Insightful classification of crystal structures using distance functions

Peter Mahler Larsen

CAMD, Department of Physics, DTU

TRATTORIA DEI MATERIALI

Bruschetta	*???*
Risotto con asparagi	*???*
Linguine All'arrabbiata	*???*
Gnocchi ai quattro formaggi	*???*
Melanzane alla parmigiana	*???*

No proper menus for materials!

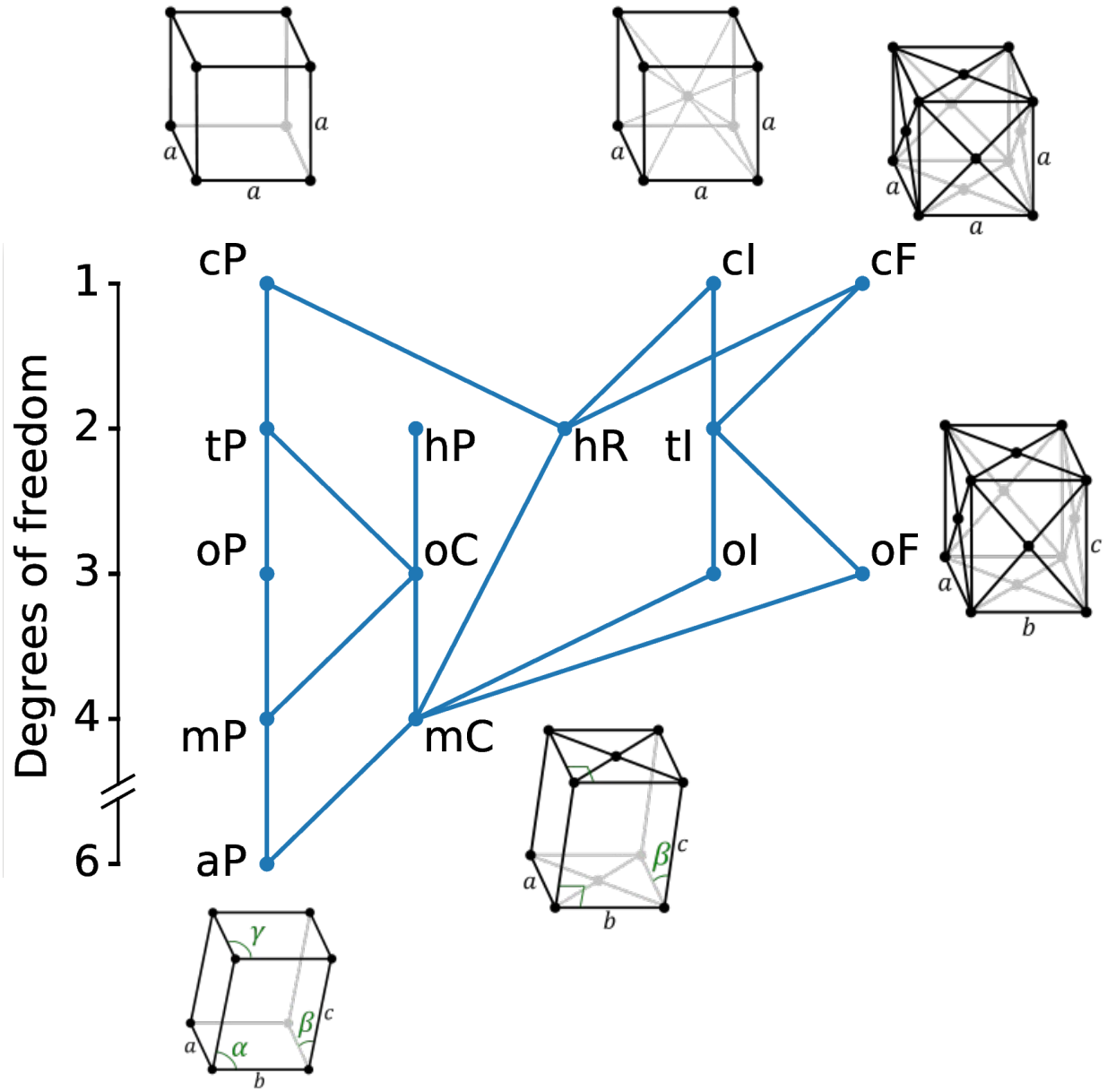
Simulate a structure. Then:

- Determine the Bravais lattice
- Find its dimensionality
- Find the primitive unit cell
- Does it have inversion symmetry?

What are the costs? Give me a menu!

Bravais lattice classification

Given a lattice basis,
what Bravais type is it?



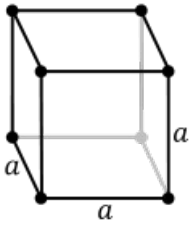
Traditional approach - tolerances

Lattice vectors: $\vec{a}$, $\vec{b}$, and $\vec{c}$,

Metric tensor:

$$\mathbf{G} = \begin{bmatrix} g_{11} & g_{12} & g_{13} \\ g_{12} & g_{22} & g_{23} \\ g_{13} & g_{23} & g_{33} \end{bmatrix} = \begin{bmatrix} \vec{a}\cdot\vec{a} & \vec{a}\cdot\vec{b} & \vec{a}\cdot\vec{c} \\ \vec{a}\cdot\vec{b} & \vec{b}\cdot\vec{b} & \vec{b}\cdot\vec{c} \\ \vec{a}\cdot\vec{c} & \vec{b}\cdot\vec{c} & \vec{c}\cdot\vec{c} \end{bmatrix}$$

Primitive cubic conditions:



$$g_{11} = g_{22} = g_{33}$$
$$g_{23} = g_{13} = g_{12} = 0$$

Approximate fulfilment of conditions:

$$|g_{11} - g_{22}| \leq \epsilon \quad |g_{11} - g_{33}| \leq \epsilon \quad |g_{22} - g_{33}| \leq \epsilon$$
$$|g_{23}| \leq \epsilon \quad |g_{13}| \leq \epsilon \quad |g_{12}| \leq \epsilon$$

How should we pick tolerance?

A distance function on lattices

Lattice bases: $\mathbf{A} \in \mathbb{R}^{3\times 3}$ and $\mathbf{B} \in \mathbb{R}^{3\times 3}$

Linear map: $\mathbf{A} = \mathbf{F}\mathbf{B}$

Deformation gradient

Polar decomposition: $\mathbf{U} = \sqrt{\mathbf{F}^T\mathbf{F}}$

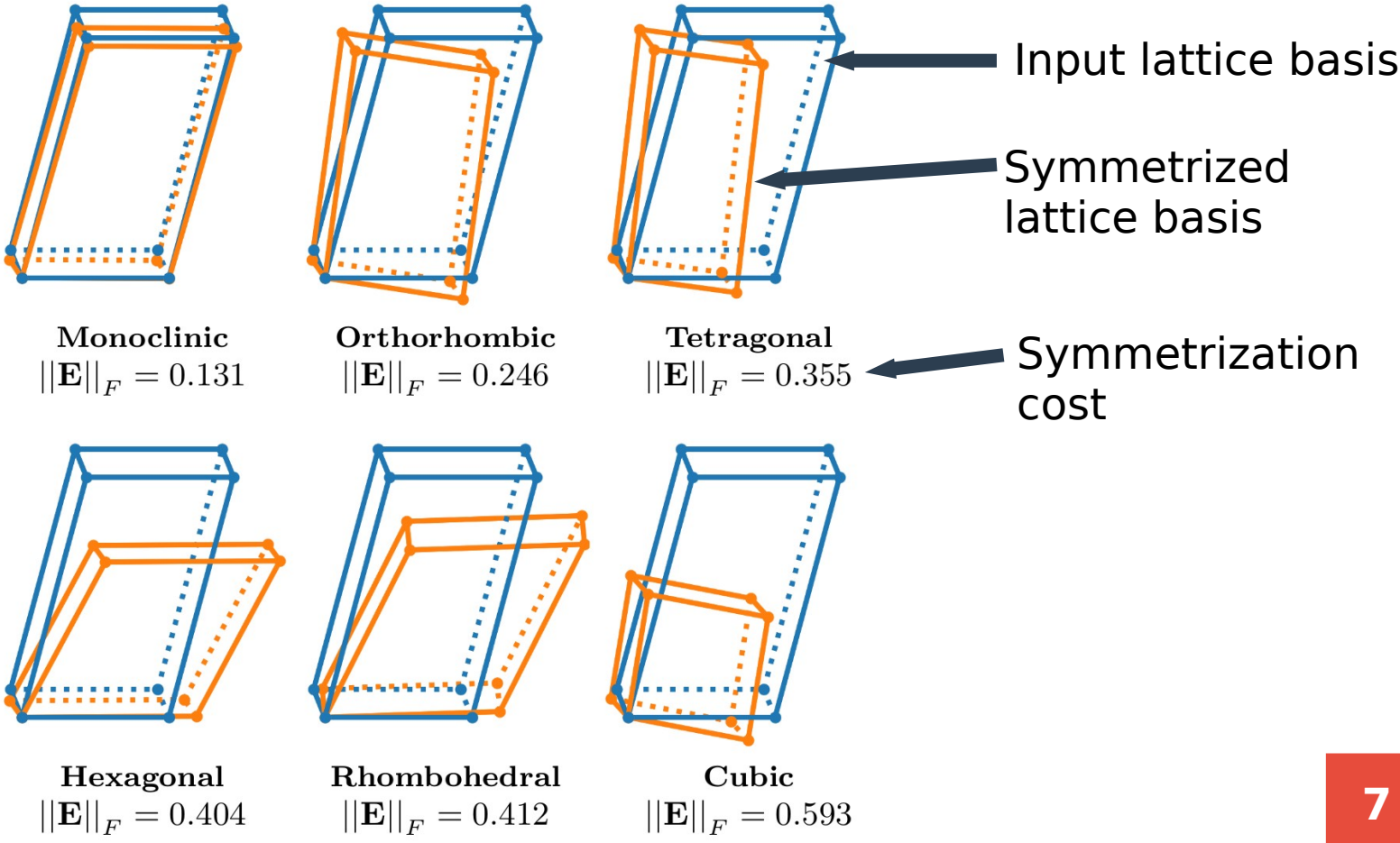
Stretch tensor

Strain tensor: $\mathbf{E} = \mathbf{U} - \mathbf{I} = \begin{bmatrix} \epsilon_{11} & \epsilon_{12} & \epsilon_{13} \\ \epsilon_{12} & \epsilon_{22} & \epsilon_{23} \\ \epsilon_{13} & \epsilon_{23} & \epsilon_{33} \end{bmatrix}$

Distance function: $d(\mathbf{A},\mathbf{B}) = ||\mathbf{E}||_F = \left|\left|\sqrt{\mathbf{B}^{-T}\mathbf{A}^T\mathbf{A}\mathbf{B}^{-1}} - \mathbf{I}\right|\right|_F$

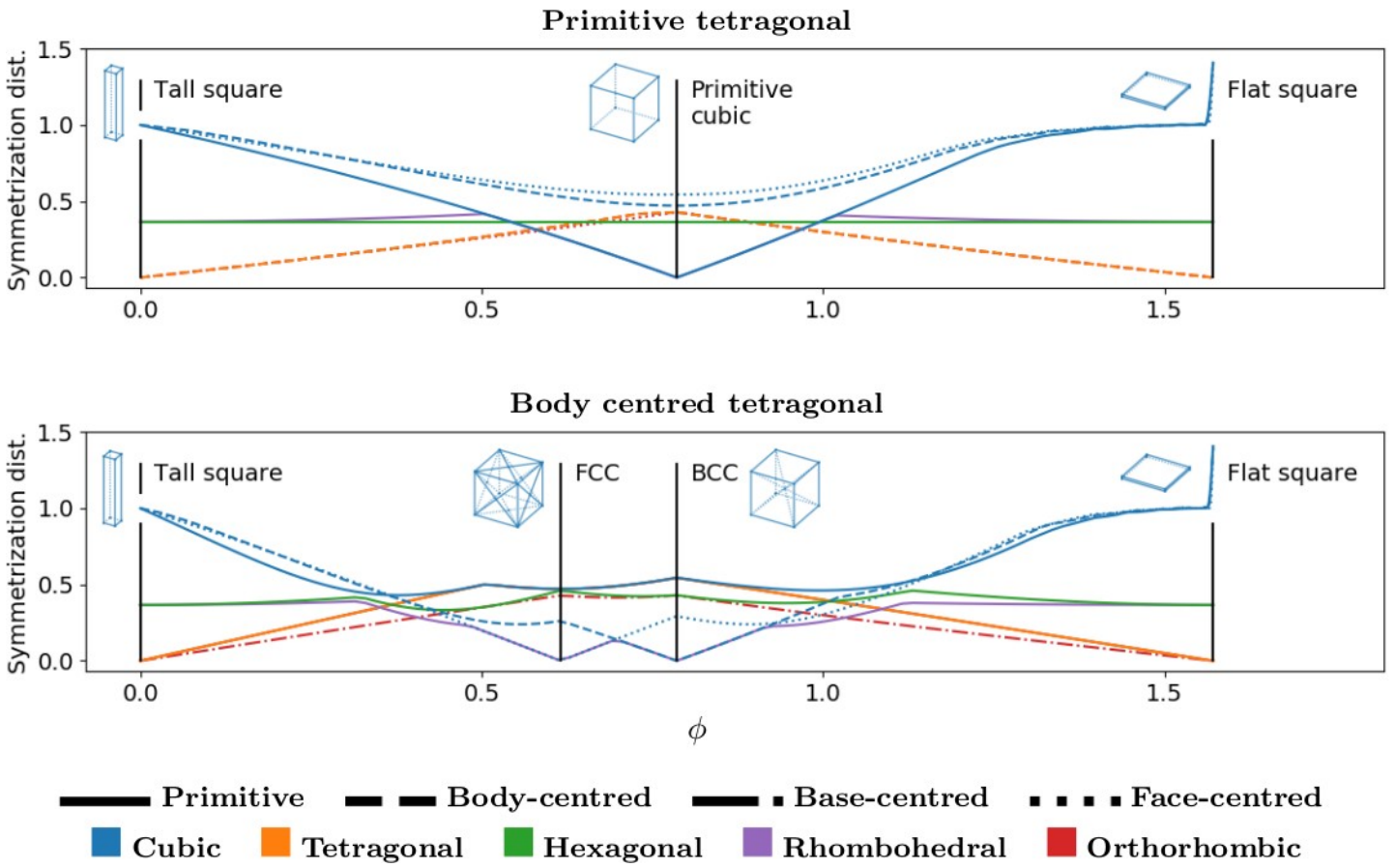
A menu for Bravais lattices

Symmetrize by minimizing distance function *d(A, B)*
B is observed lattice basis *(input)*
A is variable lattice which respects symmetry of target Bravais type



Varying the Bravais geometry

Calculate a vector of symmetrization distances



Cell parametrization

$$\begin{bmatrix} \sin\phi & 0 & 0 \\ 0 & \sin\phi & 0 \\ 0 & 0 & \cos\phi \end{bmatrix}$$

Cell parametrization

$$\begin{bmatrix} -\sin\phi & \sin\phi & \sin\phi \\ \sin\phi & -\sin\phi & \sin\phi \\ \cos\phi & \cos\phi & -\cos\phi \end{bmatrix}$$

Visualizing the Bravais lattice space

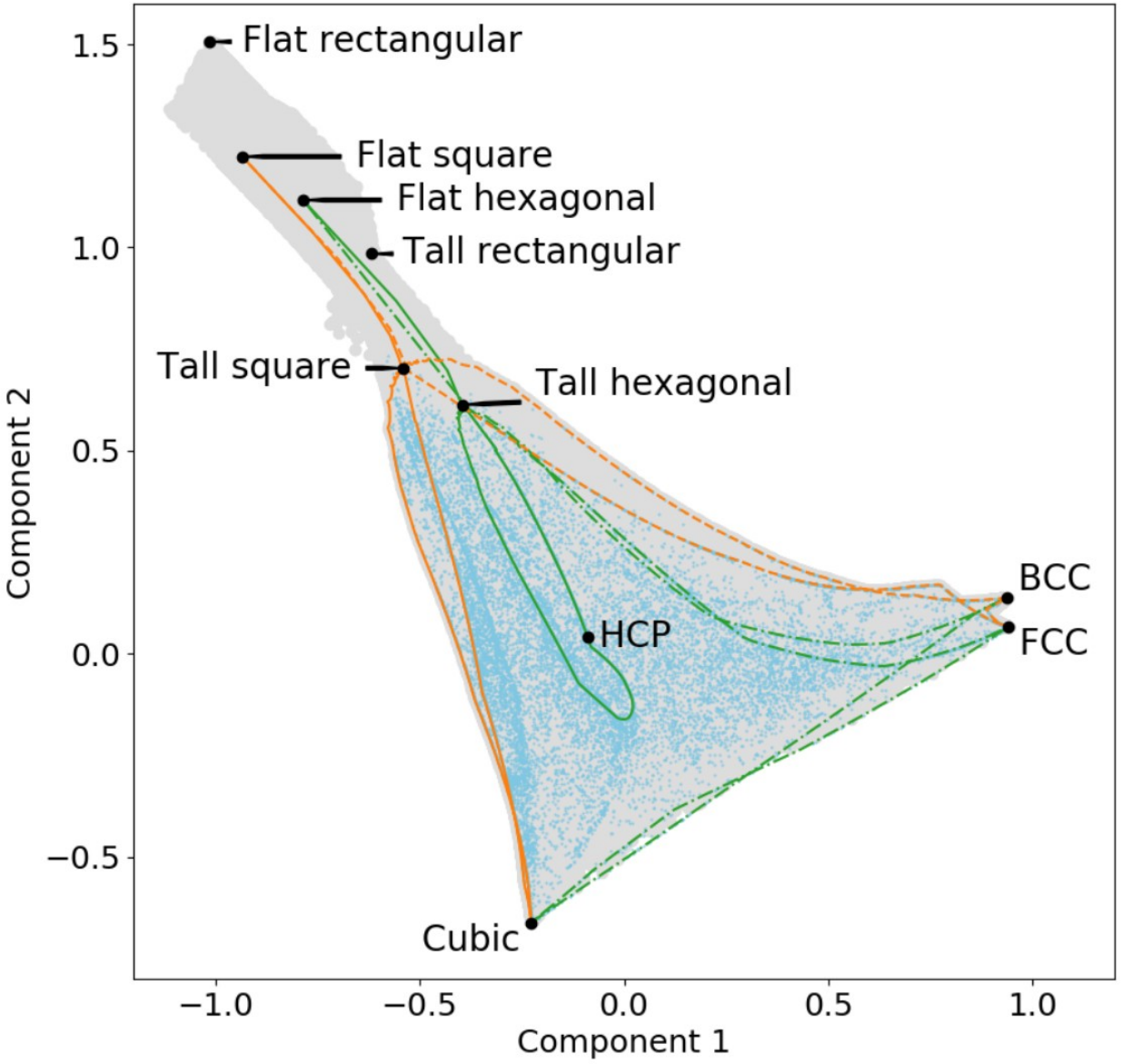
-All structures in Crystallography Open Database

(blue dots)

-PCA projection of distance vectors

-Distance function

aids understanding



Using the auguste module

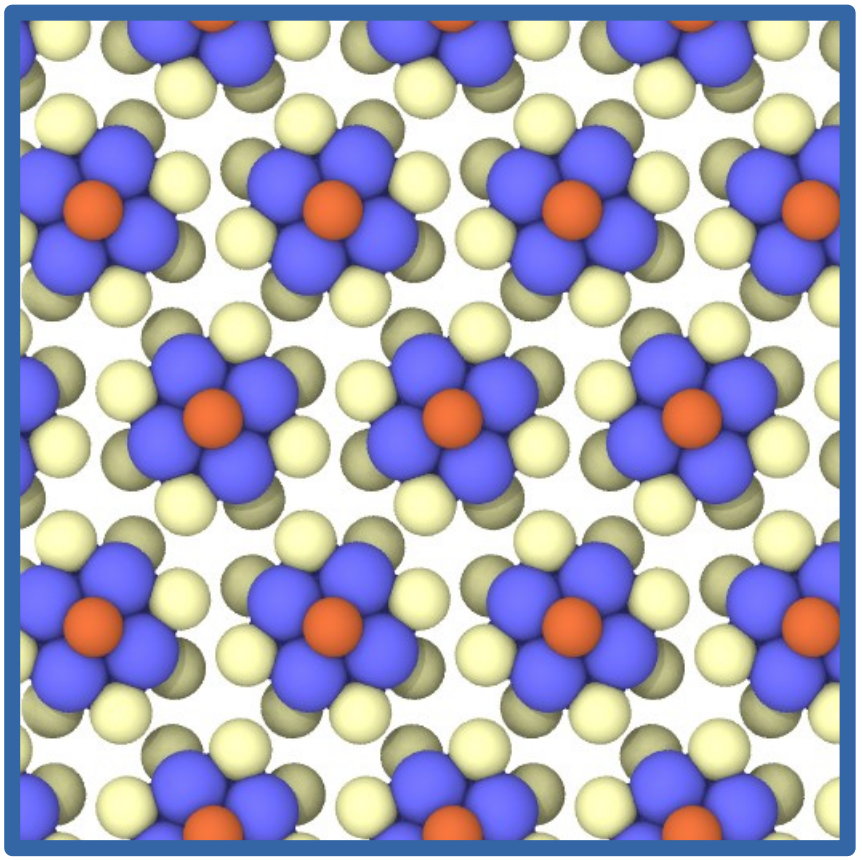
```
pmla@laptop: pip install --user auguste
pmla@laptop: python3
>>> import numpy as np
>>> import auguste
>>> cell = np.diag([10, 11, 30])
>>> auguste.symmetrize_lattice(cell, “tetragonal”)
(0.0672672793963138, array([[10.45248869,  0.        ,  0.        ],
                            [ 0.        , 10.45248869,  0.        ],
                            [ 0.        ,  0.        , 30.        ]]))
>>> auguste.calculate_vector(cell)
array([0.        , 0.        , 0.        , 0.        , 0.06726728,
       0.17328325, 0.18304468, 0.06726728, 0.18486719, 0.39230133,
       0.37133982, 0.63619127, 0.75280573, 0.75781598])
```

https://github.com/pmla/auguste/

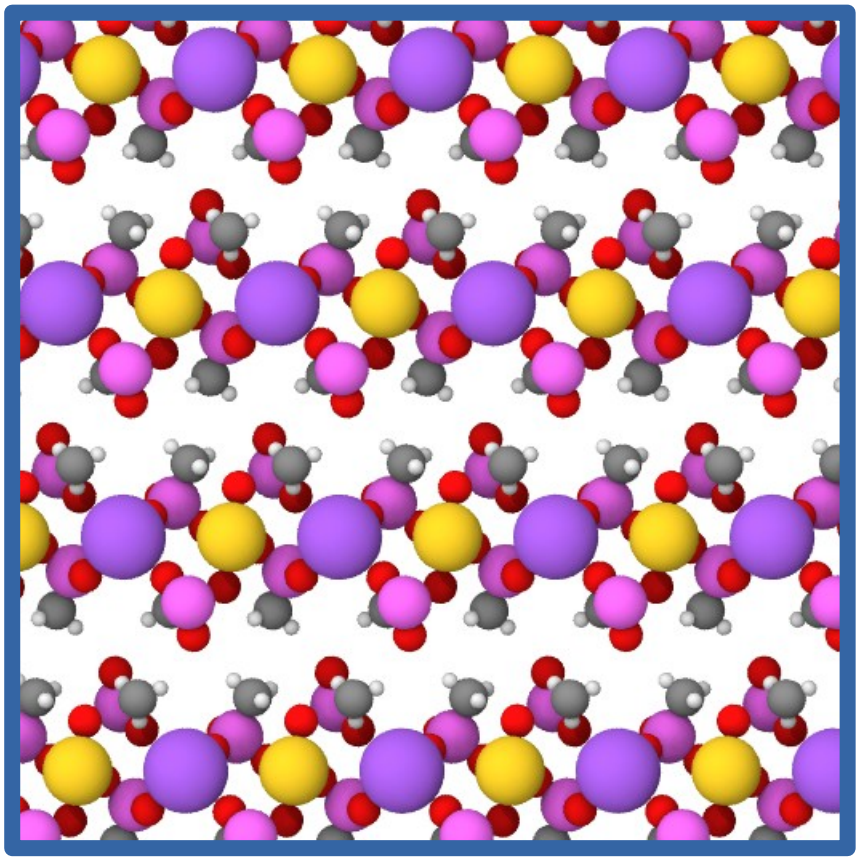
Material Dimensionality

Finding low-dimensional structures in bulk systems

$SiTa_4Te_4$

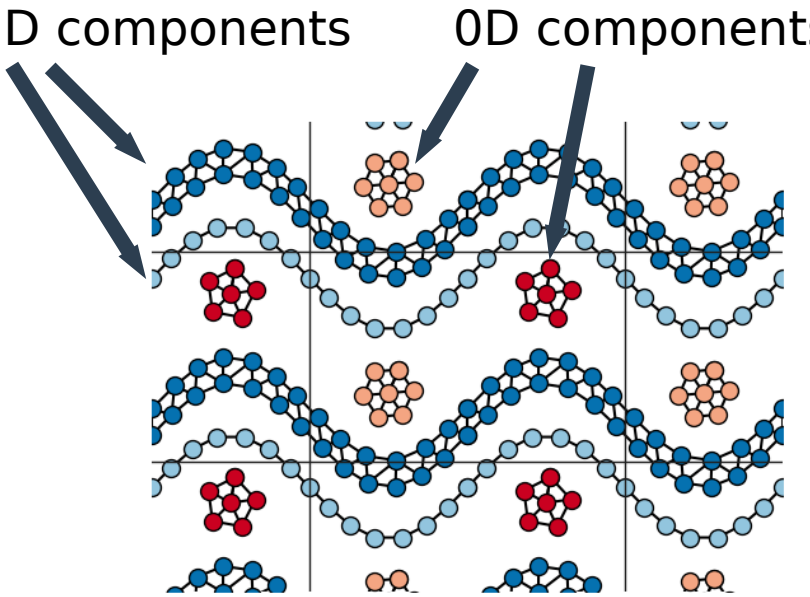


$Na(Au(CH_3SO_3)_4)$

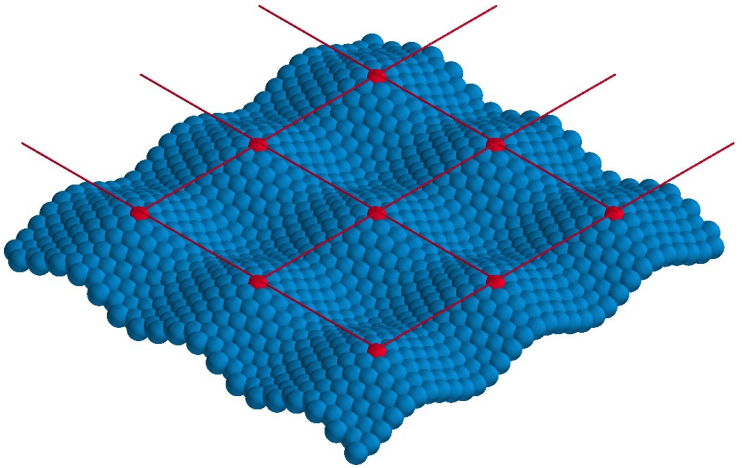


Defininging Dimensionality

Identify dimensionality of connected components

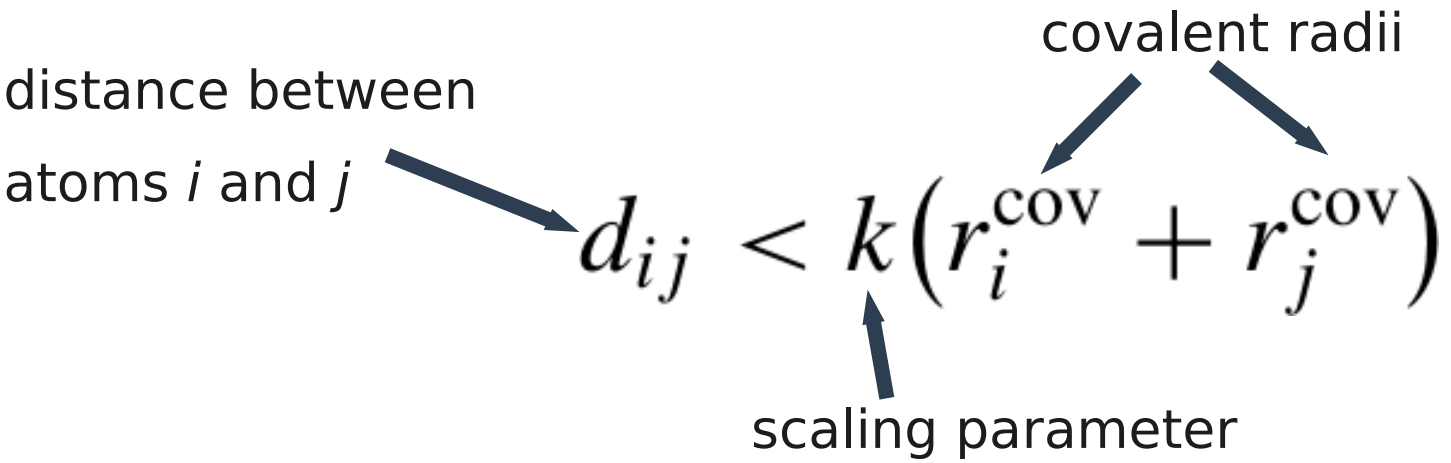


Dimensionality is determined by rank of vector space spanned by periodic images



Defining a bond

Bonding criterion of Ashton *et al.*:



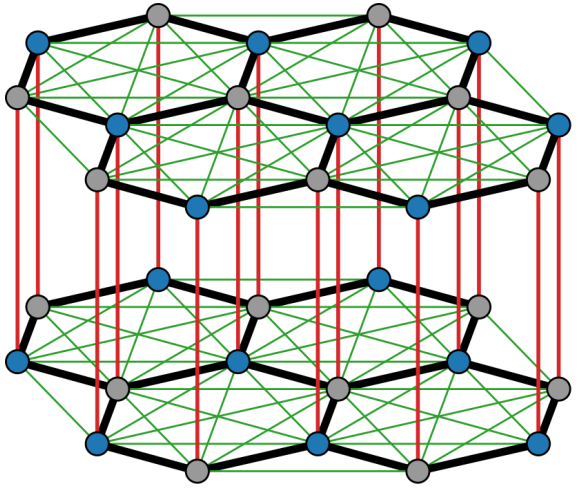
How should we select *k*?

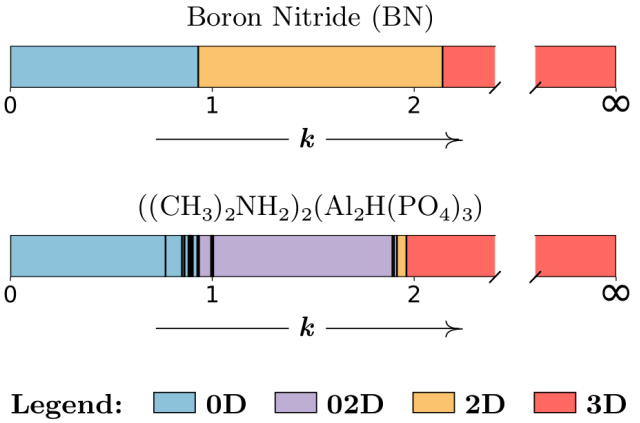
k=0 No atoms are bonded.
Each atom is its own 0D component.

k=∞ All atoms are bonded.
Structure is a single 3D component.

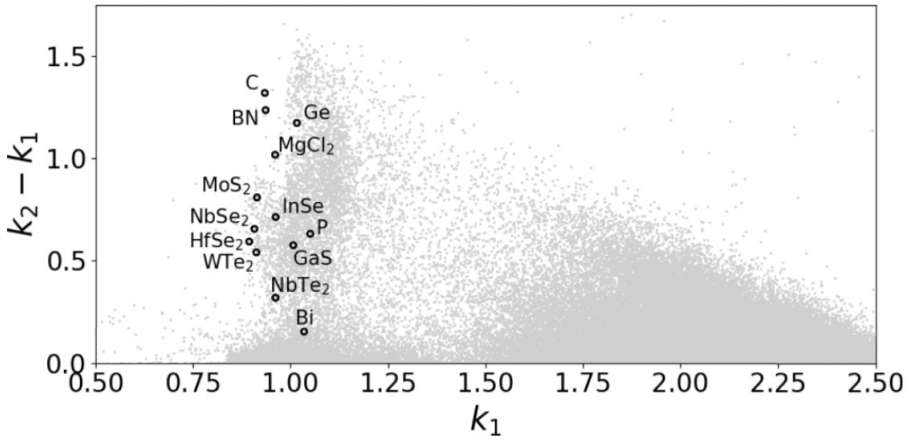
Barcodes

Dimensionality changes at discrete *k*-values:

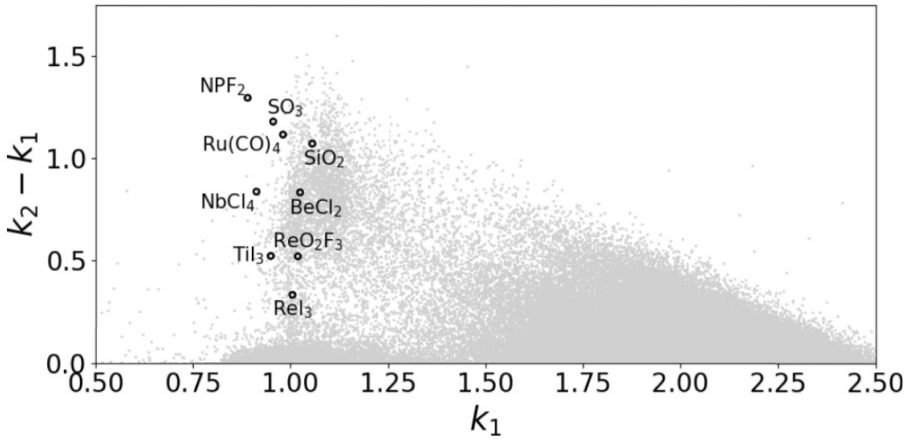




2D *k*-intervals



1D *k*-intervals



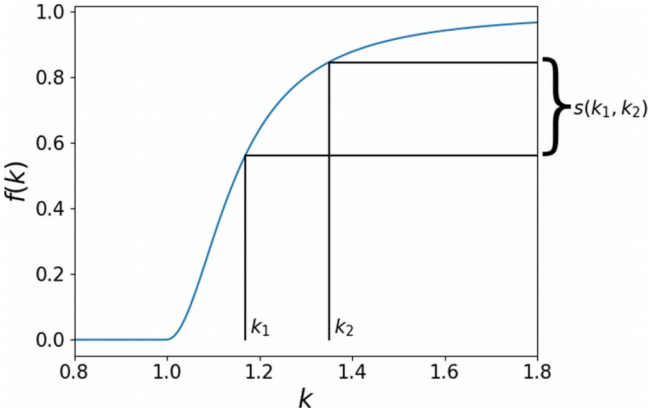
Dimensionality scoring parameter

Define a score on k-intervals:

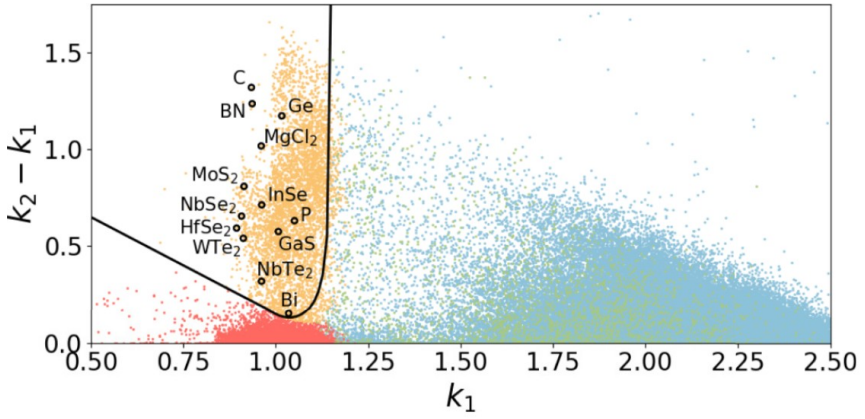
$$s(k_1, k_2) = f(k_2) - f(k_1)$$

$$f(x) = \frac{c \times \max(0, x-1)^2}{1 + c \times \max(0, x-1)^2}$$

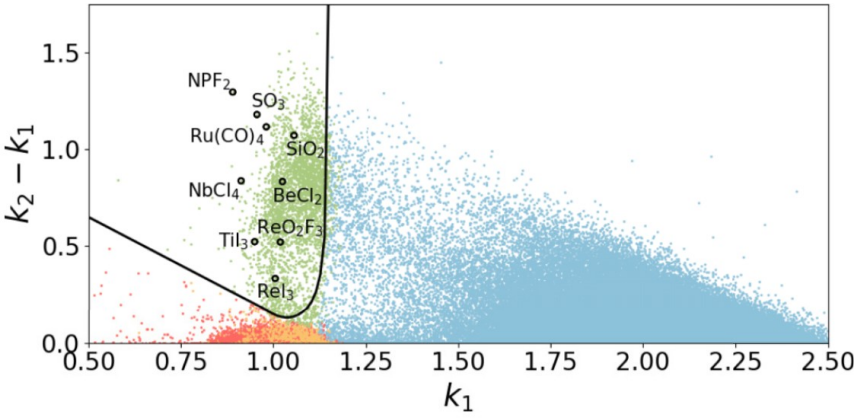
$$c = 1/0.15^2$$



2D *k*-intervals



1D *k*-intervals



0D 1D 2D 3D

Computational materials repository

Displaying rows 1-25 out of 8843 Rows: 25 Add Column

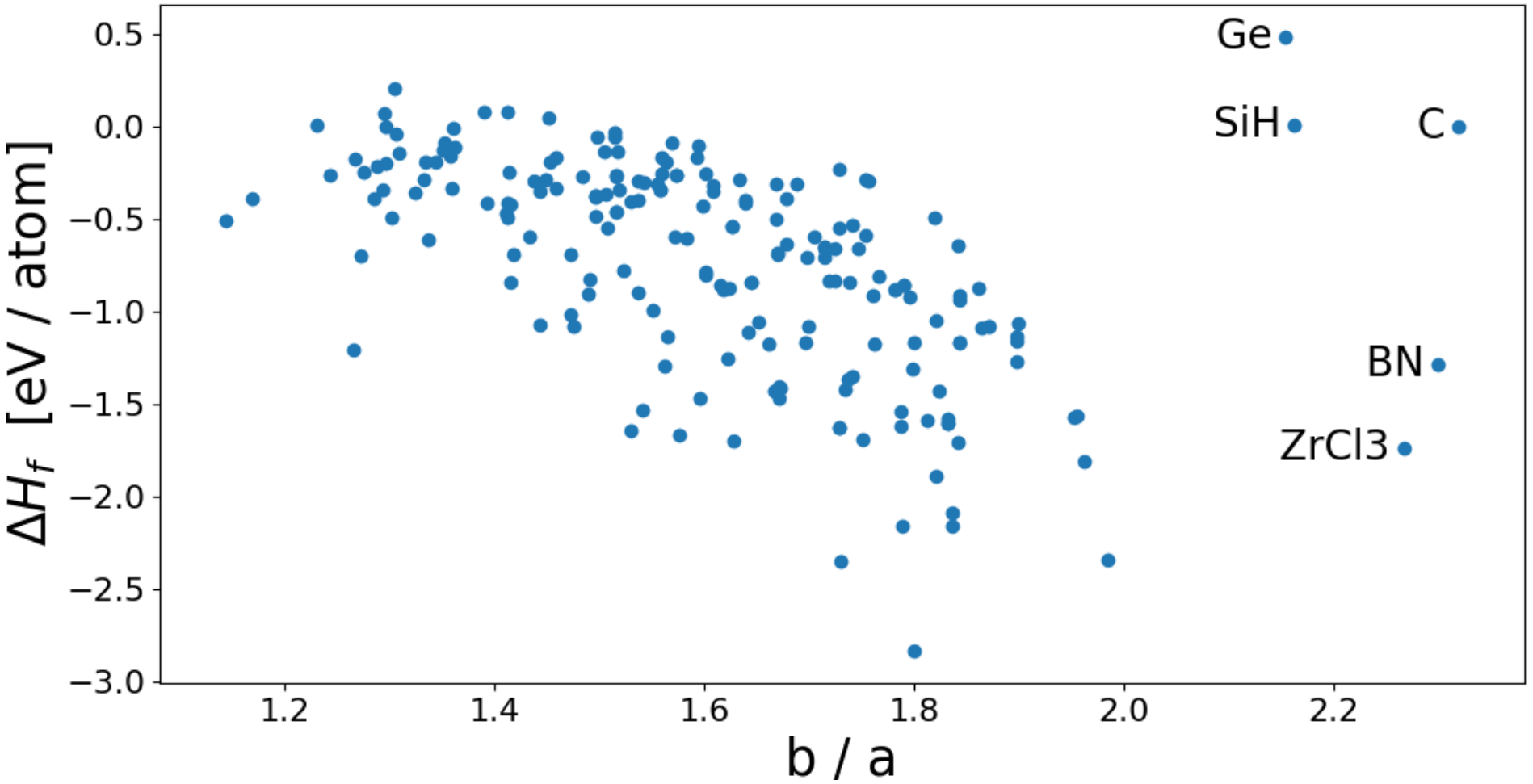
Formula ×	Source ×	ID # ×	↑ 2D score ×	publication ×
C_8	COD	9000046	0.986	The problem of the graphite structure; Kukesh, J. S. and Pauling, L.; American Mineralogist, vol. 35, pp. 125-125, 1950
B_2N_2	ICSD	27987	0.984	-
C_2	COD	9012705	0.984	The structure of graphite Locality: Ceylon, Bavaria Note: phase: graphite 3R; Lipson, H. and Stokes, A. R.; Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences, issue 984, vol. 181, pp. 101-105, 1942
B_2N_2	COD	1010602	0.983	An X-Ray Examination of the Structure of Boron Nitride; Brager, A; Acta Physicochimica (USSR), vol. 7, pp. 699-706, 1937
B_2N_2	COD	2016170	0.983	Rhombohedral boron subnitride, $B_{13}N_2$, by X-ray powder diffraction; Kurakevych, Oleksandr O. and Solozhenko, Vladimir L.; Acta Crystallographica Section C, issue 9, vol. 63, pp. i80-i82, 2007
Mg_3Cl_6	COD	1534338	0.977	An interpretation of the structures of alkaline earth chlorides in terms of interionic forces; Busing, W.R.; Transactions of the American Crystallographic Association, vol. 6, pp. 57-72, 1970

https://cmrdb.fysik.dtu.dk/lowdim/

The ASE dimensionality module

```
>>> from ase.io import read
>>> from ase.geometry.dimensionality import analyze_dimensionality
>>> atoms = read("2016170.cif")
>>> intervals = analyze_dimensionality(atoms)
>>> intervals[0]
Kinterval(dimtype='2D',
          score=0.983,
          a=0.931,
          b=2.141,
          h=(0, 0, 2, 0),
          components=array([0, 1, 1, 0]),
          cdim={0: 2, 1: 2})
```

Structure-property correlations

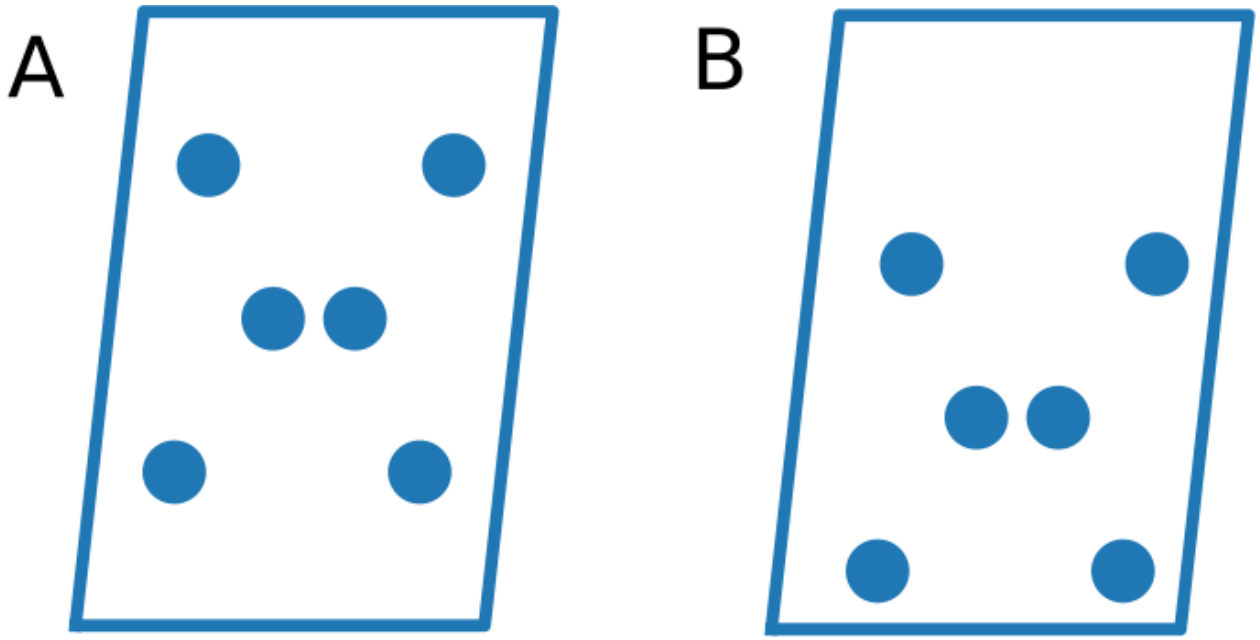


(intercomponent distance)

Tolerances don’t give this type of information!

Crystal structure analysis

When are two *crystals* identical?

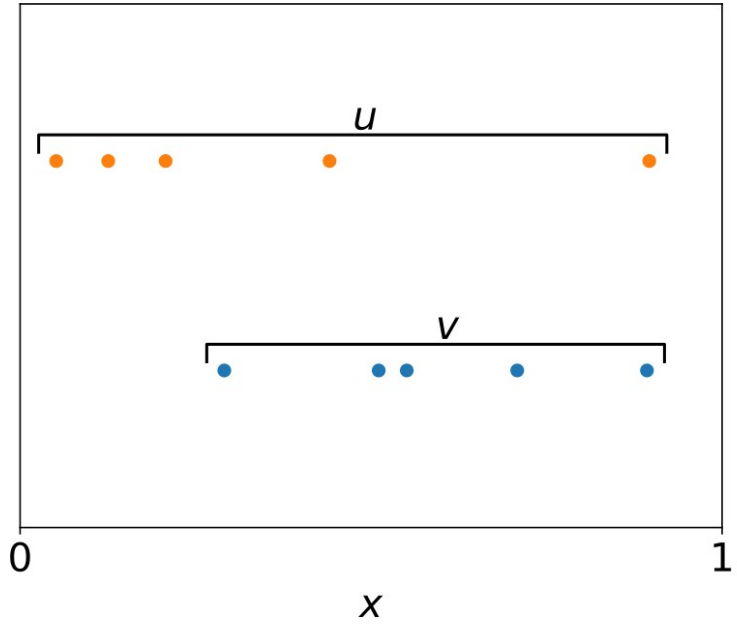


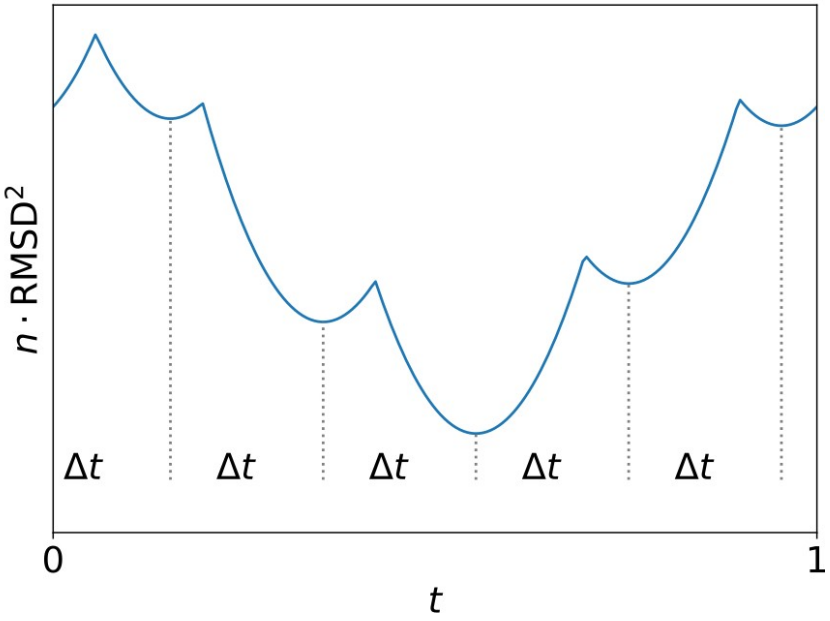
Traditional approach – tolerances!

Can we make a menu?

A distance function on crystals

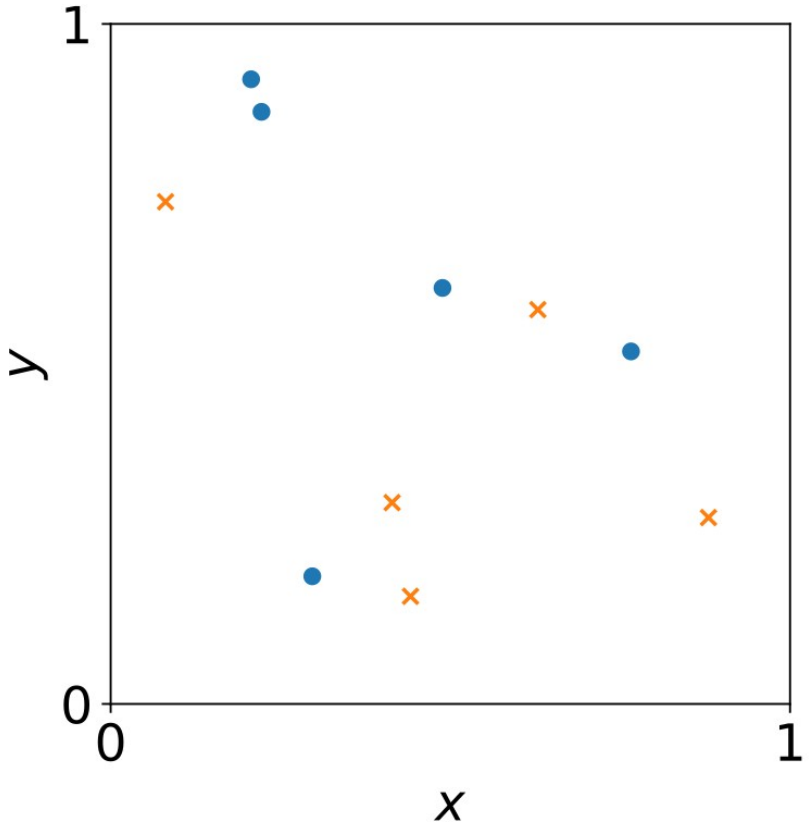
$$\mathrm{RMSD}\left(\mathbf{M}, \vec{t}\right) = \sqrt{\frac{1}{n} \sum_{(\vec{u},\vec{v}) \in \mathbf{M}} \left|\left|\vec{u} + \vec{t} - \vec{v}\right|\right|^2}$$

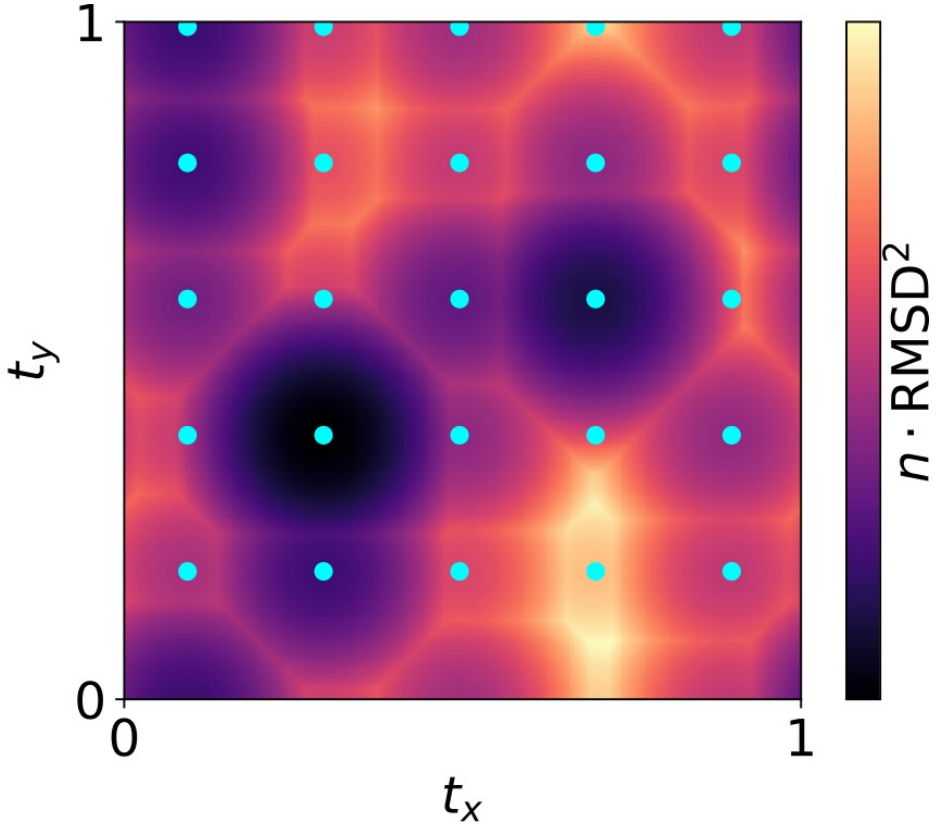




Calculating the RMSD

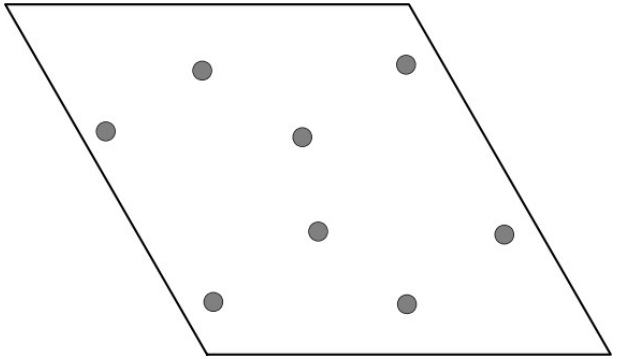
General case: Aurenhammer diagram with n^d basins





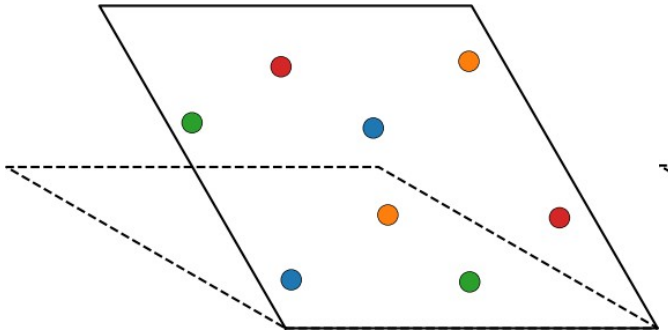
Crystal reduction

Graphene + perturbations

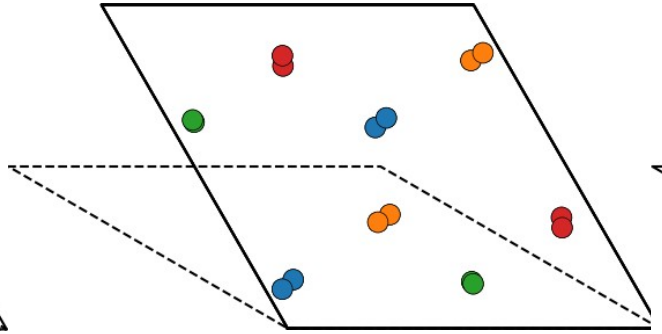


1st reduction
RMSD: 0.124

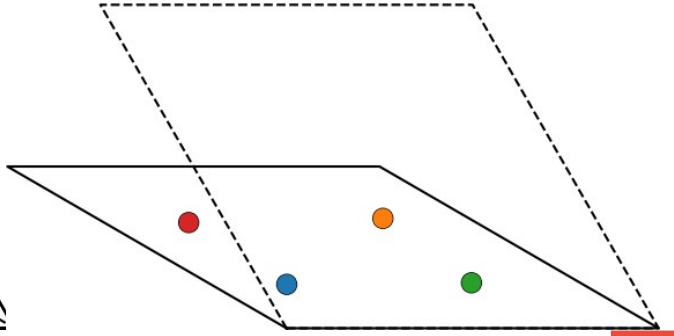
1. Group atoms



2. Apply translational symmetry

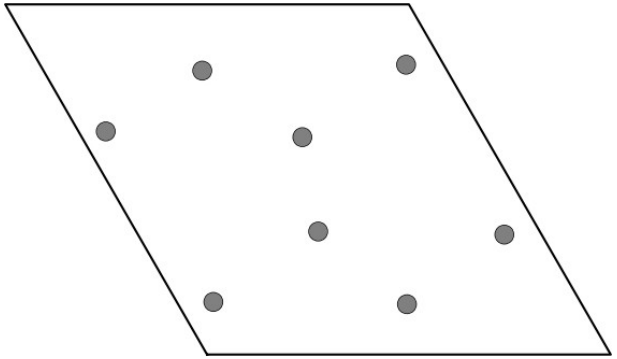


3. Calculate average positions



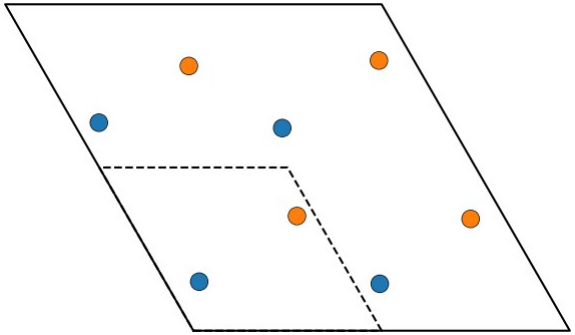
Crystal reduction

Graphene + perturbations

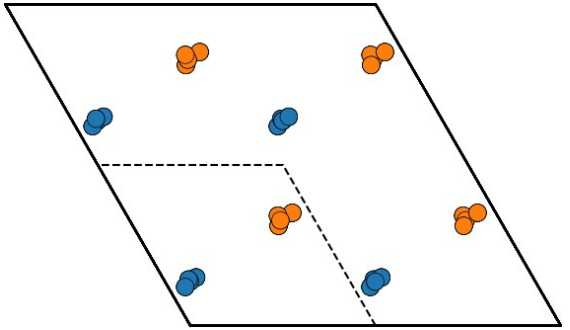


2^{nd} reduction
RMSD: 0.146

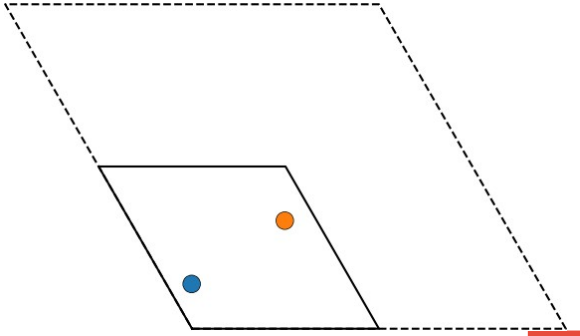
1. Group atoms



2. Apply translational symmetry



3. Calculate average positions



The `rmsd` module

```
>>> from ase.geometry.rmsd import find_crystal_reductions
>>> atoms = read("2016170.cif")
>>> result = find_crystal_reductions(atoms)
>>> result[0]
ReducedCrystal(rmsd=0.124,
               factor=2,
               atoms=Atoms(...),
               components=array([0, 1, 2, 1, 0, 2]))
```

In branch `crystal-reduction` (not merged)

Conclusion

Guessing costs is tedious.

Menus are good.

A good distance function provides many benefits.

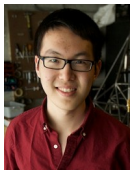
Acknowledgements

Mohnish Pandey

Mikkel Strange

Karsten Wedel Jacobsen

CAMD, Department of Physics, DTU



Edward Lee Pang

Pablo Parillo

Massachusetts Institute of Technology

CAMd

Center for Atomic-scale Materials Design

DANMARKS FRIE FORSKNINGSFOND

INDEPENDENT RESEARCH FUND DENMARK

DTU