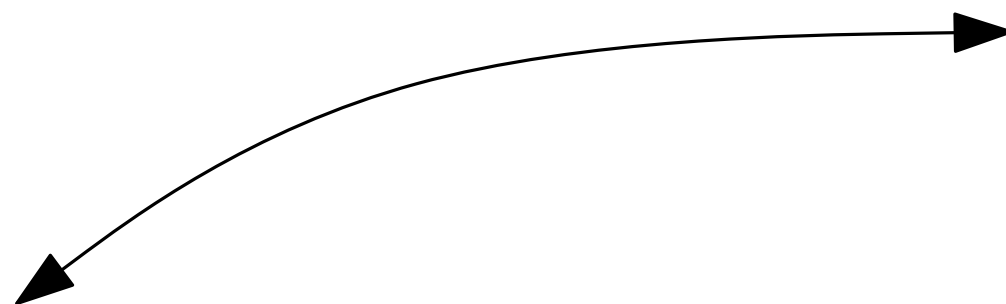
Isomap



Topological Data Analysis (TDA)

topological invariants for classification

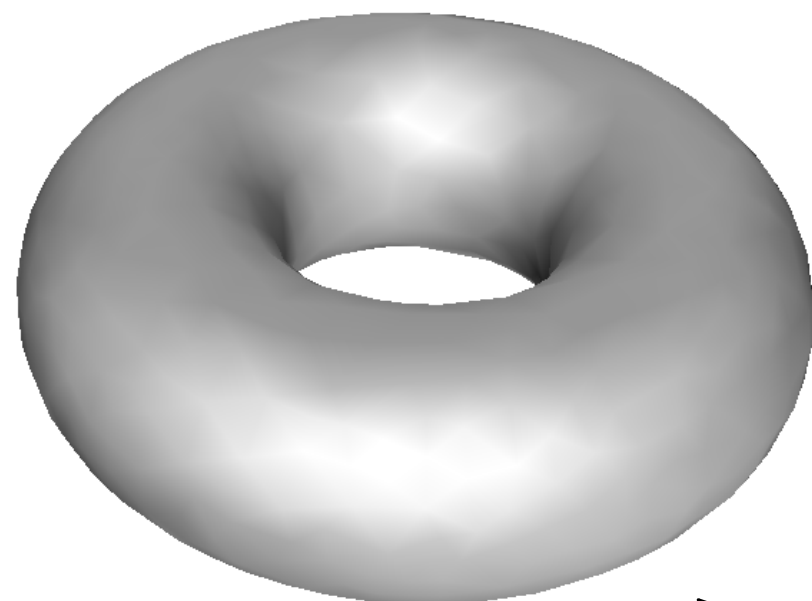
$$\beta_0 = \beta_2 = 1$$
$$\beta_1 = 2$$



triangulation

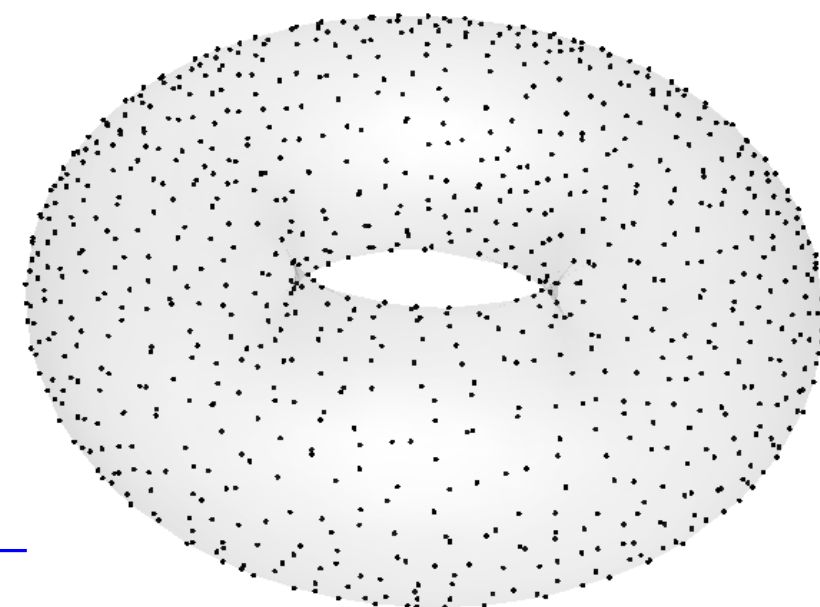
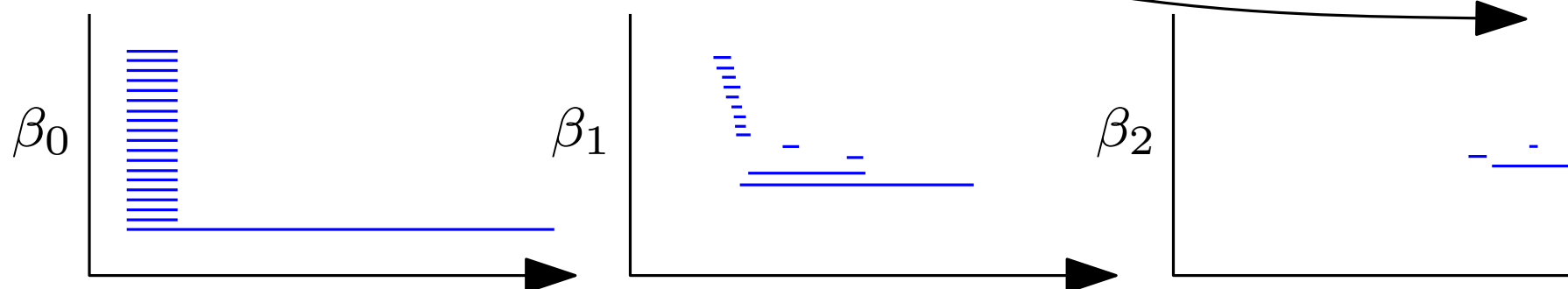
Algebraic topology

Applied algebraic topology



compact set

topological descriptors for inference and comparison



point cloud

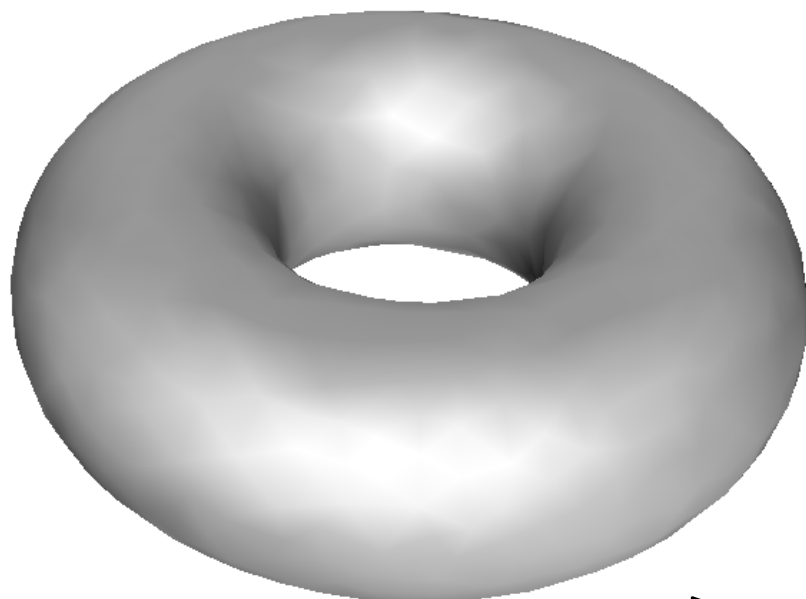
Topological Data Analysis (TDA)

Properties of topological descriptors:



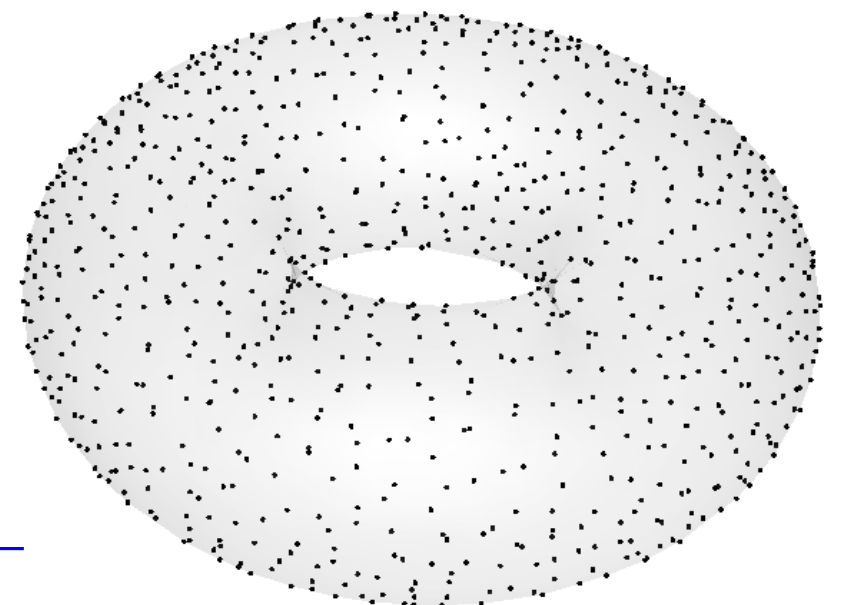
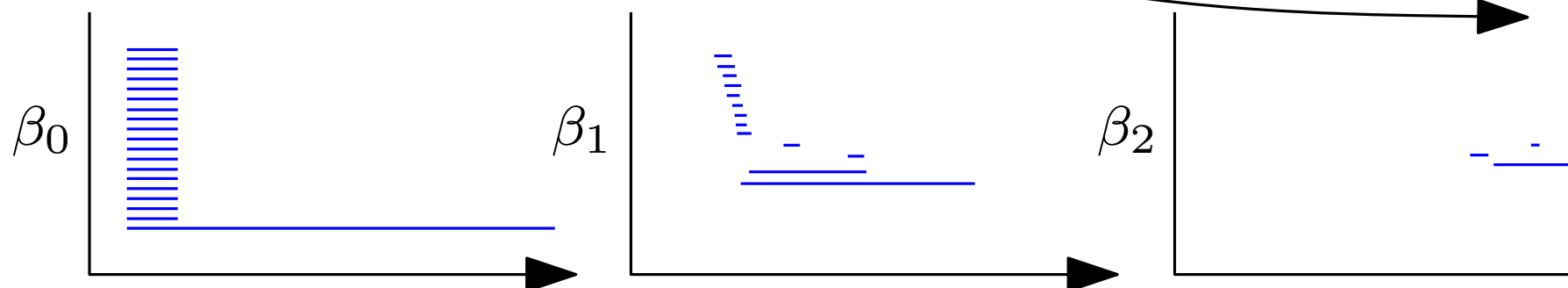
- invariant under coordinate changes
- stable with respect to perturbations
- informative
- generic

Applied algebraic topology



compact set

topological descriptors for inference and comparison



point cloud

Course outline

- Session 1: clustering (mode-seeking) + lab
- Sessions 2-3: homology theory + exercises
- Session 4-5: persistence theory + lab (**graded**)
- Session 5: topological inference + lab
- Session 6: topological descriptors + lab
- Session 7: learning with topological descriptors + exercises
- Session 8: statistics with topological descriptors + exercises
- Session 9: Reeb graphs and Mapper + lab
- **Final written exam**