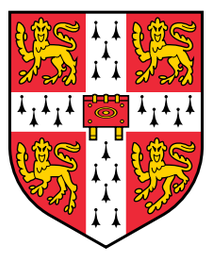
Lecture 5: fitting interatomic potentials II

Gábor Csányi



University of Cambridge
Engineering Laboratory

Interatomic potential

- Interaction of atoms when electrons are in their ground state

$$H = \langle \Psi | \mathcal{H} | \Psi \rangle = \sum_i \frac{p_i^2}{2m_i} + \sum_{ij} \frac{Z_i Z_j}{|q_i - q_j|} + \underbrace{V_{\mathrm{el}}(q_1, q_2, \ldots)}_{\text{“Interatomic potential”}}$$

difficult electronic problem

$V(q_1, q_2, \ldots)$

- This is used in *ab initio* molecular dynamics: expensive
  - Electronic problem solved explicitly
  - Potential is “global”, every atom interacts with every other

Traditional ideas for functional forms

- Pair potentials: Lennard-Jones, RDF-derived, etc.
- Three-body terms: Stillinger-Weber, MEAM, etc.
- Embedded Atom (no angular dependence)
- Bond Order Potential (BOP)
  Tight-binding-derived attractive term with pair-potential repulsion
- ReaxFF: kitchen-sink + hundreds of parameters
- Charge flow problem

$$\varepsilon_i = \frac{1}{2}\sum_j V_2(|r_{ij}|) + \sum_{jk} k(\theta_{ijk}-\theta_0)^2$$

$$\varepsilon_i = \Phi\left(\sum_j \rho(|r_{ij}|)\right)$$

Representation is implicit

These are NOT THE CORRECT functions.
Limited accuracy, not systematic

given by

GOAL: potentials ~~based on~~ quantum mechanics

The locality assumption

- Assume that potential is separable:

$$E_{\mathrm{tot}} = \sum_{\mathrm{atoms}\ i} \varepsilon(\mathbf{r}_1 - \mathbf{r}_i, \mathbf{r}_2 - \mathbf{r}_i, \ldots)$$

Finite range *atomic energy* function

$$E_{\mathrm{tot}} = \sum_{\mathrm{atoms}\ i} \varepsilon(\mathbf{r}_1 - \mathbf{r}_i, \mathbf{r}_2 - \mathbf{r}_i, \ldots) + \frac{1}{2}\sum_{i,j} \hat{L}_i \hat{L}_j \frac{1}{r_{ij}} + \sum_{i,j} \frac{\sigma_{ij}}{|r_{ij}|^6}$$

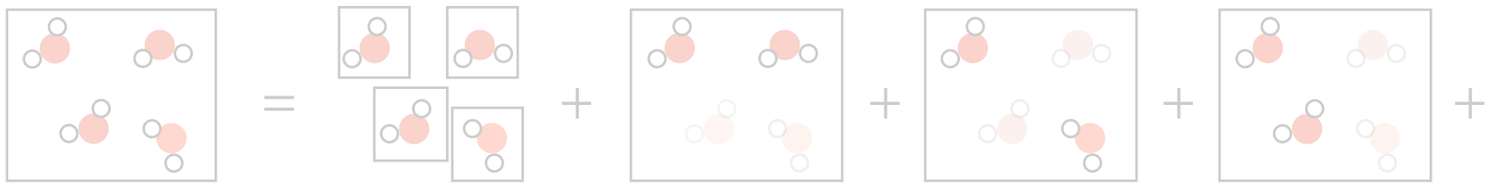
$$\hat{L}_i = q_i + \mathbf{p}_i \cdot \nabla_i + \ldots$$

Coulomb

pair dispersion

- Uncontrolled approximation
- Validity can be studied numerically
- Beginnings of rigorous treatment (C Ortner et al 2015)

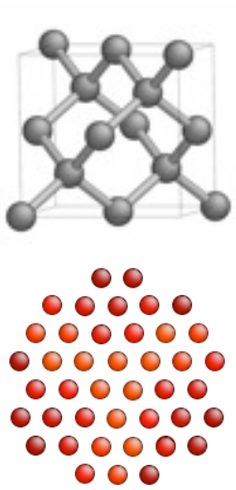
Weak interactions



Many-body
(cluster) expansion

+ higher multi-body terms... + analytic long range terms

Strong interactions

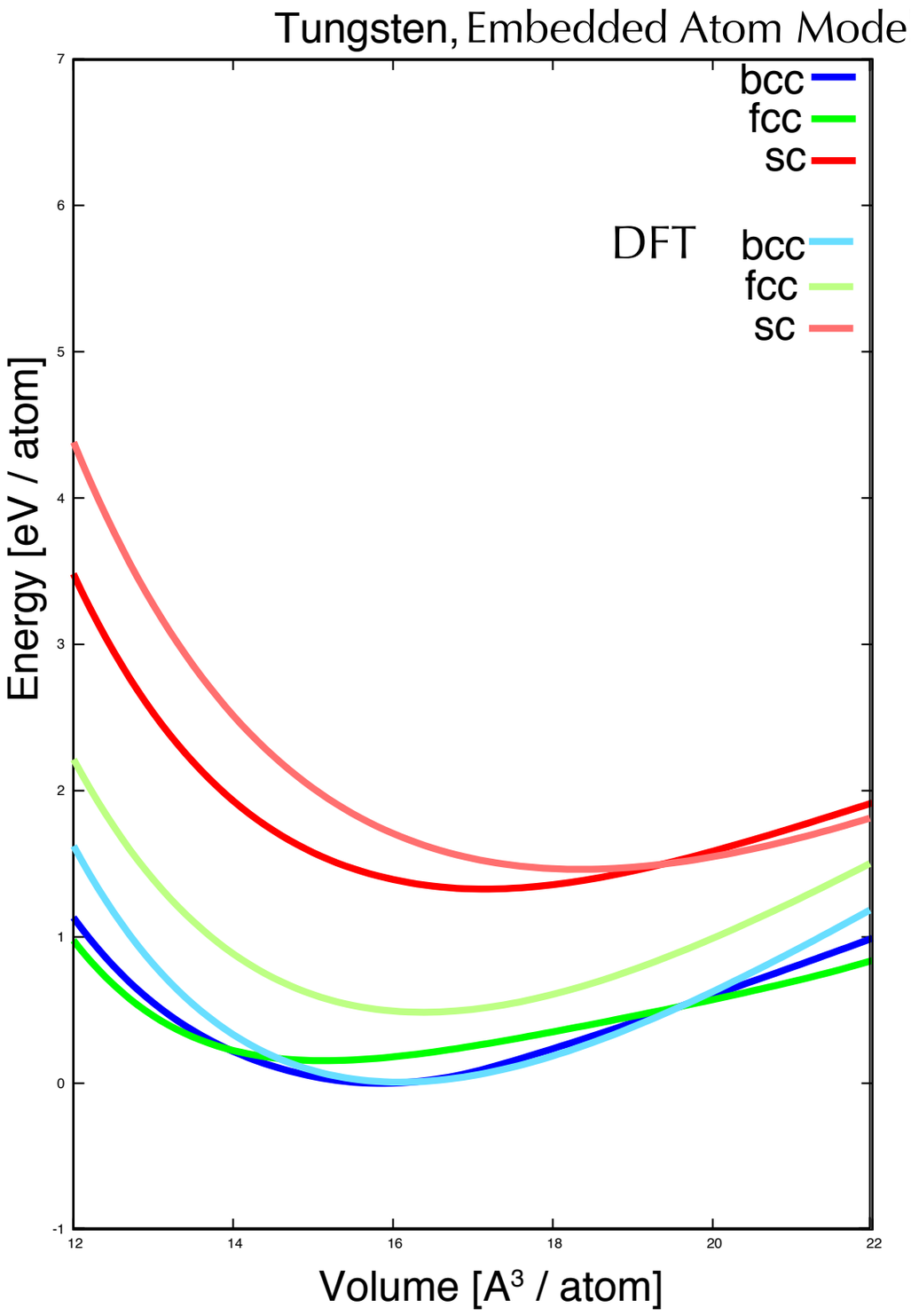


$$E = \sum_i \varepsilon(q_1^{(i)}, q_2,^{(i)} \ldots, q_M^{(i)})$$

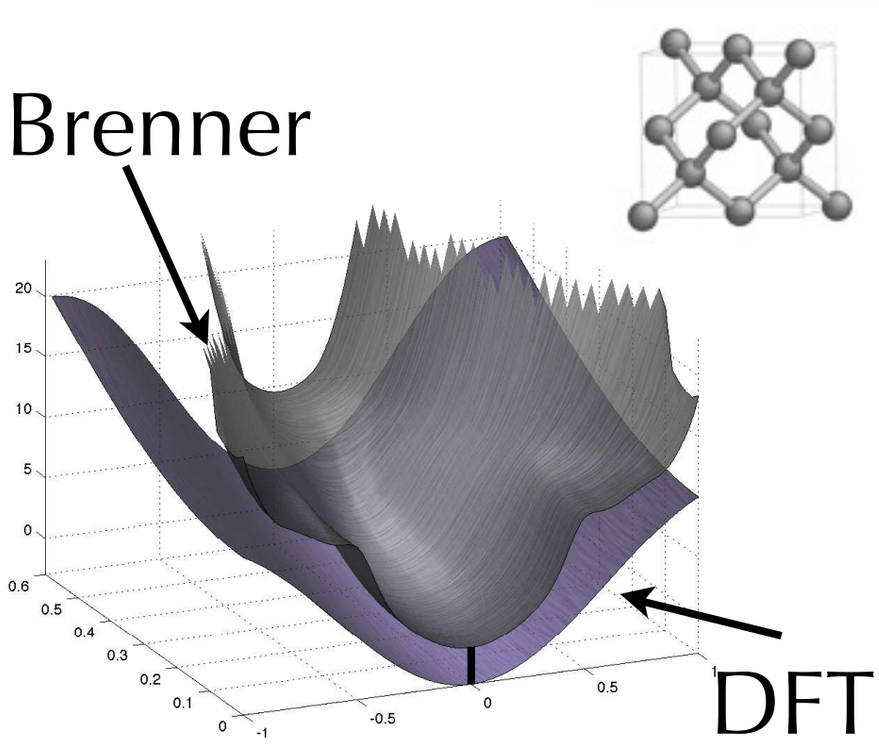
+ analytic long range terms

Need a good kernel between full atomic neighbourhoods!

Fundamental limitation of analytic functional form



Carbon (diamond)



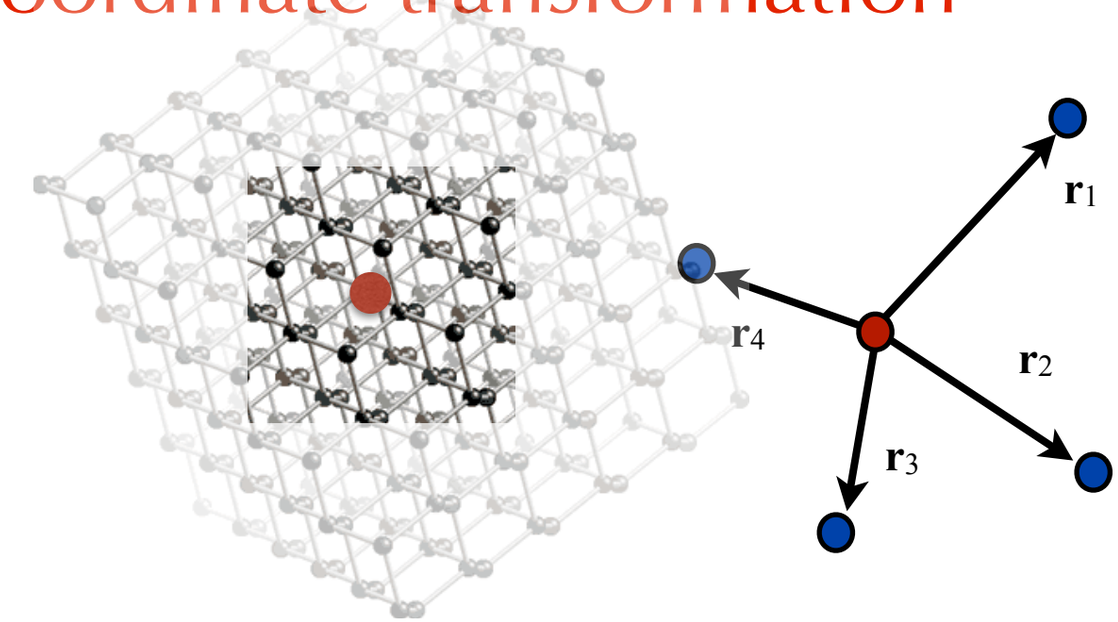
Elastic constants:

	DFT	Brenner
C_{11}	1118	1061
C_{12}	151	133
C_{44}^0	610	736
C_{44}	603	717

Representing the atomic neighbourhood in strongly bound materials

- What are the arguments of the atomic energy ε ?
  Need a *representation,* i.e. a coordinate transformation

$$\{\mathbf{r}_i\} \rightarrow \mathbf{q}$$



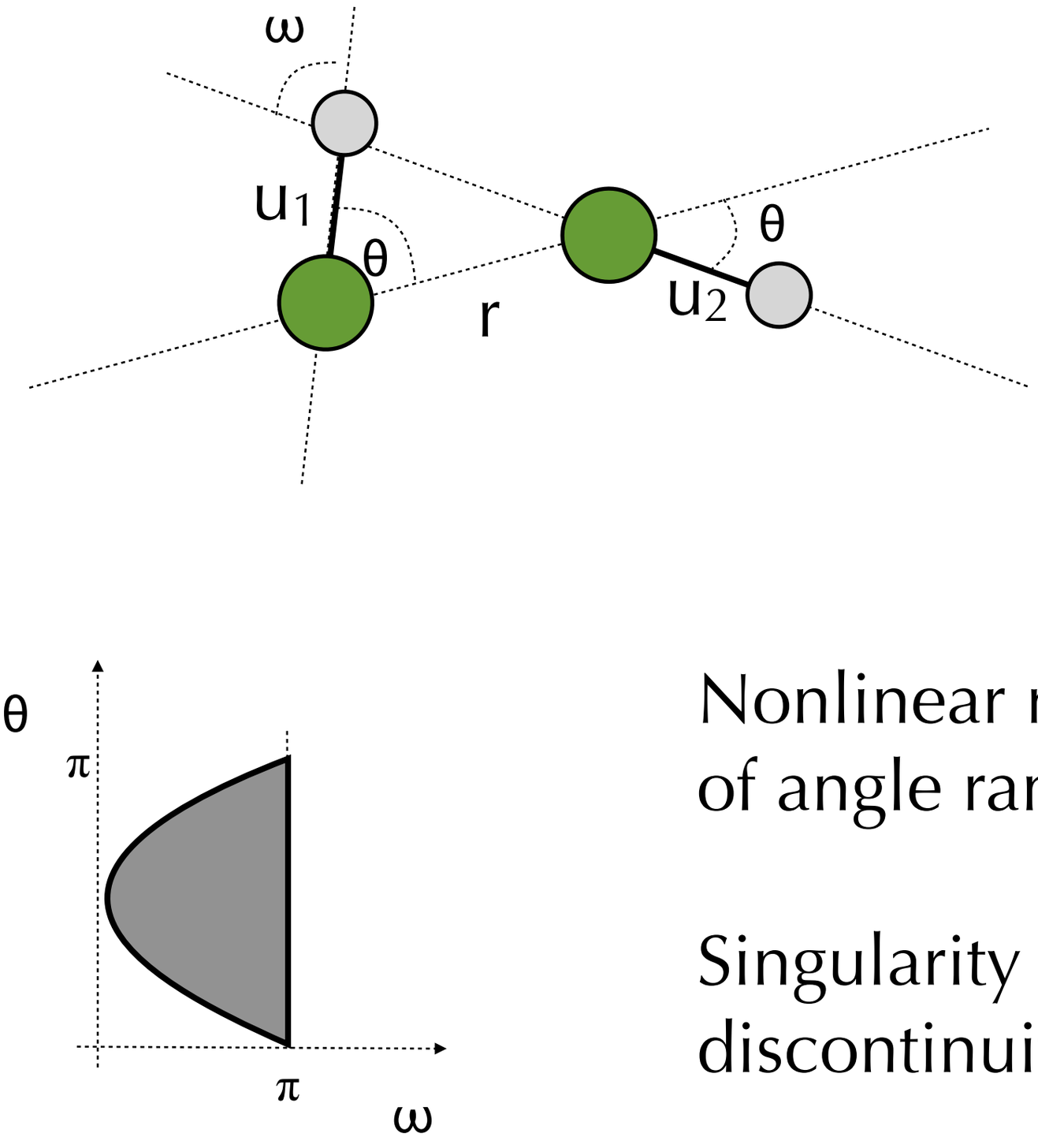
  - Exact symmetries:
    - Global Translation
    - Global Rotation
    - Reflection
    - Permutation of atoms
  - Faithful: different configurations correspond to different representations
  - Continuous, differentiable, and smooth (i.e. slowly changing with atomic position) ("Lipschitz diffeomorphic")

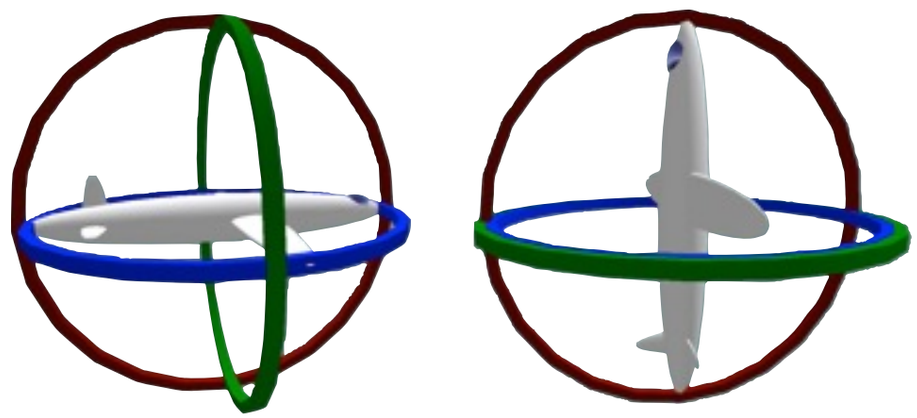
- Rotational invariance by itself is easy: $\mathbf{q} \equiv R_{ij} = \mathbf{r}_i \cdot \mathbf{r}_j$ (Weyl)

$$\begin{bmatrix} \mathbf{r}_1\cdot\mathbf{r}_1 & \mathbf{r}_1\cdot\mathbf{r}_2 & \cdots & \mathbf{r}_1\cdot\mathbf{r}_N \\ \mathbf{r}_2\cdot\mathbf{r}_1 & \mathbf{r}_2\cdot\mathbf{r}_2 & \cdots & \mathbf{r}_2\cdot\mathbf{r}_N \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{r}_N\cdot\mathbf{r}_1 & \mathbf{r}_N\cdot\mathbf{r}_2 & \cdots & \mathbf{r}_N\cdot\mathbf{r}_N \end{bmatrix}$$

  - Complete, but not invariant permutationally
  - Not continuous with changing number of neighbours

Example: pair of dimers



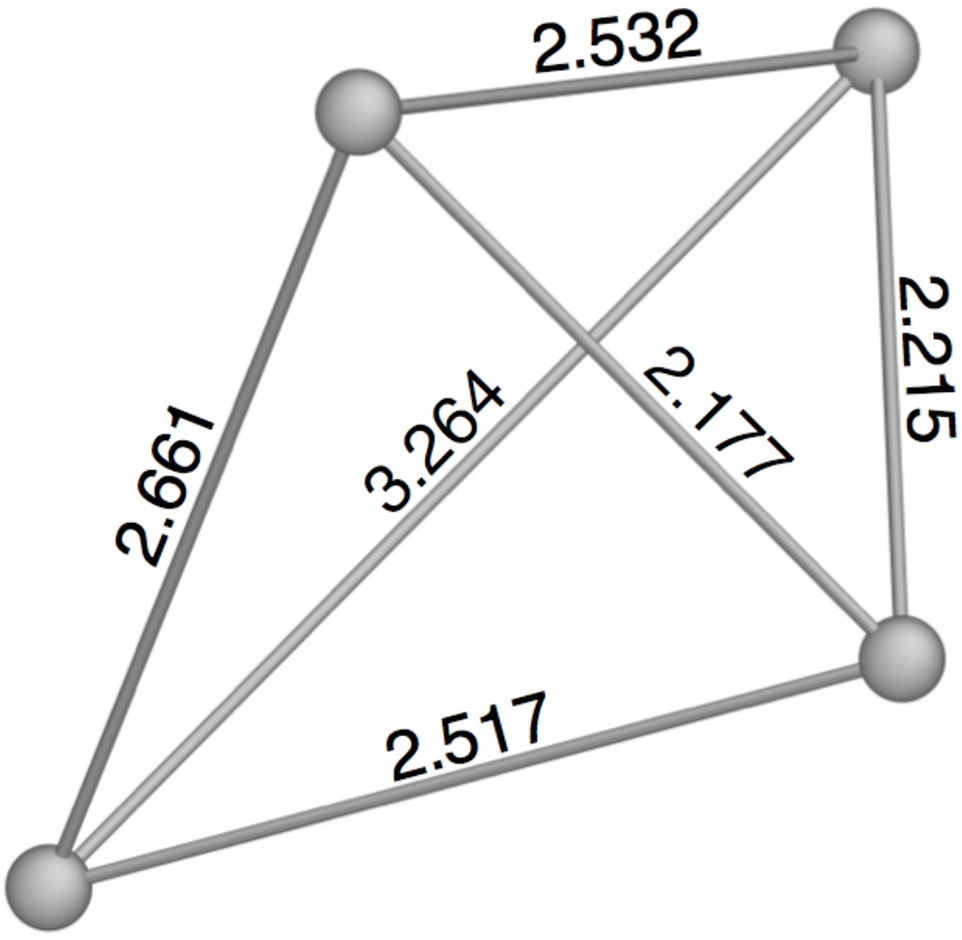


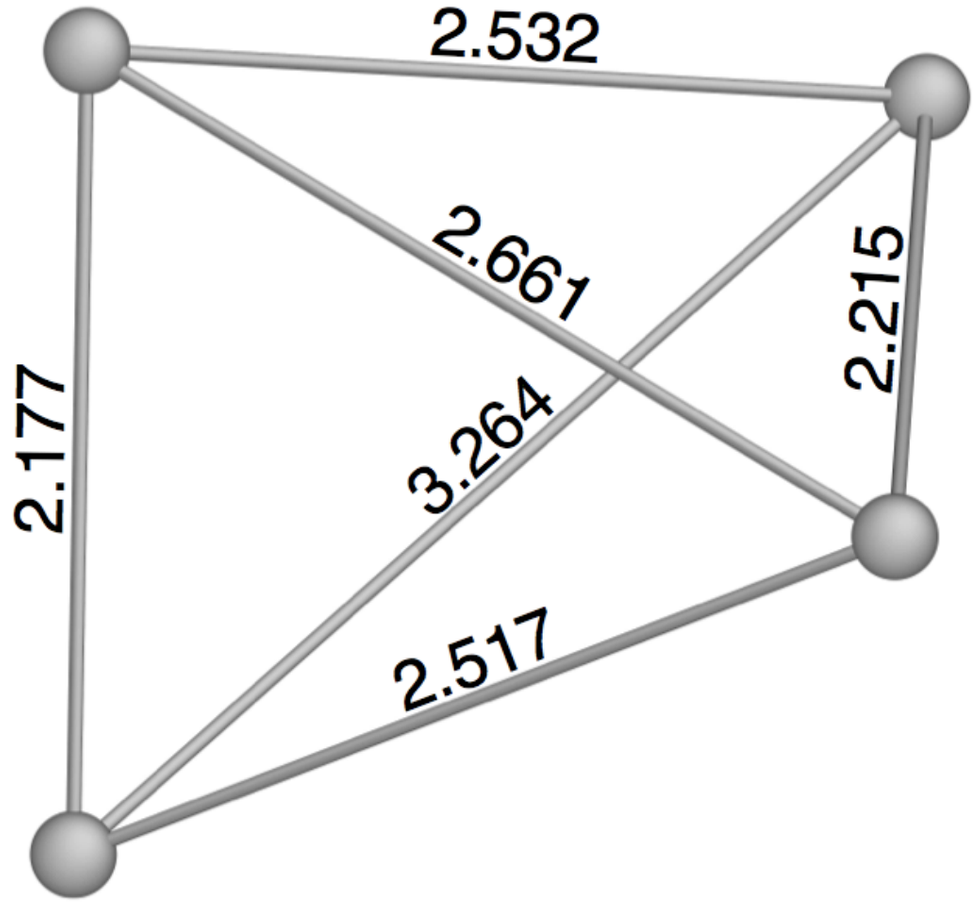
"Gimbal lock"

Nonlinear restriction
of angle ranges

Singularity at the "poles":
discontinuity in descriptor derivatives

Drop atom ordering ?



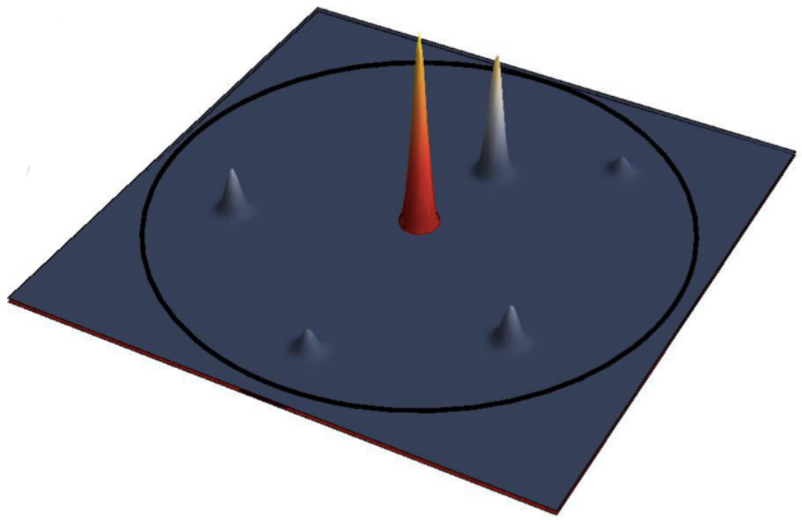


Not unique!

Atomic neighbour density function

$$\varepsilon(\mathbf{r}_1 - \mathbf{r}_i, \mathbf{r}_2 - \mathbf{r}_i, \ldots) \equiv \varepsilon[\rho_i(\mathbf{r})]$$

$$\rho_i(\mathbf{r}) = s\delta(\mathbf{r}) + \sum_j \delta(\mathbf{r} - \mathbf{r}_{ij}) f_{\mathrm{cut}}(|\mathbf{r}_{ij}|)$$



- Translation ✓
- Permutation ✓
- Need rotationally invariant features of $\rho(\mathbf{r})$

Rotational Invariance

$$\rho(\hat{\mathbf{r}}) = \sum_{lm} c_{lm} Y_{lm}(\hat{\mathbf{r}})$$

$$\rho(\hat{R}\hat{\mathbf{r}}) = \sum_{lm} c_{lm} Y_{lm}(\hat{R}\hat{\mathbf{r}}) = \sum_{lm} c_{lm} \sum_{m'} D^{l}_{mm'}(\hat{R}) Y_{lm'}(\hat{\mathbf{r}})$$

$$= \sum_{lm'} \left(\sum_{m} D^{l}_{mm'} c_{lm} \right) Y_{lm'}(\hat{\mathbf{r}})$$

Transformation of coefficients:

$$\mathbf{c}_l \xrightarrow{\hat{R}} \mathbf{D}^l \mathbf{c}_l$$

- Wigner matrices are unitary:

$$\mathbf{D}^{-1} = \mathbf{D}^{\dagger}$$

- Construct invariants:

$$p_l = \mathbf{c}_l^{\dagger} \mathbf{c}_l \rightarrow \left(\mathbf{c}_l^{\dagger} \mathbf{D}^{l\,\dagger} \right) \left(\mathbf{D}^l \mathbf{c}_l \right) = \mathbf{c}_l^{\dagger} \mathbf{c}_l$$

- Equivalent to Steinhardt bond order Q_2 , Q_4 , Q_6 ... parameters
- Higher order invariants possible, (generalisation of Steinhardt W)

First few terms of power spectrum for 3 atoms

$$p_0 = \frac{9}{4\pi}$$

$$p_1 = \frac{3}{4\pi}\left(\sum_{jk}\cos\theta_{ijk} + 3\right)$$

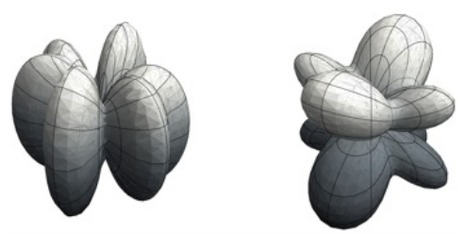
$$p_2 = \frac{5}{4\pi}\left(\frac{3}{2}\sum_{jk}\cos^2\theta_{ijk} + 6\right)$$

$$p_3 = \frac{7}{4\pi}\left(\frac{5}{2}\sum_{jk}\cos^3\theta_{ijk} - \frac{3}{2}\sum_{jk}\cos\theta_{ijk} + 3\right)$$

$$p_4 = \frac{9}{16\pi}\left(\frac{35}{2}\sum_{jk}\cos^4\theta_{ijk} - 15\sum_{jk}\cos^2\theta_{ijk} + 13\right)$$

- Polynomials of Weyl invariants
- Highly oscillatory for large *l : not smooth*
- Different *l* channels uncoupled

$f_1 = Y_{22} + Y_{2-2} + Y_{33} + Y_{3-3}$



$f_2 = Y_{21} + Y_{2-1} + Y_{32} + Y_{3-2}$

Several recent proposals

- Behler-Parrinello "symmetry functions" (2007)

$$G_i^2 = \sum_{j \neq i} e^{-\eta(r_{ij} - r_s)^2}$$

$$G_i^4 = 2^{1-\zeta} \sum_{k > j \neq i} (1 + \lambda \cos\theta_{ijk})^\zeta e^{-\eta(r_{ij}^2 + r_{ik}^2)}$$

- Bispectrum *(A. P. Bartók et al. 2010)*
- Class of very similar invariant descriptors
- Systematically refinable, complete

$$d_{n,l}^i = \sum_{j,k} g_n(r_{ij}) g_n(r_{ik}) \cos(l\theta_{ijk})$$

Bispectrum

$$\mathbf{c}_{l_1} \otimes \mathbf{c}_{l_2} \rightarrow (\mathbf{D}_{l_1} \otimes \mathbf{D}_{l_2})\, \mathbf{c}_{l_1} \otimes \mathbf{c}_{l_2}$$

$$\mathbf{D}_{l_1} \otimes \mathbf{D}_{l_2} = \mathbf{C}^{\dagger}_{l_1,l_2} \left[\bigoplus_{l=|l_1-l_2|}^{l_1+l_2} \mathbf{D}_l \right] \mathbf{C}_{l_1,l_2}$$

$$\mathbf{C}_{l_1 l_2} \mathbf{c}_{l_1} \otimes \mathbf{c}_{l_2} = \bigoplus_{l=|l_1-l_2|}^{l_1+l_2} \mathbf{g}_{l,l_1,l_2}$$

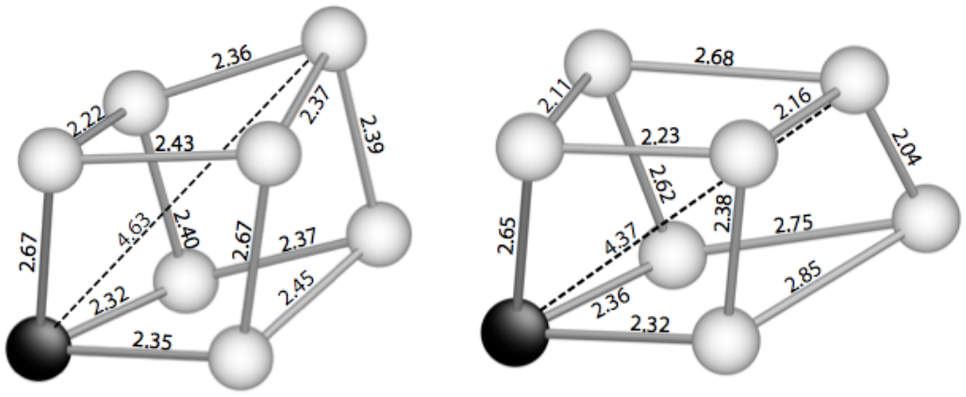
$$\mathbf{C}_{l_1 l_2} \mathbf{c}_{l_1} \otimes \mathbf{c}_{l_2} \rightarrow \left[\bigoplus_{l=|l_1-l_2|}^{l_1+l_2} \mathbf{D}_l \right] \mathbf{C}_{l_1 l_2} \mathbf{c}_{l_1} \otimes \mathbf{c}_{l_2} = \bigoplus_{l=|l_1-l_2|}^{l_1+l_2} \mathbf{D}_l \mathbf{g}_{l,l_1,l_2}$$

$$\mathbf{g}_{l,l_1,l_2} \rightarrow \mathbf{D}_l \mathbf{g}_{l,l_1,l_2}$$

$$B_{l,l_1,l_2} = \mathbf{c}_l^{\dagger} \mathbf{g}_{l,l_1,l_2} \qquad B_{l,l_1,l_2} \rightarrow \mathbf{c}_l^{\dagger} \mathbf{D}_l^{\dagger} \mathbf{D}_l \mathbf{g}_{l,l_1,l_2} = B_{l,l_1,l_2}$$

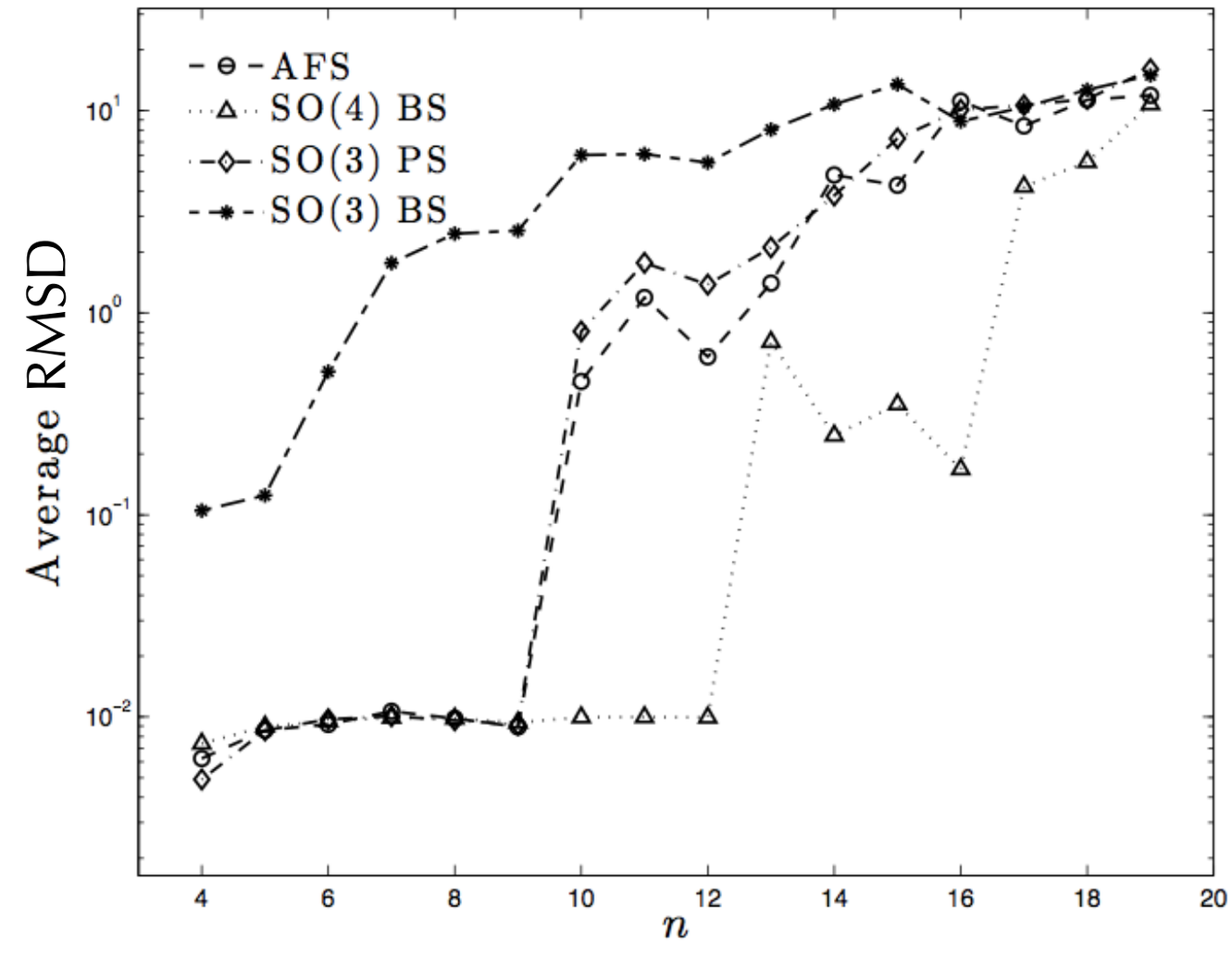
Radial part ? $B^{n,n_1,n_2}_{l,l_1,l_2}$ ...huge number of invariants

Uniqueness: environment reconstruction tests of *truncated* invariant set



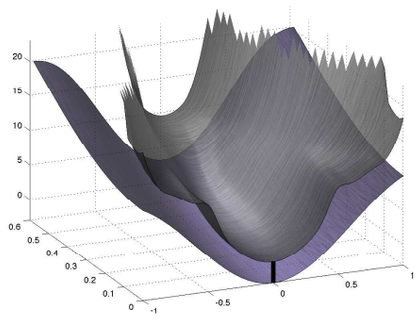
1. Select target configuration
2. Randomise neighbours
3. Try to reconstruct target by matching descriptors

Error in distinguishing

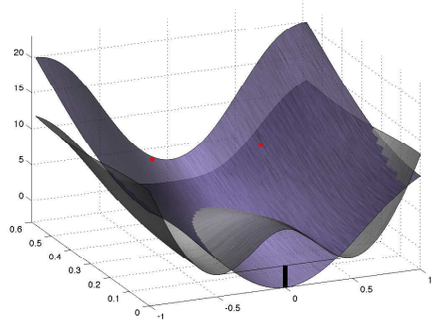


Number of neighbours

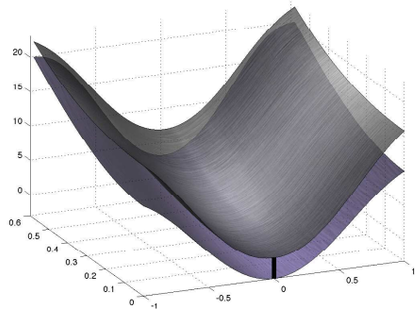
Diamond



Brenner potential



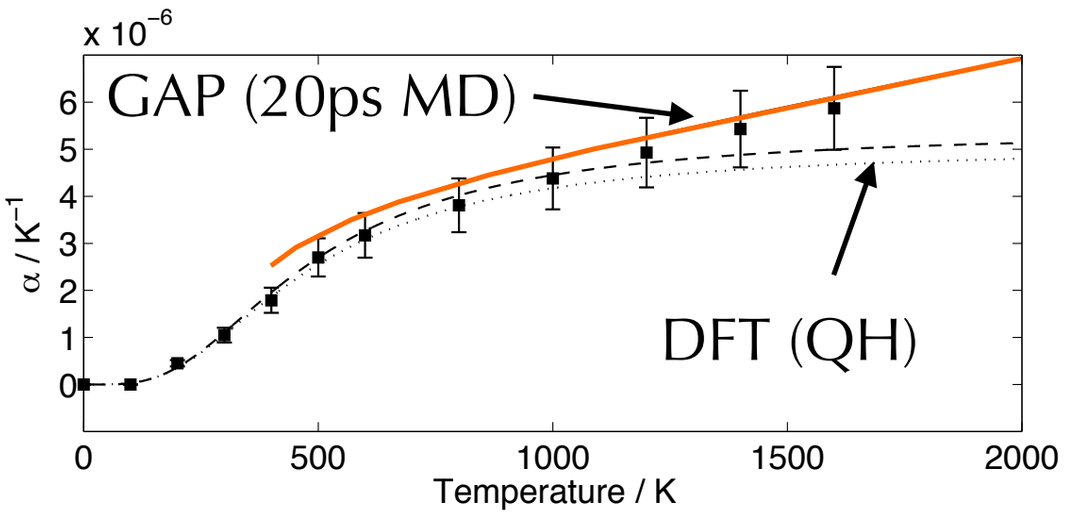
1 data point



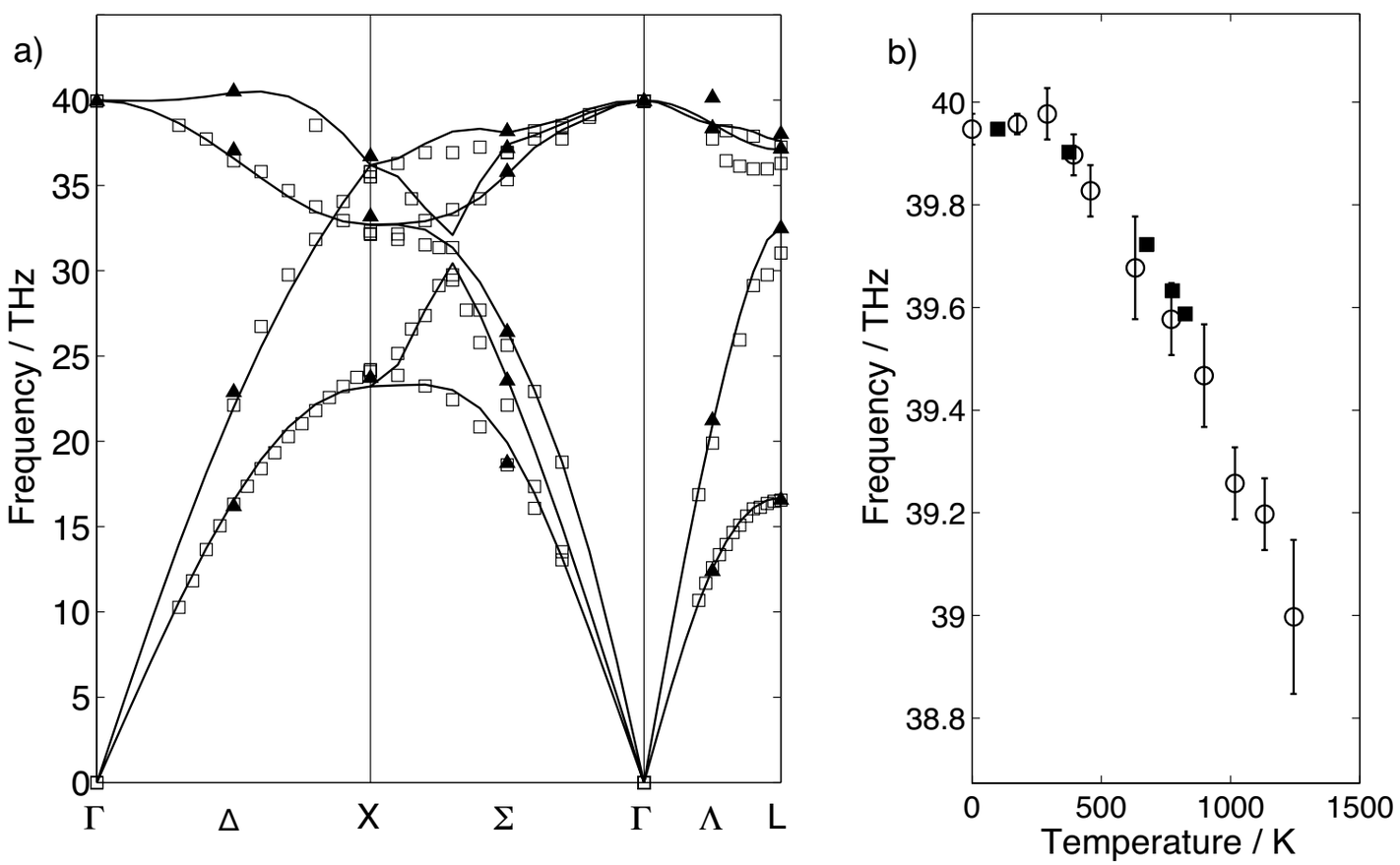
50 random data points

300 configurations from ab initio MD:

	DFT	GAP	Brenner
C_{11}	1118	1081	1061
C_{12}	151	157	133
C^0_{44}	610	608	736
C_{44}	603	601	717

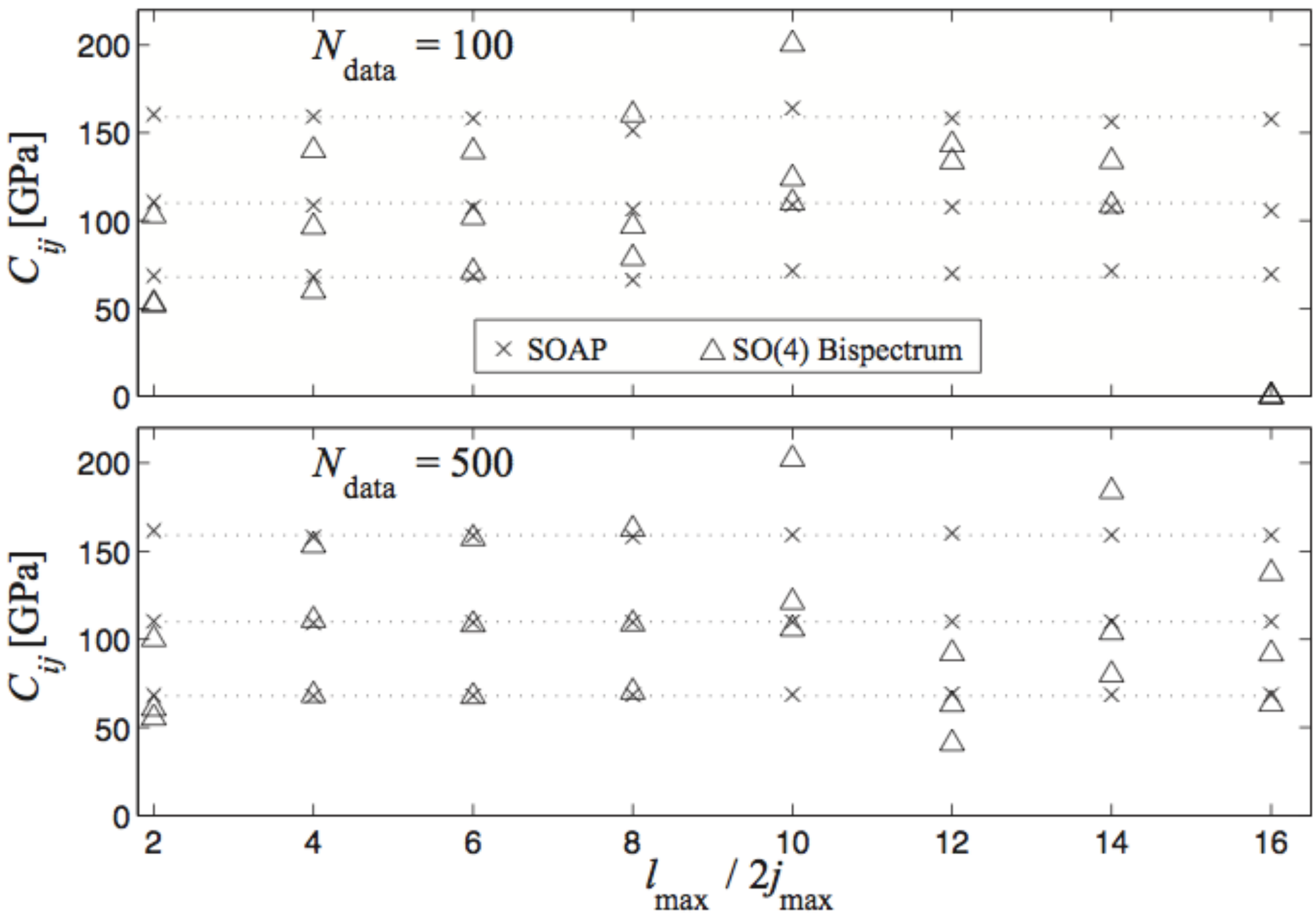


Thermal expansion



Phonon spectrum

Stability of fit: elastic constants of Si



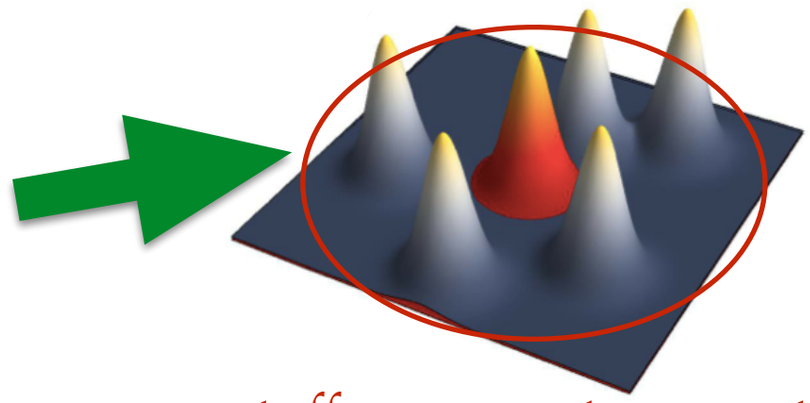
Order of angular momentum expansion

SOAP: Smooth Overlap of Atomic Positions

$$\rho_i(\mathbf{r}) = \sum_j \exp\left(-|\mathbf{r} - \mathbf{r}_{ij}|^2/2\sigma^2\right) = \sum_j \sum_{lm} c_{nlm}^{(i)j} g_n(r) Y_{lm}(\hat{\mathbf{r}})$$

- Overlap integral

$$S(\rho_i, \rho_{i'}) = \int \rho_i(\mathbf{r})\rho_{i'}(\mathbf{r})d\mathbf{r},$$



cutoff: compact support

- Integrate over all 3D rotations:

$$k(\rho_i, \rho_{i'}) = \int \left|S(\rho_i, \hat{R}\rho_{i'})\right|^2 d\hat{R} = \int d\hat{R} \left| \int \rho_i(\mathbf{r})\rho_{i'}(\hat{R}\mathbf{r})d\mathbf{r} \right|^2$$

- After LOTS of algebra: SOAP kernel

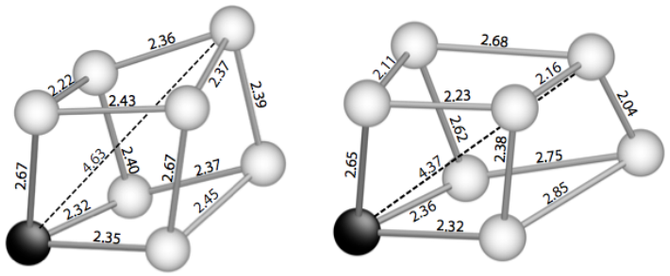
$$k(\rho_i, \rho_{i'}) = \sum_{n,n',l} p_{nn'l}^{(i)} p_{nn'l}^{(i')}$$

$$p_{nn'l} = \mathbf{c}_{nl}^{\dagger}\mathbf{c}_{n'l}$$

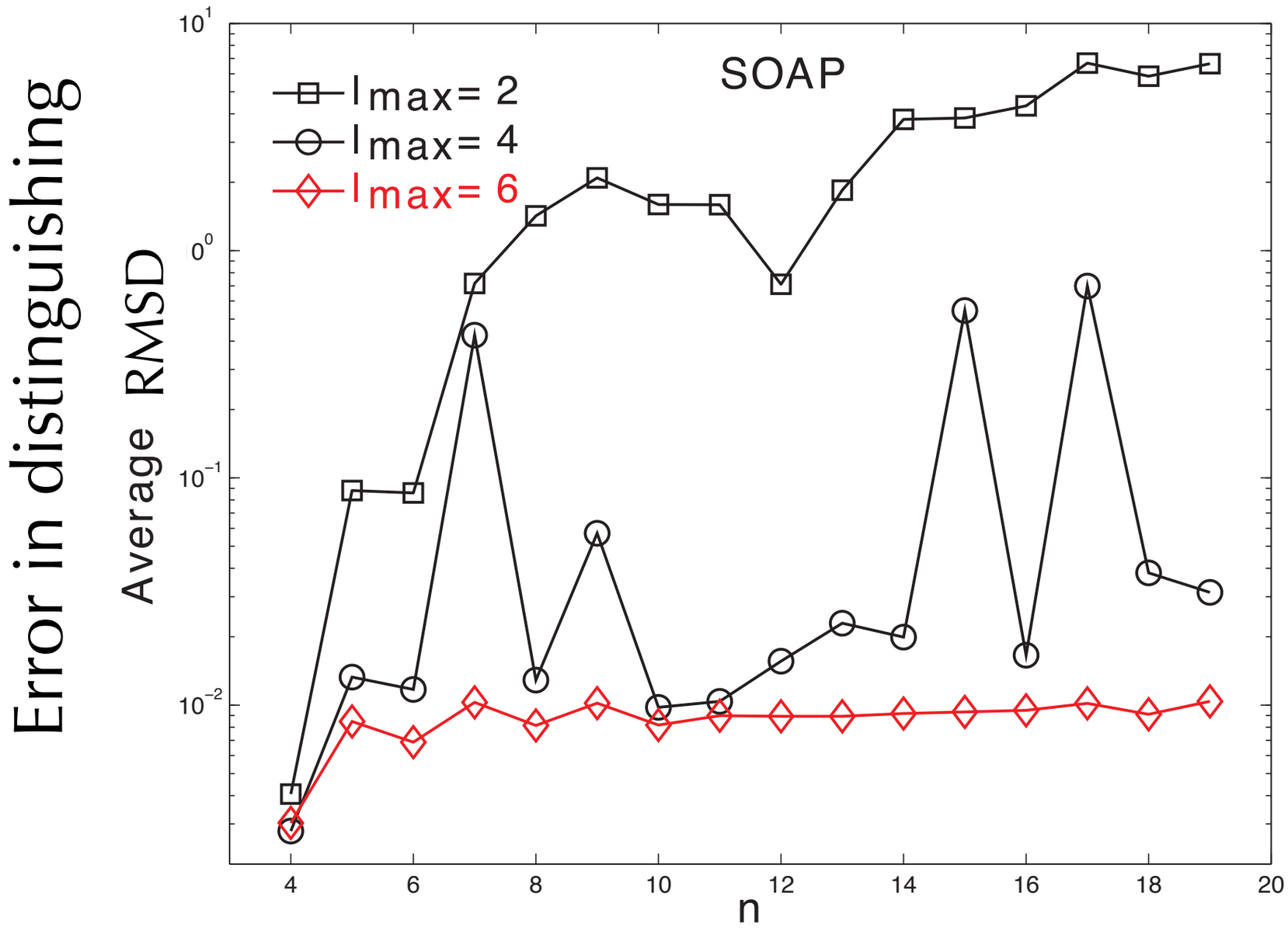
$\propto$ Steinhardt Q_l

$$\mathbf{K}_{ij} \propto |k(\rho_i, \rho_j)|^{\xi}$$

SOAP kernel is unique



1. Select target configuration
2. Randomise neighbours
3. Try to reconstruct target by matching descriptors



How to generate databases?

- Target applications: large systems
- Capability of full quantum mechanics (QM): small systems

QM MD on "representative" small systems:

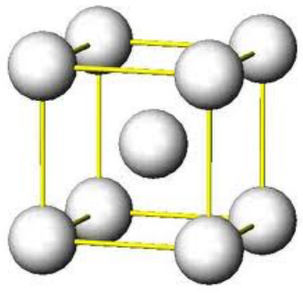
sheared primitive cell → elasticity
large unit cell → phonons
surface unit cells → surface energy
gamma surfaces → screw dislocation
vacancy in small cell → vacancy
vacancy @ gamma surface → vacancy near dislocation

Iterative refinement

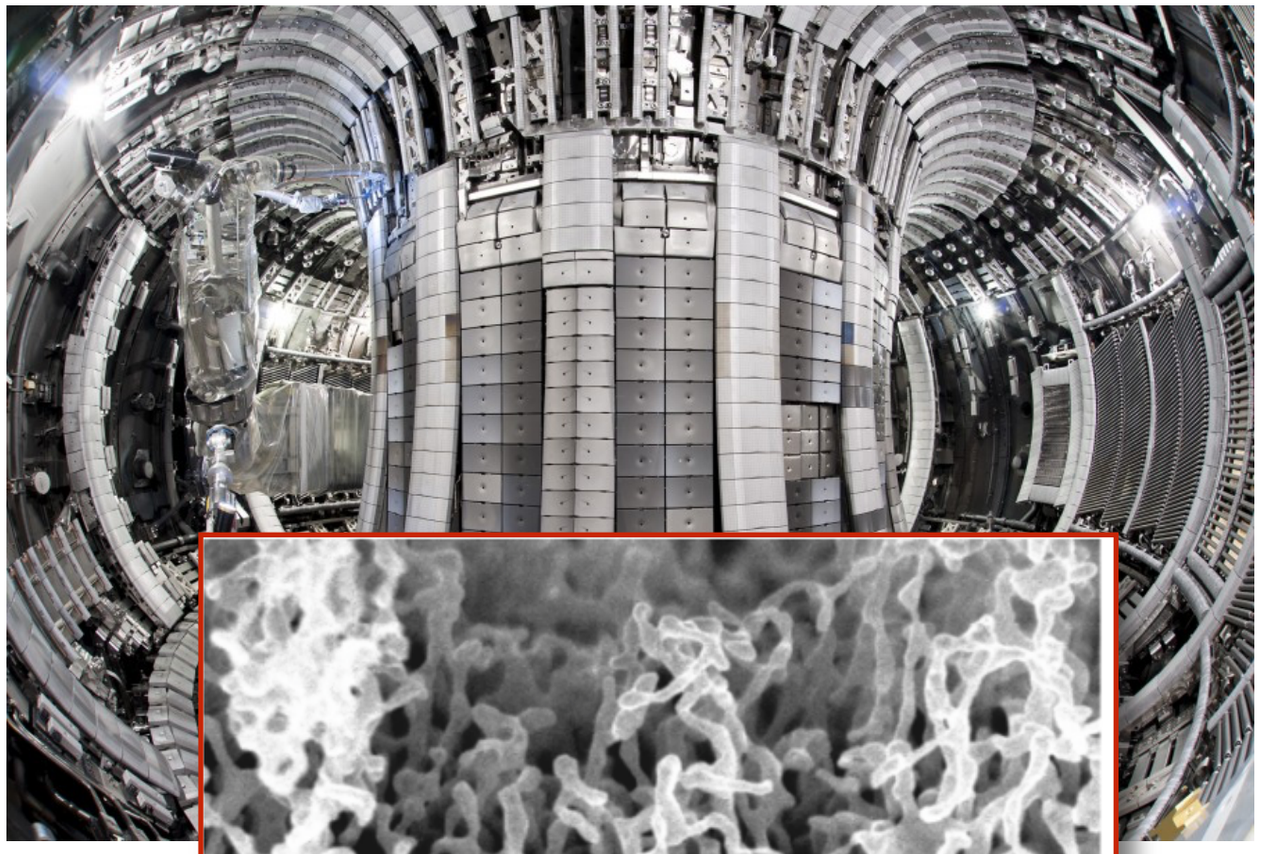
1. QM MD → Initial database
2. Model MD
3. QM → Revised database

What is the acceptable validation protocol?
How far can the domain of validity be extended?

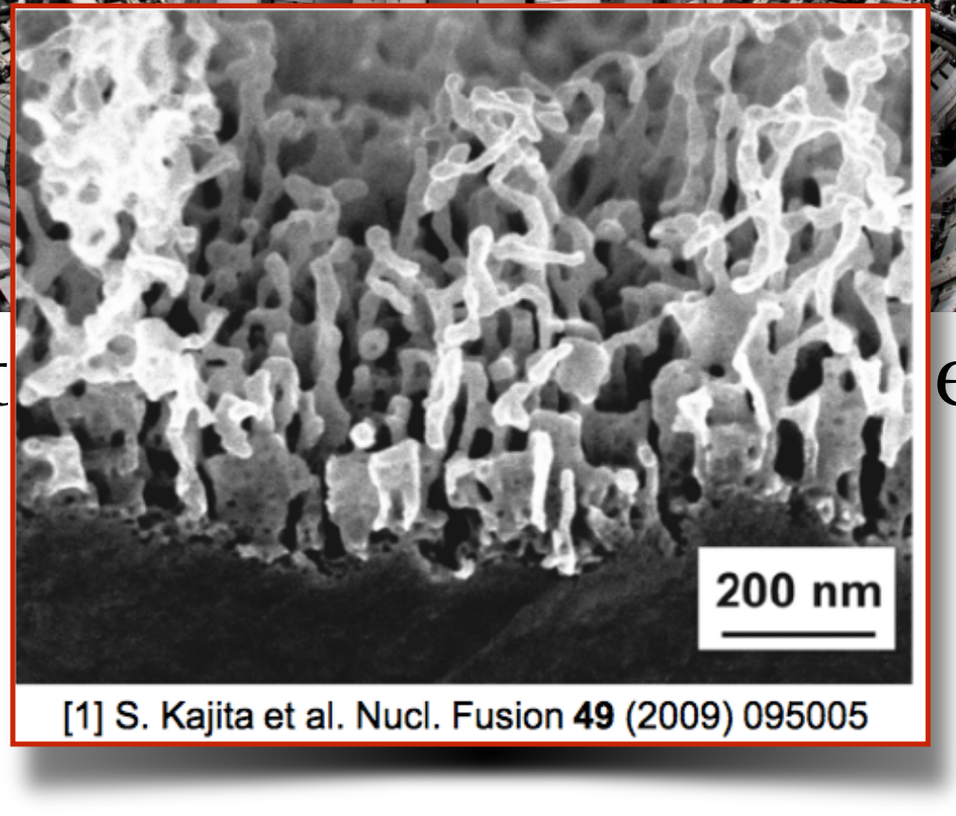
Example: tungsten



Melting point: 3422°C

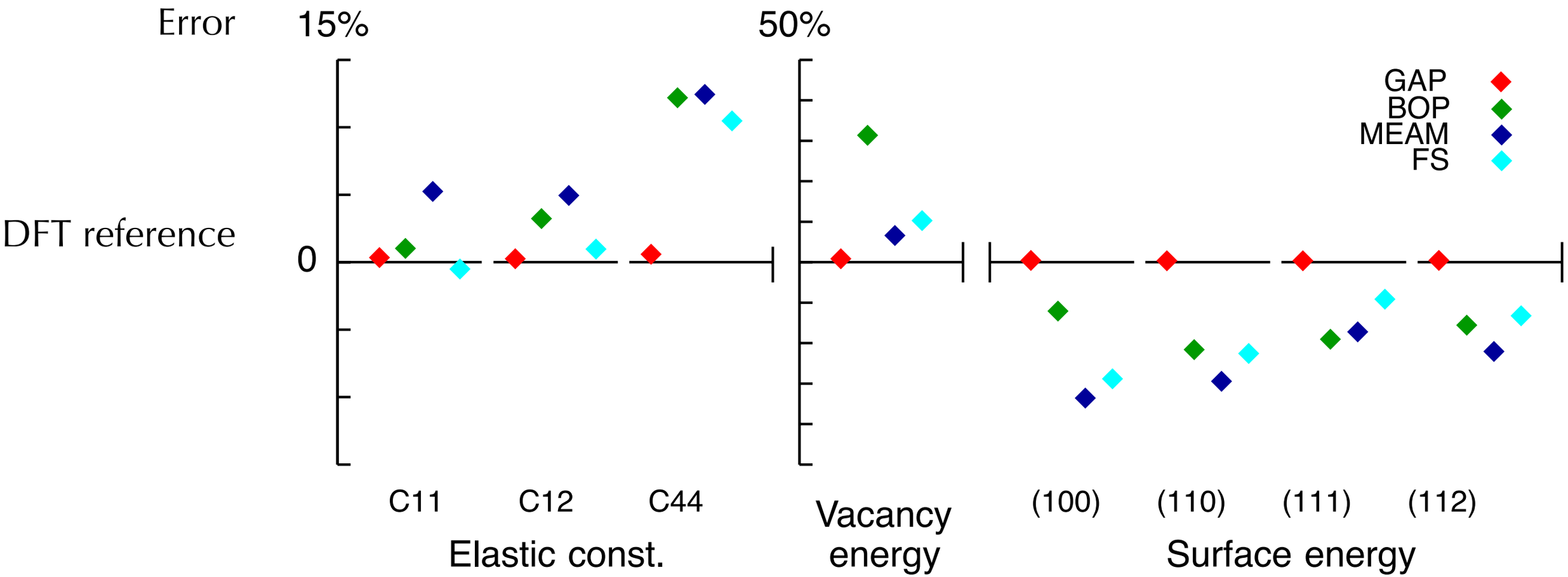


Int e



[1] S. Kajita et al. Nucl. Fusion **49** (2009) 095005

Potentials for tungsten



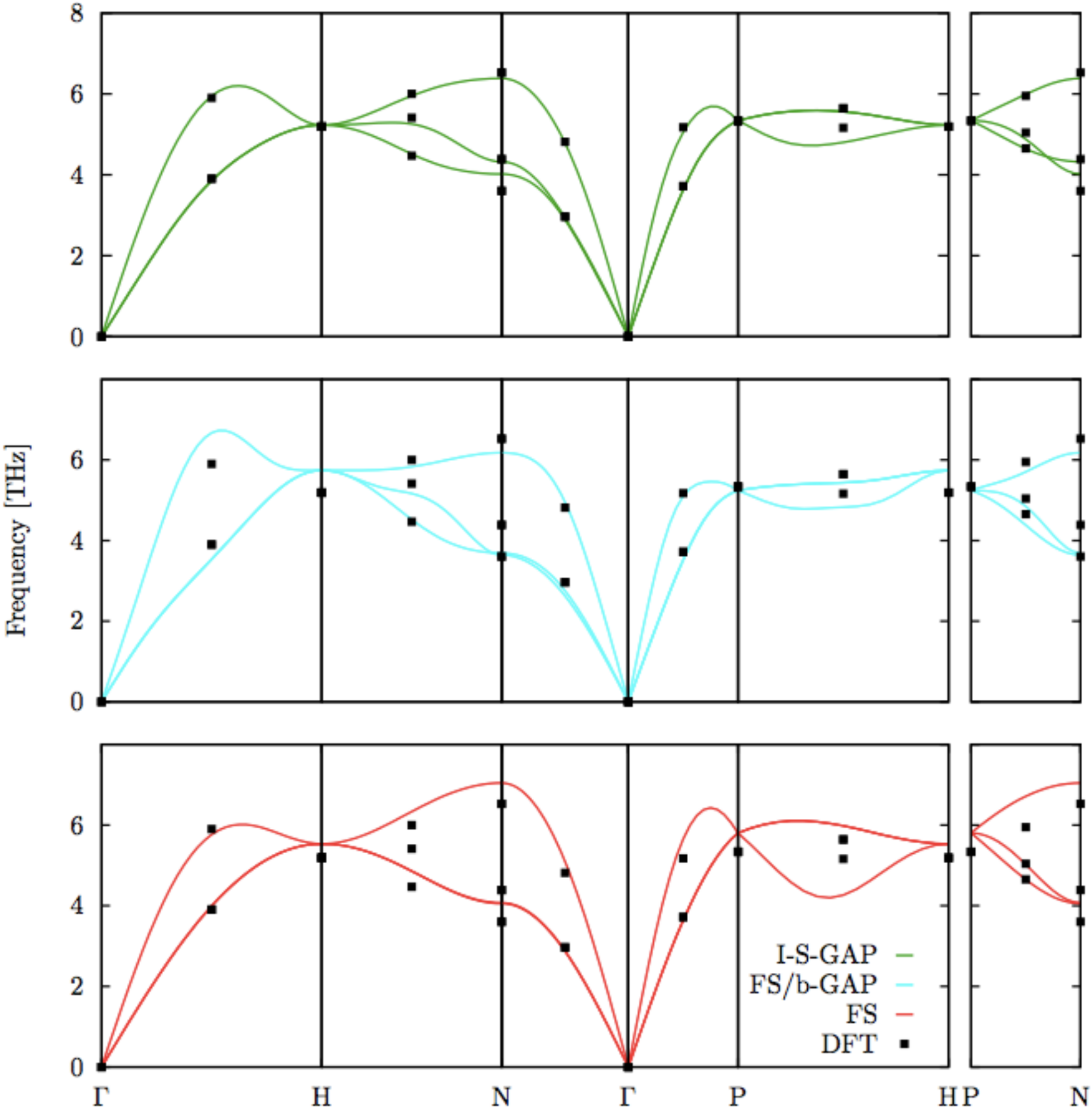
DFT code	CASTEP [37] (version 6.01)
Exchange-correlation functional	PBE
Pseudopotential	Ultrasoft (valence $5s^2 5p^6 5d^4 6s^2$)
Plane-wave energy cutoff	600 eV
Maximum k-point spacing	0.015 Å^{-1}
Electronic smearing scheme	Gaussian
Smearing width	0.1 eV

Building up databases for tungsten (W)

	Database:	Computational cost[a] [ms/atom]	Elastic constants[b] [GPa]	Phonon spectrum[b] [THz]	Vacancy formation[c] [eV]	Surface energy[b] [eV/Å^2]	Dislocation structure[d] [Å^{-1}]	Dislocation-vacancy binding energy [eV]	Peierls barrier [eV/b]
GAP_1 :	2000 × primitive unit cell with varying lattice vectors	24.70	0.623	*0.583*	*2.855*	*0.1452*	*0.0008*		
GAP_2 :	GAP_1 + 60 × 128 atom cell	51.05	0.608	0.146	*1.414*	*0.1522*	*0.0006*		
GAP_3 :	GAP_2 + vacancy in: 400 × 53 atom cell, 20 × 127 atom cell	63.65	0.716	0.142	0.018	*0.0941*	*0.0004*		
GAP_4 :	GAP_3 + (100), (110), (111), (112) surfaces 180 × 12 atom cell; (110), (112) gamma surfaces 6183 × 12 atom cell	86.99	0.581	0.138	0.005	0.0001	0.0002	*-0.960*	*0.108*
GAP_5 :	GAP_4 + vacancy in: (110), (112) gamma surface 750 × 47 atom cell	93.86	0.865	0.126	0.011	0.0001	0.0002	-0.774	*0.154*
GAP_6 :	GAP_5 + $\frac{1}{2}\langle 111\rangle$ dislocation quadrupole 100 × 135 atom cell	93.33	0.748	0.129	0.015	0.0001	0.0001	-0.794	0.112

[a] Time on a single CPU core of Intel Xeon E5-2670 2.6GHz, [b] RMS error, [c] formation energy error, [d] RMS error of Nye tensor over the 12 atoms nearest the dislocation core, cf. Figure 2.

Comparison of phonon spectra

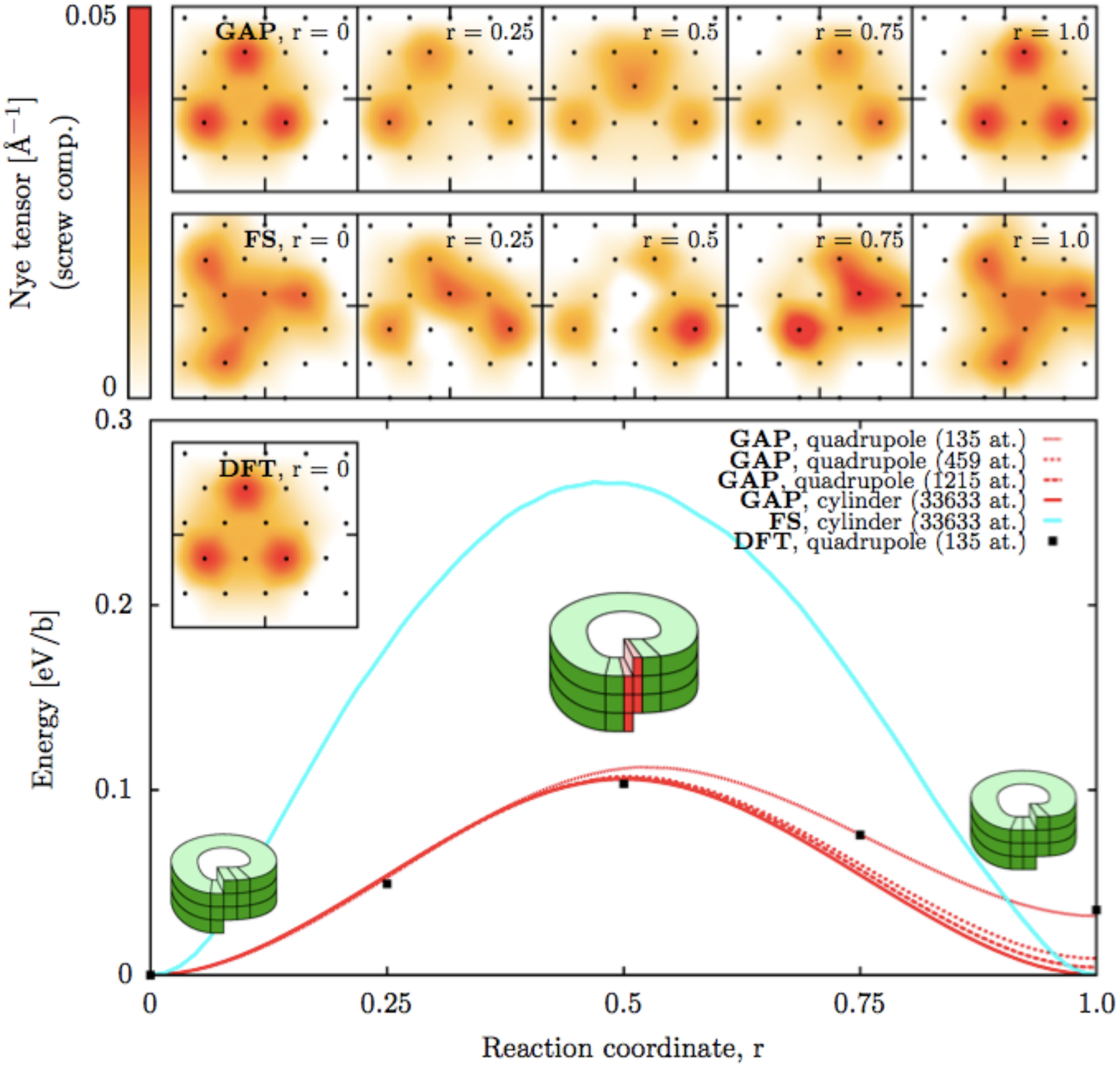


SOAP-based potential

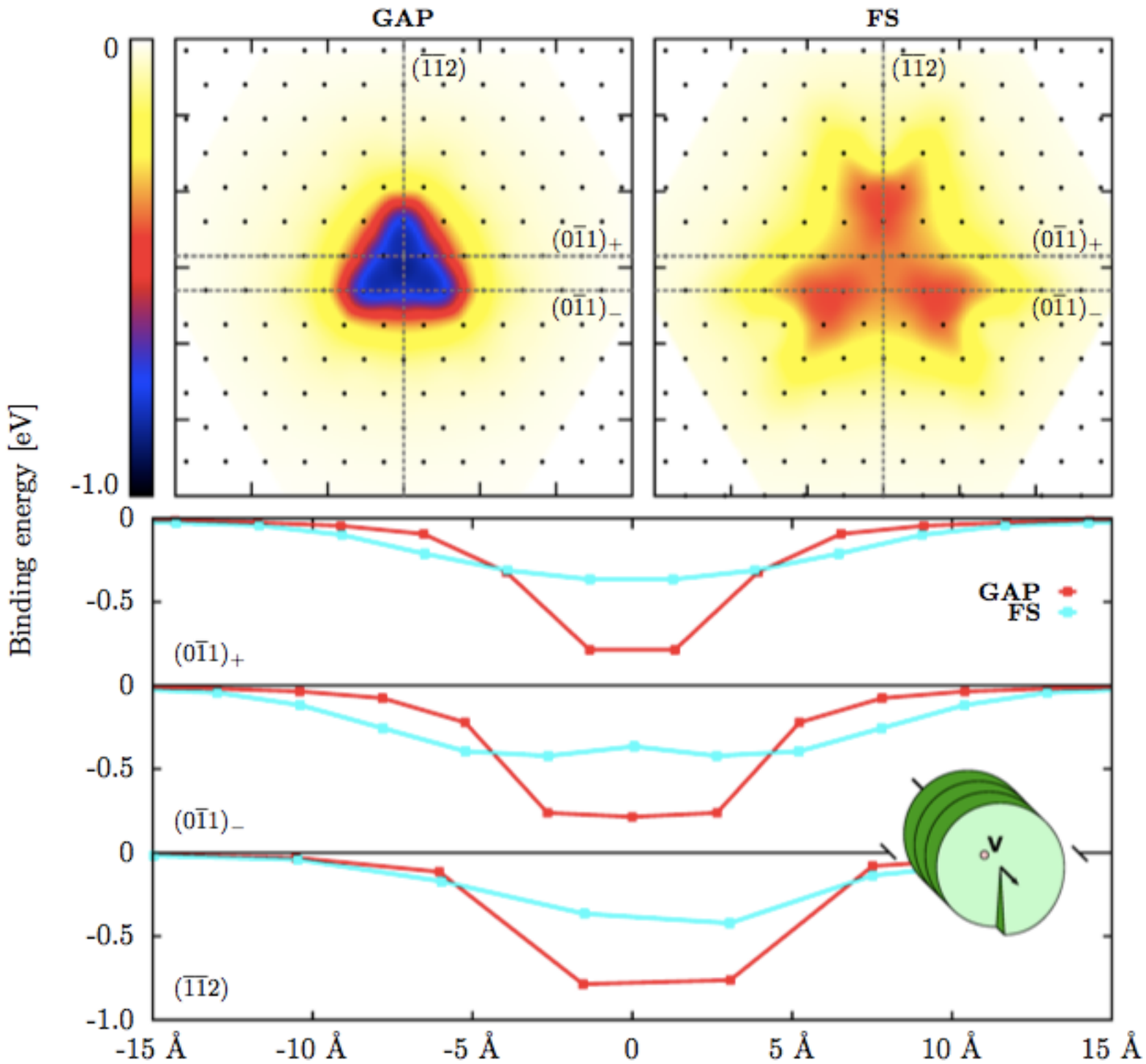
Bispectrum-based potential

Empirical potential

Peierls barrier for screw dislocation glide



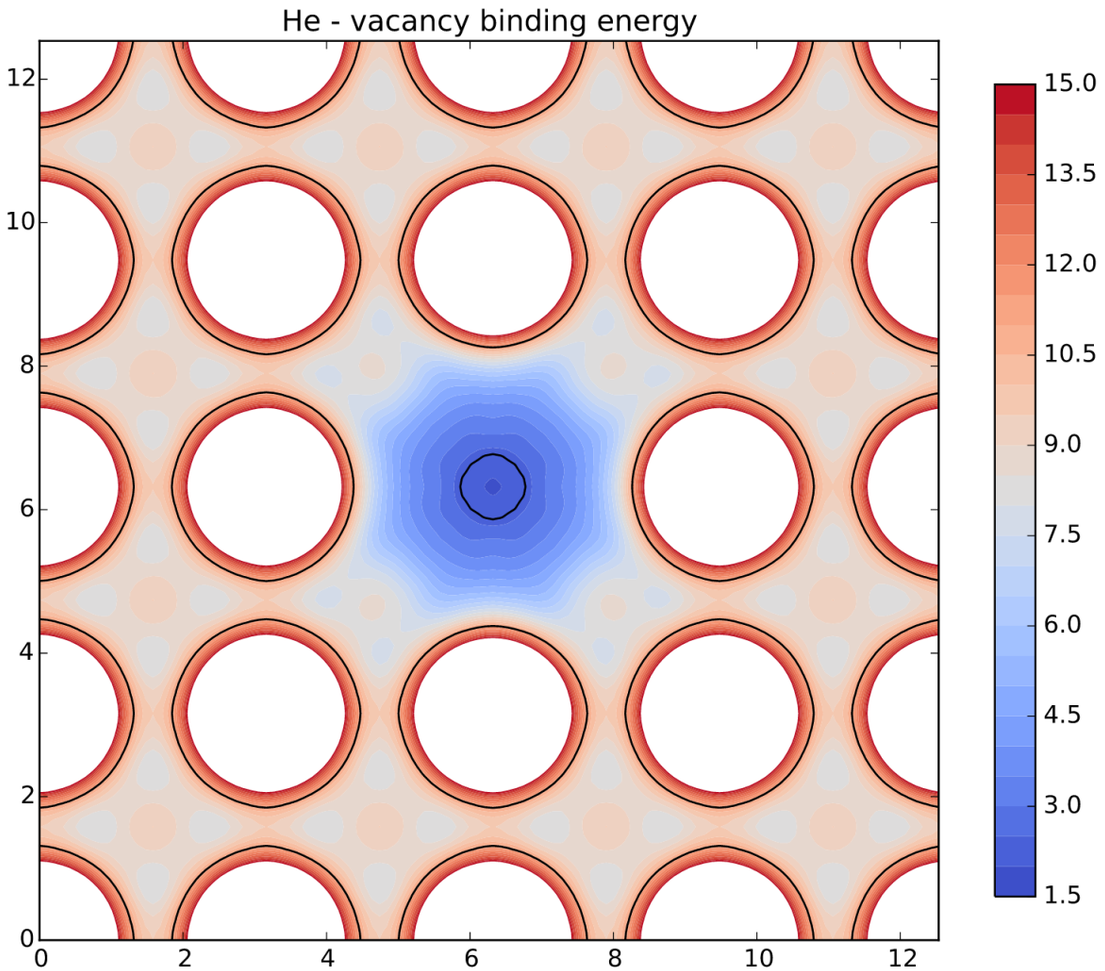
Vacancy-dislocation binding energy



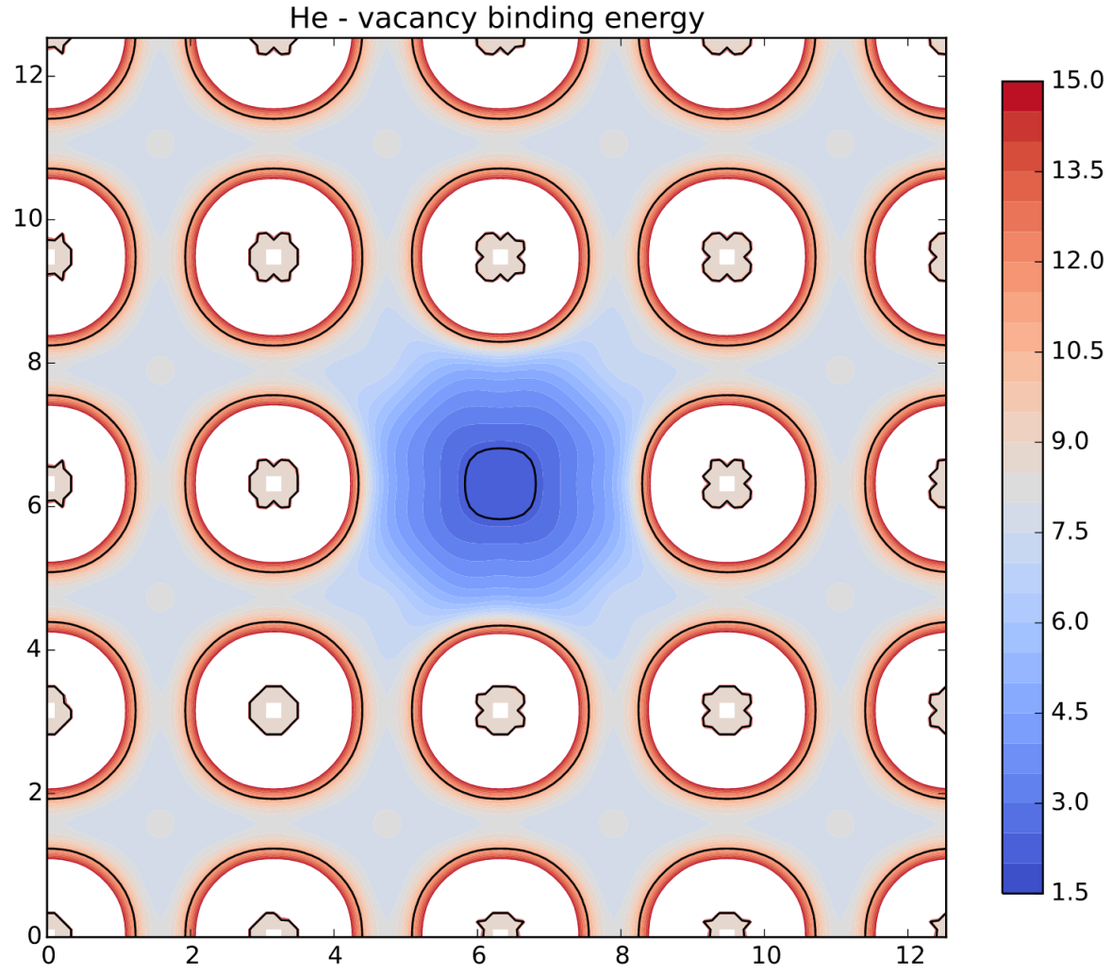
(~100,000 atoms in 3D simulation box)

He-Vacancy interaction in tungsten

2-body



SOAP-GAP

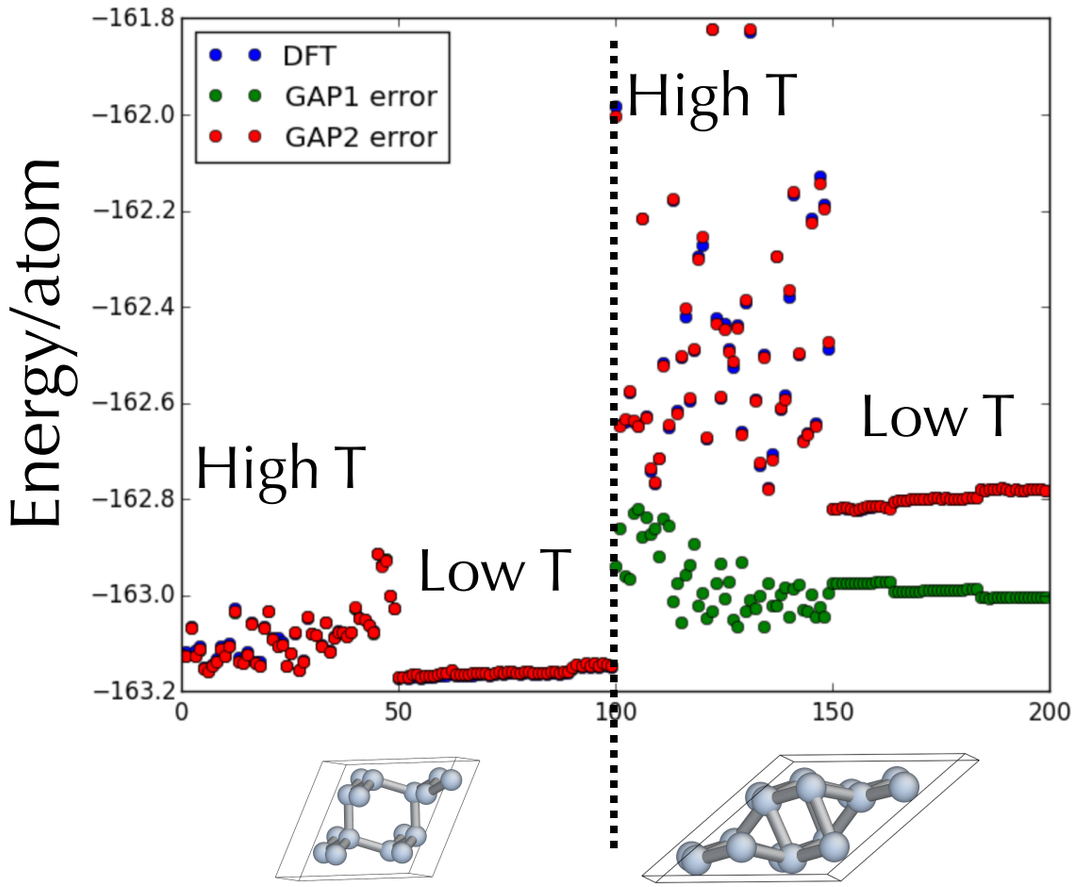


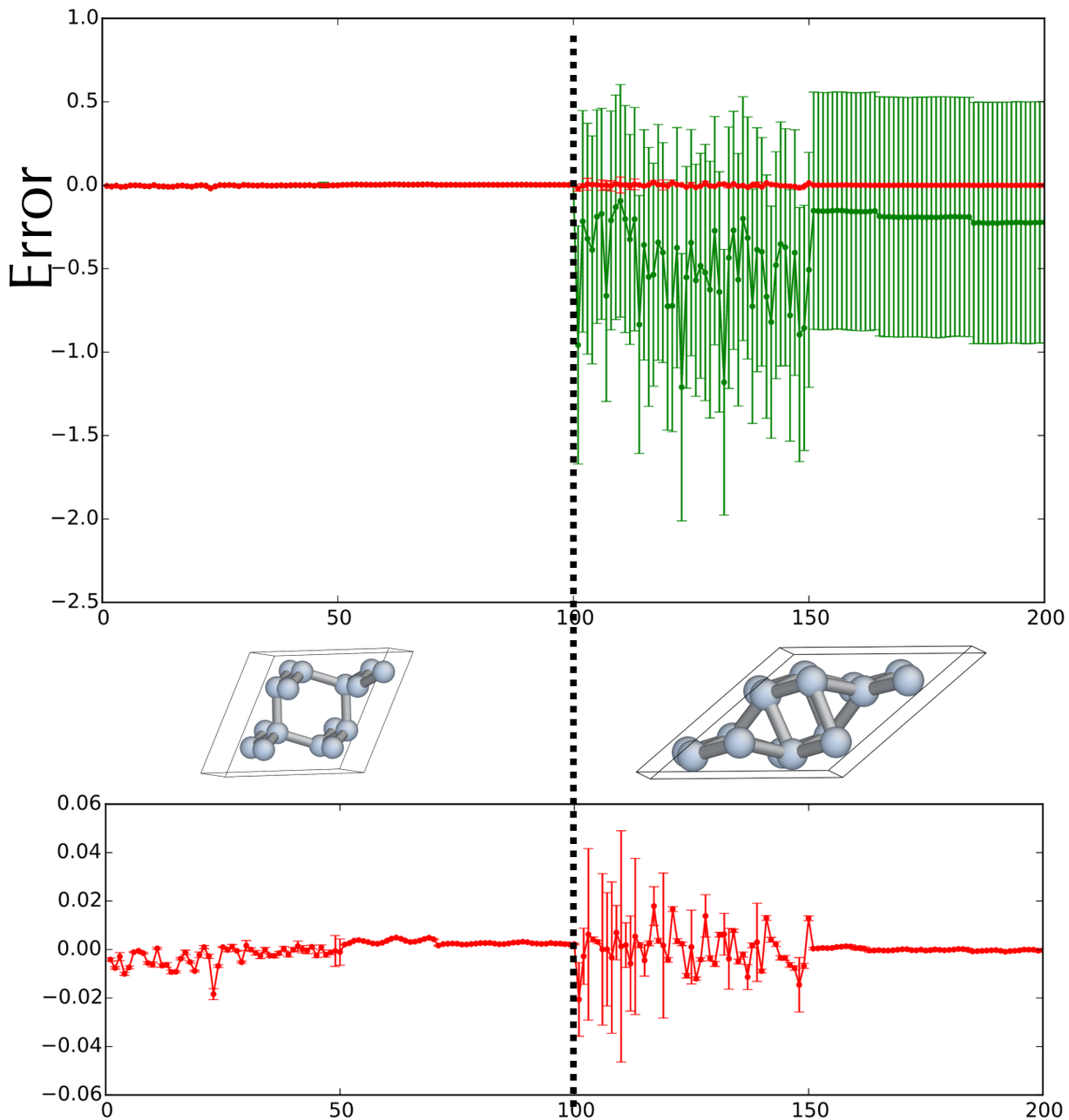
- Collaboration with CCFE (Duc Nguyen-Manh)
- Add W-He interaction on top of pure W potential
- Do the same with H@W (Takuji Oda)

Self-aware potentials: predicting the error

- Bulk silicon: Diamond and β-tin phases
- Predicted error correctly signals where GAP model is unreliable

$$\sigma_*^2 = C(\mathbf{x}_*, \mathbf{x}_*) - \mathbf{k}_*^T(\mathbf{C}_M^{-1} - (\mathbf{C}_M + \mathbf{C}_{MN}(\mathrm{Diag}(\mathbf{C}_N - \mathbf{C}_{NM}\mathbf{C}_M^{-1}\mathbf{C}_{MN}) + \sigma^2\mathbf{I})^{-1}\mathbf{C}_{NM})^{-1})\mathbf{k}_* + \sigma^2$$

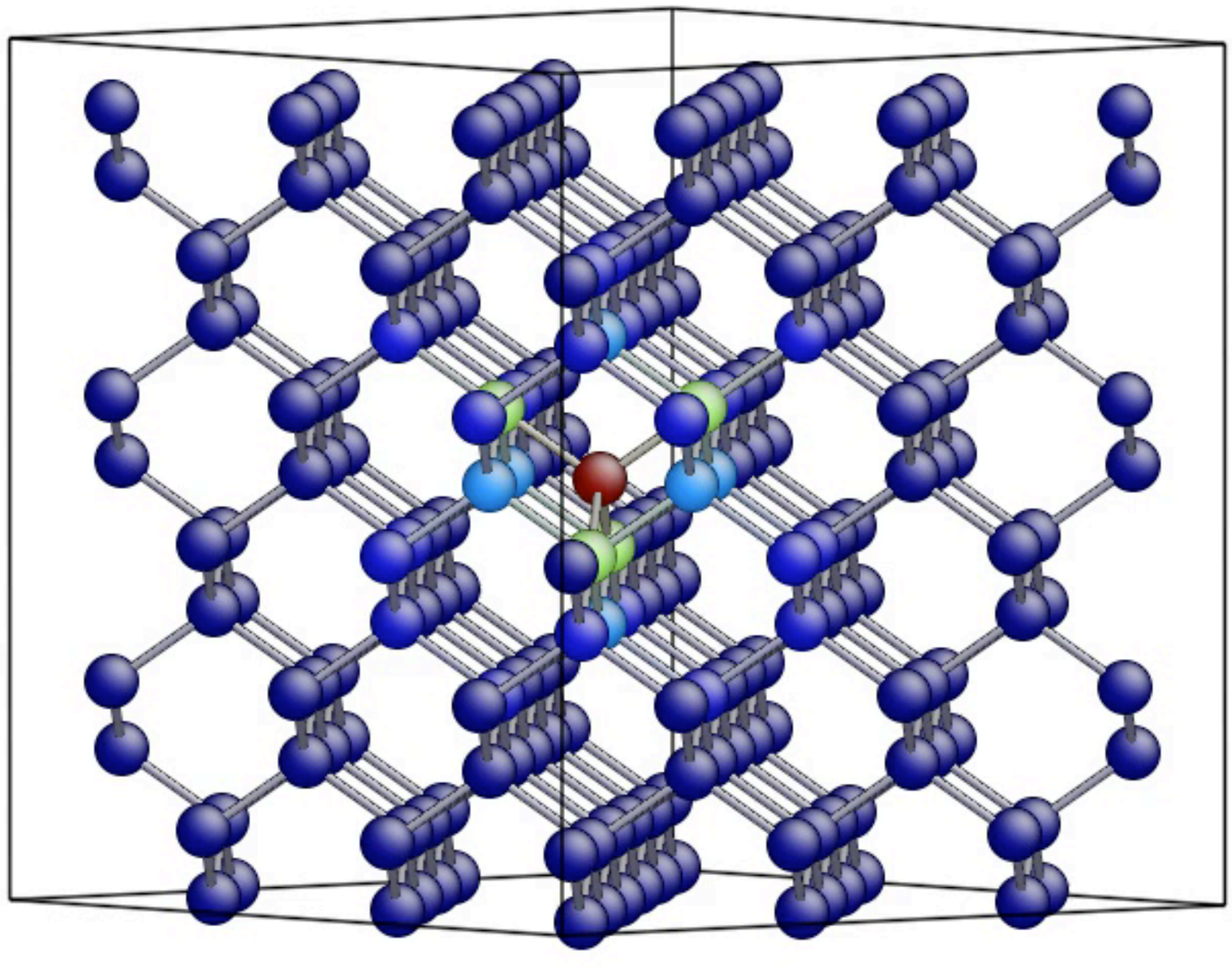




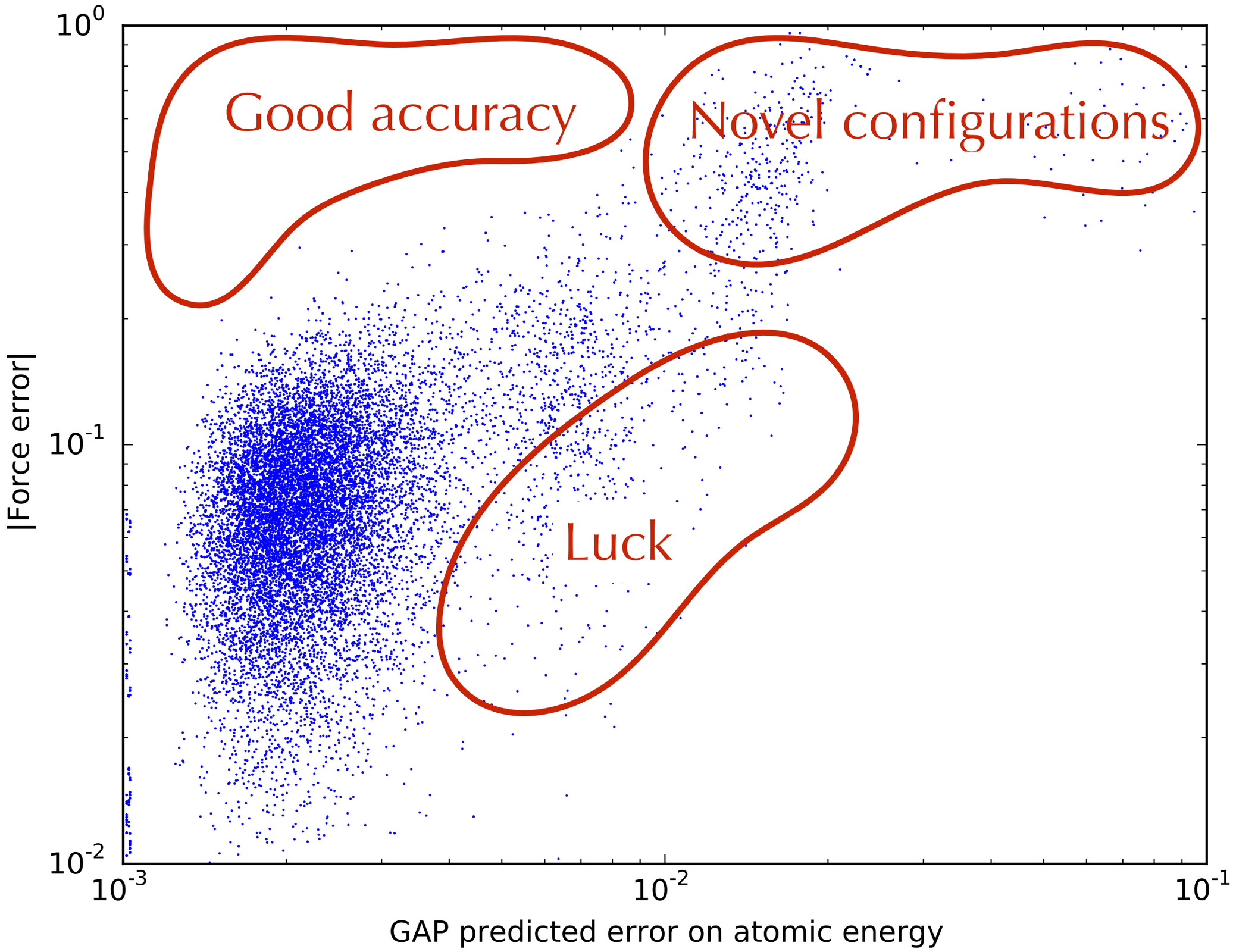
GAP is a "self-aware" potential

- Train a potential on bulk Si phases
- Introduce defect, colour by ***predicted*** error

$$\hat{\sigma}^2(x) = K(x,x) - K(x,\mathbf{x})^T(\mathbf{K}^{-1} + \sigma_\nu^2\mathbf{I})K(x,\mathbf{x})$$



Correlation with actual error

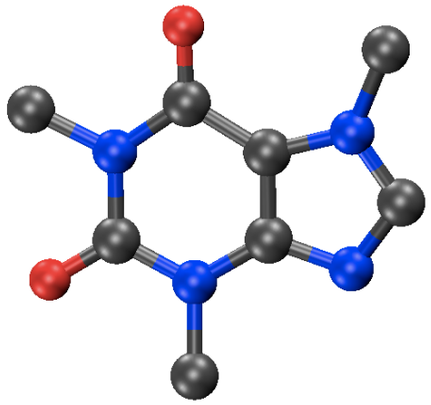


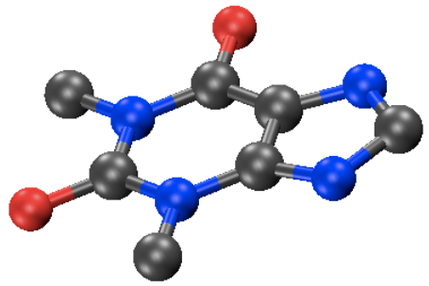
Summary of potential fitting

- Free up functional forms: Gaussian process (GP)
- Take care with the kernels: SOAP
- Build databases systematically (How big?)

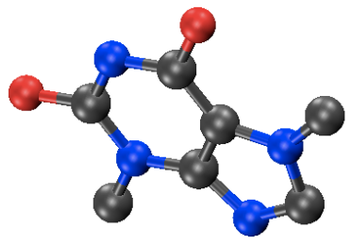
- Transferability? “Extrapolation” ?
- Long range terms ? Polarisability ?
- Error prediction
- Multiple species: random alloys ?

What else can we do with a good kernel? Let's compare molecules!





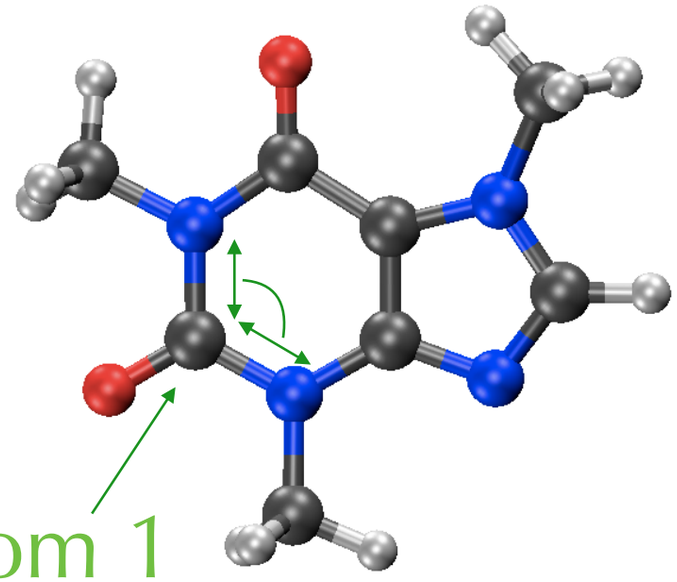
- Large databases
  - Molecular properties
  - Crystal structures
- Statistical analysis in drug design
  - Quantitative Structure Activity Relationship (QSAR)
  - Absorption Distribution Metabolism Extraction Toxicity (ADMET)
- Machine learning and data mining
  - Fitting functions to molecular properties
  - Create very accurate force fields

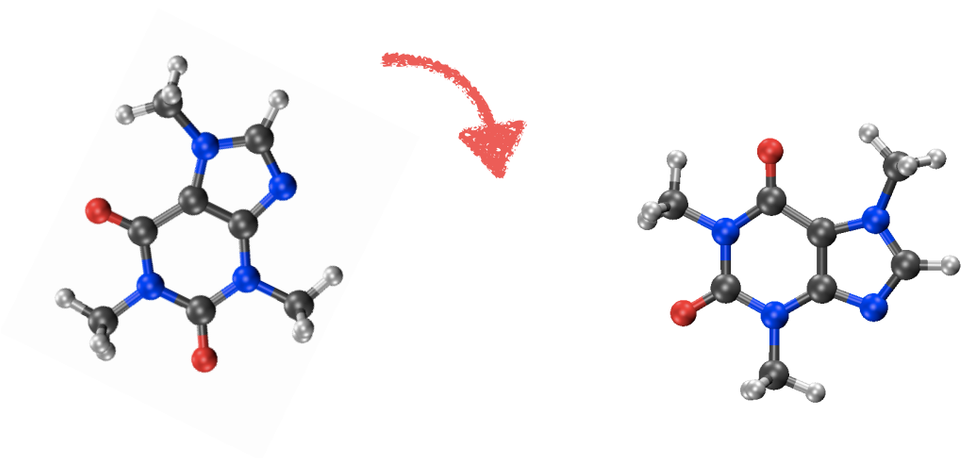


Molecular similarity

Traditional molecular geometry descriptors

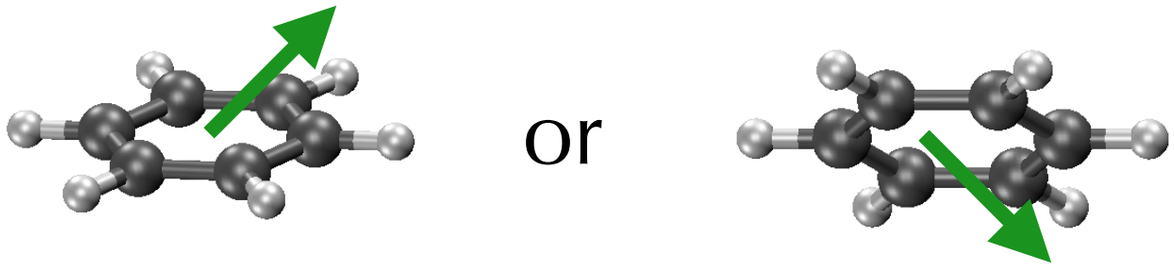
- A huge number of ad-hoc descriptors (aka fingerprints)
  - Size, chemical groups
  - Electronic and electrostatic properties
- Bond lengths and angles (“internal coordinates”)
- “Standard orientation” alignment”



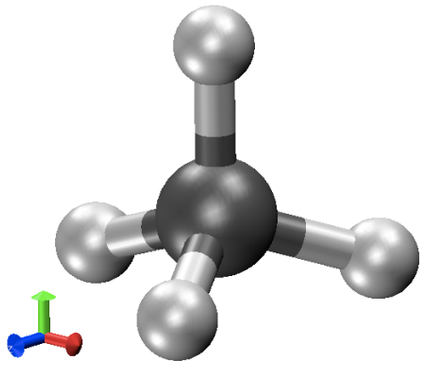


Alignment leads to discontinuity

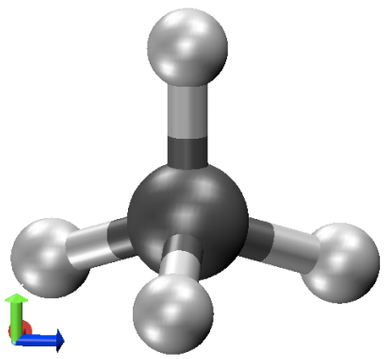
- Symmetric molecules



or

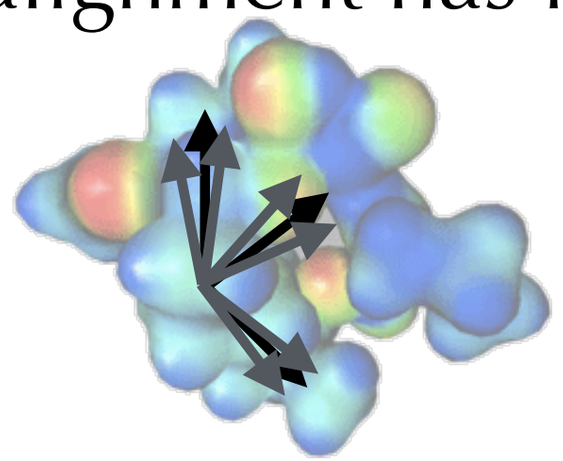


or



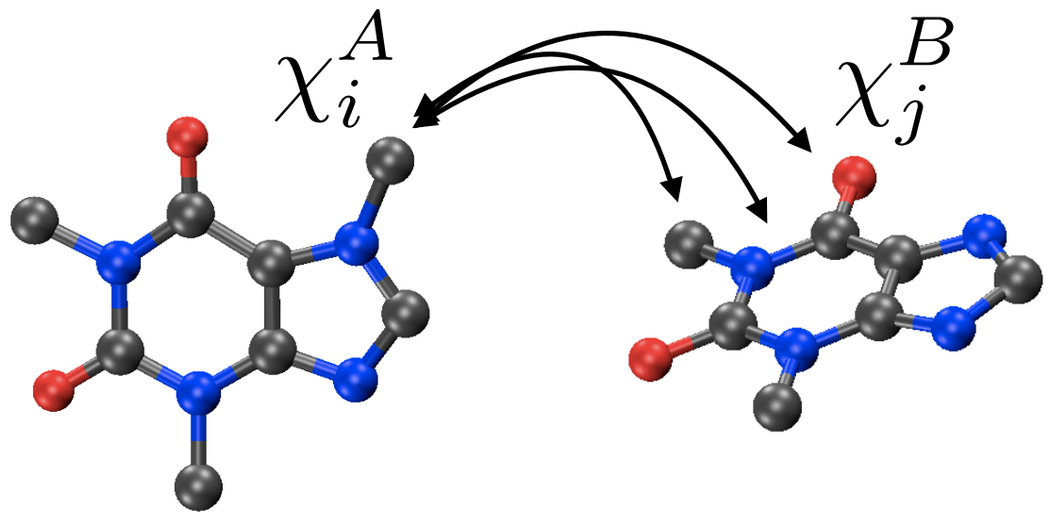
(ordering by bond length is also discontinuous)

- Field-based alignment has multiple local minima



We can use SOAP to build a kernel between entire structures, e.g. molecules

- Compute kernel between all pairs of atom environments in the two structures
- Decide what to do about matching atoms between structures



$$k(\mathcal{X}, \mathcal{X}') = \hat{\mathbf{p}}(\mathcal{X}) \cdot \hat{\mathbf{p}}(\mathcal{X}')$$

$$C_{ij}(A, B) = k\left(\mathcal{X}_i^A, \mathcal{X}_j^B\right)$$

- Average kernel:

$$\bar{K}(A, B) = \frac{1}{N^2} \sum_{ij} C_{ij}(A, B)$$

- Best Match kernel

$$\hat{K}(A, B) = \frac{1}{N} \max_{\pi} \sum_{i} C_{i\pi_i}(A, B)$$

- Regularized Match kernel

$$\hat{K}^{\gamma}(A, B) = \operatorname{Tr} \mathbf{P}^{\gamma} \mathbf{C}(A, B),$$

$$\mathbf{P}^{\gamma} = \underset{\mathbf{P} \in \mathcal{U}(N,N)}{\operatorname{argmin}} \sum_{ij} P_{ij} \left(1 - C_{ij} - \gamma \ln P_{ij}\right) \qquad \sum_j P_{ij} = 1/N$$

Multiple species and distances

- Extension of SOAP to multiple species

$$\rho^{\alpha}_{\mathcal{X}}(\mathbf{r}) = \sum_{i \in \mathcal{X}\alpha} \exp\left(-\frac{(\mathbf{x}_i - \mathbf{r})^2}{2\sigma^2}\right)$$

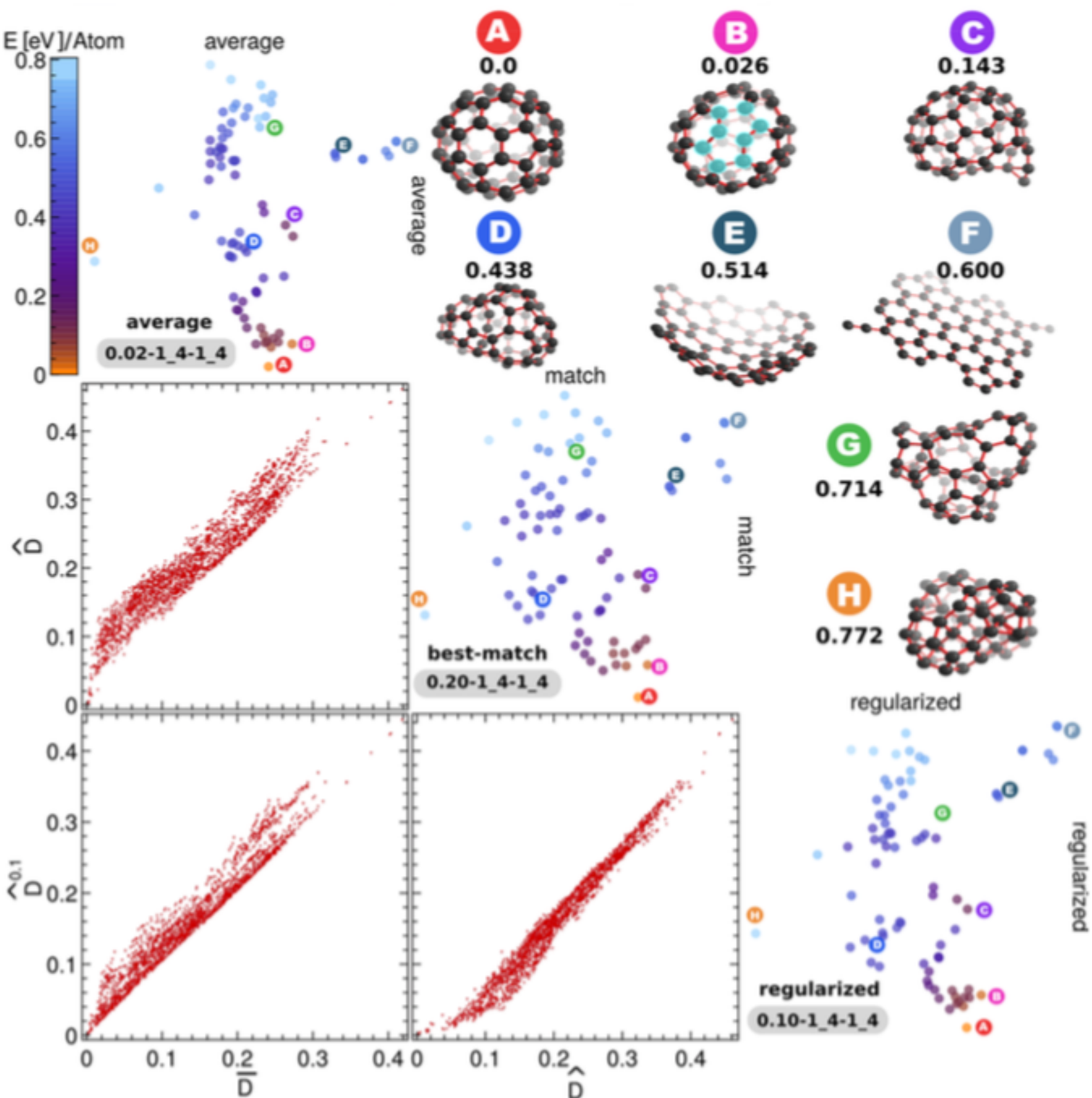
$$p(\mathcal{X})^{\alpha\beta}_{b_1 b_2 l} = \sum_m c^{\alpha\dagger}_{b_1 l m} c^{\beta}_{b_2 l m}$$

$$\tilde{k}(\mathcal{X}, \mathcal{X}') = \int \mathrm{d}\hat{R} \left| \int \sum_{\alpha} \rho^{\alpha}_{\mathcal{X}}(\mathbf{r}) \rho^{\alpha}_{\mathcal{X}'}(\hat{R}\mathbf{r}) \mathrm{d}\mathbf{r} \right|^2 = \sum_{\alpha\beta} \mathbf{p}_{\alpha\beta}(\mathcal{X}) \cdot \mathbf{p}_{\alpha\beta}(\mathcal{X}')$$

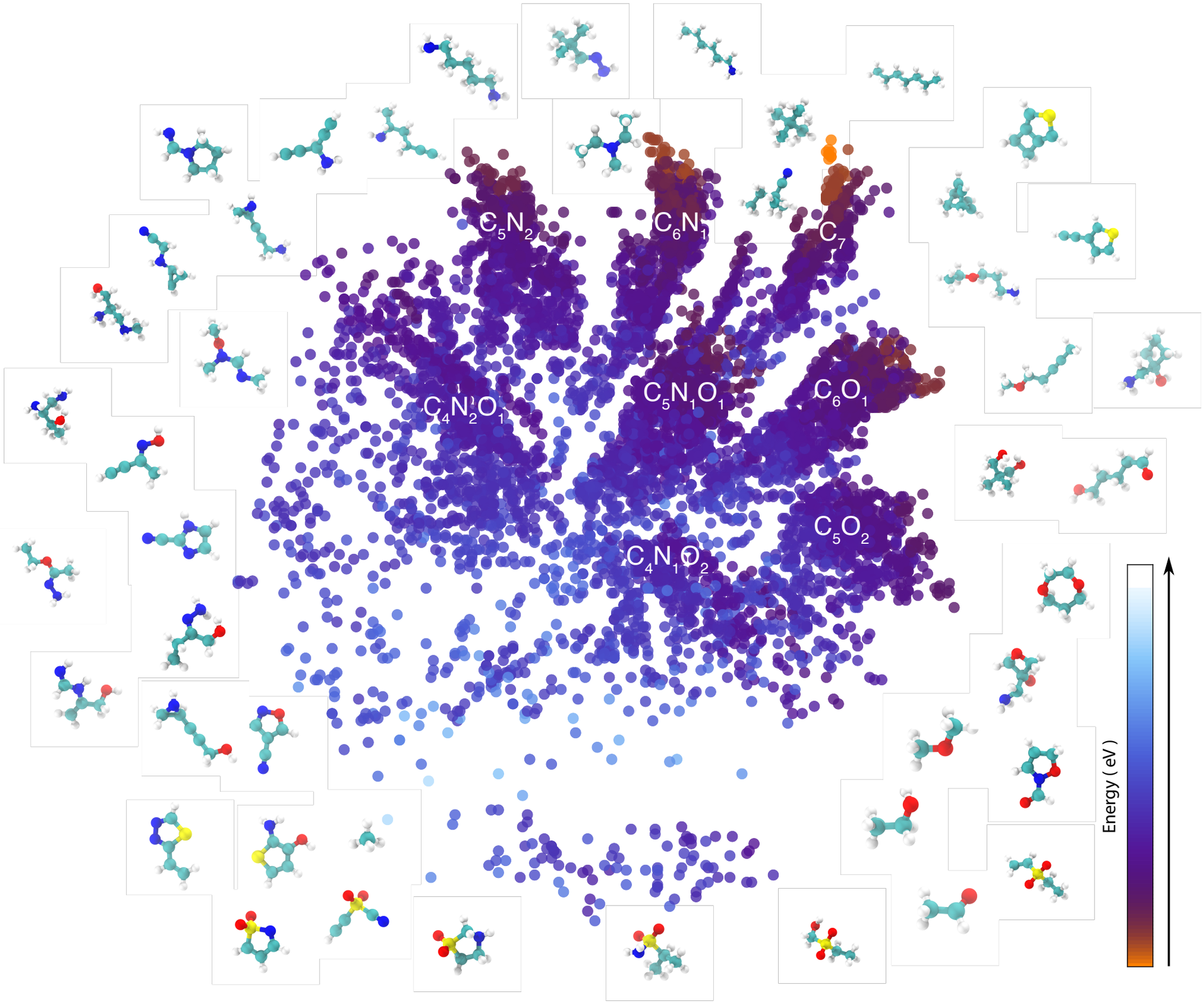
- Define a “distance” between structures using the kernel:

$$D(A, B)^2 = 2 - 2K(A, B)$$

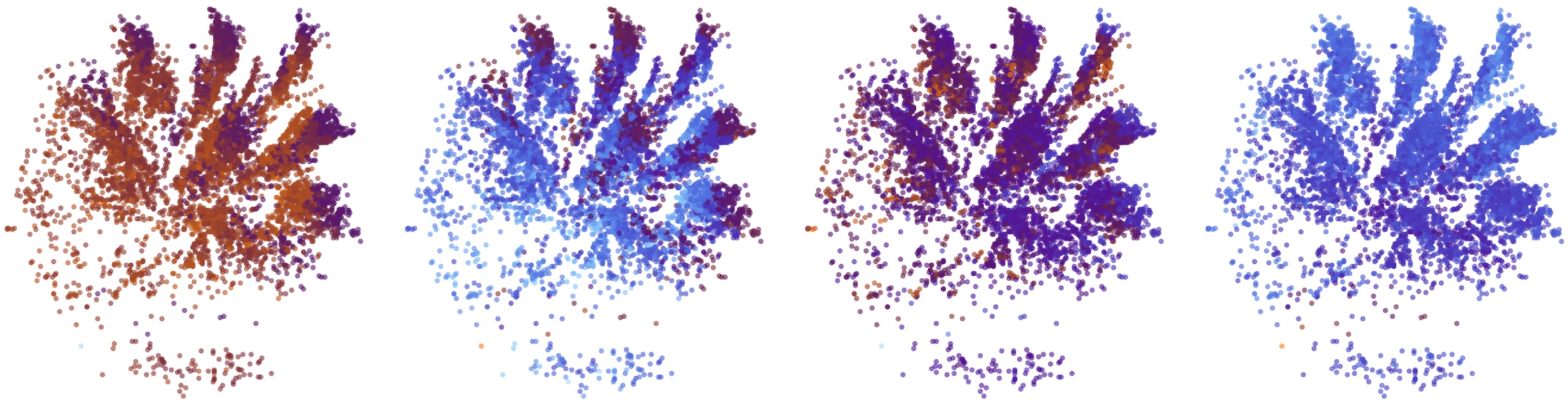
Embed structures in 2D to create maps



Map of all molecules with up to 7 heavy atoms



Properties



Learning properties with Gaussian processes

Property	SD	MAE	RMSE	MAE[24]	RMSE[24]
E (PBE0)	9.70	0.04	0.07	0.16	0.36
α (PBE0)	1.34	0.05	0.07	0.11	0.18
α (SCS)	1.47	0.02	0.04	0.08	0.12
HOMO (GW)	0.70	0.12	0.17	0.16	0.22
HOMO (PBE0)	0.63	0.11	0.15	0.15	0.21
HOMO (ZINDO)	0.96	0.13	0.18	0.15	0.22
LUMO (GW)	0.48	0.12	0.17	0.13	0.21
LUMO (PBE0)	0.68	0.08	0.12	0.12	0.20
LUMO (ZINDO)	1.31	0.10	0.15	0.11	0.18
IP (ZINDO)	0.96	0.19	0.28	0.17	0.26
EA (ZINDO)	1.41	0.13	0.18	0.11	0.18
E^*_{1st} (ZINDO)	1.87	0.18	0.41	0.13	0.31
E^*_{max} (ZINDO)	2.82	1.56	2.16	1.06	1.76
I_{max} (ZINDO)	0.22	0.08	0.12	0.07	0.12

[24] G. Montavon, M. Rupp, V. Gobre, A. Vazquez-Mayagoitia, K. Hansen, A. Tkatchenko, K.-R. Mller, and O. A. von Lilienfeld, "Machine learning of molecular electronic properties in chemical compound space," New Journal of Physics **15**, 095003 (2013).

Can do the same to periodic structures