SAND2022-6897C

Sandia National Laboratories

Low-Dimensional Representation of the Short- and Medium-Range Structure of Glassy Materials

Thomas J. Hardin (Sandia National Laboratories)

Mark Wilson (Sandia National Laboratories)

Michael Chandross (Sandia National Laboratories)

Michael Falk (Johns Hopkins University)

Michael Shields (Johns Hopkins University)

with thanks to

Christopher Schuh (Massachusetts Institute of Technology)

Tess Smidt (Massachusetts Institute of Technology)

Yannis Kevrekidis (Johns Hopkins University)

May 24, 2022 ● Glass and Optical Materials Division Annual Meeting

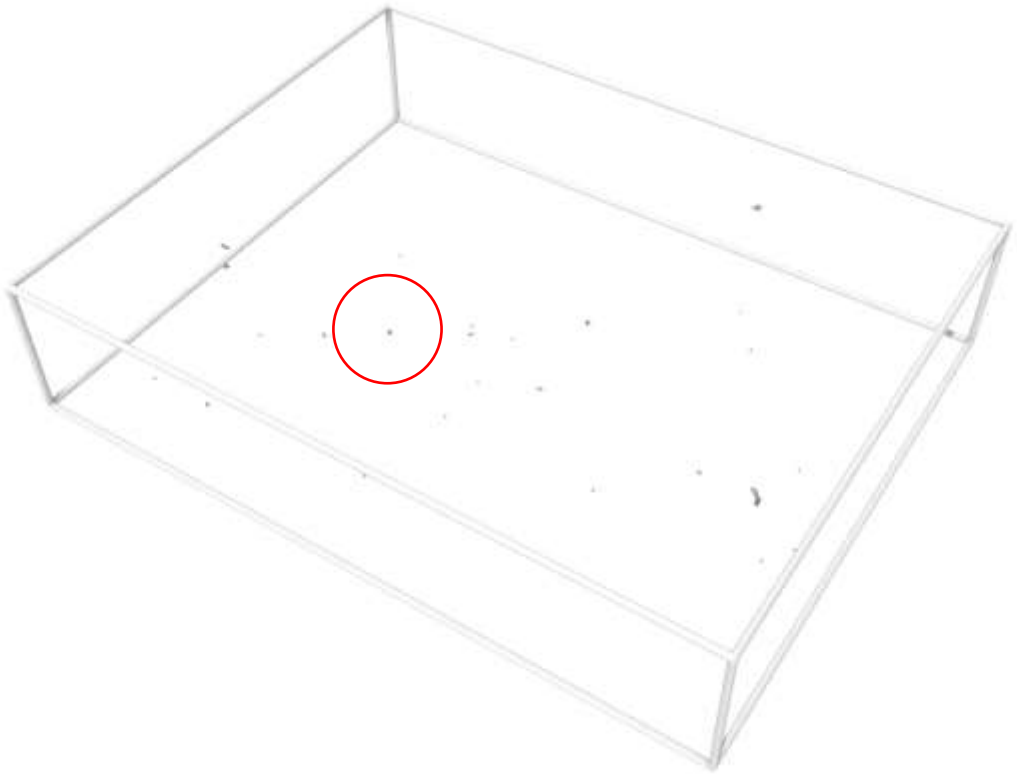
UUR - ***

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc. for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

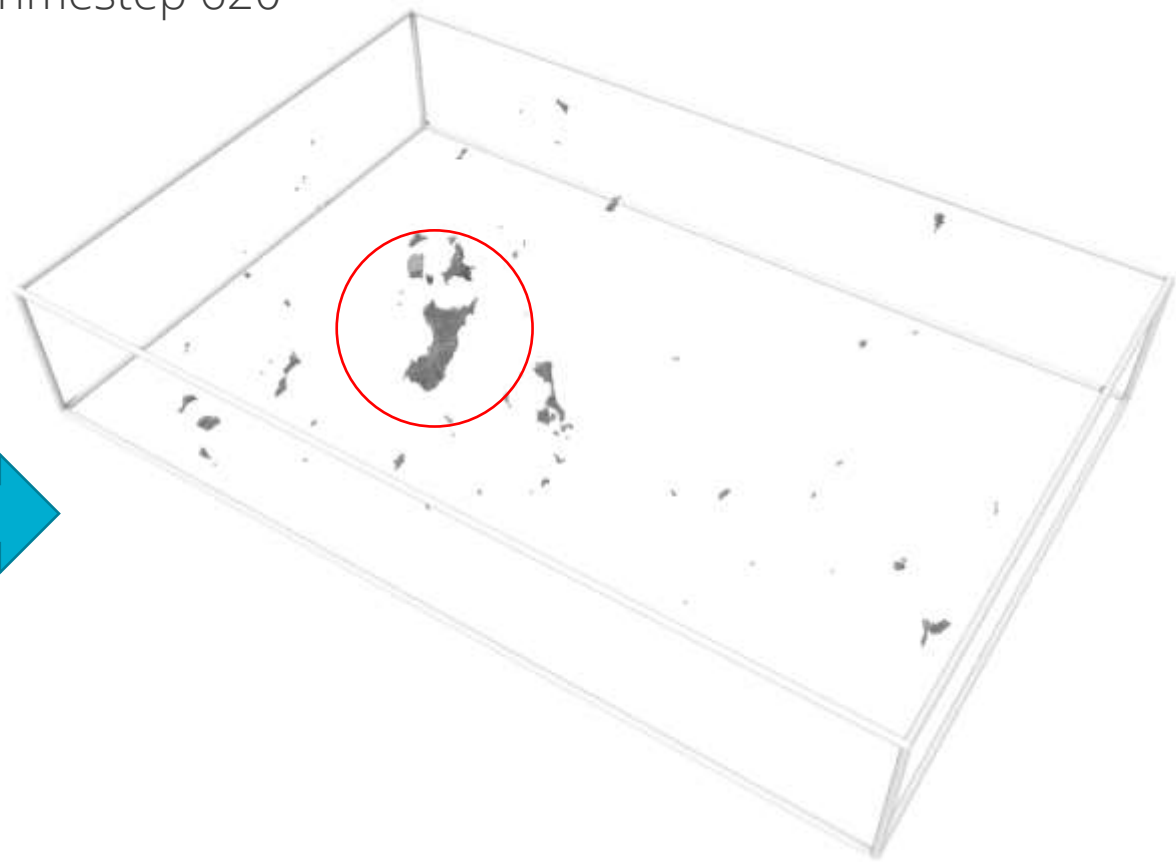
Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

Machine Learning to Predict Crack Nucleation in Silica Glass

Timestep 0



Timestep 620

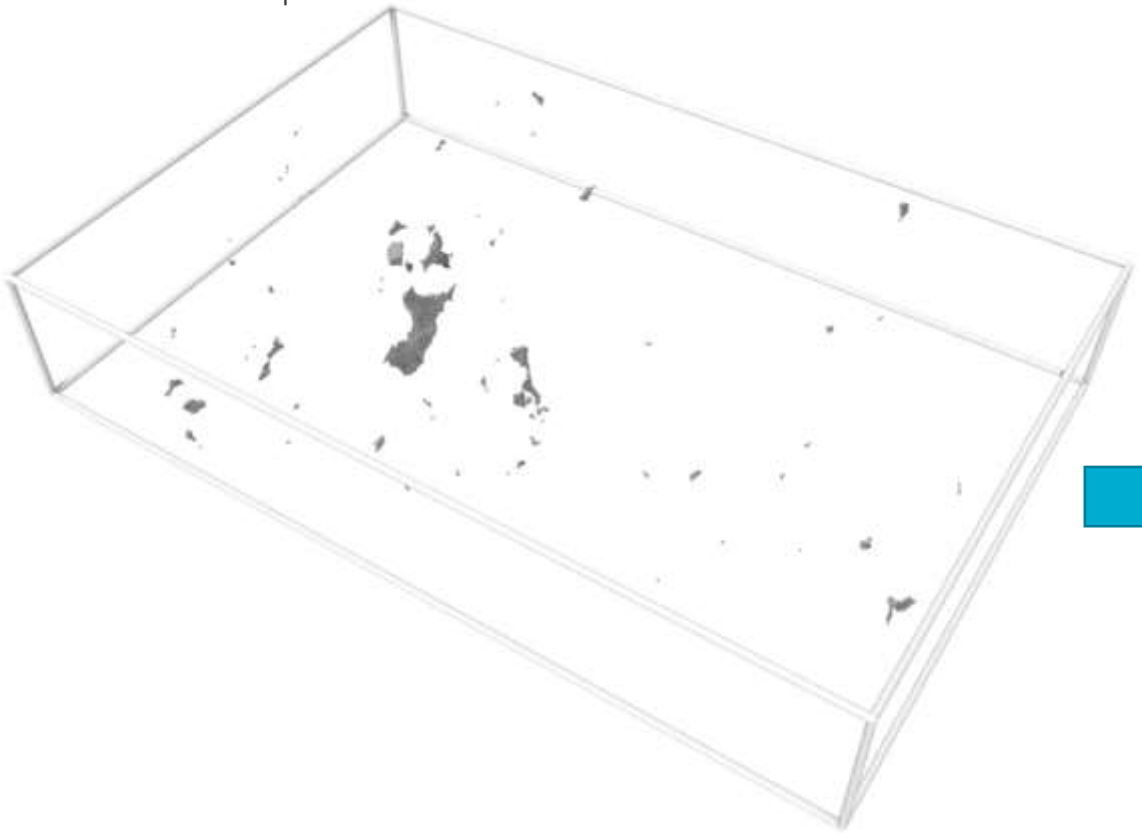


In roughly ½ of our samples, crack nuclei could be traced back to a void at step 0.

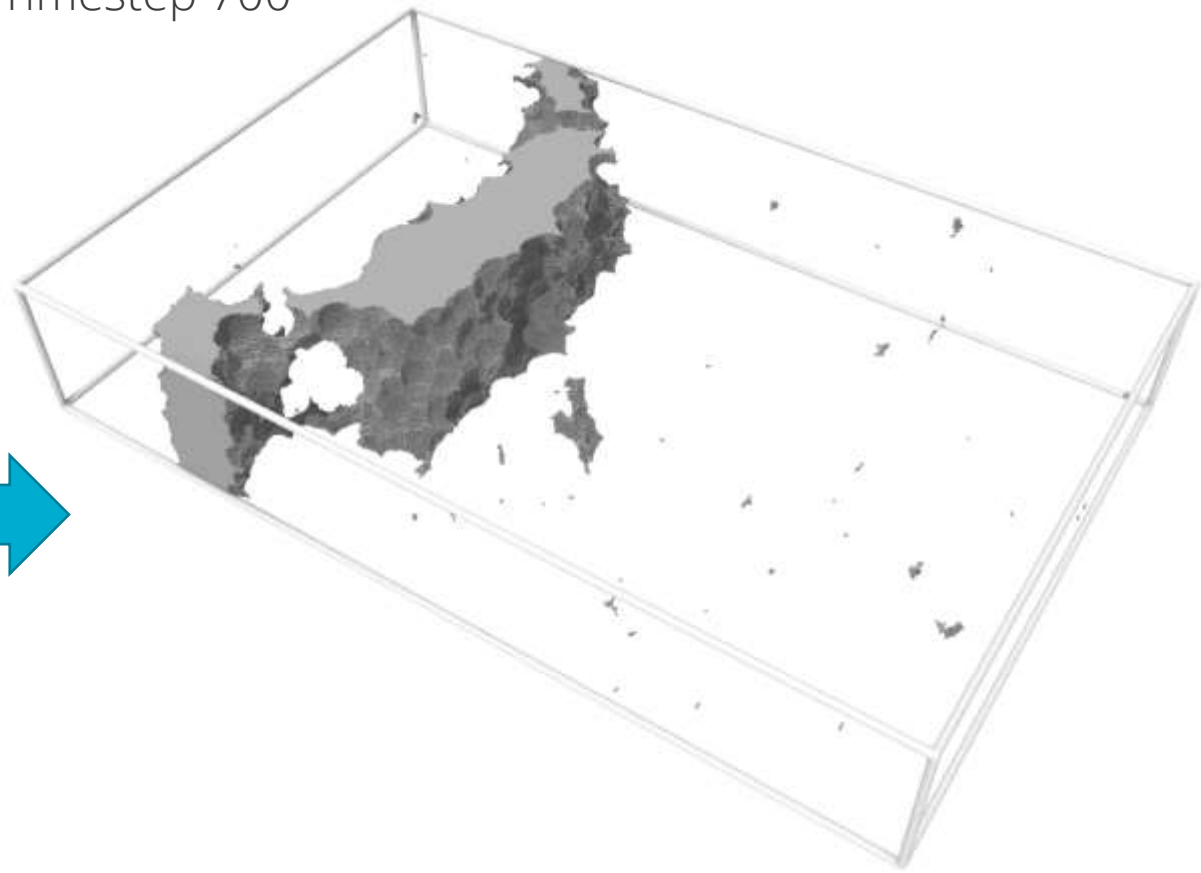
Thanks to Mark Wilson (Sandia) and Allon Percus (Claremont Graduate University)

Machine Learning to Predict Crack Nucleation in Silica Glass

Timestep 620



Timestep 700

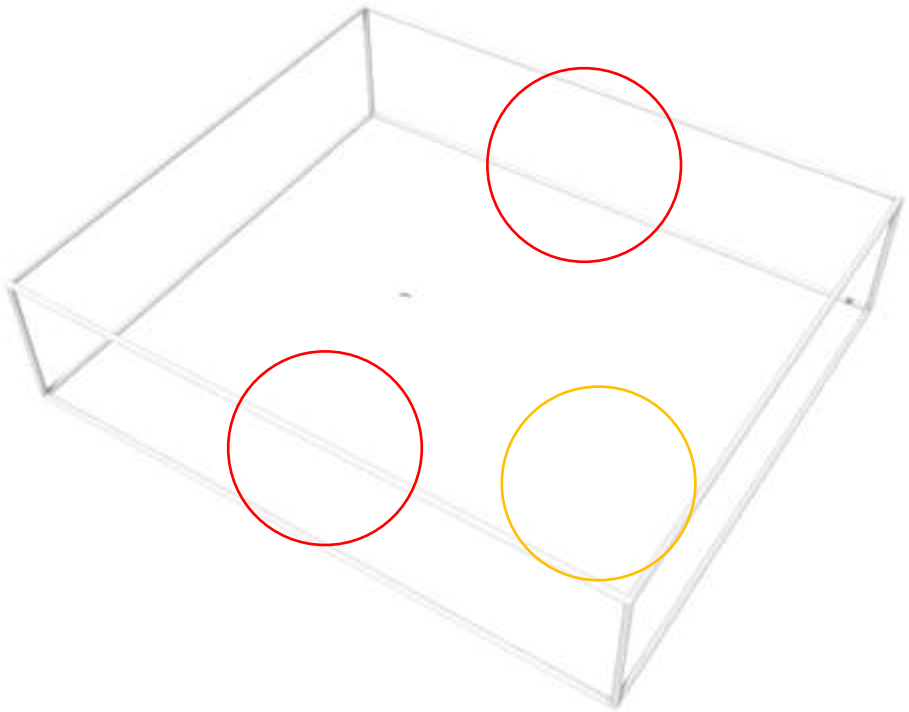


In roughly ½ of our samples, crack nuclei could be traced back to a void at step 0.

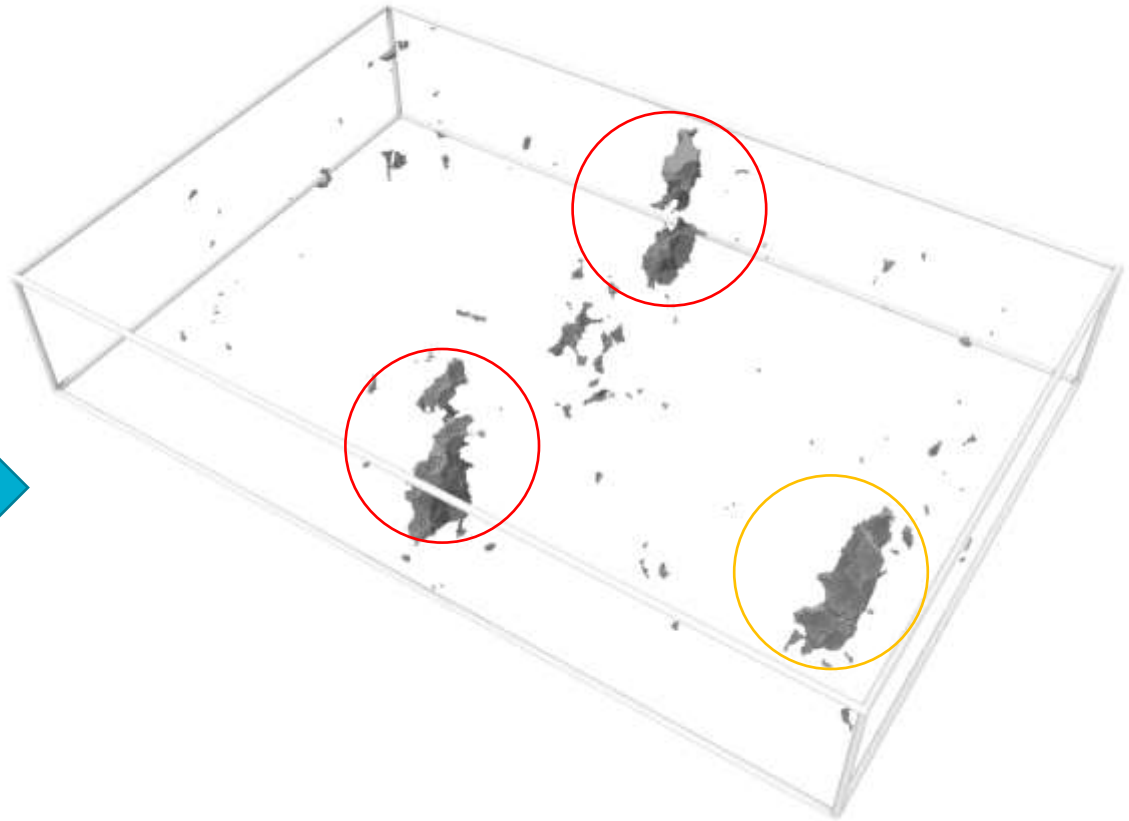
Simulations by Mark Wilson (Sandia), graph analysis by Allon Percus (Claremont Graduate University)

Machine Learning to Predict Crack Nucleation in Silica Glass

Timestep 0



Timestep 640



In the other ½ of our samples, crack nuclei formed away from discernible voids.

Machine Learning to Predict Crack Nucleation in Silica Glass

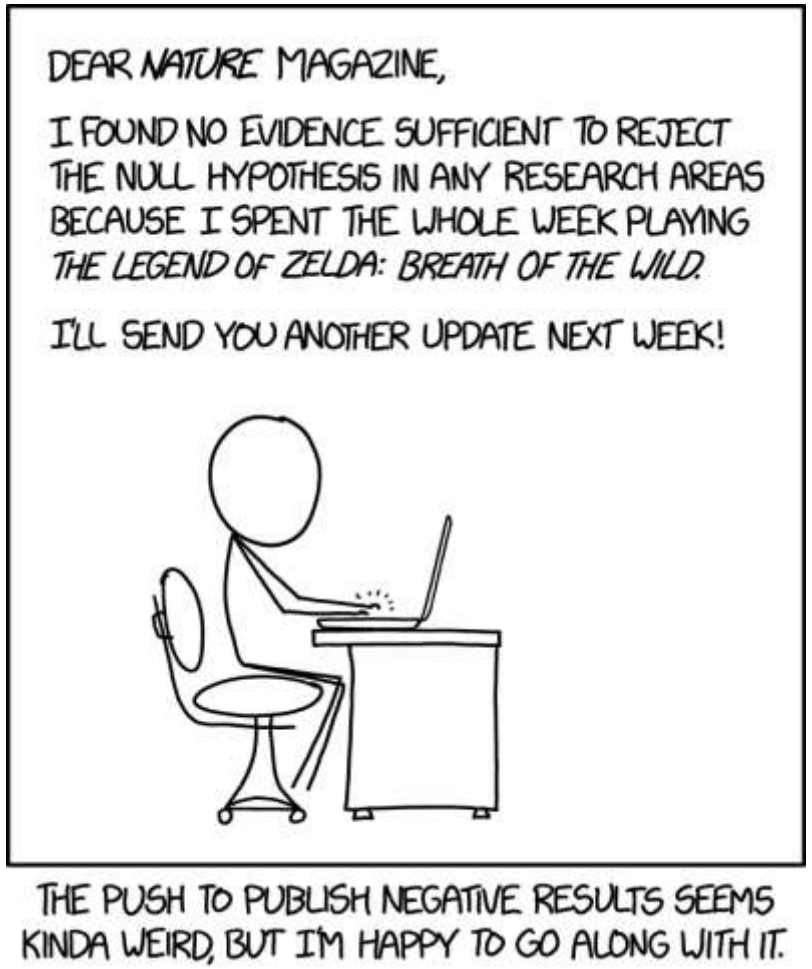
We tried all the intuitive and mathematically elegant features we could think of →

- Kinetic energy
- Potential energy
- Strain energy
- Stress components
- Voronoi volume
- # nearest neighbors (1nn-4nn)
- Average degree (1nn-4nn)
- Min bond angle
- Max bond angle
- Coordination number
- # bridging oxygens
- Non-affine displacement
- Graph centrality
- ...and many others

Machine Learning to Predict Crack Nucleation in Silica Glass

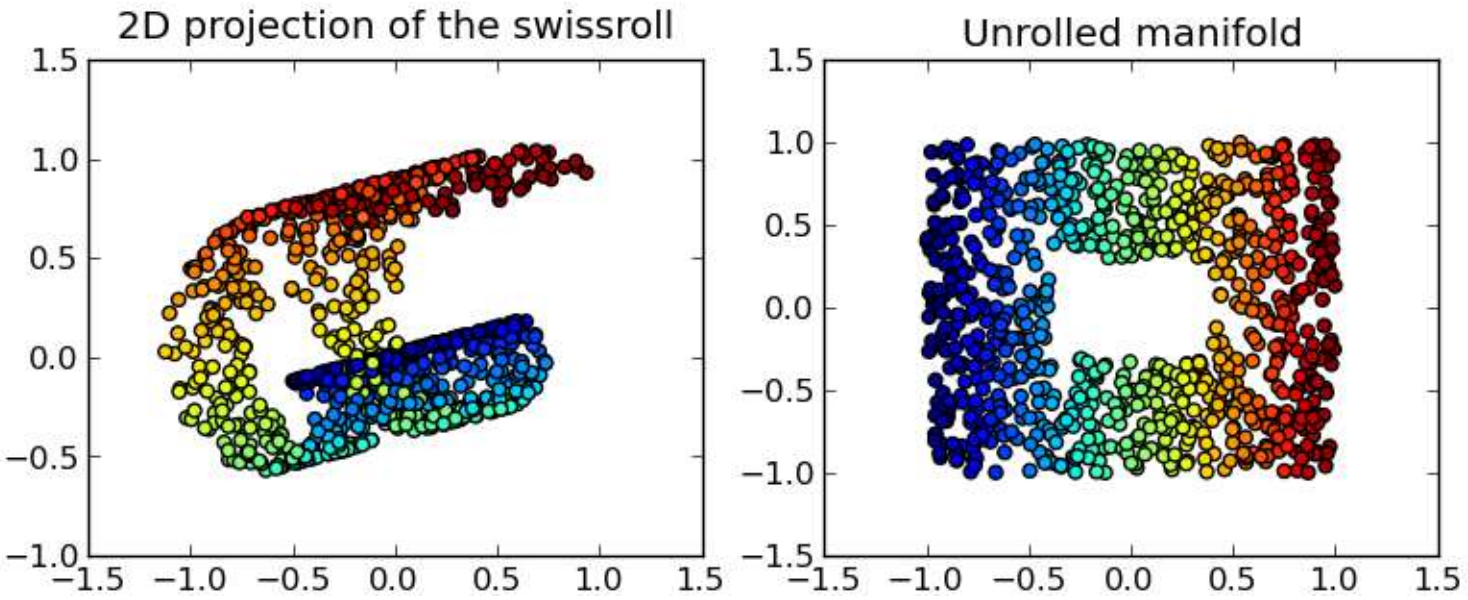
- The set of obvious, easy, or elegant observables in a glass is NOT necessarily adequate to predict, model, or understand the material.

- We have turned to a data-driven structural description of glass, sacrificing human intuition for descriptive completeness.



https://xkcd.com/2020/

Concept: Manifold Learning / Dimensionality Reduction

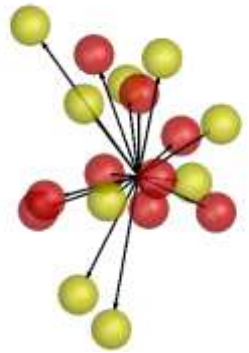


Given a set of points $\mathcal{X} \in \mathbb{R}^n$,
learn a function $f: \mathbb{R}^n \longrightarrow \mathbb{R}^q$ where $q < n$

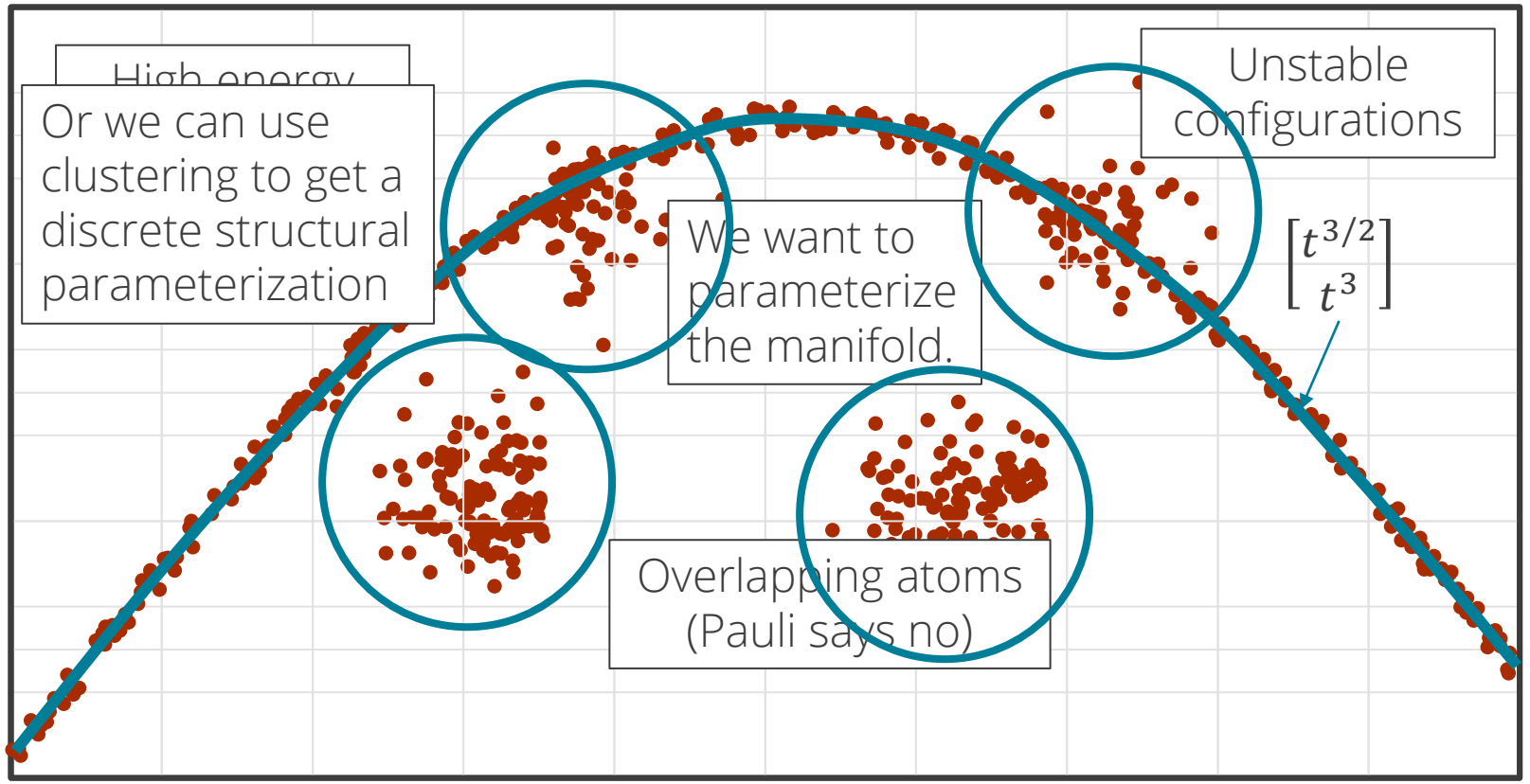
- where neighboring points in $\mathcal{X}$ are neighbors in $f(\mathcal{X})$
- where $f(\mathcal{X})$ captures the important information in $\mathcal{X}$

Concept: Manifold Learning / Dimensionality Reduction

We need 3n real numbers to naively quantify an atomic configuration, plus n integers to encode species information.

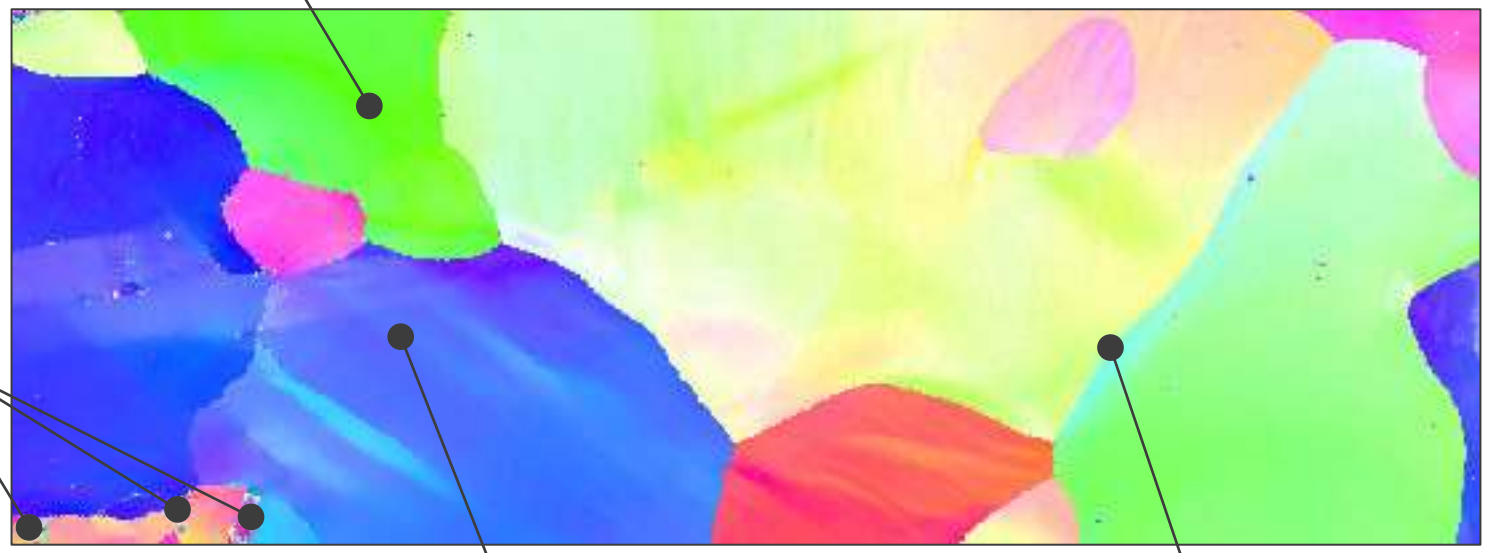


So imagine we're plotting configs cut out of some material on axes spanning $\mathbb{R}^{3n} \times \mathbb{N}$.



We only observe atomic configurations in certain regions of $\mathbb{R}^{3n} \times \mathbb{N}$.
Enthalpy pulls atomic configurations onto a manifold.
Entropy and kinetics spread atomic configurations out on that manifold.

A Wildly Successful Example of Dimensionality Reduction



Dimensionality reduction strategy

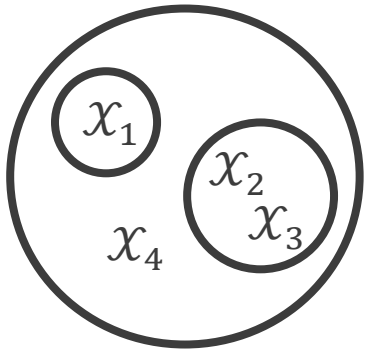
Cut out environments

$$\mathcal{X} = \begin{bmatrix} \boldsymbol{X}_1 & s_1 \\ \vdots & \vdots \\ \boldsymbol{X}_n & s_2 \end{bmatrix}$$

This is the real technical challenge →

Compute **generalized distances** between pairs of environments

$$\boldsymbol{D} = \begin{bmatrix} d(\mathcal{X}_1, \mathcal{X}_1) & \cdots & d(\mathcal{X}_N, \mathcal{X}_1) \\ \vdots & \ddots & \vdots \\ d(\mathcal{X}_1, \mathcal{X}_N) & \cdots & d(\mathcal{X}_N, \mathcal{X}_N) \end{bmatrix}$$



Hierarchical Clustering *discrete parameterization of manifold*

Diffusion Maps *continuous parameterization of manifold*

$$f\left(\begin{bmatrix} \boldsymbol{X}_1 & s_1 \\ \vdots & \vdots \\ \boldsymbol{X}_n & s_2 \end{bmatrix}\right) = \begin{bmatrix} y_1 \\ \vdots \\ y_q \end{bmatrix} \qquad (q \ll 4n)$$

Properties of a Generalized Distance/Dissimilarity Function

Continuity and **Smoothness**: the generalized distance is stable with respect to small atomic perturbations.

- An atom jumping across the cutoff radius won't dramatically change the result

Completeness: the generalized distance between two configurations is zero iff the two configurations are equivalent.

Rotation Invariance: the generalized distance is the same regardless of frame.

Permutation Invariance: the distance is the same regardless of order of atoms.

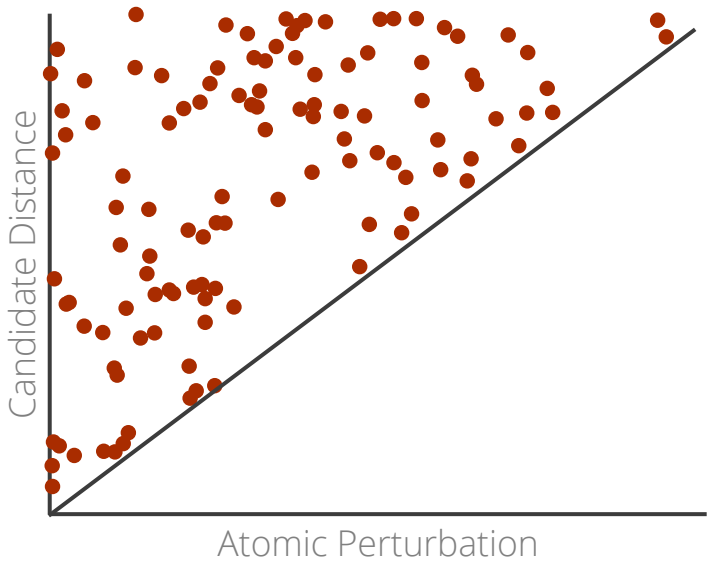
Tolerates Variable Numbers of Atoms: calculates meaningful distances between environments with different numbers of atoms.

Differentiable: the distance can be differentiated with respect to atomic positions. Important for empirical potentials.

Speed: the distance can be quickly calculated between a pair of atomic configurations.

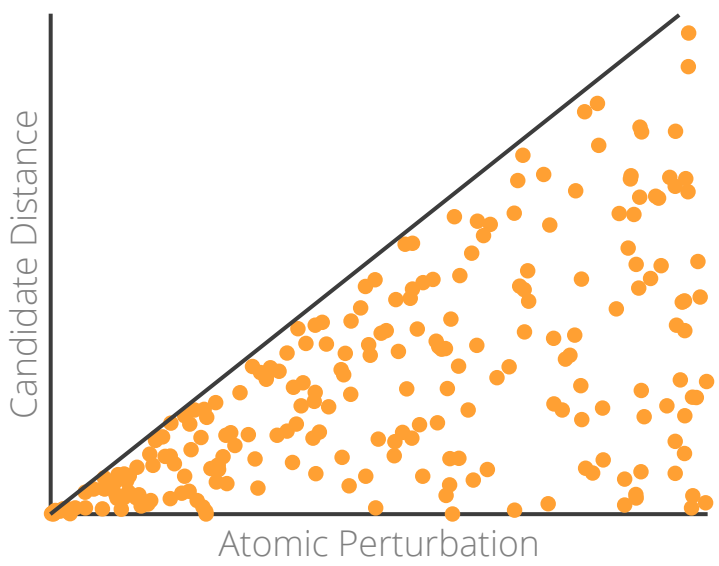
Thinking about continuity and completeness

Discontinuous



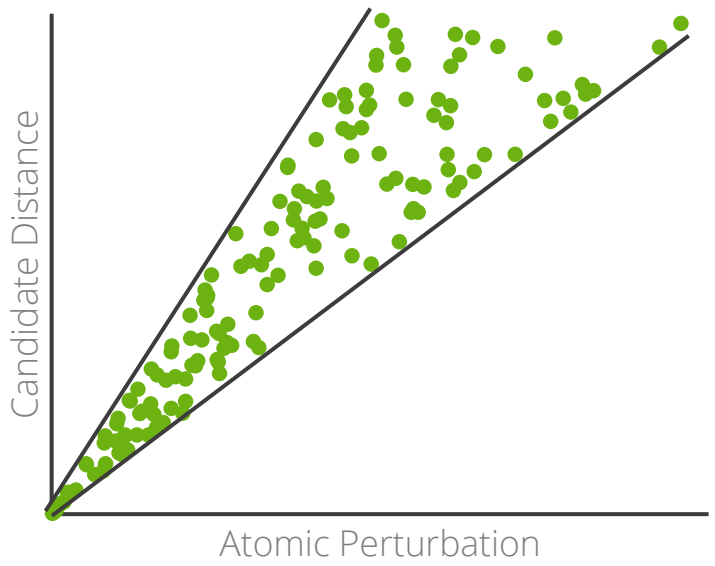
- Isograph
- Careless cutoff radius handling
- Careless rotation invariance
- Careless permutation invariance

Incomplete



- Bispectrum
- SOAP
- ACE

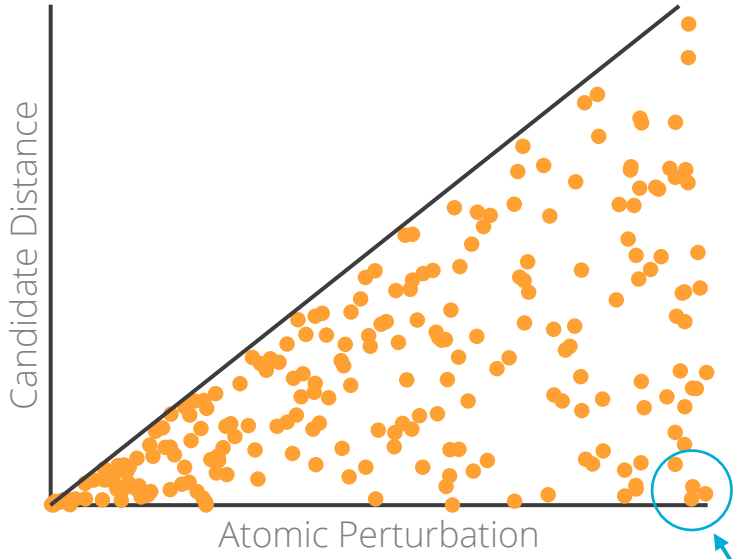
Continuous + Complete



- GIIP

The trouble with an incomplete distance function

Demo- the vegetable cutting mat of science



Fails to distinguish between dissimilar atomic configurations.

Gaussian Integral Inner Product (GIIP) Distance

Integral inner product of two functions

(1) $$\langle a, b \rangle = \int_{\mathbb{R}^3} a(\mathbf{x}') \cdot b(\mathbf{x}') \, d\mathbf{x}'$$

Norm of a function induced by the inner product above

(2) $$||a|| = \sqrt{\langle a, a \rangle}$$

Gaussian function with standard deviation sigma, normalized to 1

(3) $$G_\sigma(\mathbf{x}) = \exp\left[-|\mathbf{x}|^2/(2\sigma^2)\right] / (\pi^{3/4}\sigma^{3/2})$$

Atomic density function consisting of weighted Gaussians centered on atomic positions

(4) $$\rho_{\mathcal{X}}(\mathbf{x}) = \sum_{\mathbf{x}' \in \mathcal{X}} w_{\mathbf{x}'} G_{\sigma_{\mathbf{x}'}}(\mathbf{x} - \mathbf{x}')$$

Gaussian Integral Inner Product (GIIP) between two configurations

(5) $$\left\langle \mathcal{X}^\alpha, \mathcal{X}^\beta \right\rangle = \left\langle \rho_{\mathcal{X}^\alpha}, \rho_{\mathcal{X}^\beta} \right\rangle$$

Distance between two configurations can be calculated with three inner products

(6) $$\left|\mathcal{X}^\alpha - \mathcal{X}^\beta\right|^2 = \left\langle \mathcal{X}^\alpha - \mathcal{X}^\beta, \mathcal{X}^\alpha - \mathcal{X}^\beta \right\rangle = \left\langle \mathcal{X}^\alpha, \mathcal{X}^\alpha \right\rangle + \left\langle \mathcal{X}^\beta, \mathcal{X}^\beta \right\rangle - 2 \cdot \left\langle \mathcal{X}^\alpha, \mathcal{X}^\beta \right\rangle$$

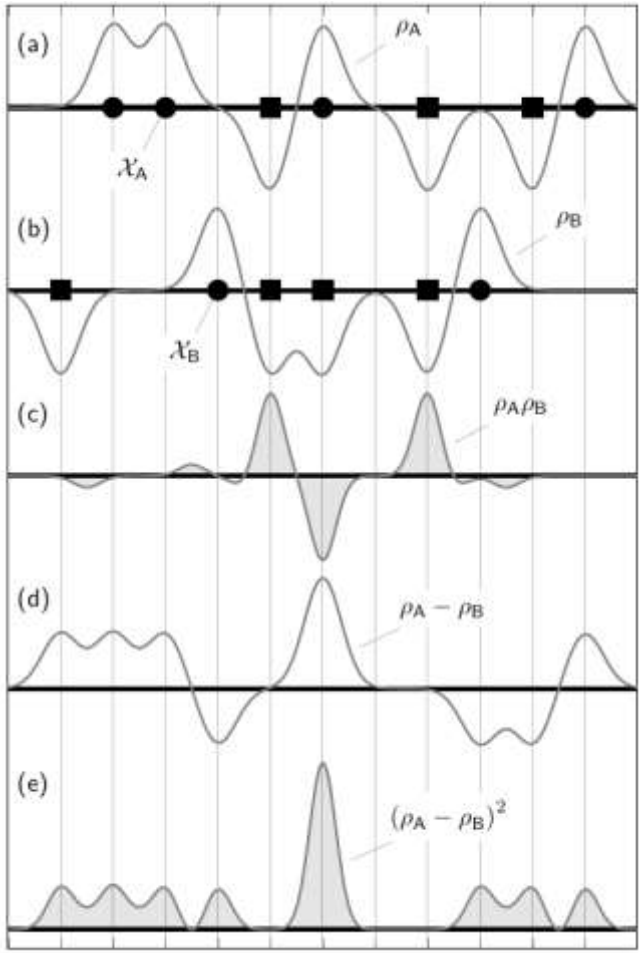
Orientation invariance by minimizing over all possible rotations/rotoinversions

(7) $$\min_{\mathbf{R} \in \mathrm{O}(3)} \left|\mathcal{X}^\alpha - \mathbf{R}\mathcal{X}^\beta\right|^2$$

GIIP is analytically tractable in a computationally convenient form

(8) $$\left\langle \mathcal{X}^\alpha, \mathcal{X}^\beta \right\rangle = 2\sqrt{2} \sum_{\mathbf{x}^\alpha \in \mathcal{X}^\alpha} \sum_{\mathbf{x}^\beta \in \mathcal{X}^\beta} w_{\mathbf{x}^\alpha} w_{\mathbf{x}^\beta} \left(\frac{\sigma_{\mathbf{x}^\alpha} \sigma_{\mathbf{x}^\beta}}{\sigma_{\mathbf{x}^\alpha}^2 + \sigma_{\mathbf{x}^\beta}^2} \right)^{3/2} \exp\left[-|\mathbf{x}^\alpha - \mathbf{x}^\beta|^2 / (2\sigma_{\mathbf{x}^\alpha}^2 + 2\sigma_{\mathbf{x}^\beta}^2)\right]$$

Understanding the GIIP Distance – One-dimensional Example



Atomic density function for atomic configuration 1

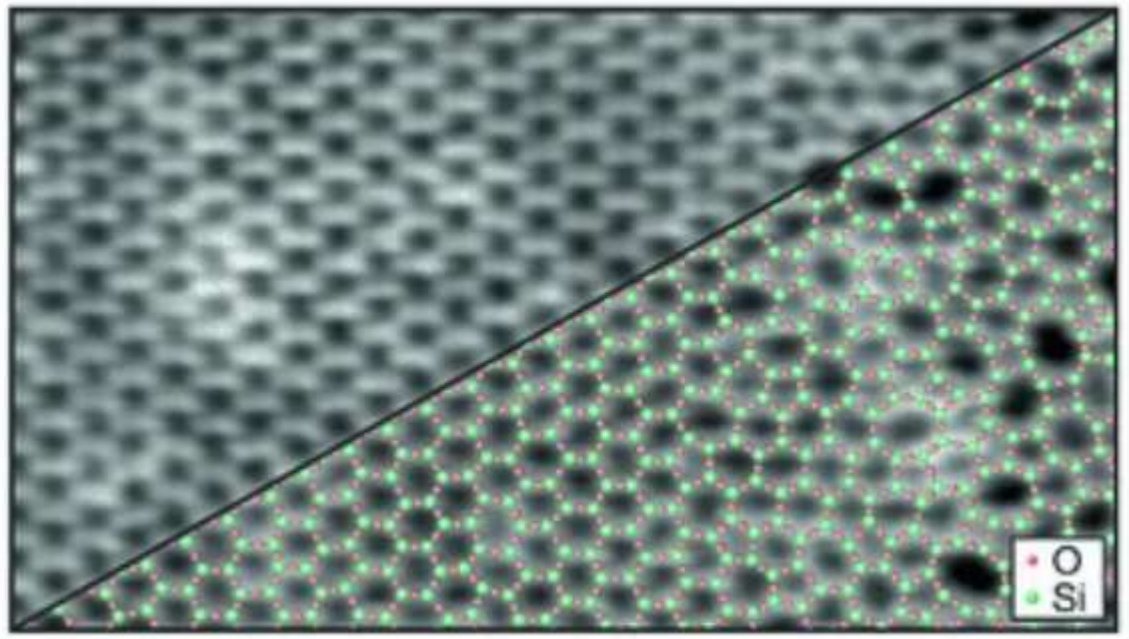
Atomic density function for atomic configuration 2

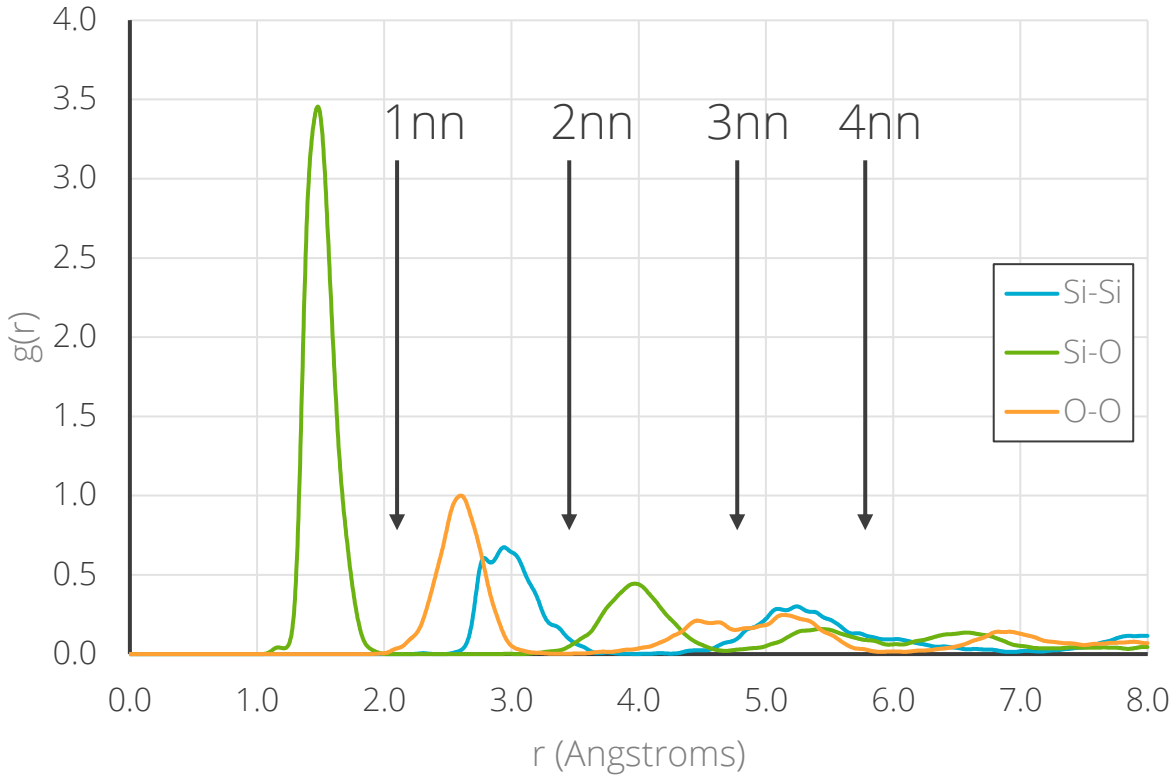
The GIIP between atomic configurations 1 and 2 is the integral of the product of their respective atomic density functions.

To find the squared GIIP distance between atomic configurations 1 and 2, take the difference between their respective atomic density functions, square it, and integrate the squared difference.

Experimental two-dimensional silica

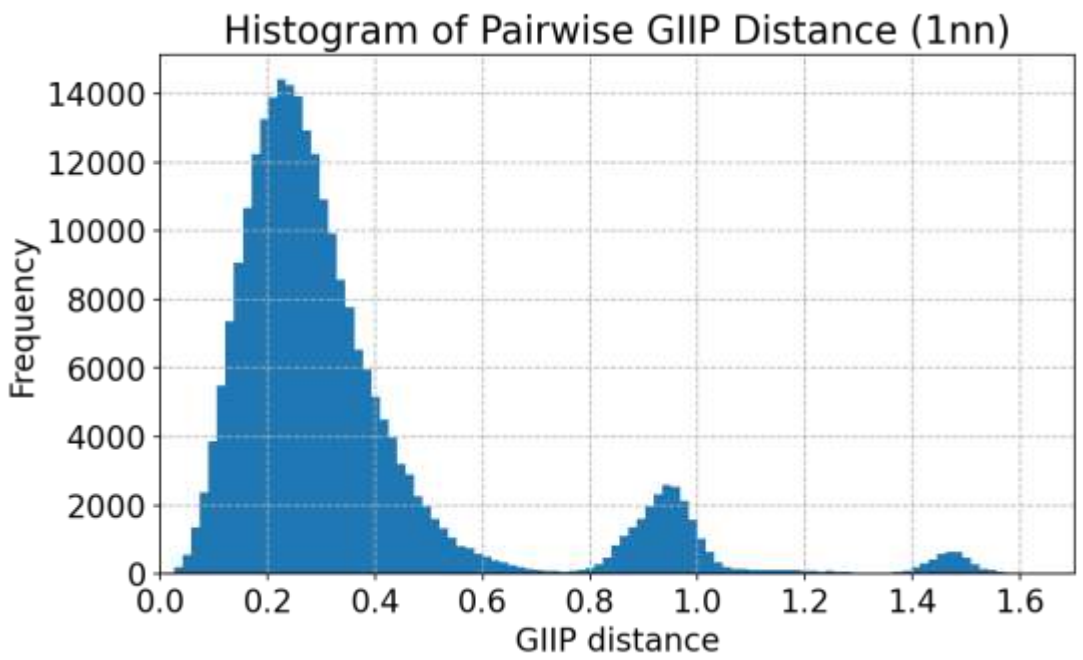
STM image of metal-supported silica bilayer

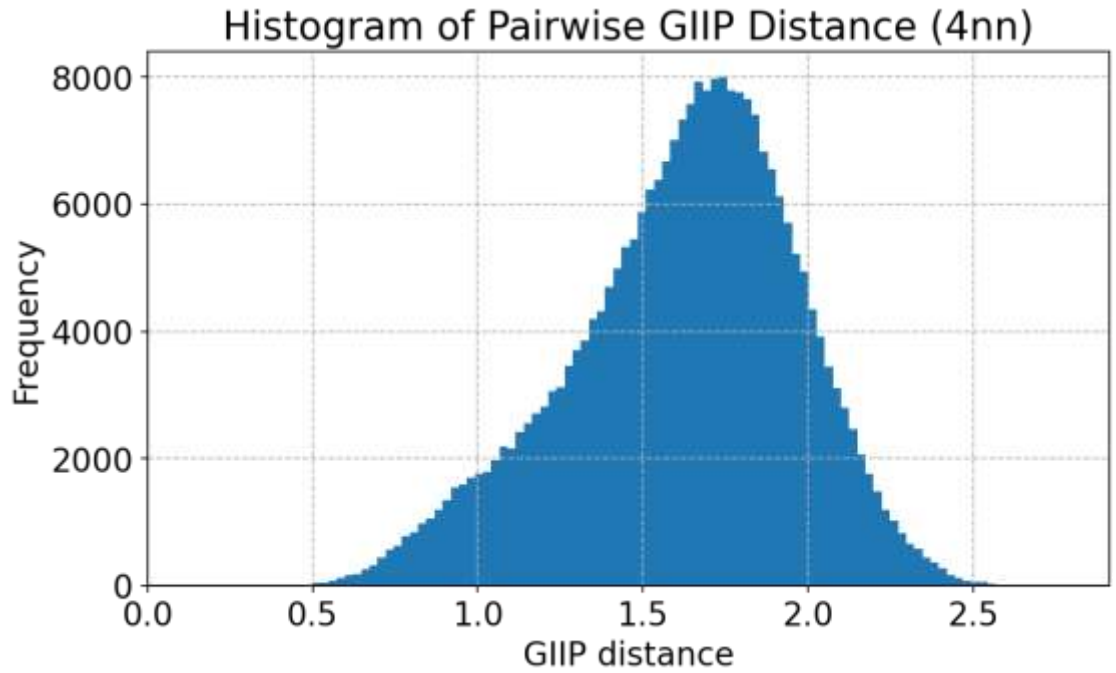




Microscopy published by Buechner, Lichtenstein, Heyde, Freund (2015)
Atomistic data extracted by Franz Bamer (RWTH Aachen University)

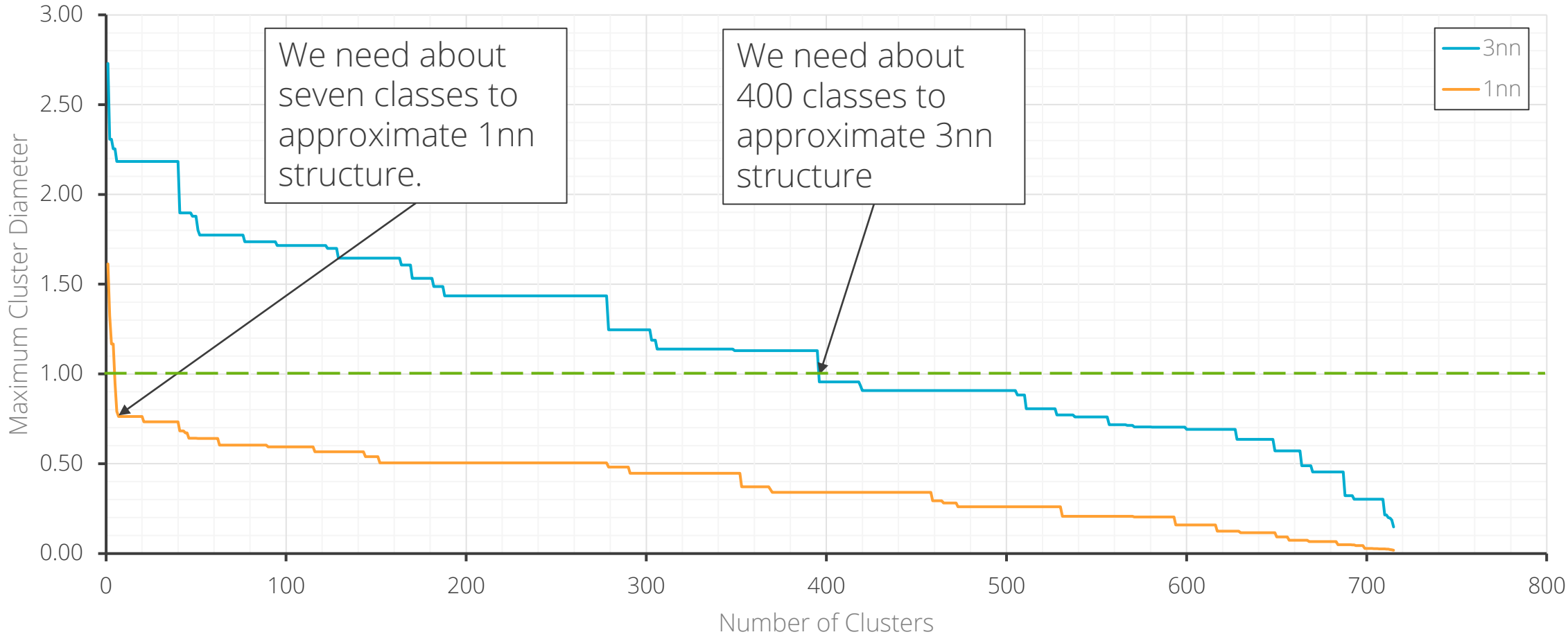
Experimental two-dimensional silica: GIIP distance histogram





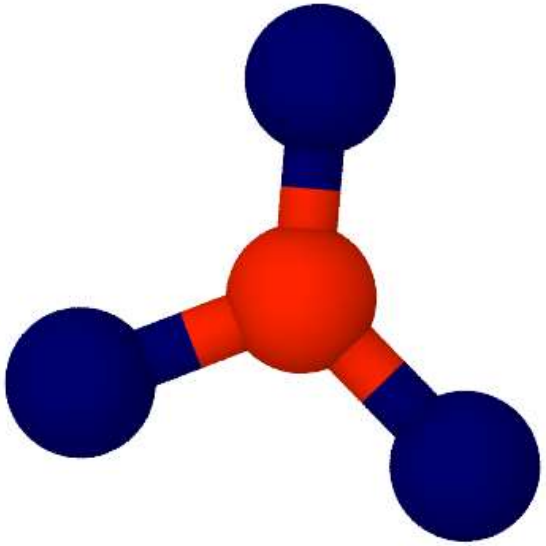
Experimental two-dimensional silica – hierarchical clustering

Hierarchical clustering lets us divide the data into any number of classes.

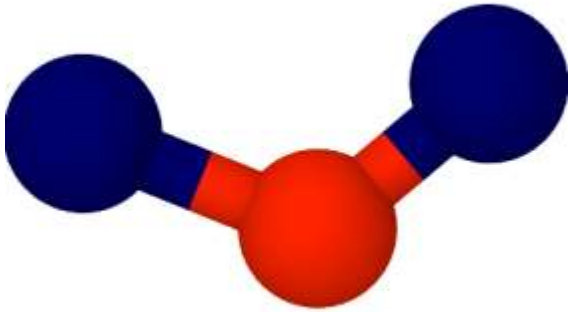


Experimental two-dimensional silica – 1nn clusters

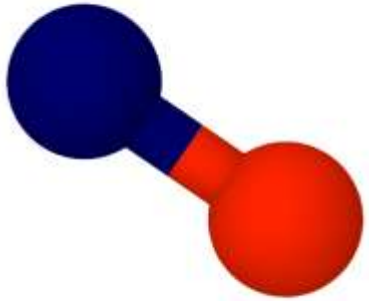
Class A: n=670



Class B: n=30



Class C: n=3

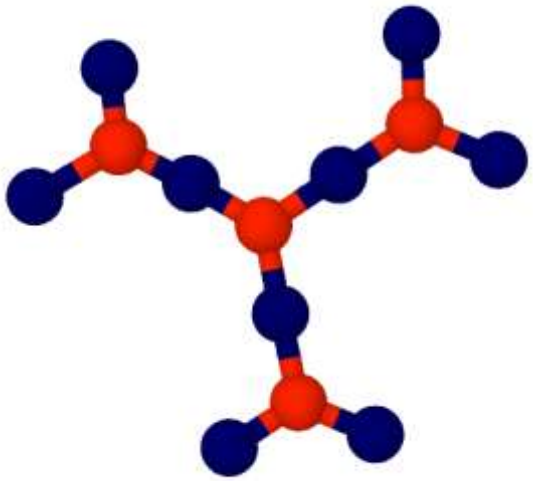


Si

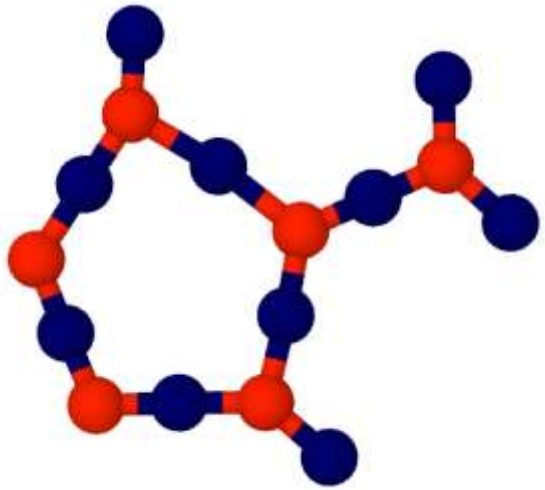
O

Experimental two-dimensional silica – 3nn clusters

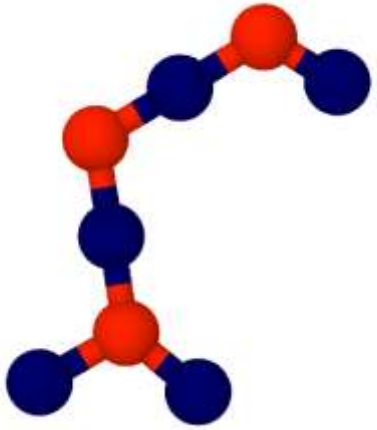
Class A: n=506



Class B: n=37



Class C: n=36

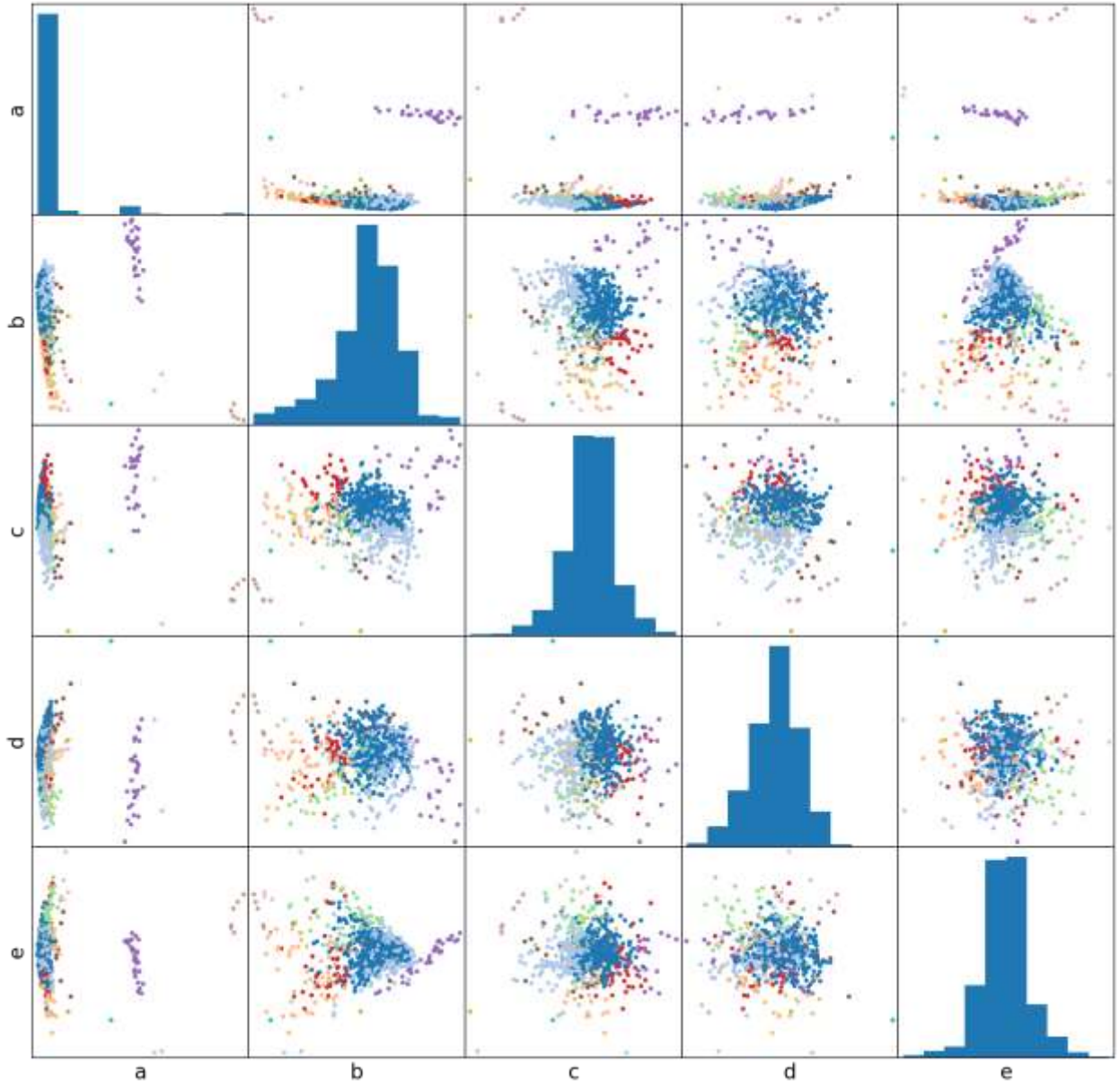


Si

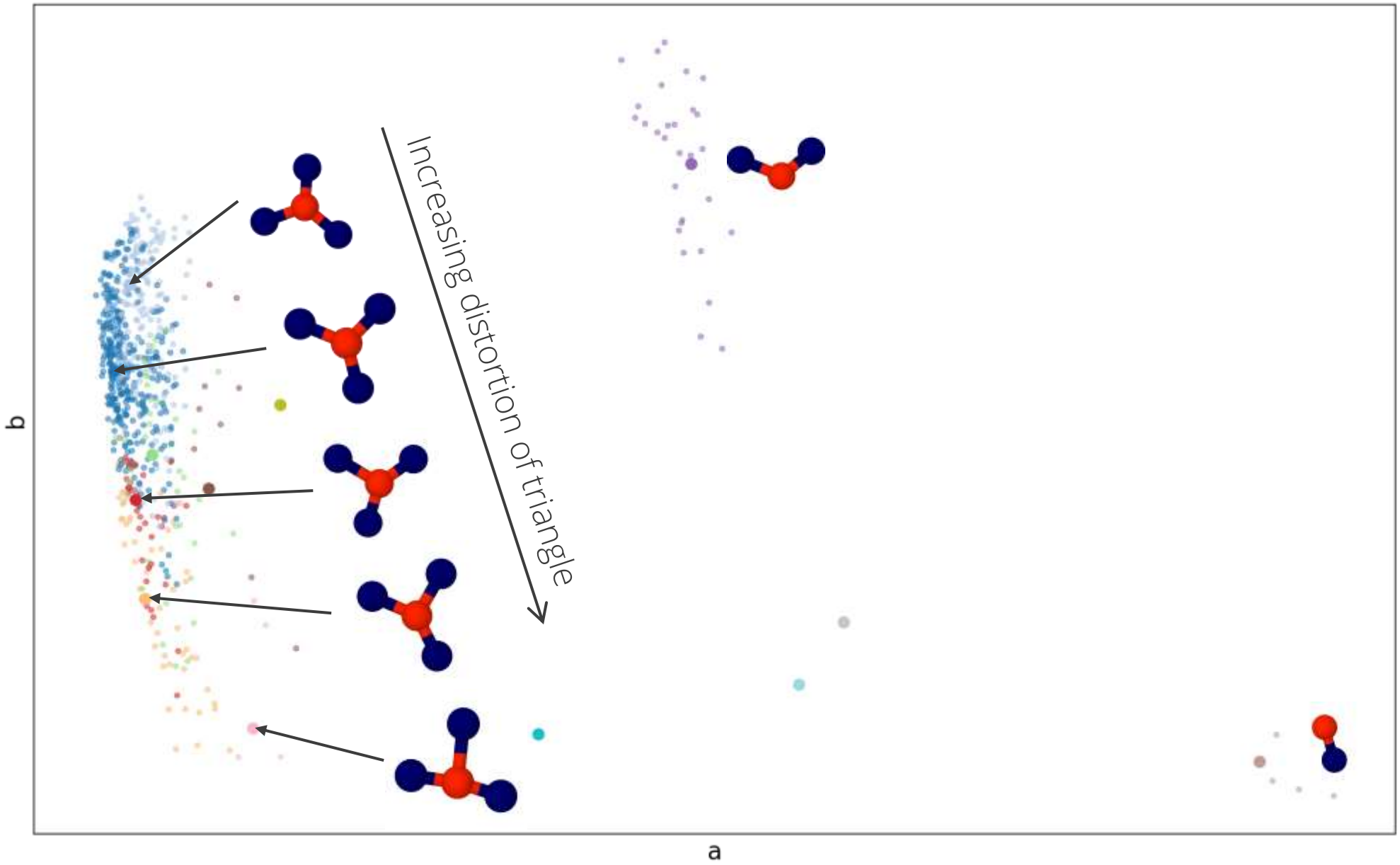
O

Experimental two-dimensional silica – diffusion coordinates (1nn)

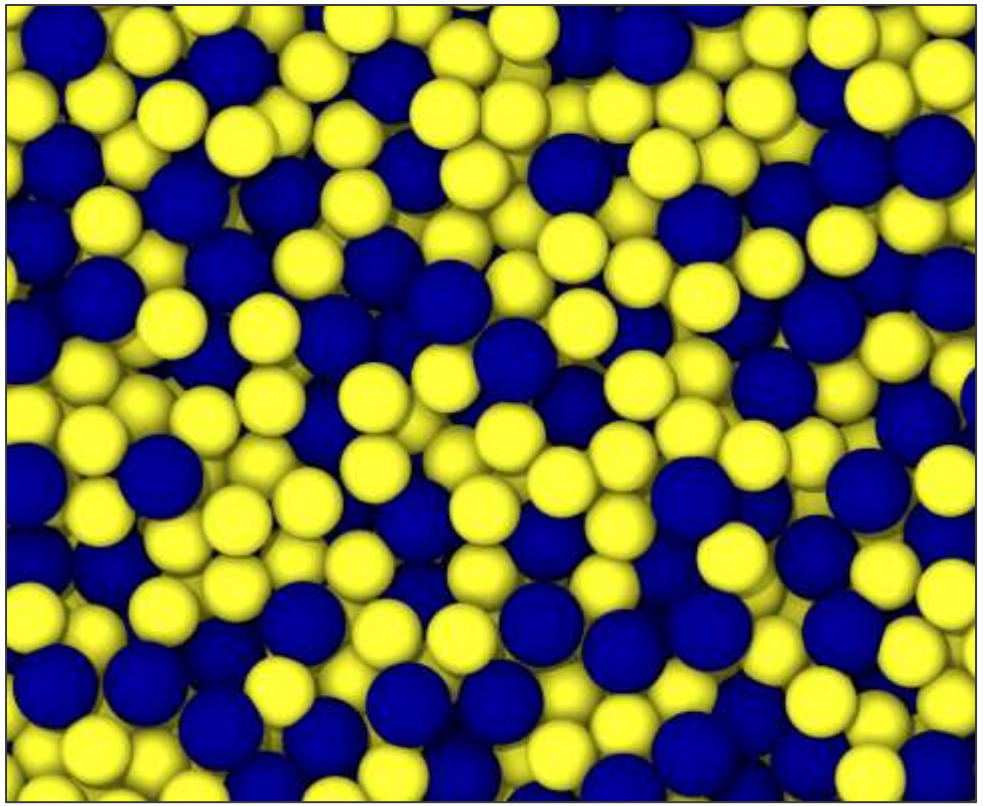
Colored by clustering class

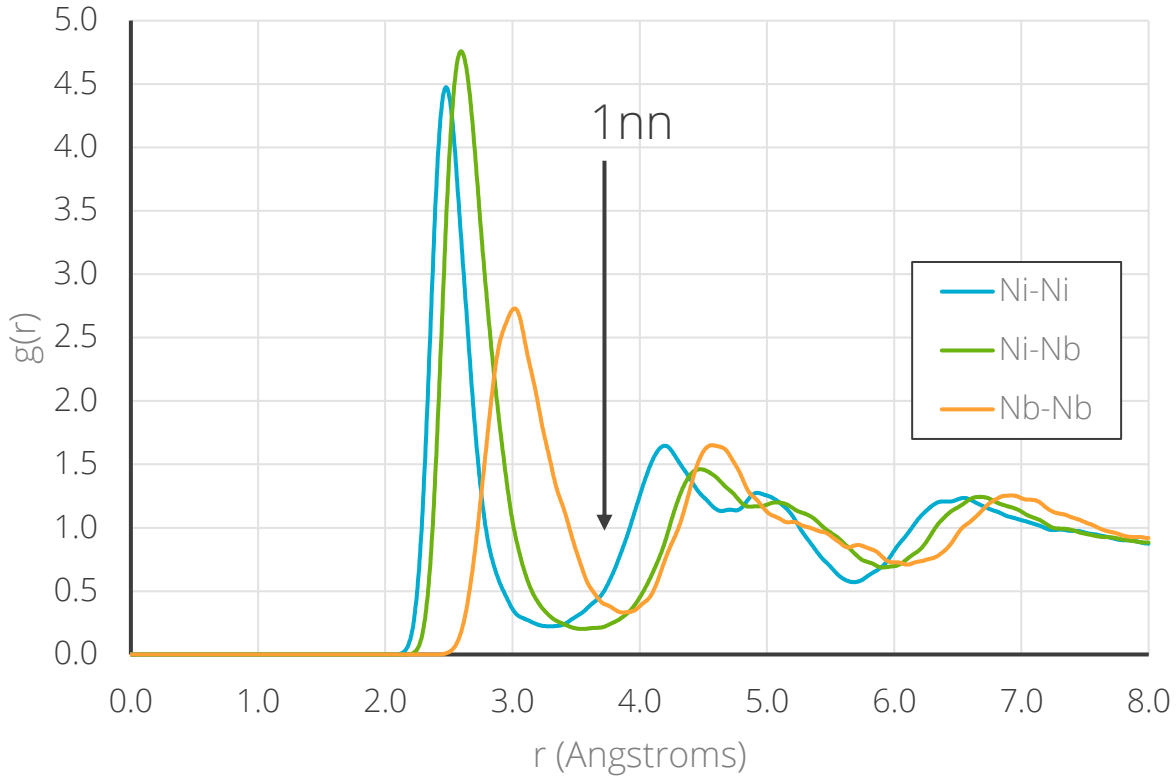


Experimental two-dimensional silica – diffusion coordinates (1nn)



Three-dimensional EAM Ni_2Nb metallic glass



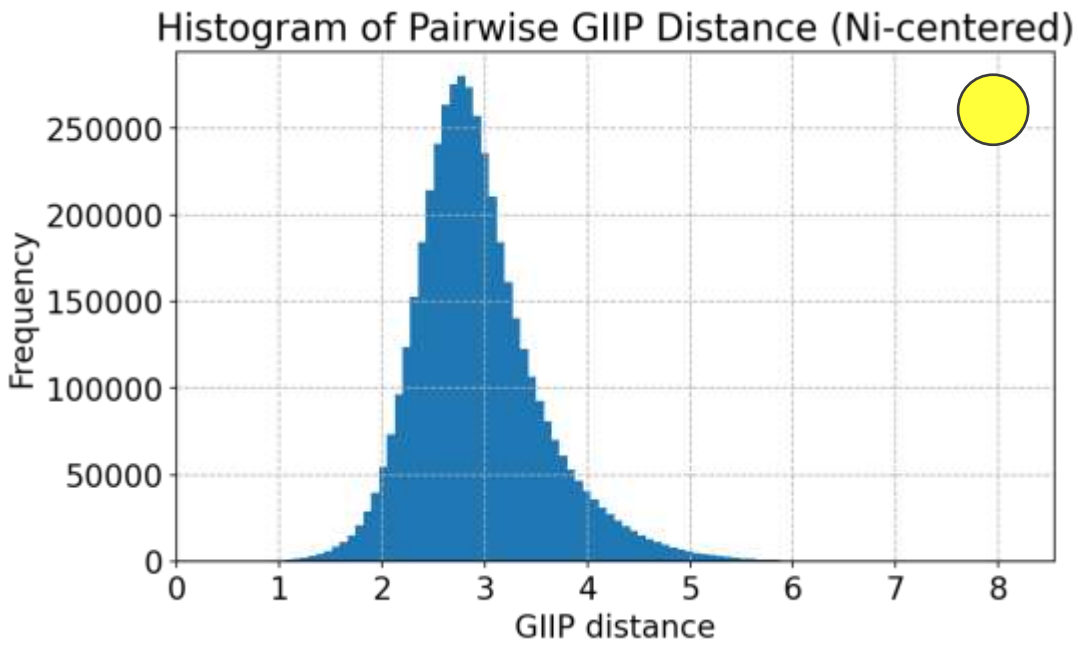


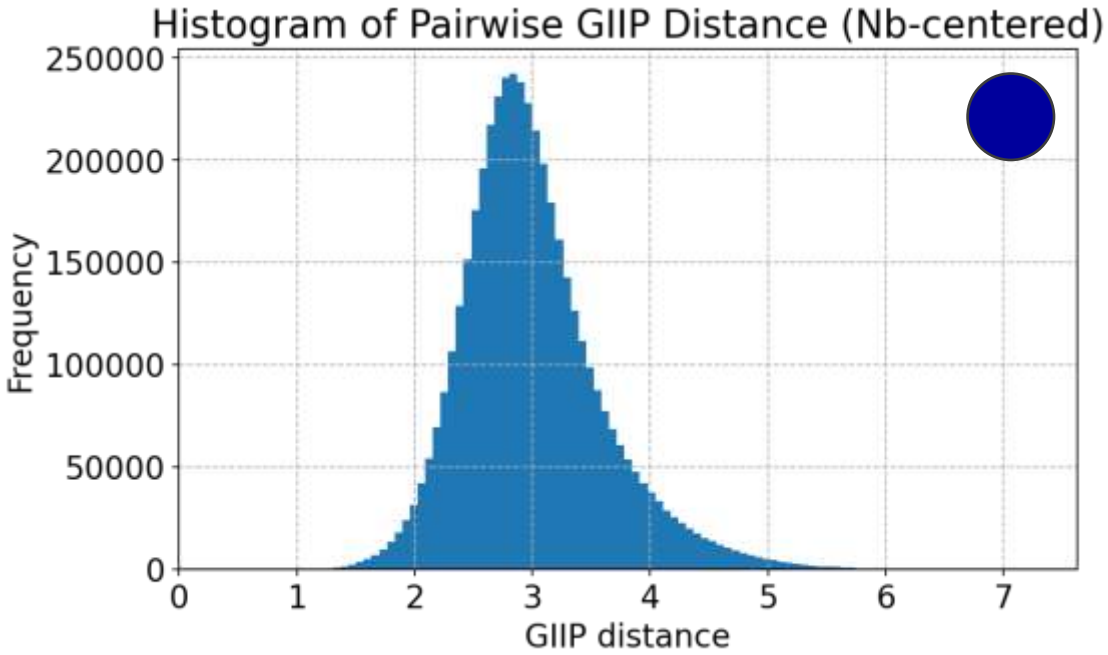
=Ni

=Nb

Simulation by Michael Chandross

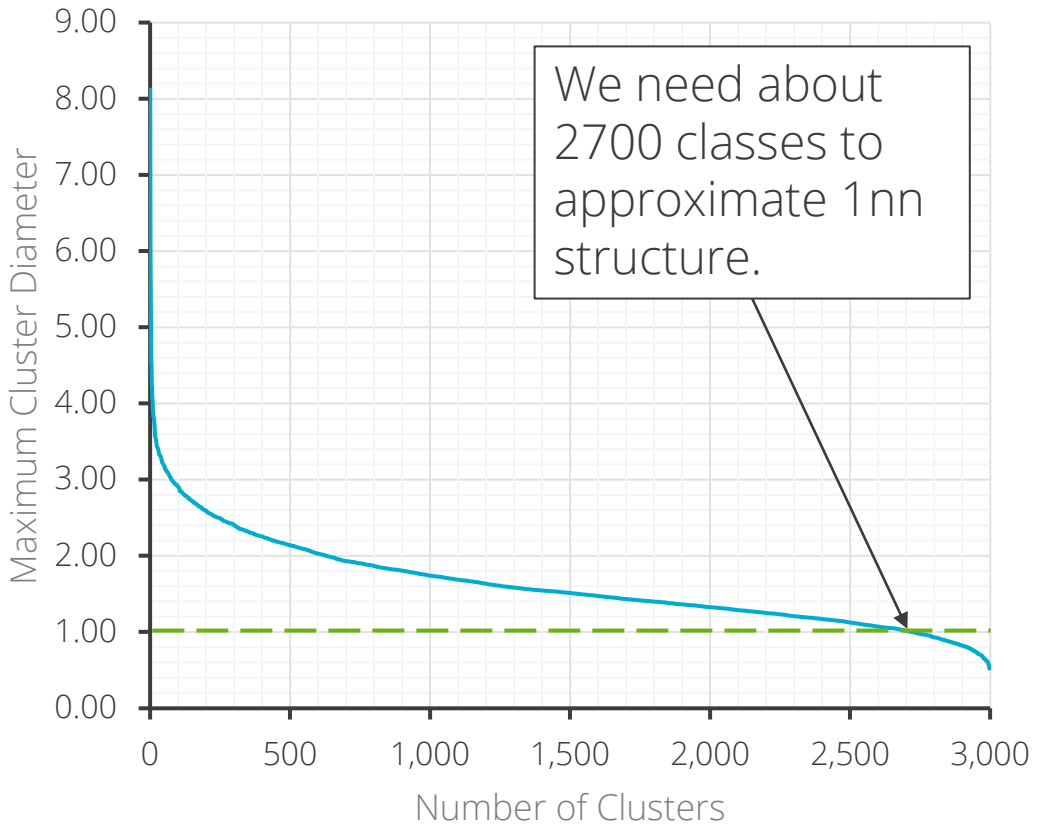
Three-dimensional EAM Ni_2Nb metallic glass – histograms



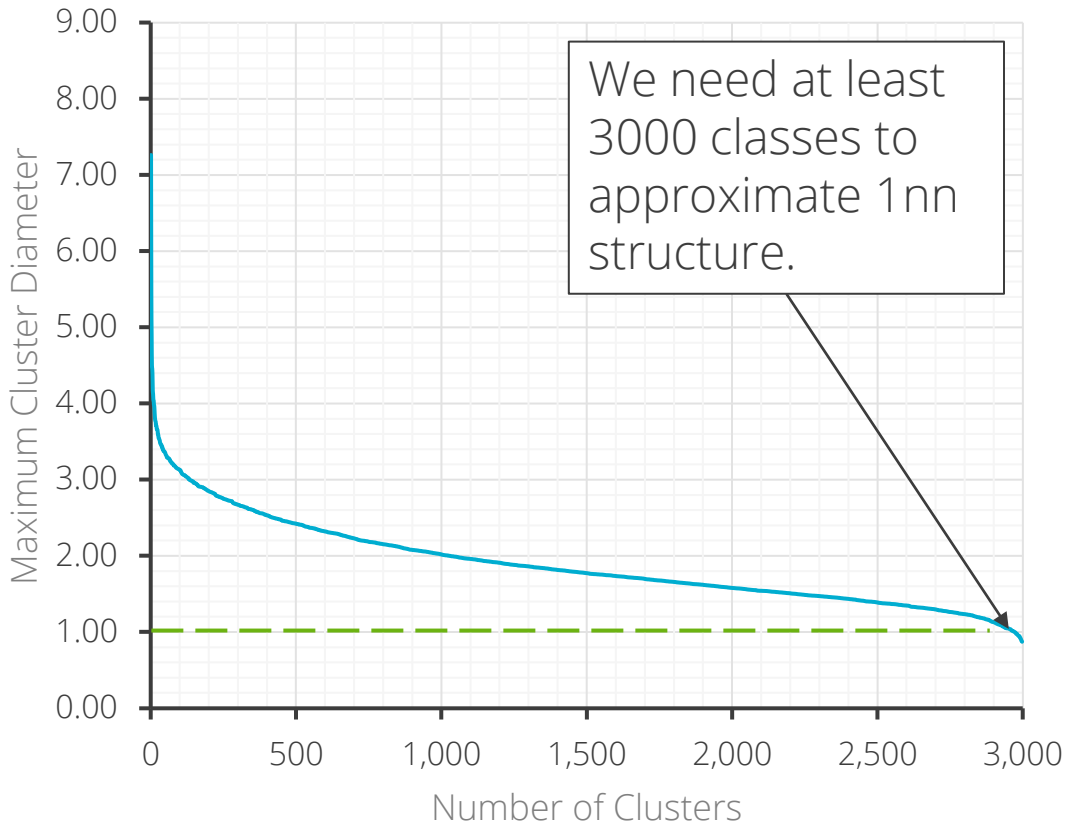


Three-dimensional EAM Ni_2Nb metallic glass – hierarchical clustering

Centered on Ni

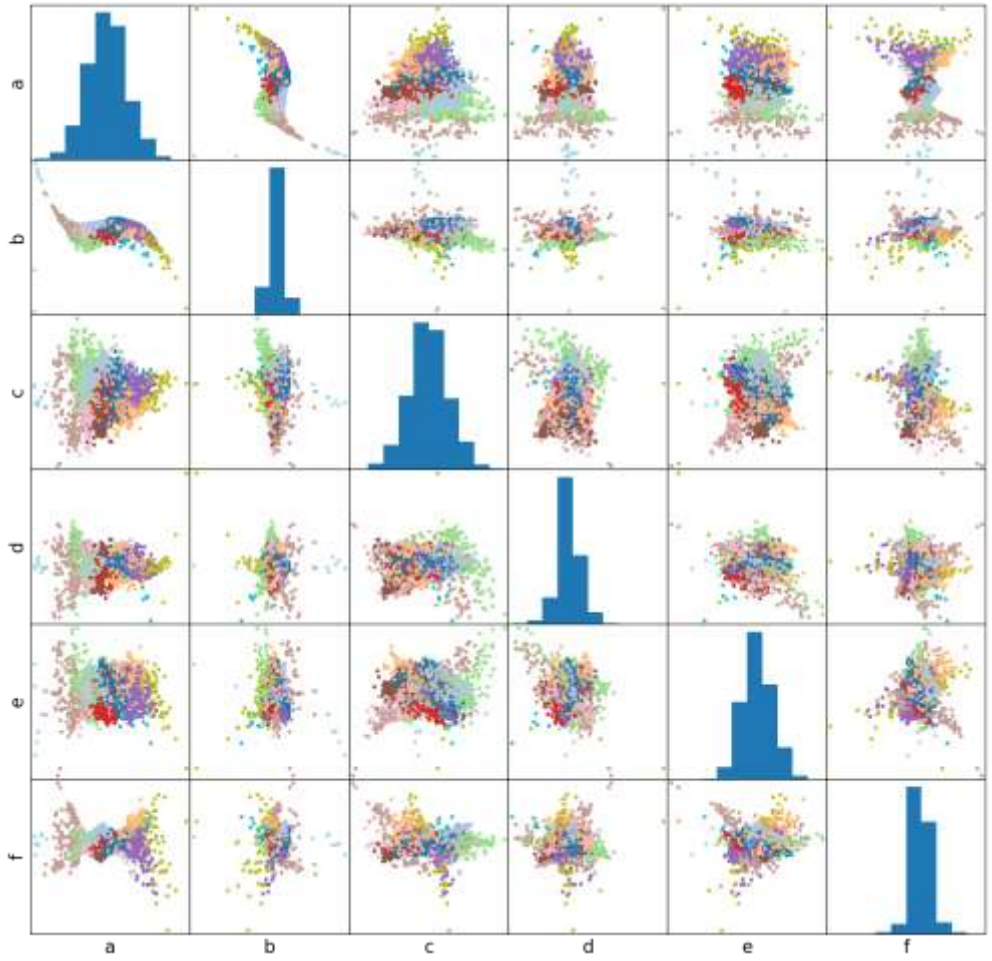


Centered on Nb

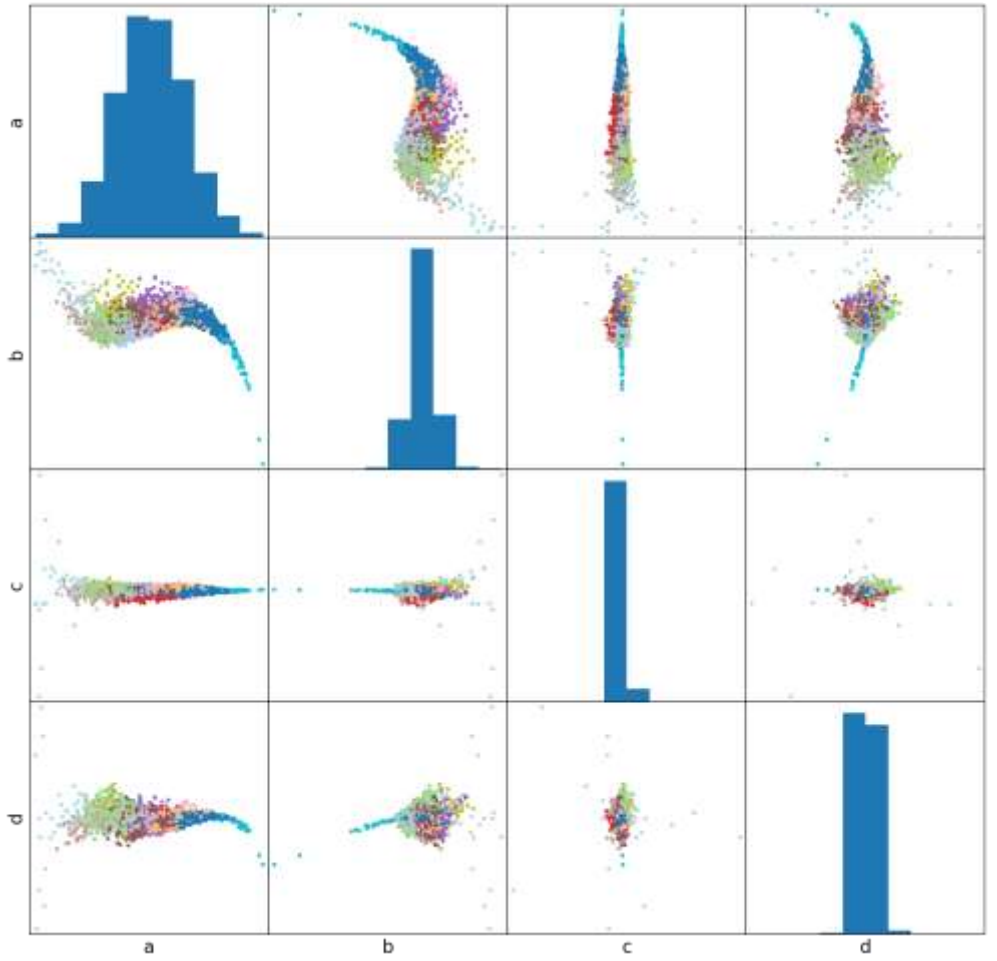


Three-dimensional EAM Ni_2Nb metallic glass – diffusion coordinates

Centered on Ni

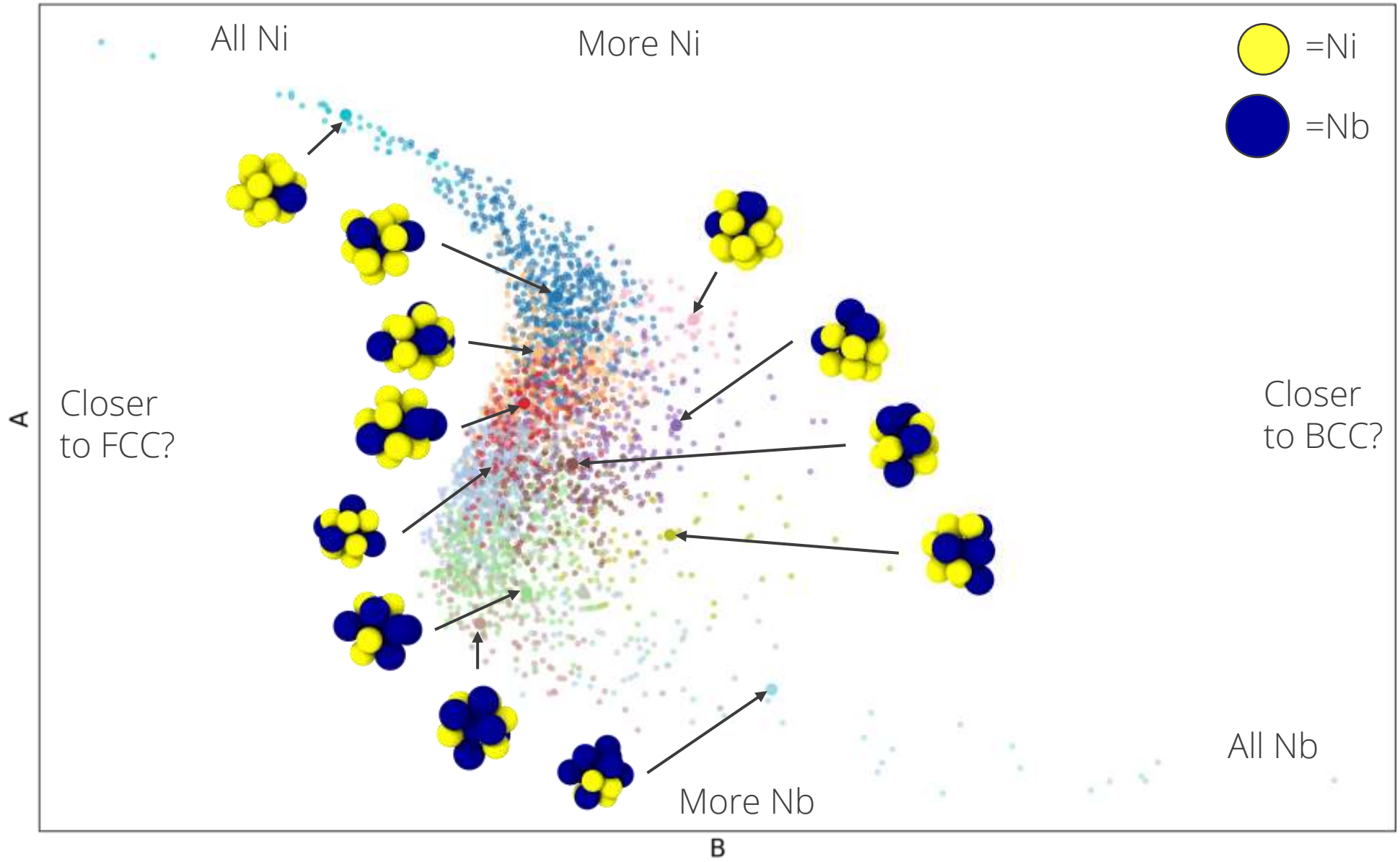


Centered on Nb

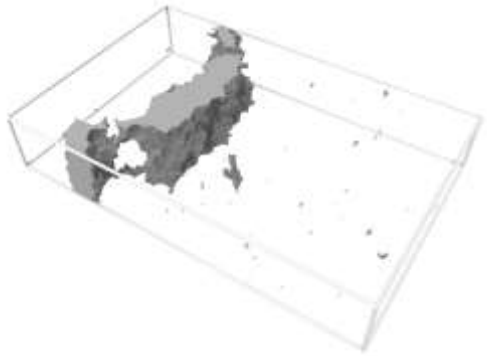


Colored by clustering class

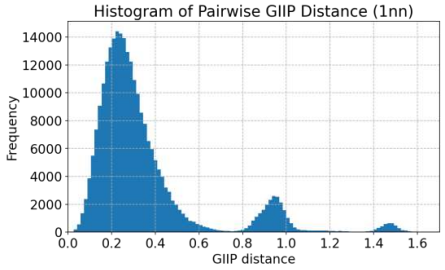
Three-dimensional EAM Ni_2Nb metallic glass – clusters



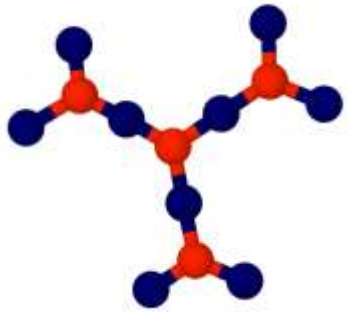
Conclusions and Looking Forward



Intuition and mathematical elegance aren't adequate to quantify the structure of glass.



So we developed a data-driven structural descriptor using the Gaussian Integral Inner Product.



The data-driven approach does a good job of capturing structure in a granular (complete) way, but it costs us intuitive interpretability.

We hope that data-driven descriptors will complement more intuitive approaches and emerging glass characterization methods.