Topological Data Analysis on Materials Science

Yasuaki Hiraoka

WPI-AIMR, Tohoku University

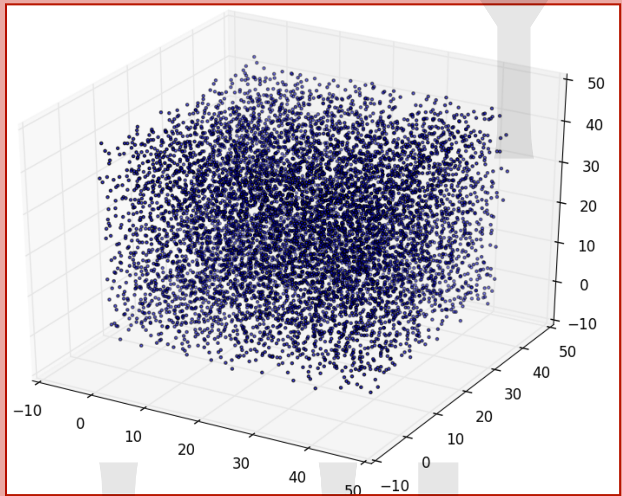
Supported by

JST CREST
SIP Structural Materials for Innovation
JST Innovation Hub MI^2I
NEDO

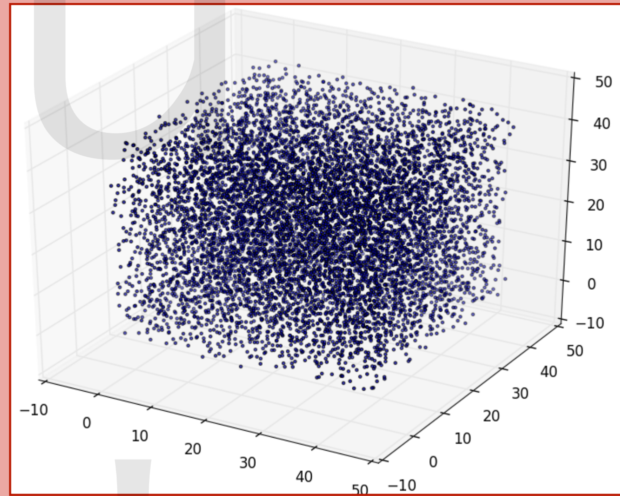
Background : Shape of Data

Data-driven science studies potential values of big and complicated data for machine learning and AI

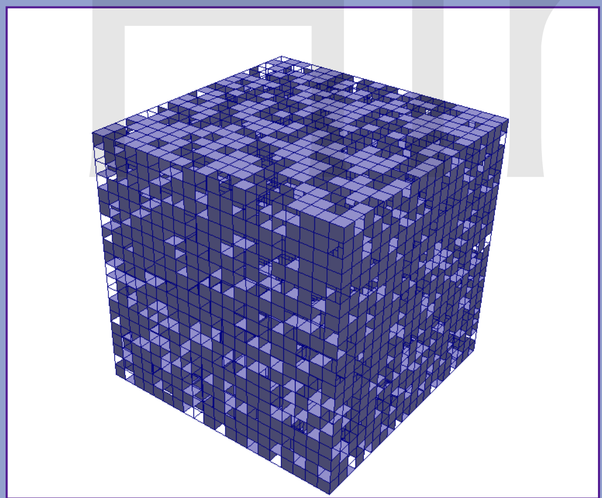
Point cloud data



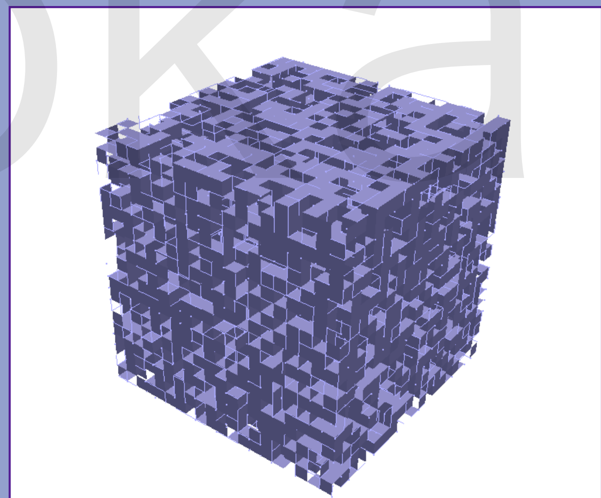
How different?



3D image data

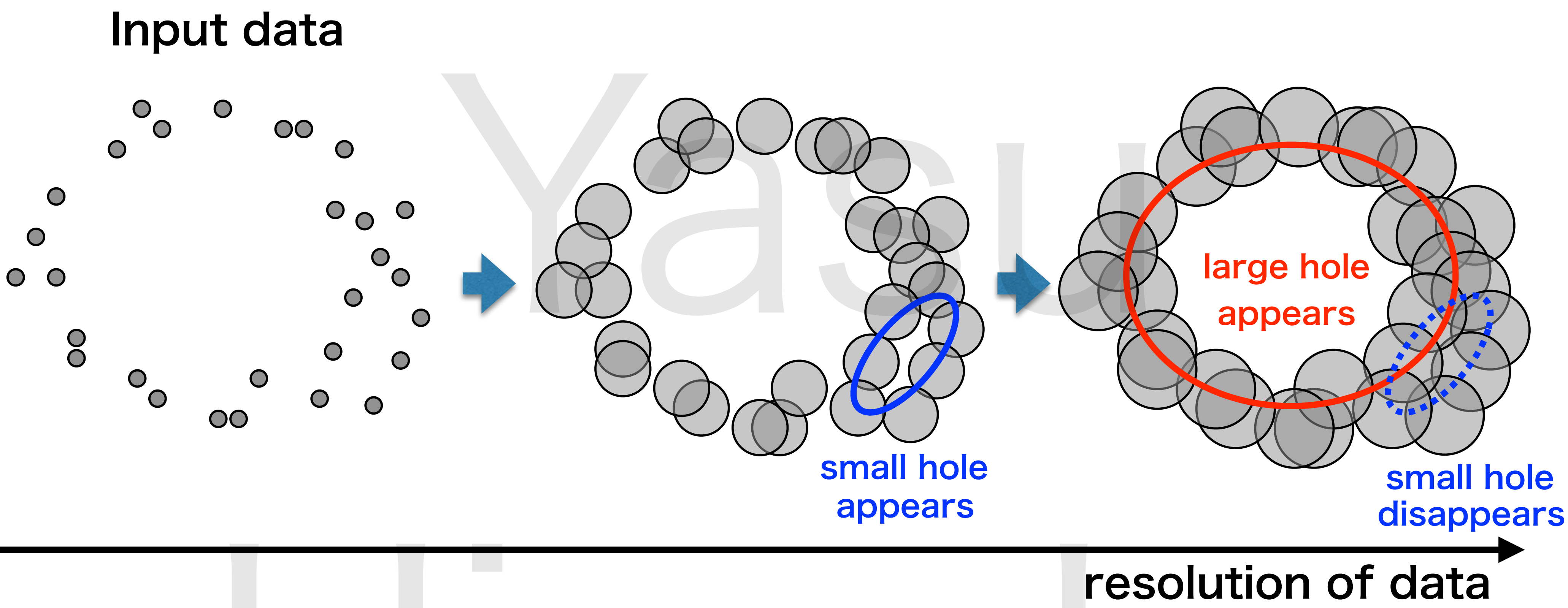


How different?

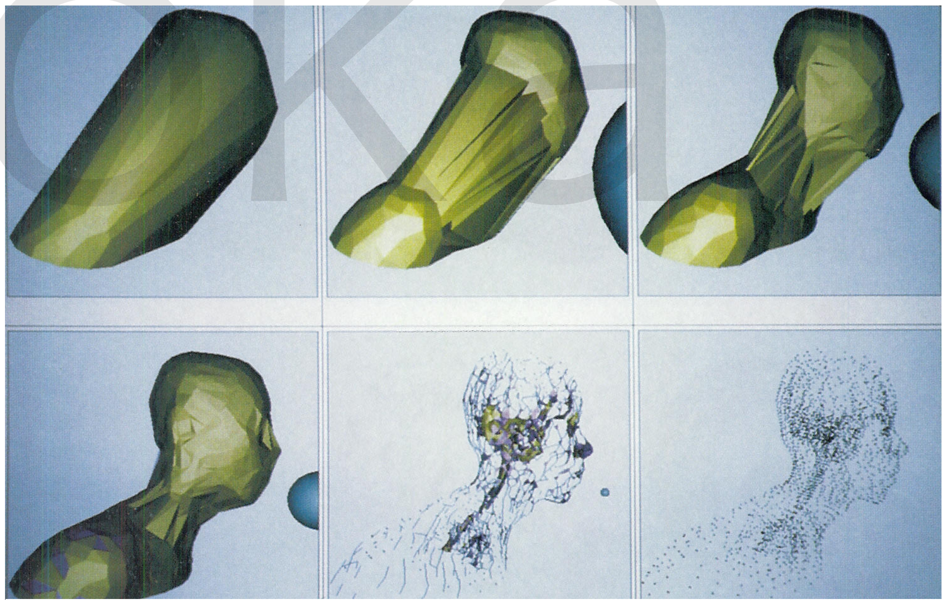


→ develop mathematical theory for shape of data

Idea : Shape of Data



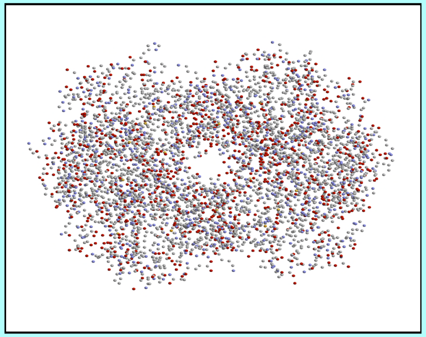
- fattening point data
- changing resolution for multi-scale analysis
- characterization using birth & death of holes



(ref. Edelsbrunner, Mucke)

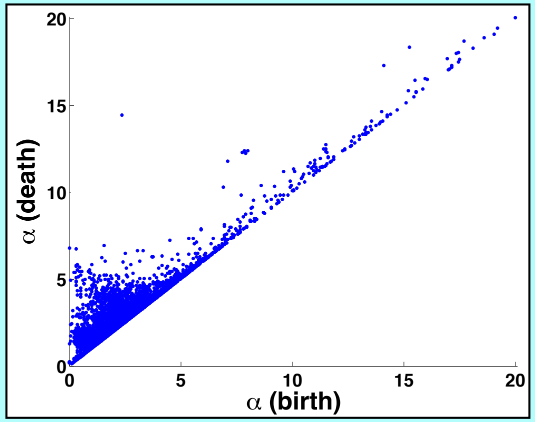
New math: Persistent homology

Input data → Persistent Homology → Shape of data



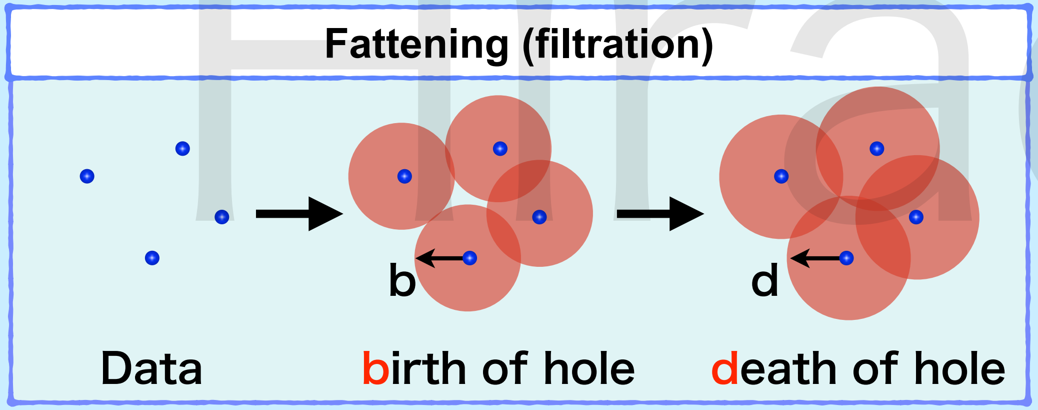
Atomic configuration of hemoglobin

- characterize holes in data
- describe number, size, and shapes
- multi-scale analysis

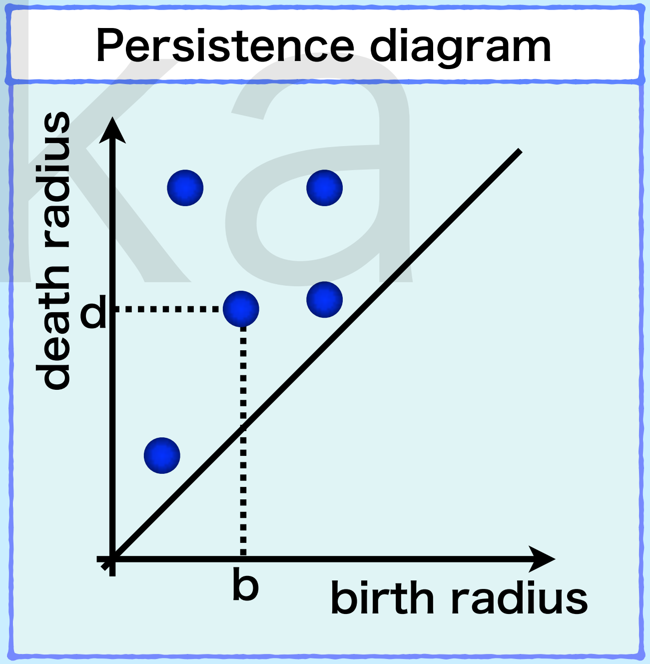


Persistence Diagram (PD)

Persistence diagram of point cloud



inverse

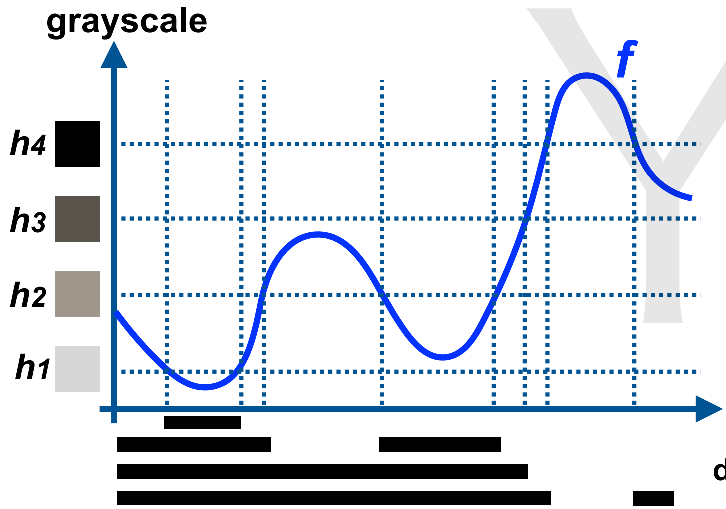


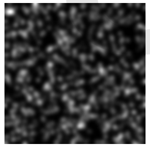
- each point (called generator) in PD expresses a hole in data
- birth & death axes measure shapes of holes
- points close to diagonal are noisy
- points away from diagonal are robust

Note: 2D histogram uncovers further geometry

Persistent homology of digital image

1. Grayscale persistence



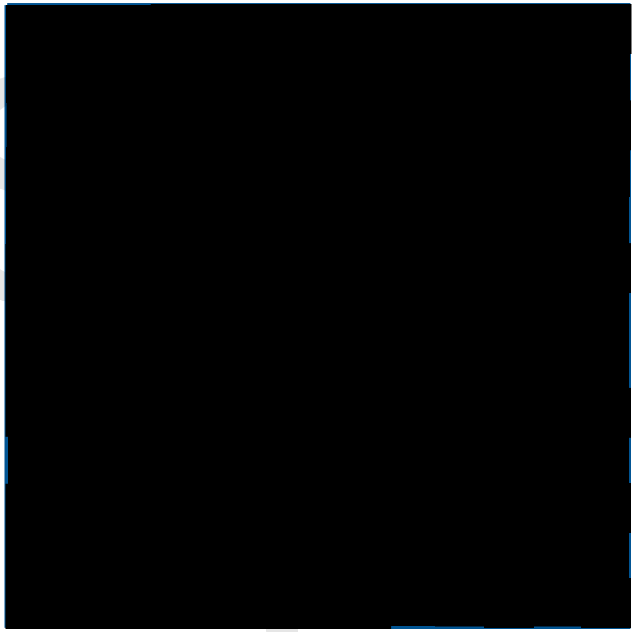


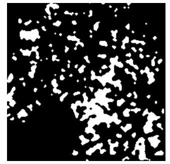
gray scale digital image

- **sub-level set** $X_h := \{x \in X \mid f(x) \le h\}$
- **fattening** $X_{h_1} \subset X_{h_2} \subset \cdots \subset X_{h_T}$

  by $h_1 \le h_2 \le \cdots \le h_T$

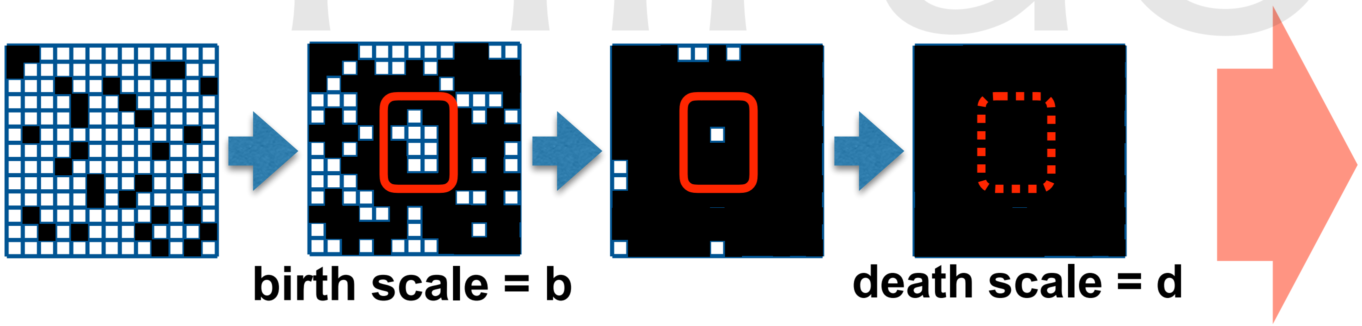
2. Spatial persistence





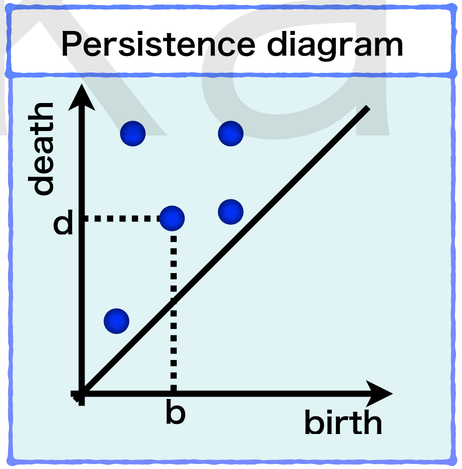
black-white image

Persistence diagram of digital images



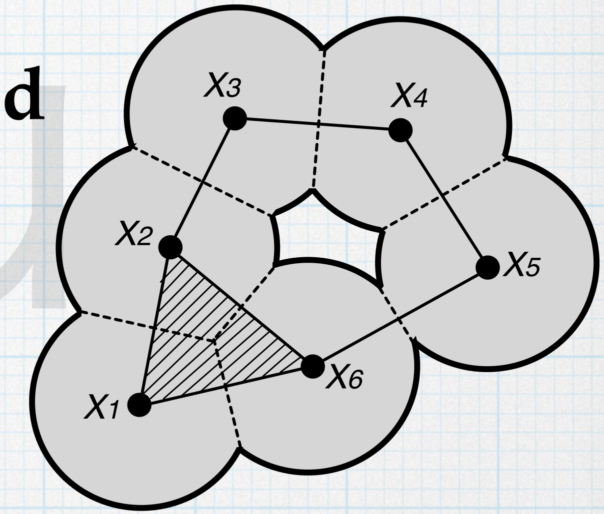
birth scale = b

death scale = d



Characterize grayscale/spatial persistent holes in images

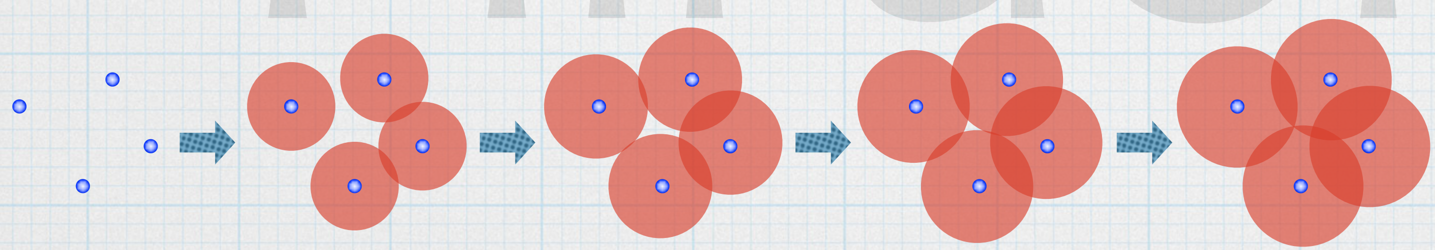
Alpha filtration

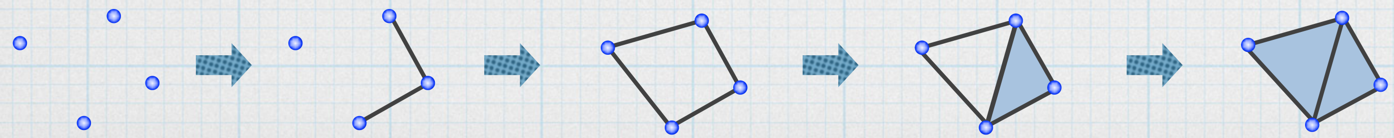


- $X = \{x_i \in \mathbf{R}^m \mid i = 1, \ldots, n\}$ **: point cloud**
- $\mathbf{R}^m = \cup_i V_i$ **: Voronoi decomp.**
- $\cup_i B_i(r) = \cup_i (B_i(r) \cap V_i)$
- **Alpha shape** $\mathcal{A}(X, r)$ **: dual of** $\{B_i(r) \cap V_i \mid i = 1, \ldots, n\}$ **(simplicial complex)**
- **Nerve theorem:** $\cup_i B_i(r) \simeq \mathcal{A}(X, r)$

easy to analyze by computers

- $\mathcal{A}(X, r) \subset \mathcal{A}(X, s)$ **for** $r < s$

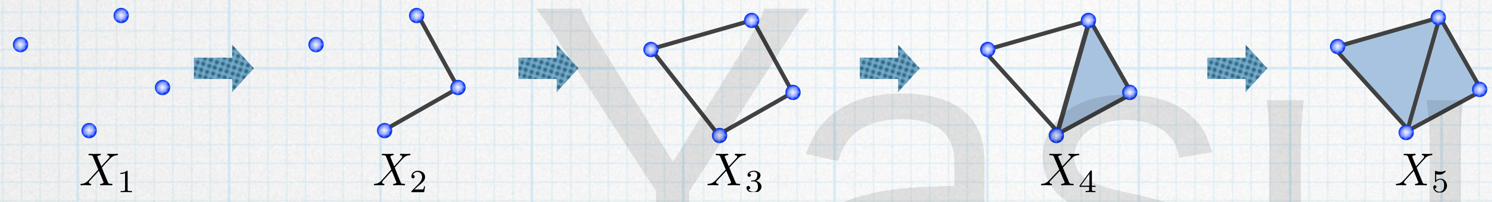




filtration:
changing resolution

Edelsbrunner, Letscher, Zomorodian, Carlsson, de Silva

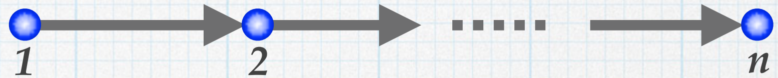
Persistent homology, diagram



$H_1(\mathcal{X}) \simeq I[3,4]$

- **filtration** $\mathcal{X} : X_1 \subset X_2 \subset \cdots \subset X_n$
- **persistent homology** $H_\ell(\mathcal{X}) : H_\ell(X_1) \to H_\ell(X_2) \to \cdots \to H_\ell(X_n)$

representations on A_n



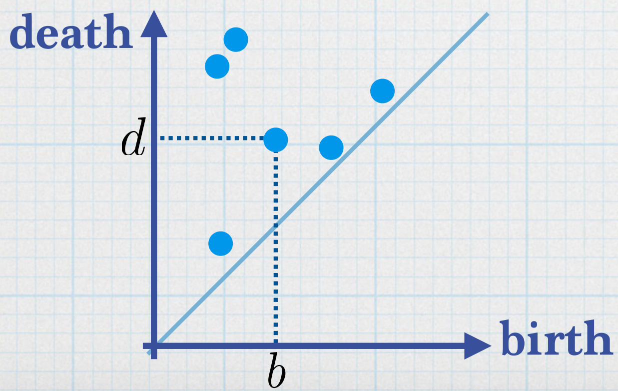
- **interval decomposition (Gabriel Thm, fin.gen. PID module)**

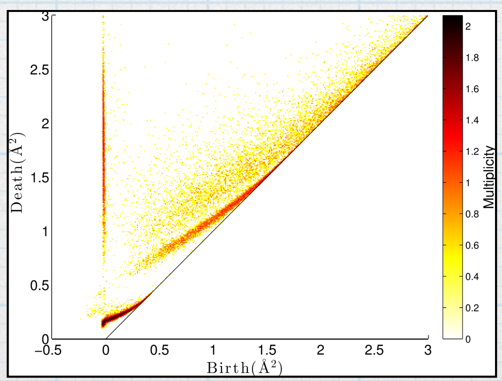
d - b : lifetime (or persistence)

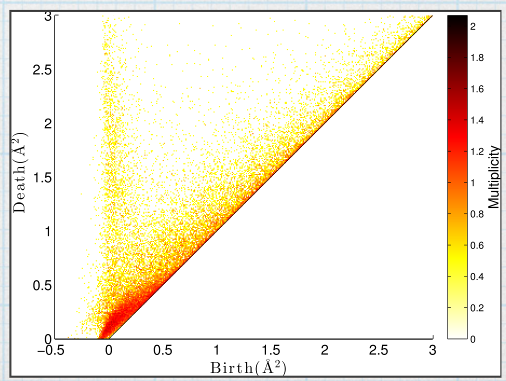
$$H_\ell(\mathcal{X}) \simeq \bigoplus_{i=1}^{s} I[b_i, d_i] \qquad I[b,d] : 0 \to \cdots \to 0 \to \underset{\text{at } X_b}{K} \to \cdots \to \underset{\text{at } X_d}{K} \to 0 \to \cdots \to 0$$

Each interval represents birth & death of a topological feature

- **persistence diagram** $D_k(\mathcal{X}) = \{(b_i, d_i) \in \mathbb{R}^2_{\geq 0} : i = 1, \ldots, p\}$



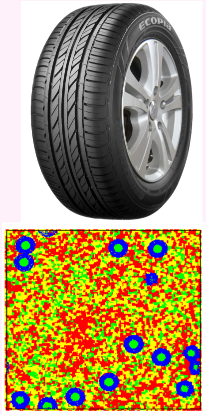




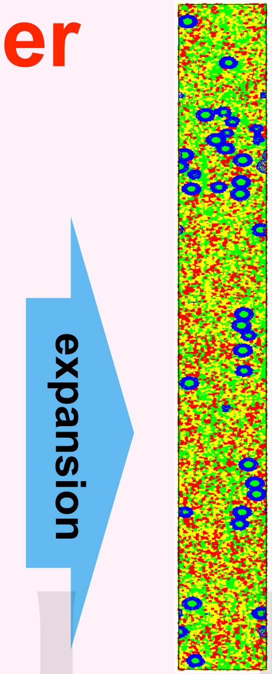
Materials TDA

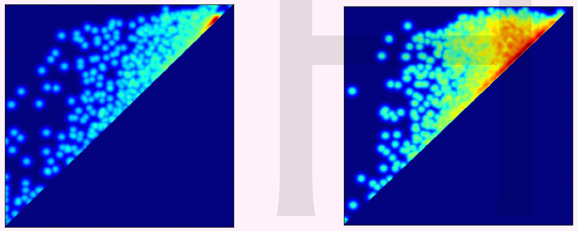
WPI-AIMR, CREST, SIP, MI^2I, NEDO

Polymer

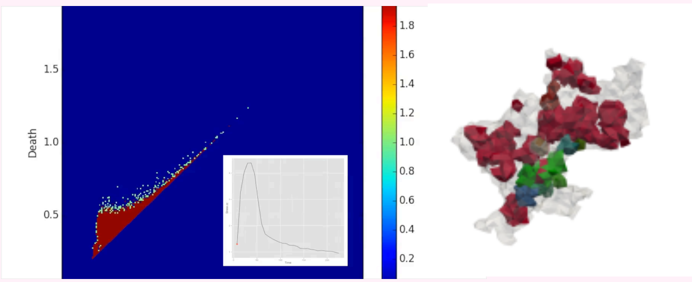


AFM image (Nakajima)





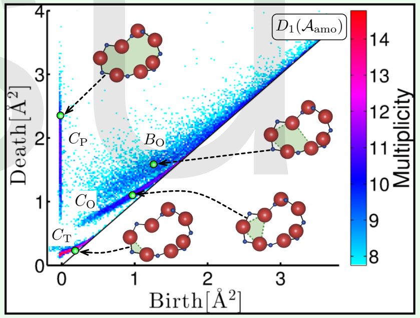
craze formation

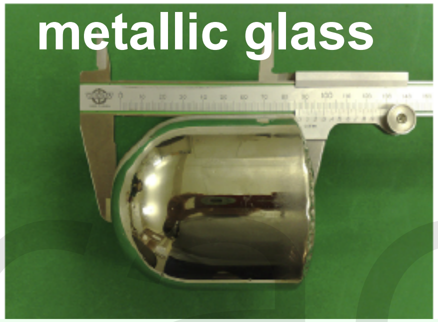


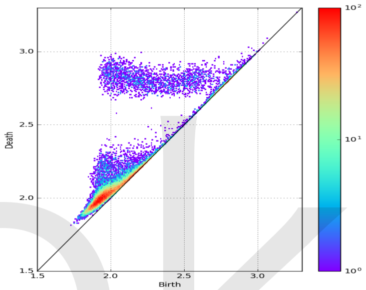
PRE (2017)

Glass

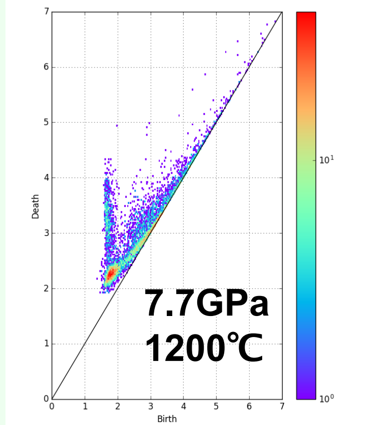
PNAS (2016)

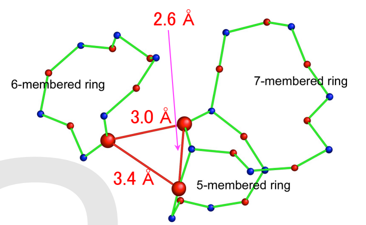






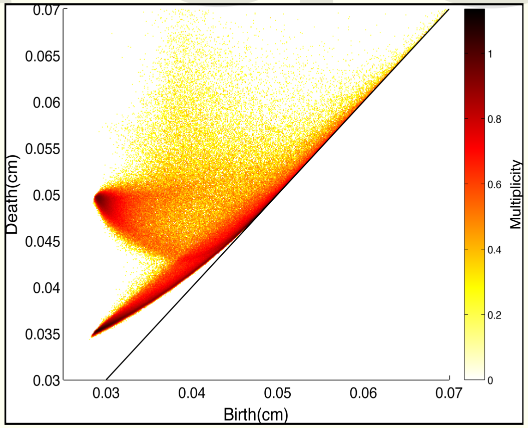
densified silica

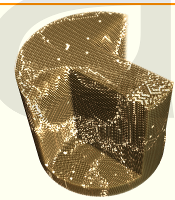




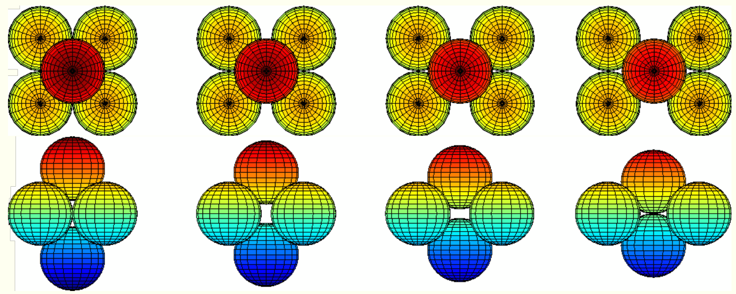
Grain

Nature Communications

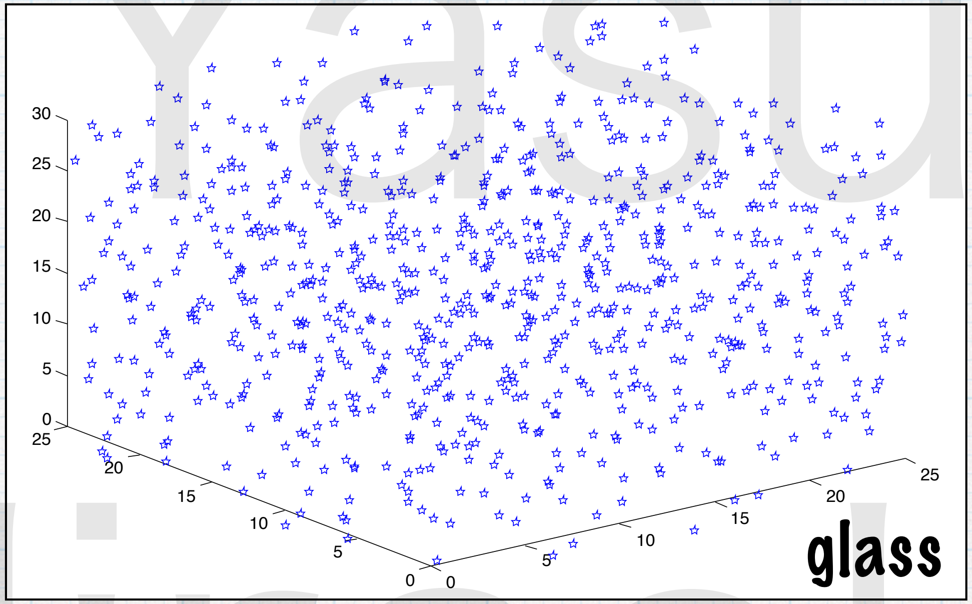




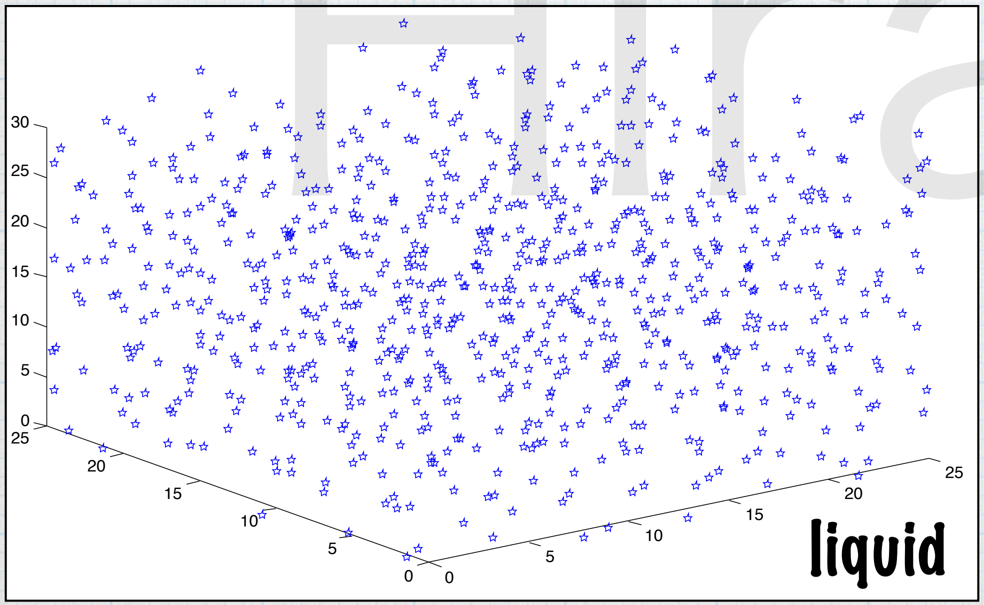
deformation of octa.

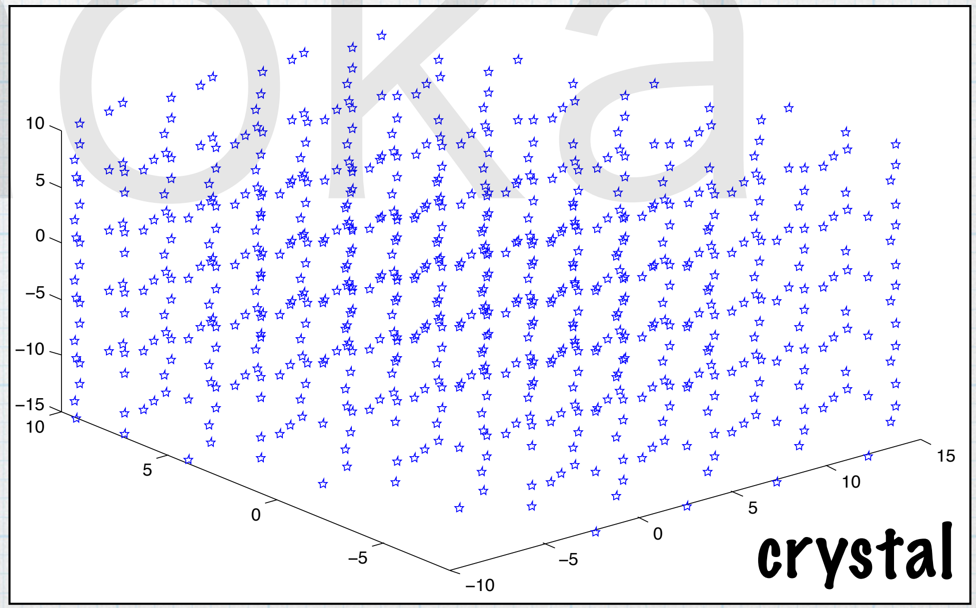


Atomic configurations of silica (SiO2)



obtained by molecular dynamics (MD) simulation

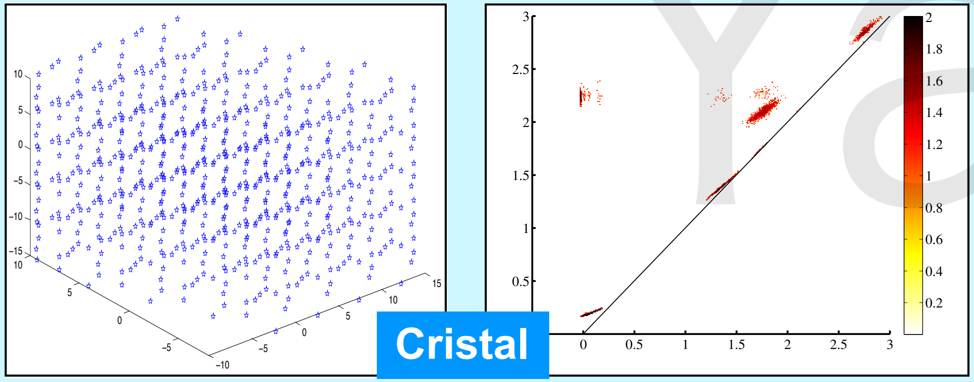


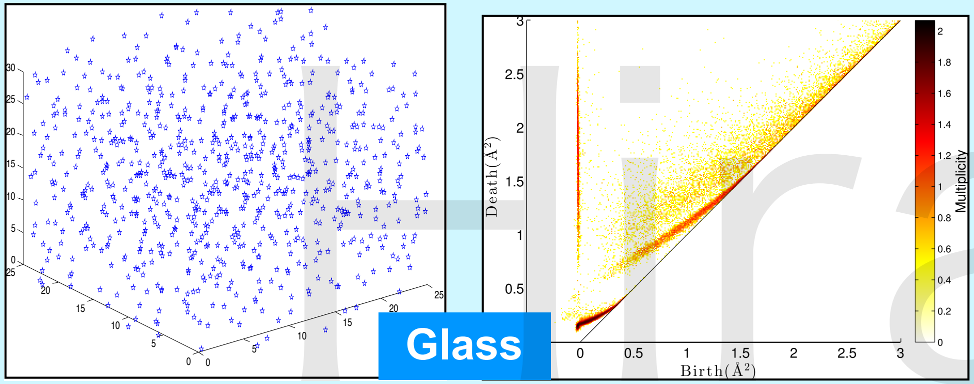


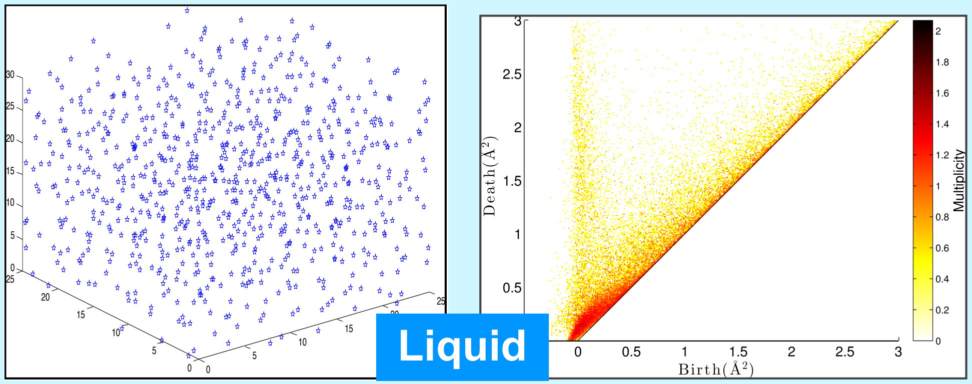
Hierarchical Structural Analysis of Silica Glass

with Nakamura, Hirata, Escolar, Matsue, Nishiura PNAS (2016) CREST TDA, SIP

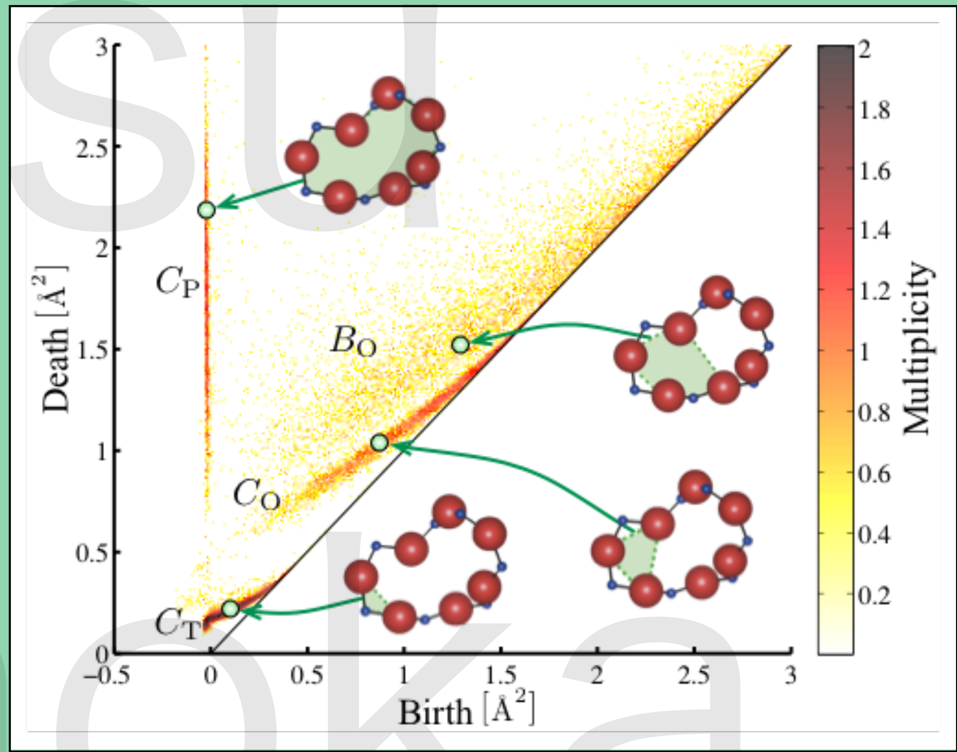
MD and PD1







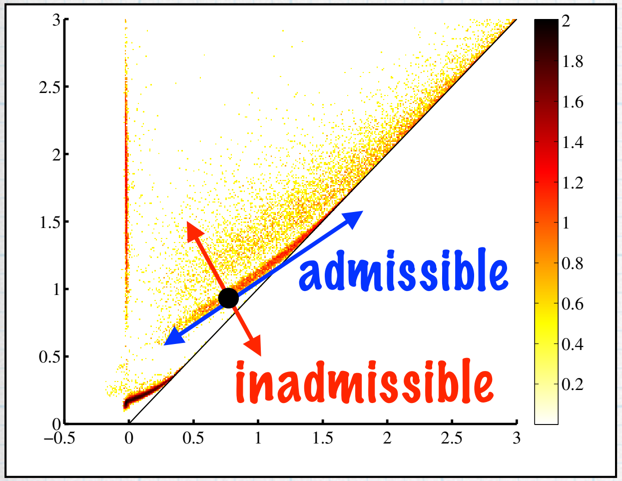
Inverse Analysis

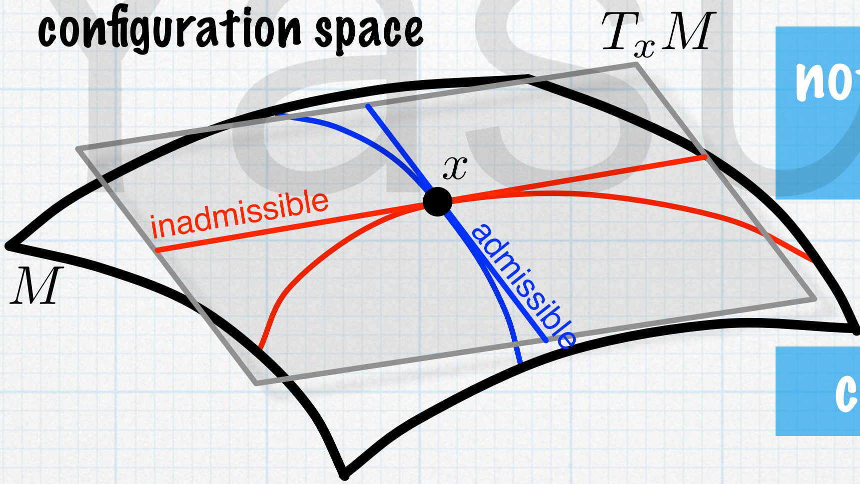


- Glass contains curves in PD
- Curves express geometric constraints (orders) of atomic configurations
- Inverse analysis reveals hierarchical ring structures
- PD multi-scale analysis characterizes inter-tetrahedral O-O orders (curve Co)
- Useful tool for structural analysis

Y.H., et al. PNAS (2016)

Curves and constrains



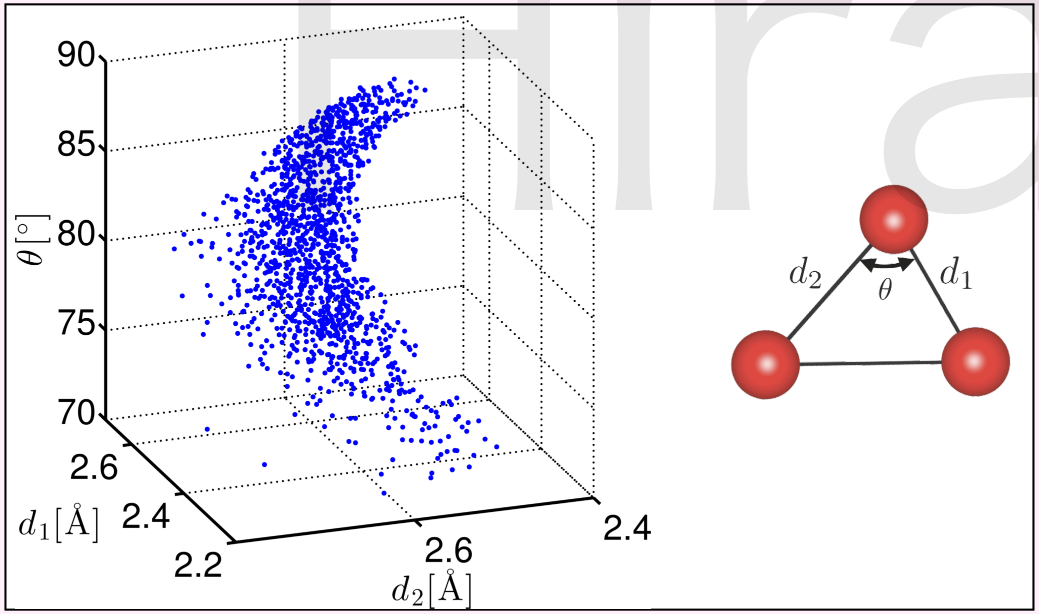


not allowed to move in the normal direction

rigidity

constrains of the rings

O-O-O structures (MRO)



- **O-O-O ring constrains are discovered**
- **necessary to study both distance and angle distributions simultaneously (conventional methods cannot detect)**