Machine learning the Born-Oppenheimer potential energy surface: from molecules to materials

Gábor Csányi
Engineering Laboratory

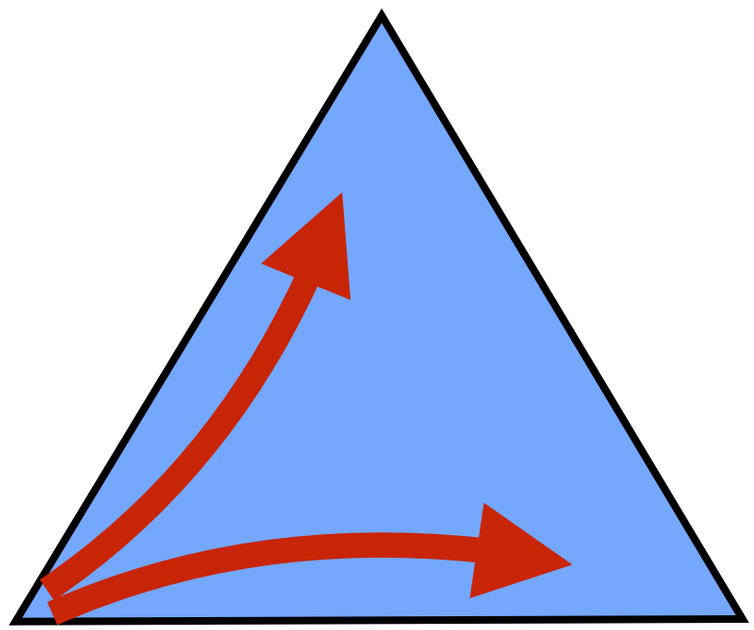
UNIVERSITY OF
CAMBRIDGE

Interatomic potentials for molecular dynamics

Transferability

biomolecular force fields (biochemistry)

AMBER
CHARMM
AMOEBA
GROMACS
OPLS
...



Accuracy

small molecules in gas phase (quantum chemistry)

Bowman
Szalewicz
Paesani
...

Reactive

solid state materials (physics & materials)

Tersoff
Brenner
EAM
BOP
...

Ingredients

- Representation of atomic neighbourhood → smoothness, faithfulness, continuity
- Interpolation of functions → flexible but smooth functional form, few sensible parameters
- Database of configurations → predictive power non-domain specific

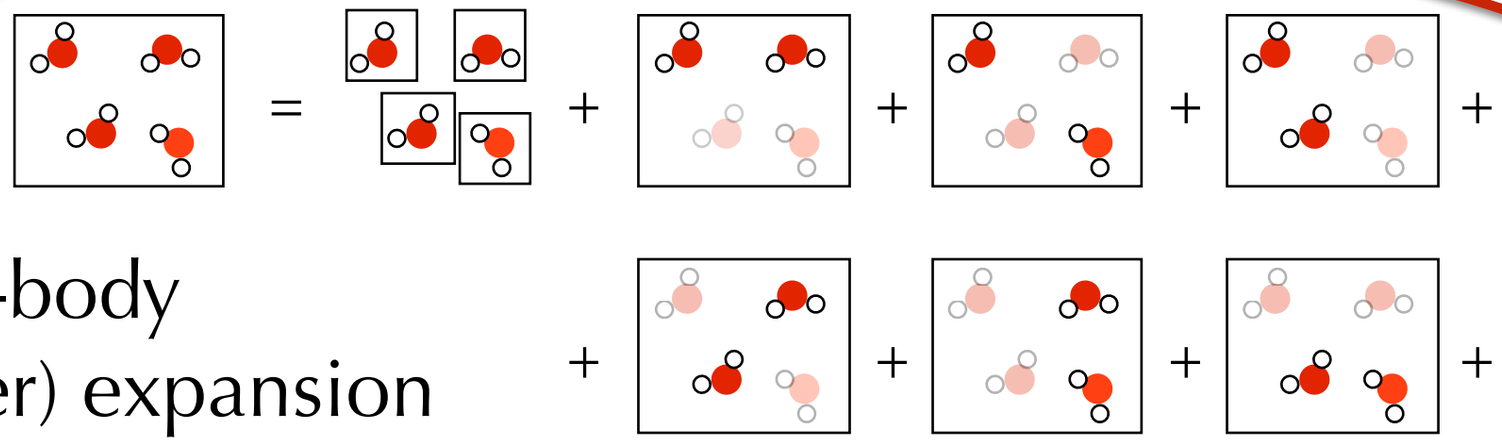
Interpolating the solution of an eigenproblem

$$-i\hbar\frac{\partial\Psi}{\partial t} = \mathcal{H}\Psi \qquad \mathcal{H}_{\rm el} = \frac{1}{2}\sum_i \nabla^2_{r_i} + \sum_{ij}\frac{Z_j(-1)}{|R_j - r_i|} + \sum_{ij}\frac{1}{|r_i - r_j|}$$

Lowest eigenstate: $E_0 \equiv V_{\rm el}(R_1, R_2, \ldots)$ has spatial locality

Weak

Many-body
(cluster) expansion



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

Strong

$$E = \sum_i \varepsilon(q_1^{(i)}, q_2,^{(i)} \ldots, q_M^{(i)})$$ + analytic long range terms

q_i are descriptors of local atomic neighbourhood

Parametrisation of the H_2O dimer

- Standard representation:

  6 atoms ⇒ 15 interatomic distances : $x = \{r_{ij}\}$

- Symmetrize kernel function:

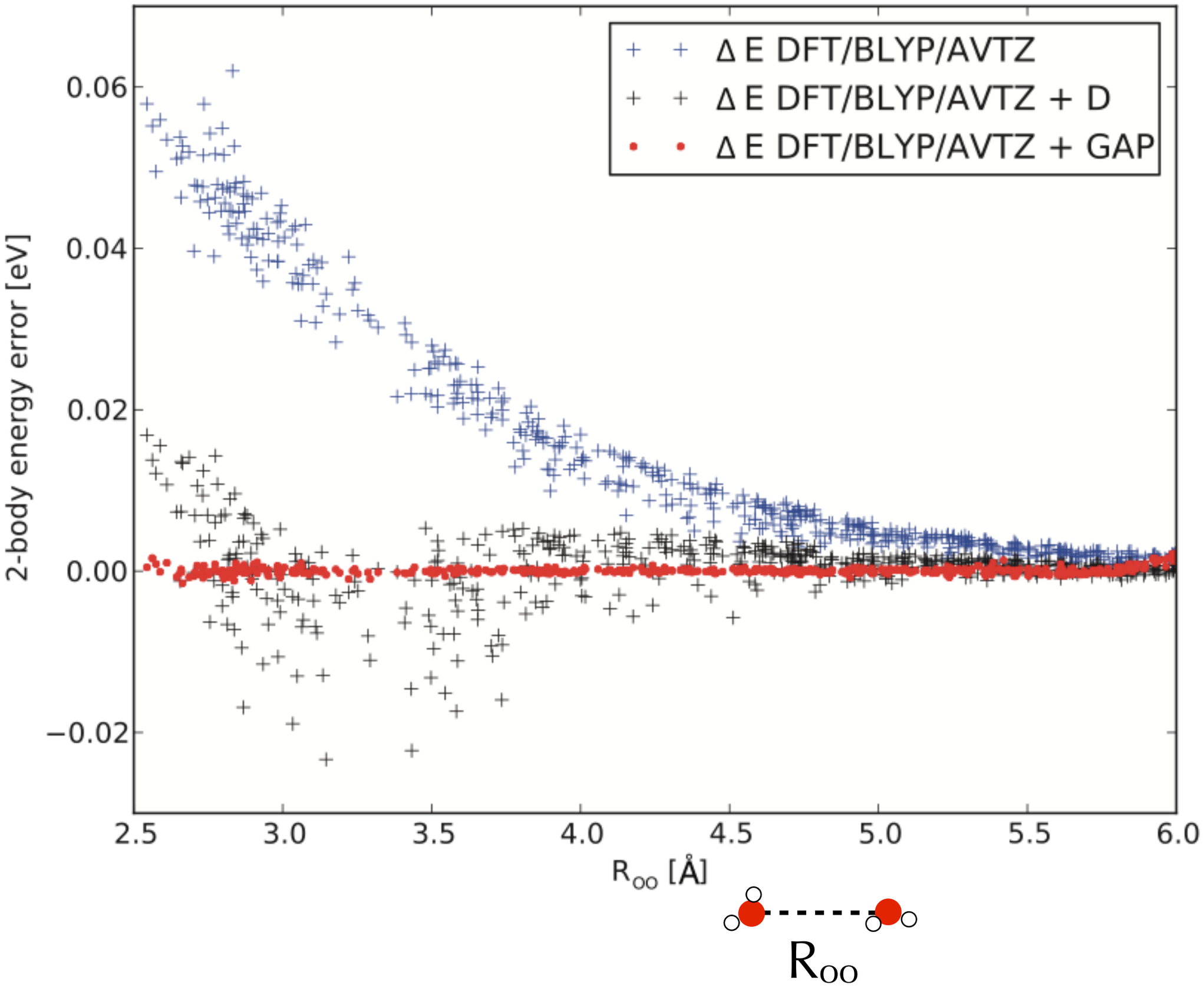
$$\tilde{K}(x, x') = \sum_{p \in S} K(p(x), x')$$

S: symmetry group of molecules

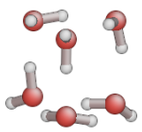
- Use analytic forces if available:

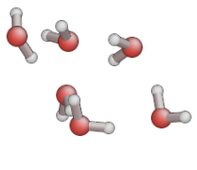
$$K'(x, x') = \frac{\partial}{\partial x} K(x, x')$$

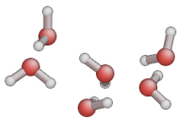
H2O dimer corrections

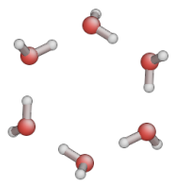


Water hexamers - relative energies - no ZPC









CCSD(T)	BLYP	DMC	Bowman	BLYP+GAP
0.00	0.00	0.00	0.00	0.00
0.24	-0.59	0.33	0.46	0.15
0.70	-2.18	0.56	2.39	0.52
1.69	-2.44	1.53	4.78	1.77

[kcal / mol]

(B3LYP optimised geometries from Tschumper)

GAP error < 0.03 kcal/mol / H_2O

Ice polymorphs

Relative energies

[meV / H_2O]	Expt	DMC±5	PBE0+vdW^{TS}	BLYP+GAP±3
Ih	0	0	0	0
II	1	-4	6	-5
IX	4		3	-3
VIII	33	30	76	30

error < 0.15 kcal/mol / H_2O

Ice polymorphs

Absolute binding energies

[meV / H_2O]	Expt	DMC ±5	PBE0+vdW^{TS}	BLYP+GAP ±3
Ih	610	605	672	667
II	609	609	666	672
IX				
VIII	577	575	596	637

BLYP+GAP overbinds *uniformly* by ~ 60 meV

⇒ 3-body terms!

Interpolating the solution of an eigenproblem

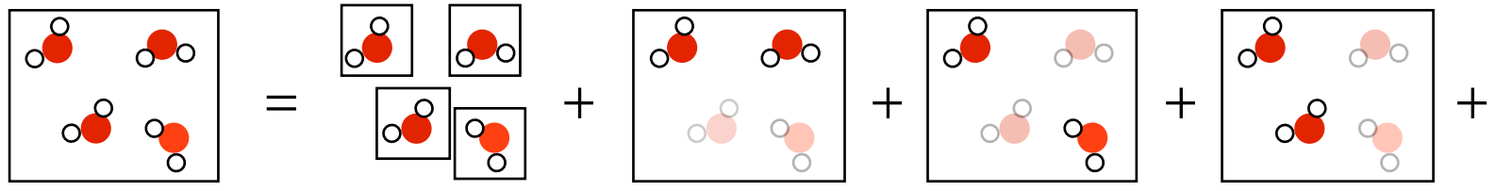
$$-i\hbar\frac{\partial\Psi}{\partial t} = \mathcal{H}\Psi \qquad \mathcal{H}_{\mathrm{el}} = \frac{1}{2}\sum_i \nabla_{r_i}^2 + \sum_{ij}\frac{Z_j(-1)}{|R_j - r_i|} + \sum_{ij}\frac{1}{|r_i - r_j|}$$

Lowest eigenstate: $E_0 \equiv V_{\mathrm{el}}(R_1, R_2, \ldots)$ has spatial locality

---

Weak

Many-body
(cluster) expansion



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

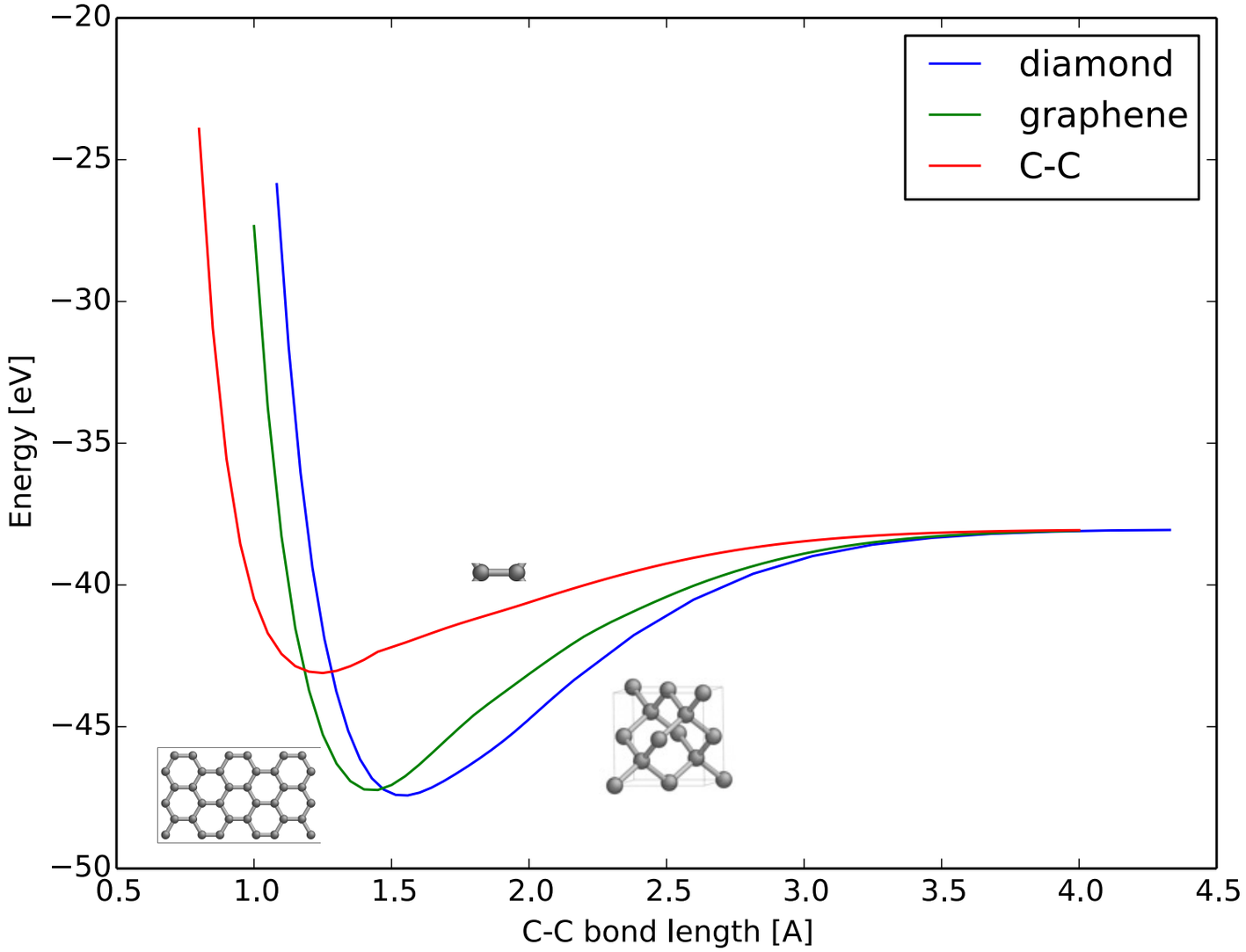
Strong

$$E = \sum_i \varepsilon(q_1^{(i)}, q_2,^{(i)} \ldots, q_M^{(i)})$$ + analytic long range terms

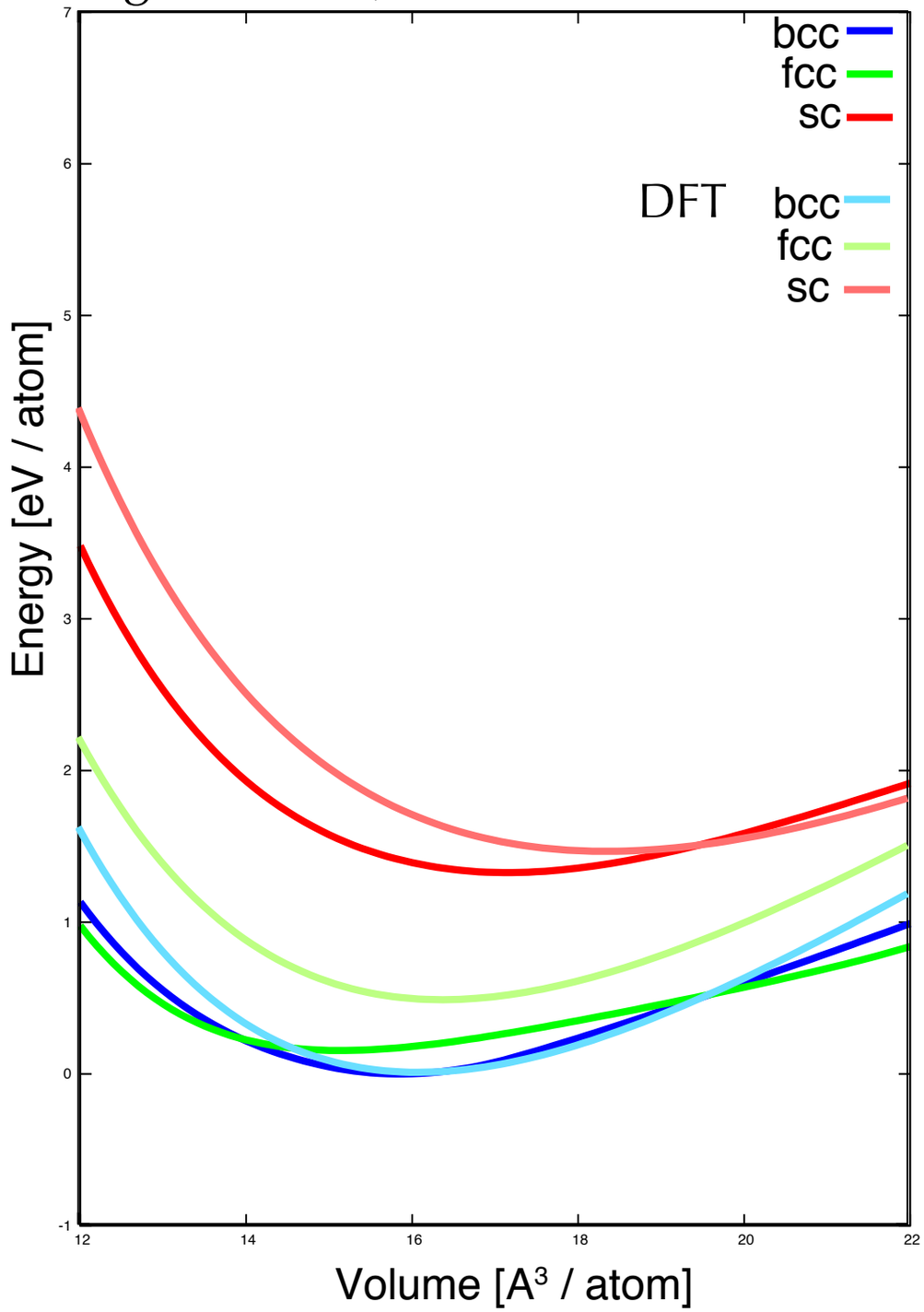
q_i are descriptors of local atomic neighbourhood

Quantum mechanics is many-body

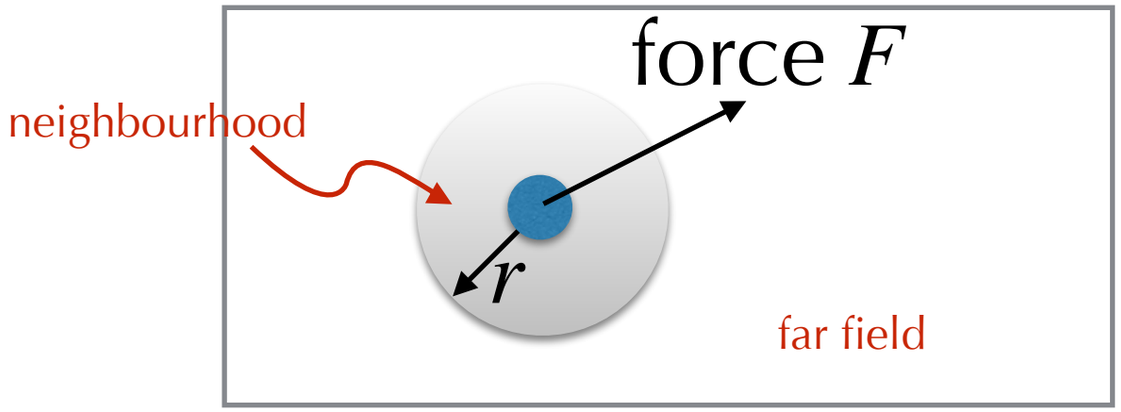
Carbon: tight binding



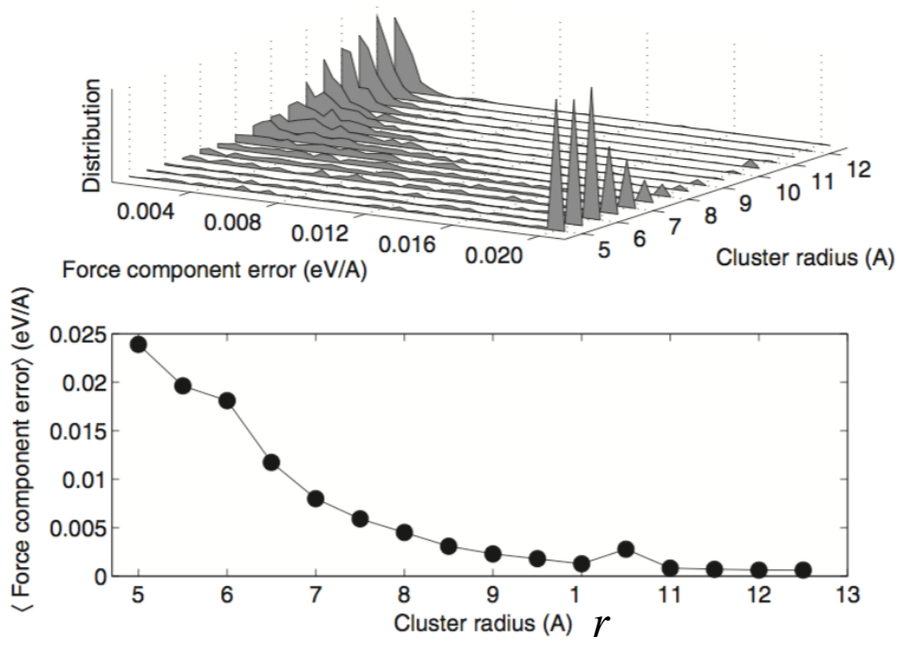
Tungsten: DFT, Embedded Atom Model



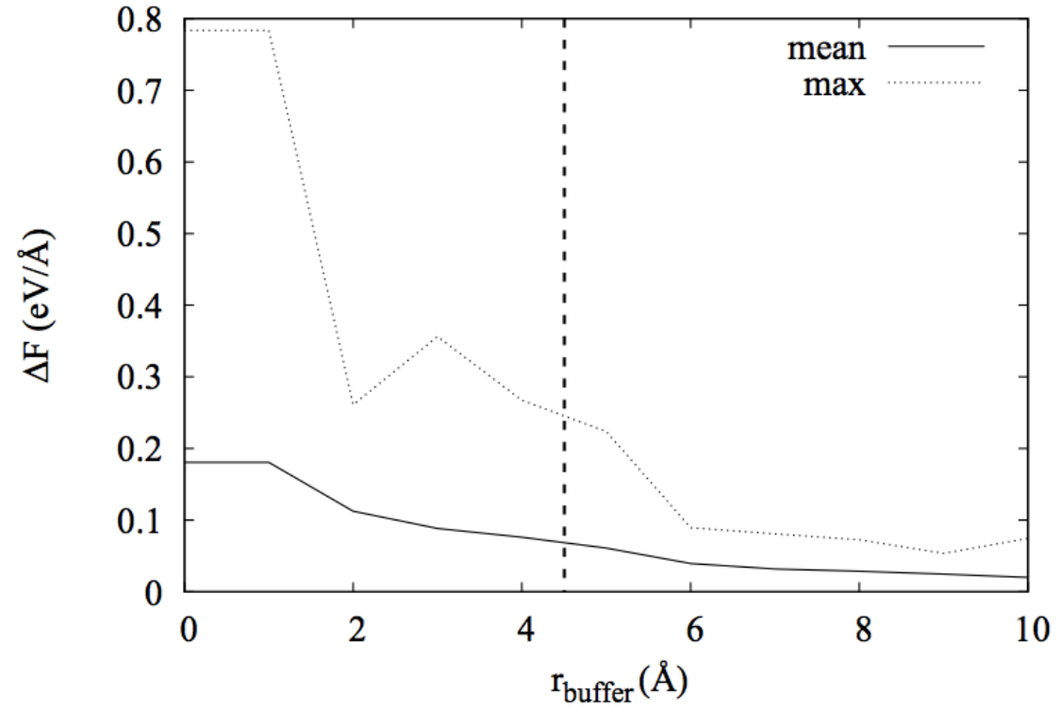
Quantum mechanics has some locality



Force errors around
Si self interstitial



Force errors around
O in water with QM/MM



Is there a finite range atomic energy function that gives the correct total energy and forces?

$$V_{\rm el}(R_1,R_2,\ldots) \approx \sum_i^{\rm atoms} \varepsilon(R_1-R_i,R_2-R_i,\ldots) + \frac{1}{2}\sum_{ij} \hat{L}_i\hat{L}_j \frac{1}{R_{ij}} + \frac{c_{ij}}{|R_{ij}|^6} \qquad \hat{L}_i = q_i + \mathbf{p}_i\cdot\nabla_i + \ldots$$

Finite range *atomic energy* function

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

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

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

Representation is implicit

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

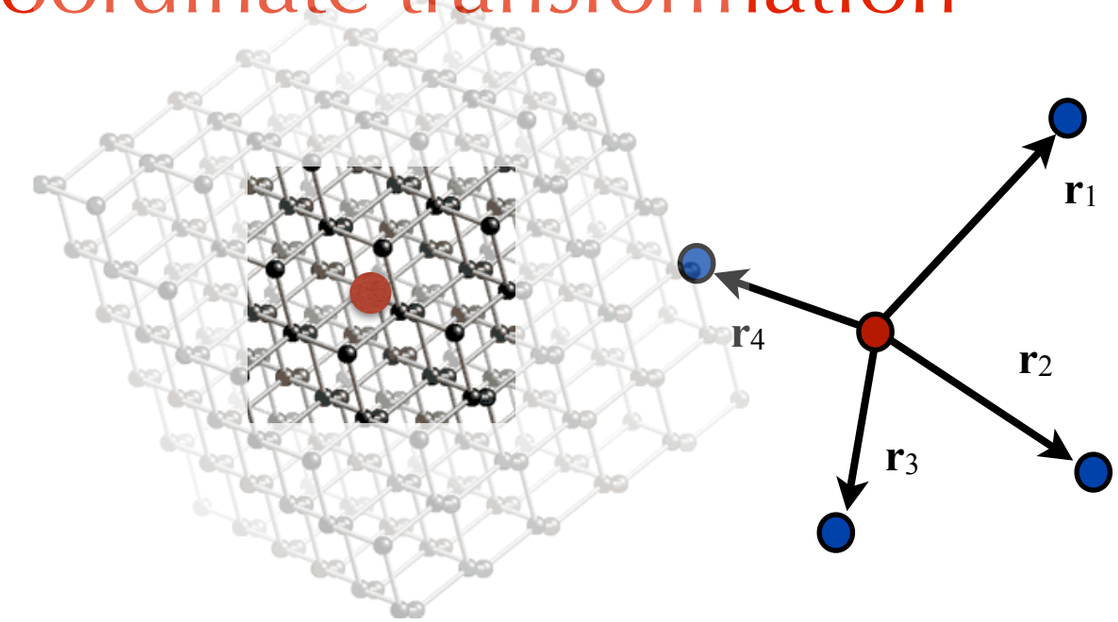
given by

GOAL: potentials ~~based on~~ quantum mechanics

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”)

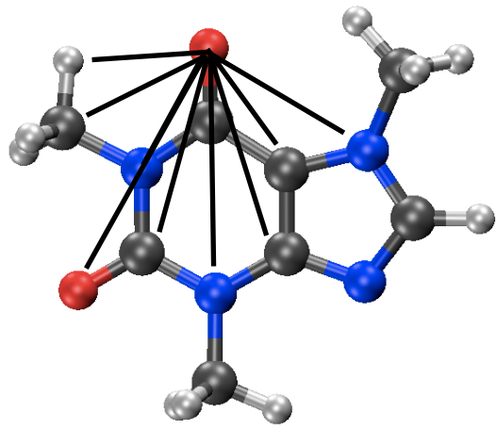
- Rotational invariance by itself is easy: $\mathbf{q} \equiv R_{ij} = \mathbf{r}_i \cdot \mathbf{r}_j$ (Weyl)

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

Matrix of interatomic distances

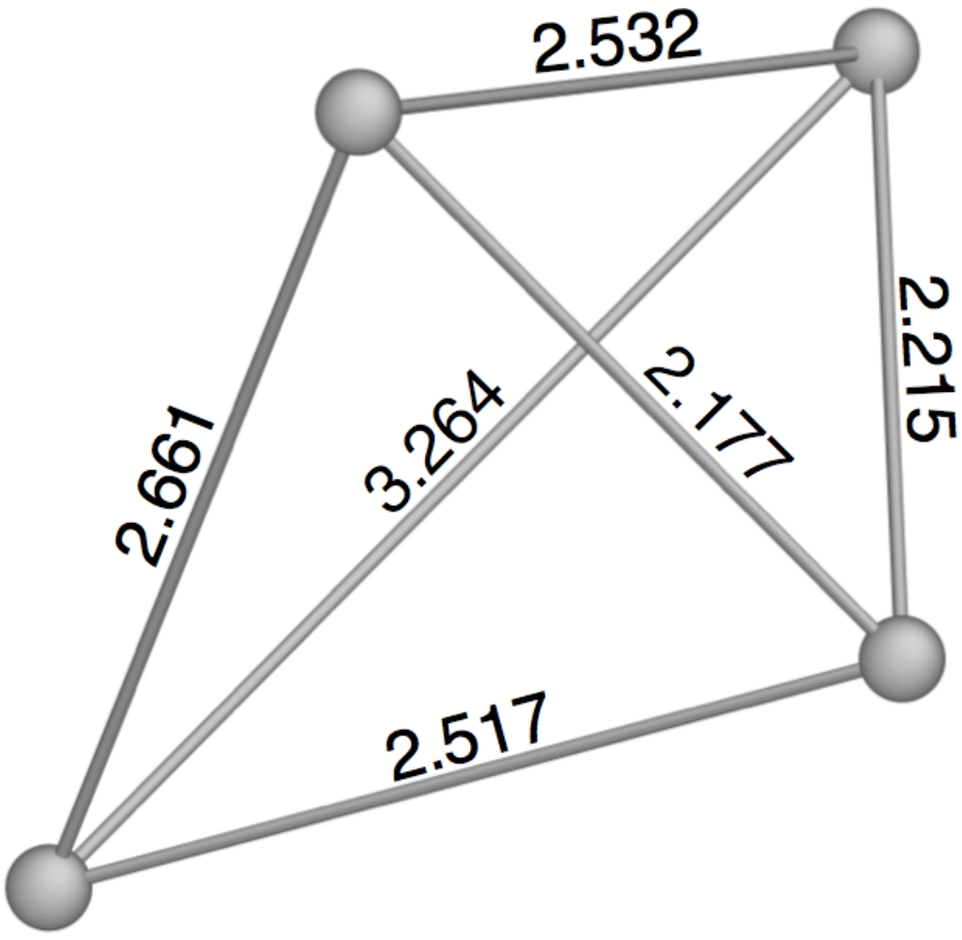
- Equivalent to Weyl matrix

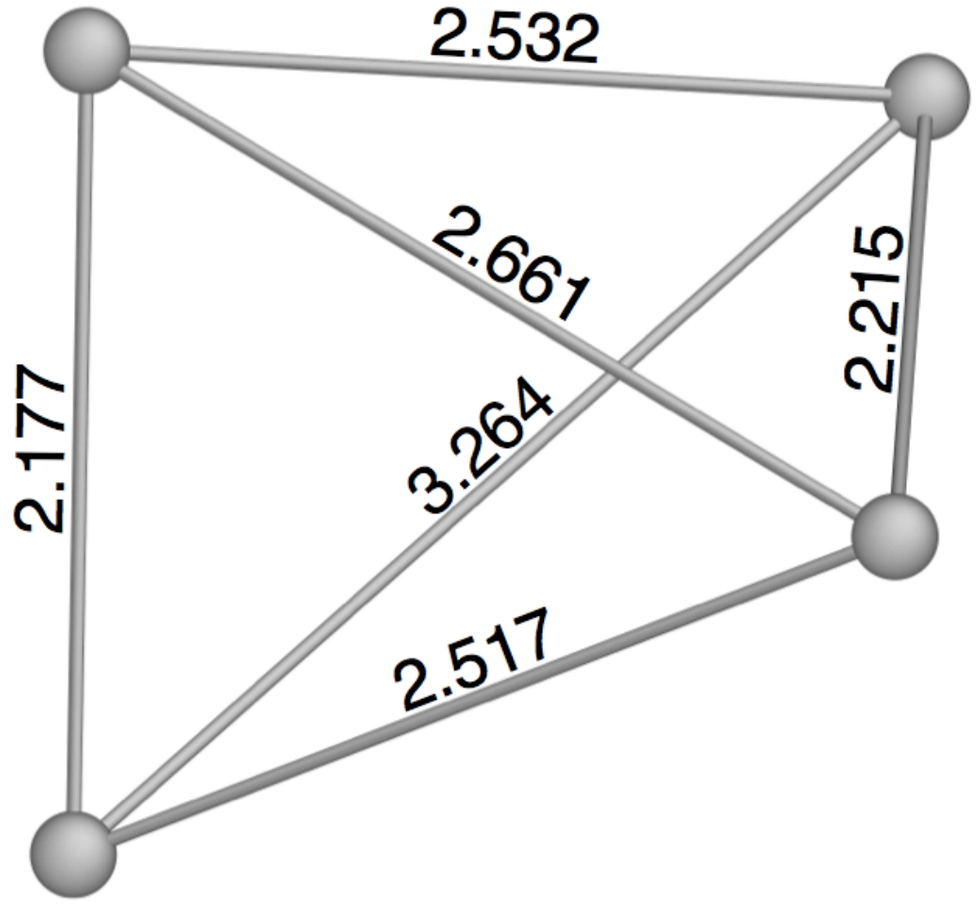
$$\begin{bmatrix} \mathbf{r}_1 \cdot \mathbf{r}_1 & \mathbf{r}_1 \cdot \mathbf{r}_2 & \cdots & \mathbf{r}_1 \cdot \mathbf{r}_N \\ \mathbf{r}_2 \cdot \mathbf{r}_1 & \mathbf{r}_2 \cdot \mathbf{r}_2 & \cdots & \mathbf{r}_2 \cdot \mathbf{r}_N \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{r}_N \cdot \mathbf{r}_1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



- Rotationally invariant: no alignment needed
- Permutation problem: Weyl matrix is not permutation invariant

Drop atom ordering ?





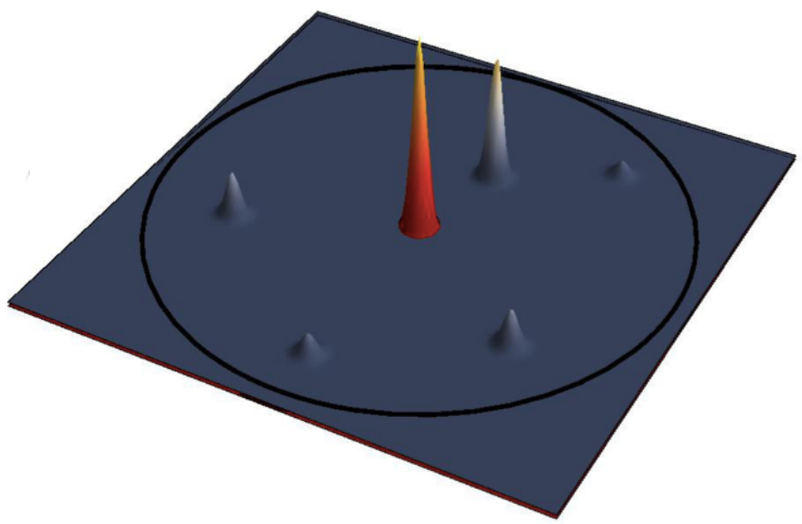
Not unique!

Atomic neighbour density function

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

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

$$\rho(r, \theta, \phi) = \sum_{n=1}^{\infty} \sum_{l=0}^{\infty} \sum_{m=-l}^{l} c_{l,n}^{m} g_n(r) Y_l^m(\theta, \phi)$$



- 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

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

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

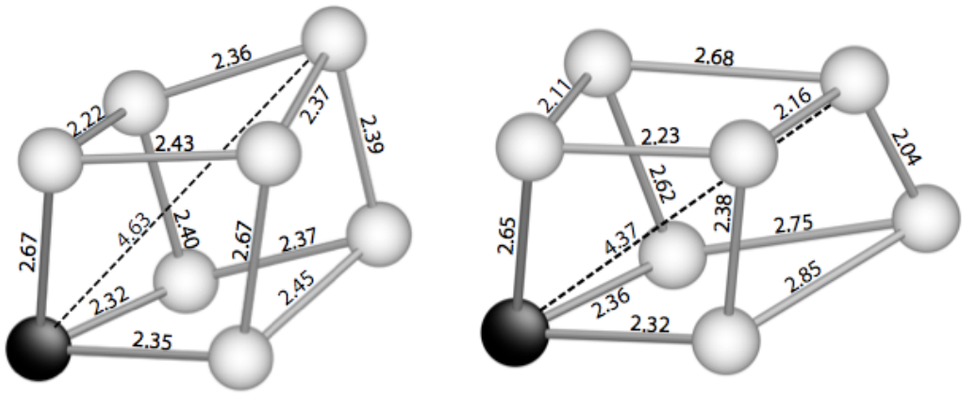
$$p_2 = \frac{5}{4\pi}\left(\frac{3}{2}\sum_{jk}\cos^2\theta_{ijk} + 6\right)$$

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

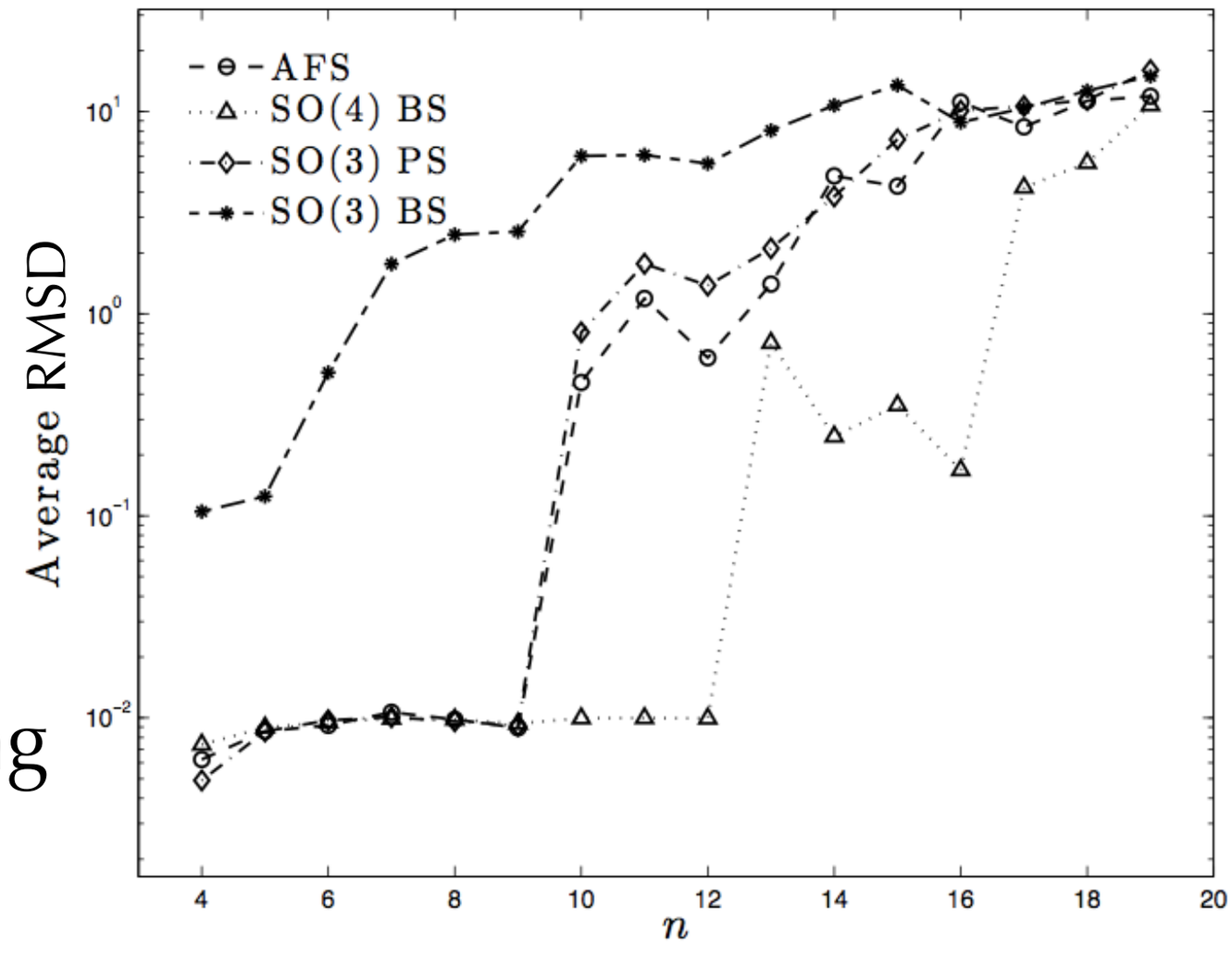
$$p_4 = \frac{9}{16\pi}\left(\frac{35}{2}\sum_{jk}\cos^4\theta_{ijk} - 15\sum_{jk}\cos^2\theta_{ijk} + 13\right)$$

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

Uniqueness: environment reconstruction tests



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



Error in distinguishing

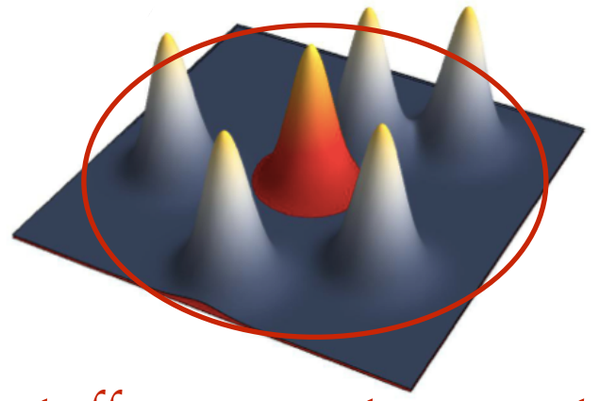
Number of neighbours

Construct smooth similarity kernel directly

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

- 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:

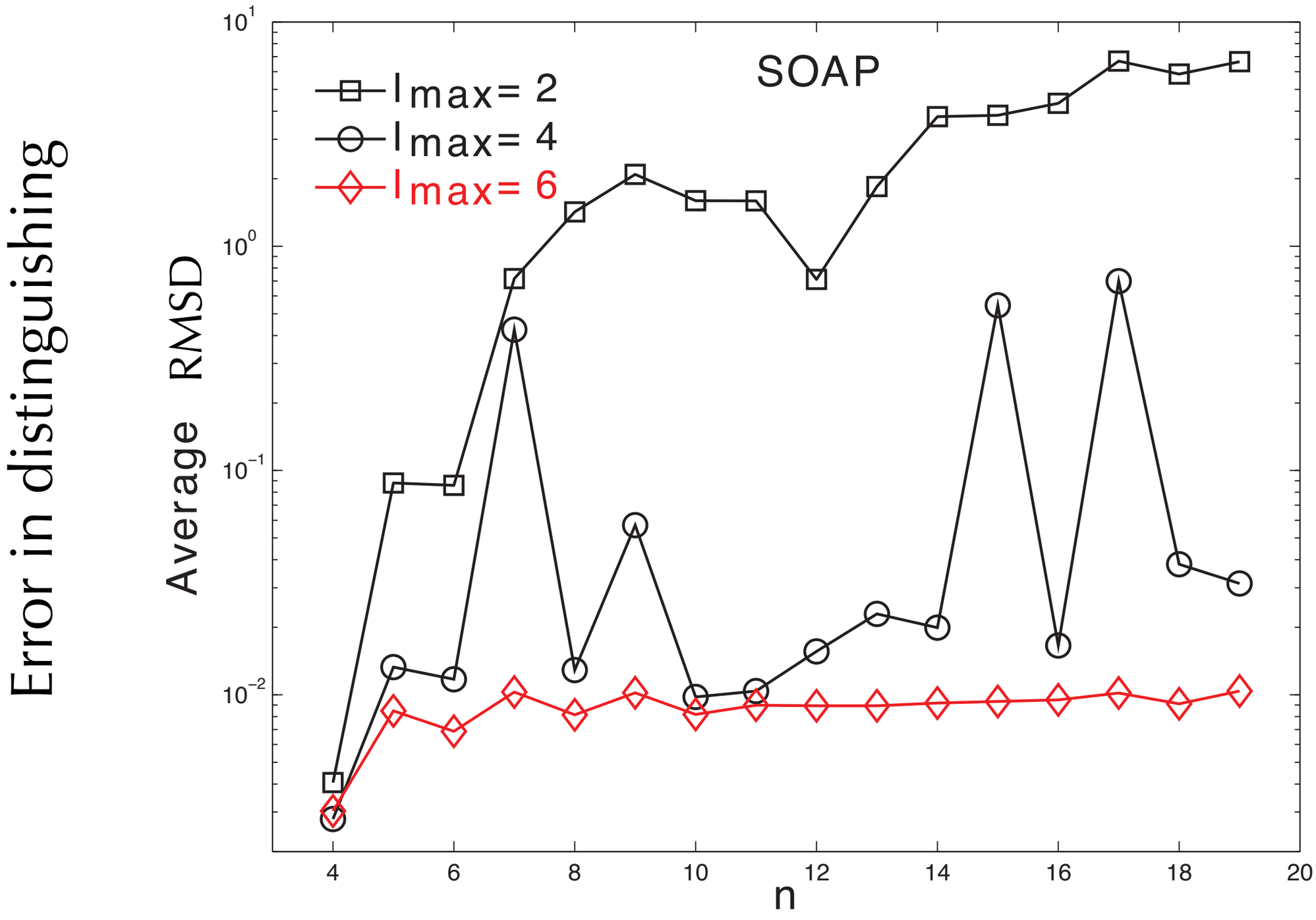
$$k(\rho_i, \rho_{i'}) = \int \left|S(\rho_i, \hat{R}\rho_{i'})\right|^2 d\hat{R} = \int d\hat{R} \left|\int \rho_i(\mathbf{r})\rho_{i'}(\hat{R}\mathbf{r})d\mathbf{r}\right|^2$$

- After LOTS of algebra: SOAP kernel

$$k(\rho_i, \rho_{i'}) = \sum_{n,n',l} p_{nn'l}^{(i)} p_{nn'l}^{(i')} \qquad p_{nn'l} = \mathbf{c}_{nl}^{\dagger}\mathbf{c}_{n'l}$$

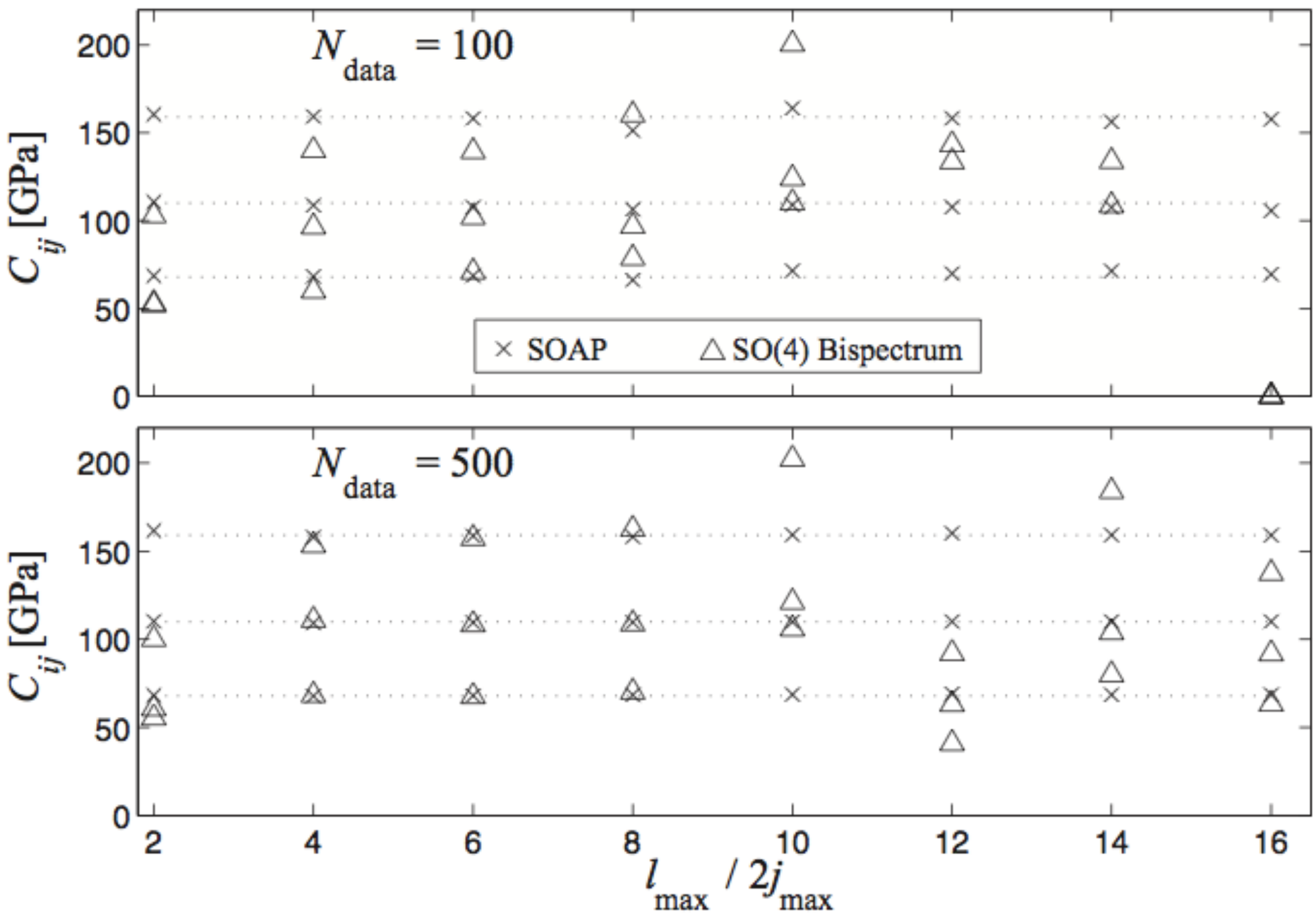
$$\boxed{K(\mathbf{q}, \mathbf{q}') \propto |k(\rho, \rho')|^{\xi}}$$

Smooth Overlap of Atomic Positions (SOAP)



Number of neighbours

Stability of fit: elastic constants of Si



Order of angular momentum expansion

How to generate databases?

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

QM MD on "representative" small systems:

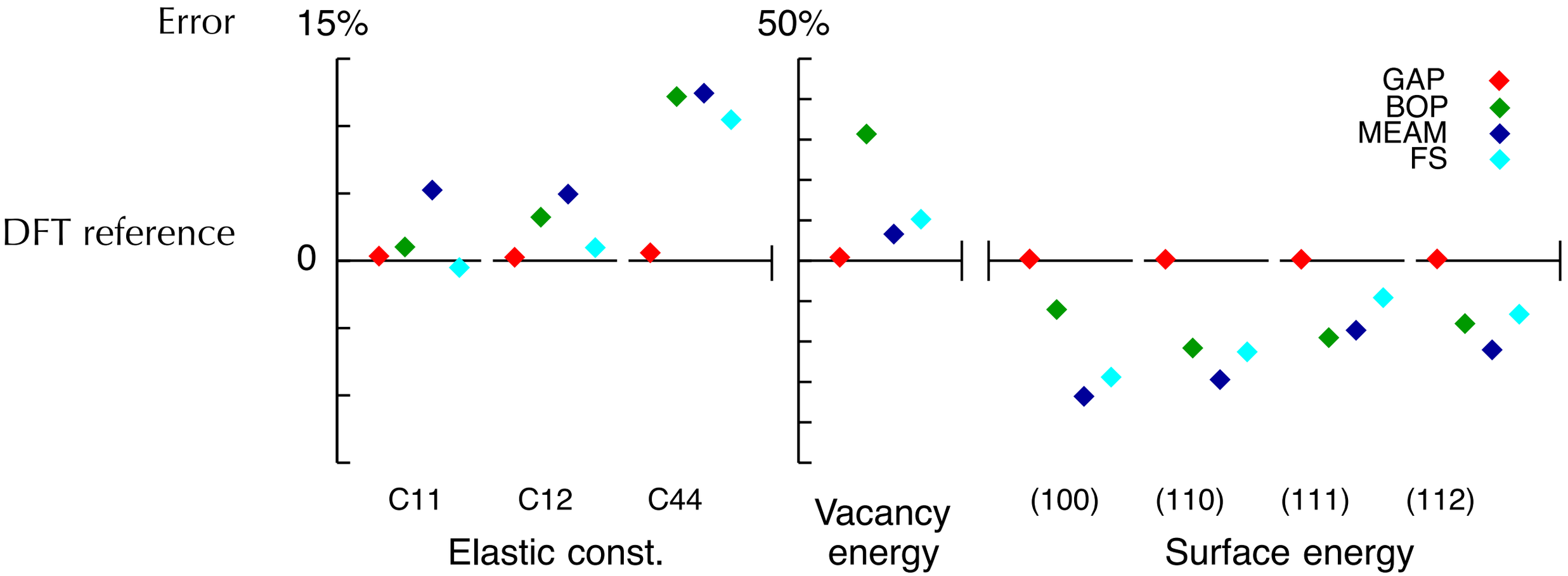
sheared primitive cell → elasticity
large unit cell → phonons
surface unit cells → surface energy
gamma surfaces → screw dislocation
vacancy in small cell → vacancy
vacancy @ gamma surface → vacancy near dislocation

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?

Existing potentials for tungsten

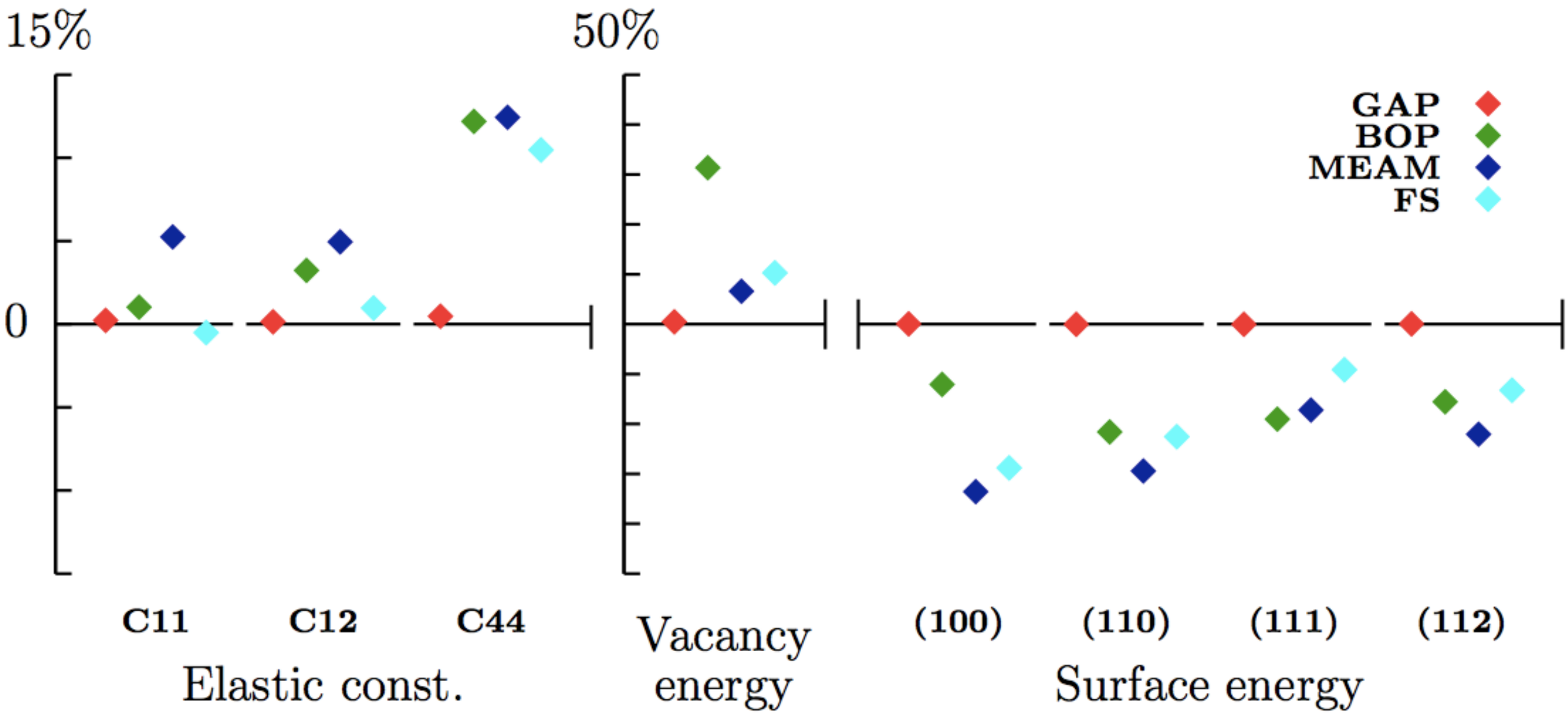


Building up databases for tungsten (W)

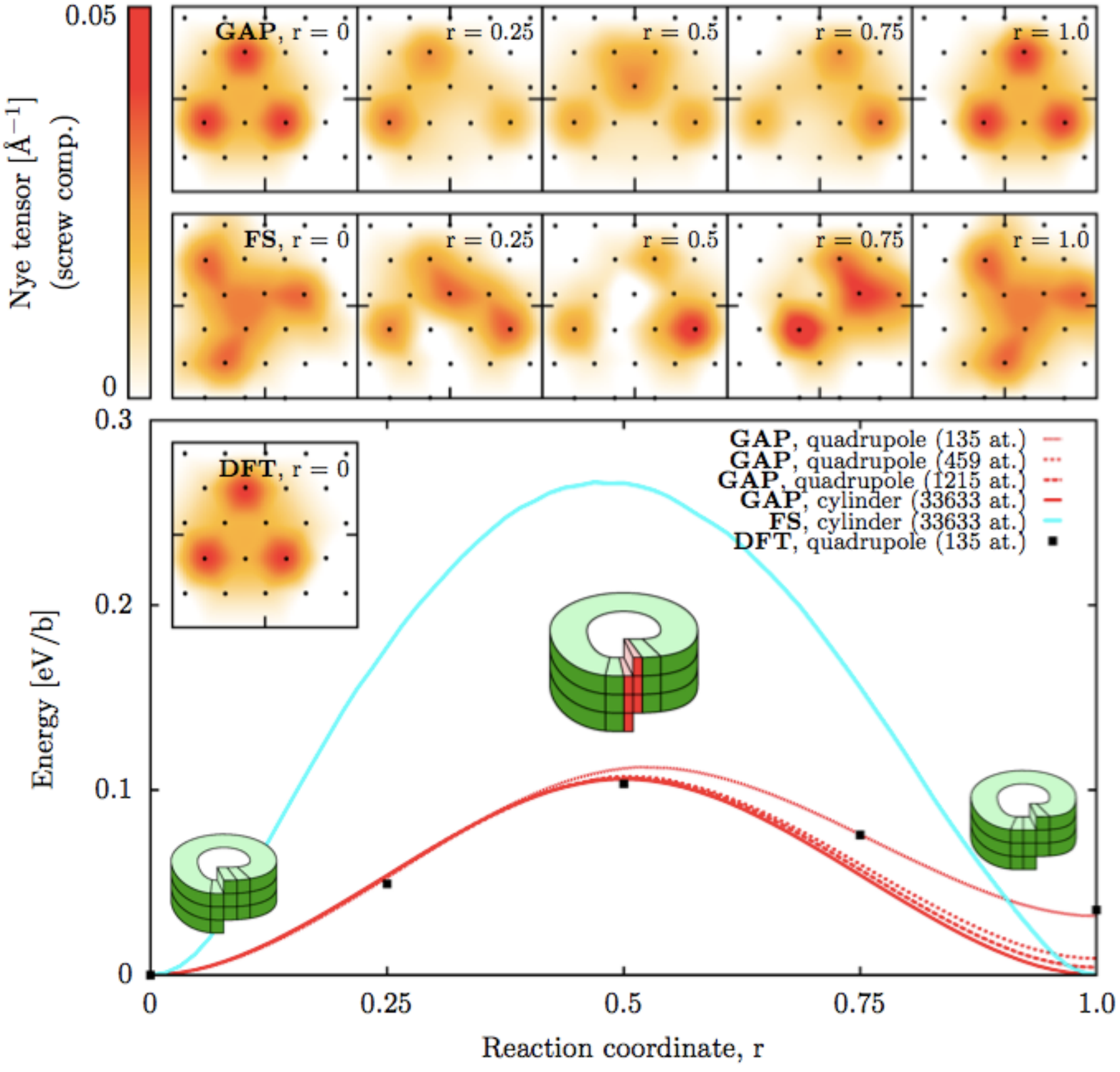
	Database:	Computational cost[a] [ms/atom]	Elastic constants[b] [GPa]	Phonon spectrum[b] [THz]	Vacancy formation[c] [eV]	Surface energy[b] [eV/Å^2]	Dislocation structure[d] [Å^{-1}]	Dislocation-vacancy binding energy [eV]	Peierls barrier [eV/b]
GAP_1 :	2000 × primitive unit cell with varying lattice vectors	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.

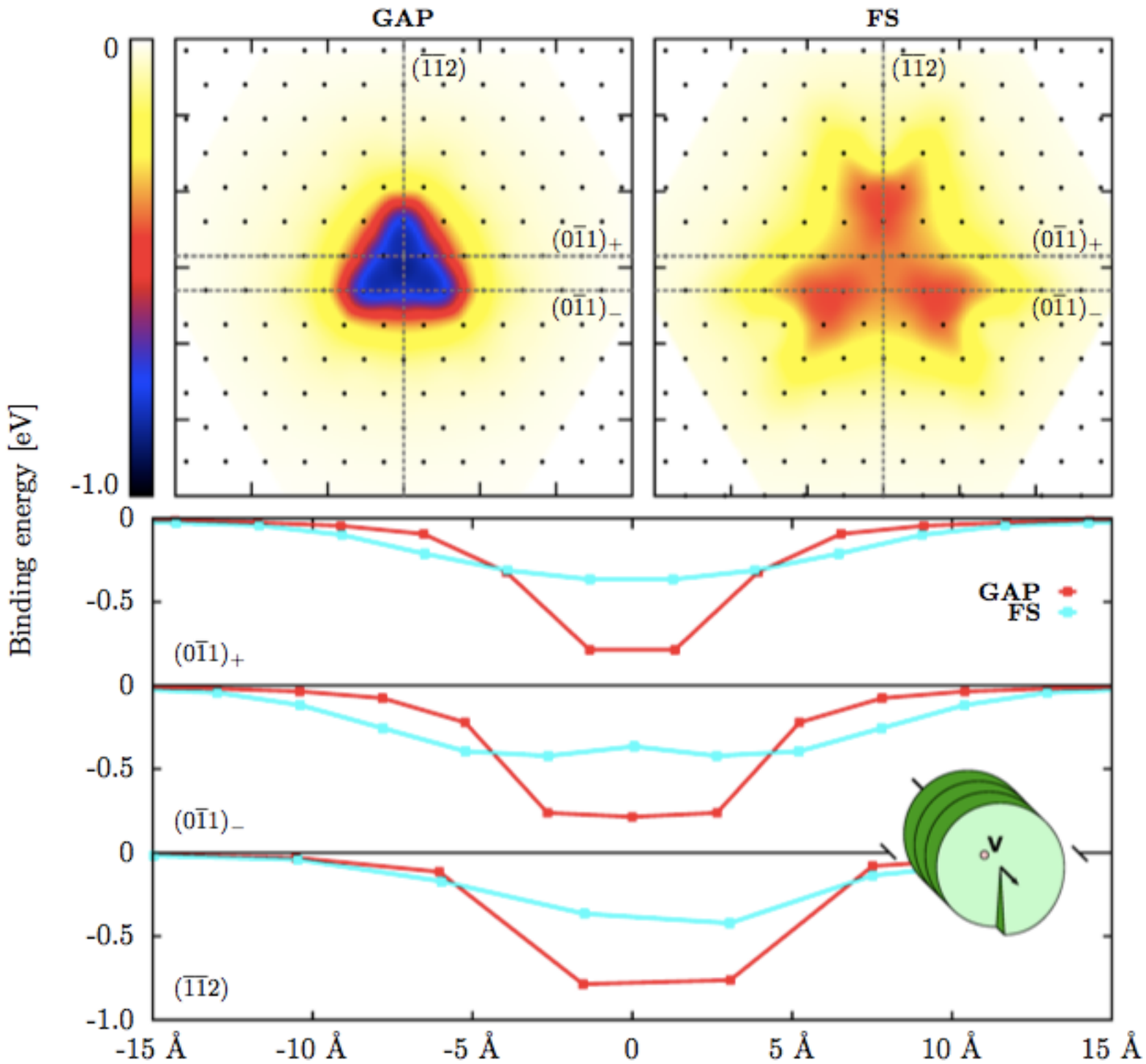
Improvement for tungsten (W)



Peierls barrier for screw dislocation glide



Vacancy-dislocation binding energy



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

Outstanding problems

- Accuracy on database ⟶ accuracy in properties?
- Database contents ⟶ region of validity ?
- Alloys - permutational complexity? Chemical variability?
- Systematic treatment of long range effects
- Electronic temperature