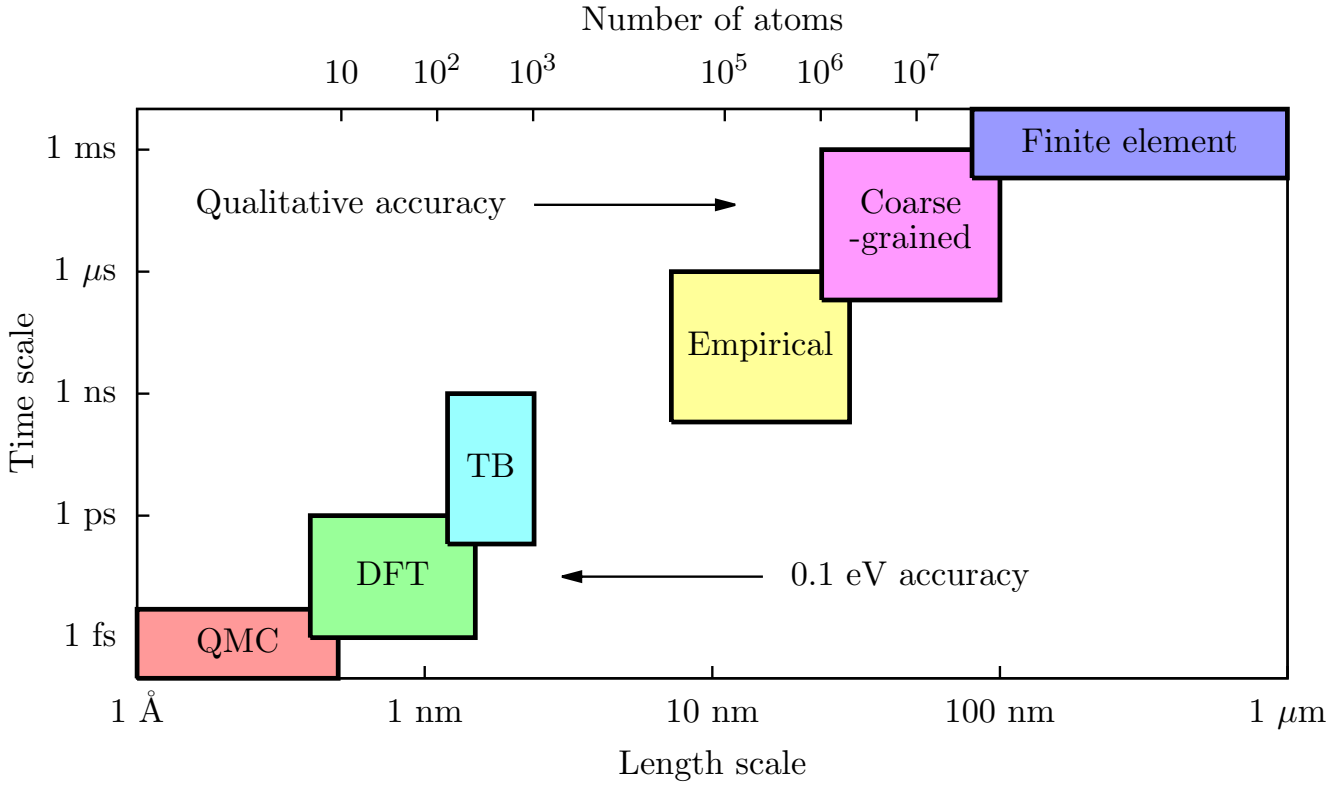
Accuracy and transferability of GAP models for tungsten

Wojciech J. Szlachta Albert P. Bartók Gábor Csányi

Engineering Laboratory
University of Cambridge

5 November 2014

Motivation



- Calculations from first principles lead to predictive capabilities that allow discovering novel materials with desired properties.
- Need for quantum mechanical accuracy, but with no electrons!

Empirical Potentials

- Typical interatomic potentials:
  - fixed functional form with a set number of adjustable parameters (i.e. empirical, analytical formula)
  - parameters fit to reproduce arbitrarily chosen target properties

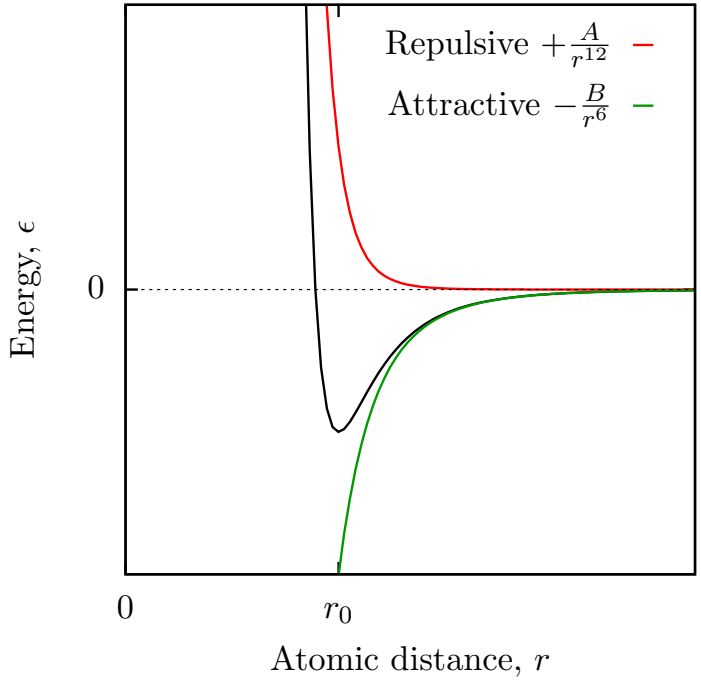
$$E = \sum_{i}^{N} V_1(\mathbf{x}^{(i)}) + \underbrace{\sum_{i}^{N} \sum_{\substack{j \\ j<i}}^{N} V_2(\mathbf{x}^{(i)}, \mathbf{x}^{(j)})}_{\text{bonds}}$$

$$+ \underbrace{\sum_{i}^{N} \sum_{\substack{j \\ j<i}}^{N} \sum_{\substack{k \\ k<j}}^{N} V_3(\mathbf{x}^{(i)}, \mathbf{x}^{(j)}, \mathbf{x}^{(k)})}_{\text{angles}} + \dots$$

Empirical Potentials

$$E = \sum_{i}^{N} \epsilon_i$$

$$\epsilon_i = \sum_{j}^{N} \left(\frac{A}{r^{12}} - \frac{B}{r^6} \right)$$



- Lennard-Jones potential:
  - choose target properties
  - fit parameters A and B to reproduce target properties

Gaussian Approximation Potential

- Apply **Bayesian probability** / **machine learning** to infer underlying function

Bayes' theorem

$$\text{posterior} = \frac{\text{likelihood} \times \text{prior}}{\text{marginal likelihood}}$$

- Start with prior distribution over functions
- Set of target values is observed
- Calculate posterior distribution for any new prediction with mean ϵ_i and variance $\sigma^2_{\epsilon_i}$

Gaussian Approximation Potential

- **Interpolation:** Gaussian process regression

  A. P. Bartók, M. C. Payne, R. Kondor, and G. Csányi, Phys. Rev. Lett. 104, 136403 (2010)

  - QM is inherently many dimensional
  - large function space, overfitting

- **Representation:** Smooth overlap of atomic positions

  A. P. Bartók, R. Kondor, and G. Csányi, Phys. Rev. B 87, 184115 (2013)

  - faithful representation of atomic environments
  - rotational & permutational invariance, smoothness

- **Database:** Tungsten interatomic potential

  W. J. Szlachta, A. P. Bartók, and G. Csányi, Phys. Rev. B 90, 104108 (2014)

  - domain specificity
  - predictive power

Gaussian Process Regression

- Optimal way of interpolating many-dimensional functions
- No fixed functional form
- Number of free parameters not fixed

$$\epsilon_i(\mathbf{q}_i) = \sum_j \alpha_j \overbrace{k(\mathbf{q}_j, \mathbf{q}_i)}^{\text{covariance function}}$$

$$\sigma^2_{\epsilon_i}(\mathbf{q}_i) = k(\mathbf{q}_i, \mathbf{q}_i) - \sum_j \beta_j k(\mathbf{q}_j, \mathbf{q}_i)$$

Covariance function $k(\mathrm{q}_j, \mathrm{q}_i)$

Prior distribution over functions is determined by choice of the covariance function.

Gaussian Process Regression

- Optimal way of interpolating many-dimensional functions
- No fixed functional form
- Number of free parameters not fixed

$$\epsilon_i(\mathbf{q}_i) = \sum_j \alpha_j \overbrace{k(\mathbf{q}_j, \mathbf{q}_i)}^{\text{covariance function}}$$

$$\sigma^2_{\epsilon_i}(\mathbf{q}_i) = k(\mathbf{q}_i, \mathbf{q}_i) - \sum_j \beta_j k(\mathbf{q}_j, \mathbf{q}_i)$$

Training data $\{\epsilon_j, \mathbf{q}_j\}_j$

Fit quality determined by both training data and choice of the covariance function.

Gaussian Process Regression

- Optimal way of interpolating many-dimensional functions
- No fixed functional form
- Number of free parameters not fixed

$$\epsilon_i(\mathbf{q}_i) = \sum_j \alpha_j \overbrace{k(\mathbf{q}_j, \mathbf{q}_i)}^{\text{covariance function}}$$

$$\sigma^2_{\epsilon_i}(\mathbf{q}_i) = k(\mathbf{q}_i, \mathbf{q}_i) - \sum_j \beta_j k(\mathbf{q}_j, \mathbf{q}_i)$$

Prediction mean $\epsilon_i(\mathbf{q}_i)$ and variance $\sigma^2_{\epsilon_i}(\mathbf{q}_i)$

Resulting potential is only as good as the training data!

Gaussian Process Regression (Notes)

$$\epsilon_i(\mathbf{q}_i) = \sum_j \alpha_j k(\mathbf{q}_j, \mathbf{q}_i)$$

- Calculation of α (and β) requires inversion of the covariance matrix $\mathbf{K}$:

$$\mathbf{K}_{ij} = k(\mathbf{q}_i, \mathbf{q}_j)$$

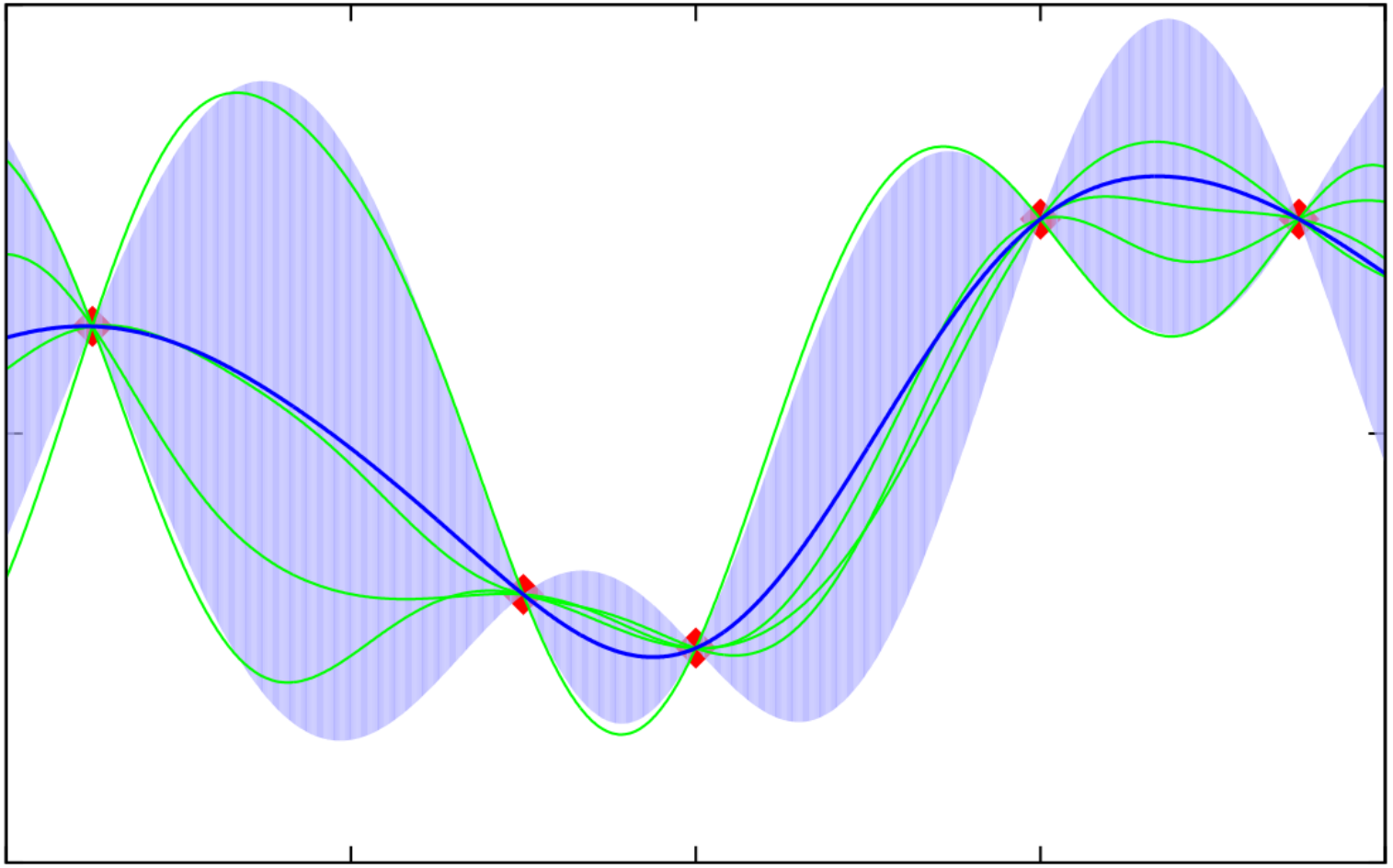
$$\alpha = (\mathbf{K} + \sigma_\nu \mathbf{I})^{-1}\mathbf{y}$$

- Linear regression can be quickly recovered with dot product covariance:

$$\mathbf{K}_{ij} = \sigma_w^2\, \mathbf{q}_i \cdot \mathbf{q}_j$$

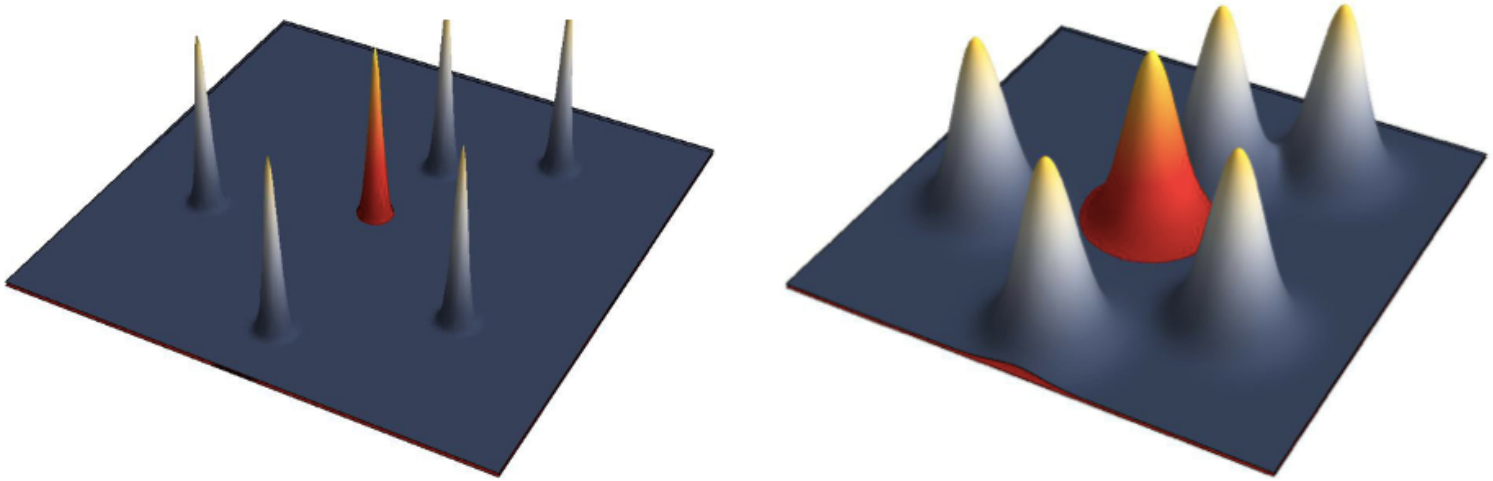
$$\epsilon_i(\mathbf{q}_i) = \alpha' \cdot \mathbf{q}_i$$

Gaussian Process Regression



Smooth Overlap of Atomic Positions

- **Need rotationally and permutationally invariant description of atomic environment!**



$$\epsilon(\{\mathbf{x}_j - \mathbf{x}_i\}_j^N) \rightarrow \epsilon(\rho_i)$$

$$\rho_i(\mathbf{r}) = \sum_j e^{-|\mathbf{r}-\mathbf{r}_{ij}|^2/2\sigma^2} f_{\mathrm{cut}}(|\mathbf{r}_{ij}|)$$

Smooth Overlap of Atomic Positions

$$\epsilon(\rho_i) \rightarrow \epsilon(\mathbf{q}_i)$$

$$\rho_i(\mathbf{r}) = \sum_{nlm} c^i_{nlm} g_n(|\mathbf{r}|) Y_{lm}(\hat{\mathbf{r}})$$

- We can now define:

$$\mathbf{q}_i = \left\{ \sum_m (c^i_{nlm})^* c^i_{n'lm} \right\}_{nn'l} \qquad \hat{\mathbf{q}}_i = \mathbf{q}_i / |\mathbf{q}_i|$$

- And we can demonstrate that:

$$\begin{aligned} K_{ij} = k(\mathbf{q}_i, \mathbf{q}_j) &= \sigma_w^2 \, |\hat{\mathbf{q}}_i \cdot \hat{\mathbf{q}}_j|^\xi \\ &\rightarrow \left| \int d\hat{R} \left| \int d\mathbf{r} \rho_i(\mathbf{r}) \rho_j(\hat{R}\mathbf{r}) \right|^2 \right|^\xi = k(\rho_i, \rho_j) \end{aligned}$$

Gaussian Approximation Potential (Caveats)

- **Atomic energies cannot be directly computed from QM data!**

Training from total energies, forces and stresses

Inferring function $\epsilon(\mathbf{q}_i)$ from linear combination of its values / partial derivatives (and atomic positions) possible.

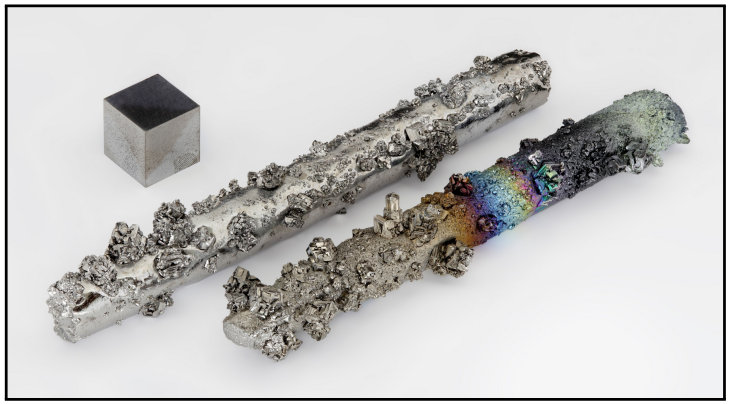
$$\textbf{total energies:} \quad E = \sum_i^N \epsilon_i$$

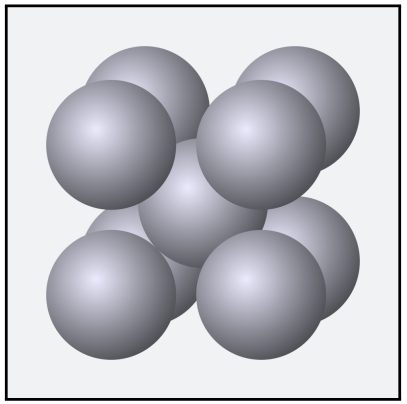
$$\textbf{atomic forces:} \quad \{\mathbf{f}^{(i)} = -\boldsymbol{\nabla}^{(i)} \sum_j^N \epsilon_j\}_i^N$$

$$\textbf{stress virials:} \quad \tau_{\alpha\beta} = -\sum_i^N x_\alpha^{(i)} \frac{\partial}{\partial x_\beta^{(i)}} \sum_j^N \epsilon_j$$

- Pseudo training points to deal with large data sets efficiently (which optimally represent the underlaying teaching data).

Tungsten Interatomic Potential



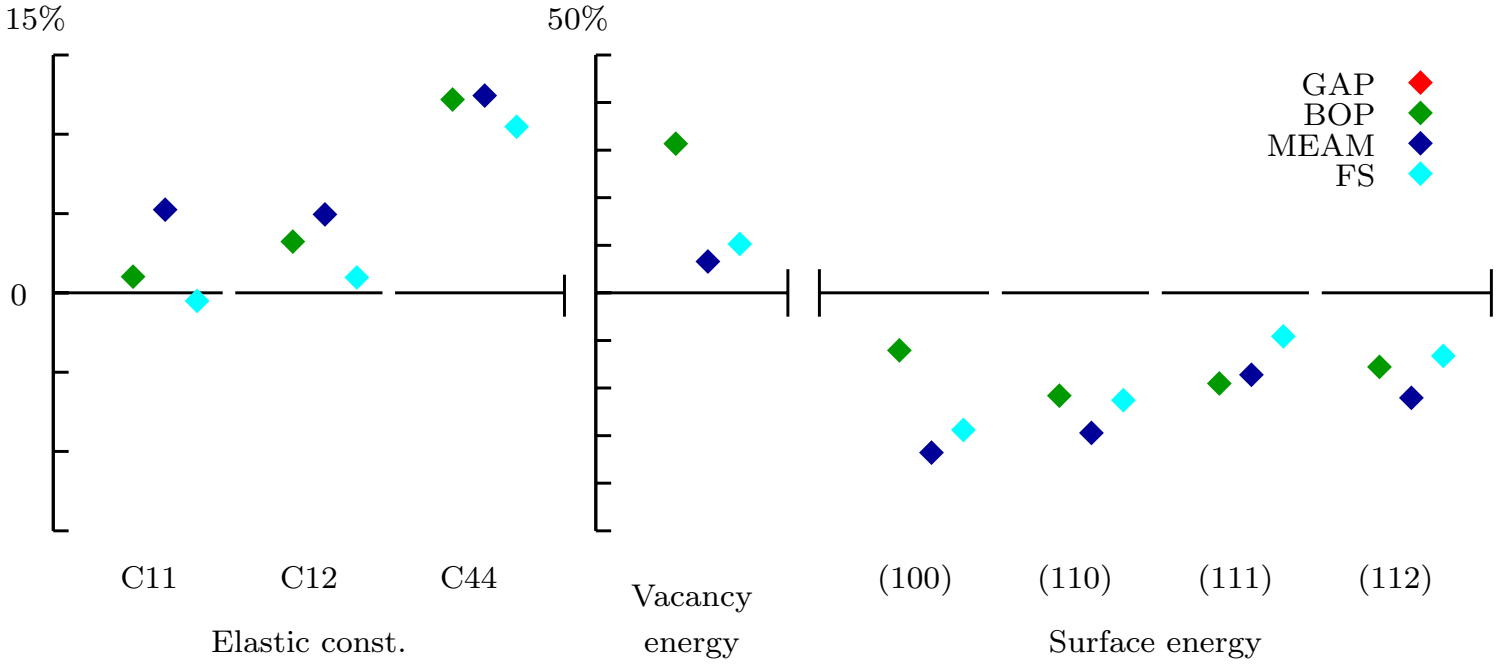


- Non-magnetic BCC transition metal with the highest melting temperature (3680 K).
- Often thought of as a prototype for this class of elements - good "toy model" before attempting iron!

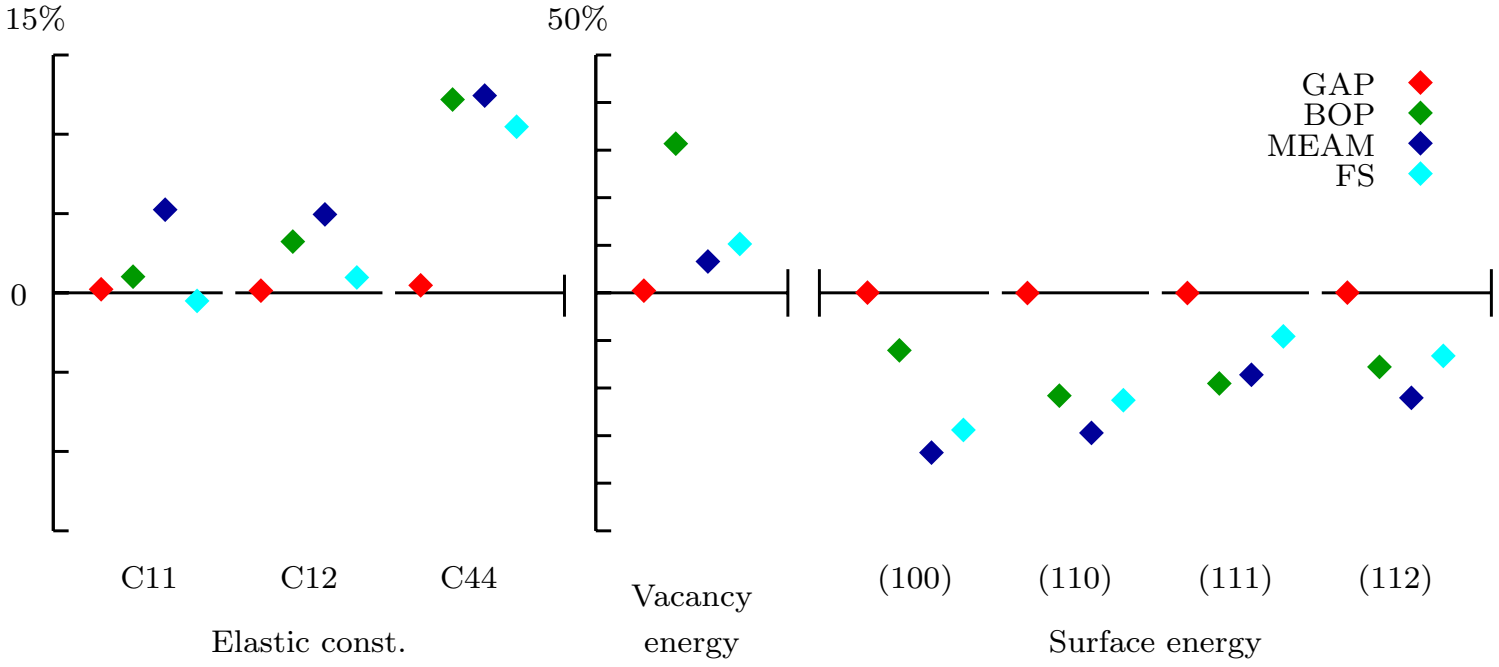
GAP Databases

	Database:		Computational cost [ms/atom]	Elastic constants [GPa]	Phonon spectrum [THz]	Vacancy formation [eV]	Surface energy [eV/Å^2]	Dislocation structure [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 unit cell	51.05	0.608	0.146	*1.414*	*0.1522*	*0.0006*		
GAP_3 :	GAP_2 +	vacancy in: 400 × 53-atom unit cell, 20 × 127-atom unit 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 unit cell (110), (112) gamma surfaces 6183 × 12-atom unit 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 unit 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 unit cell	93.33	0.748	0.129	0.015	0.0001	0.0001	-0.794	0.112

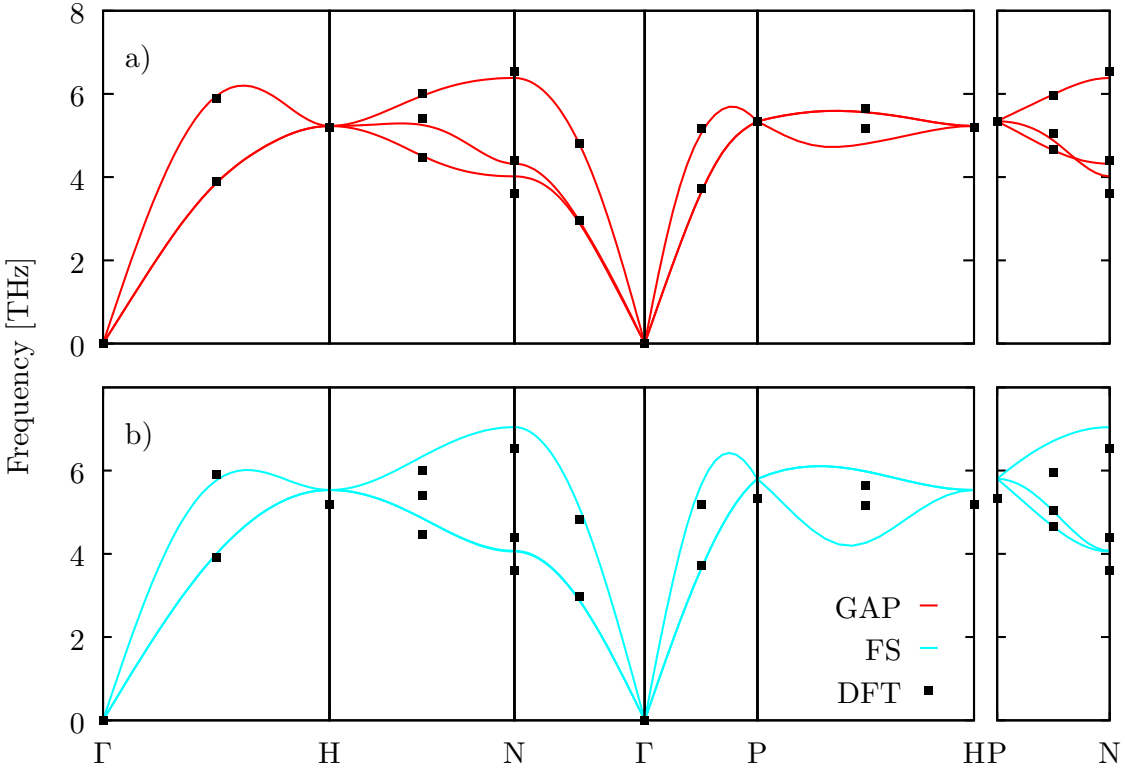
Elastic Constants & Lattice Defects



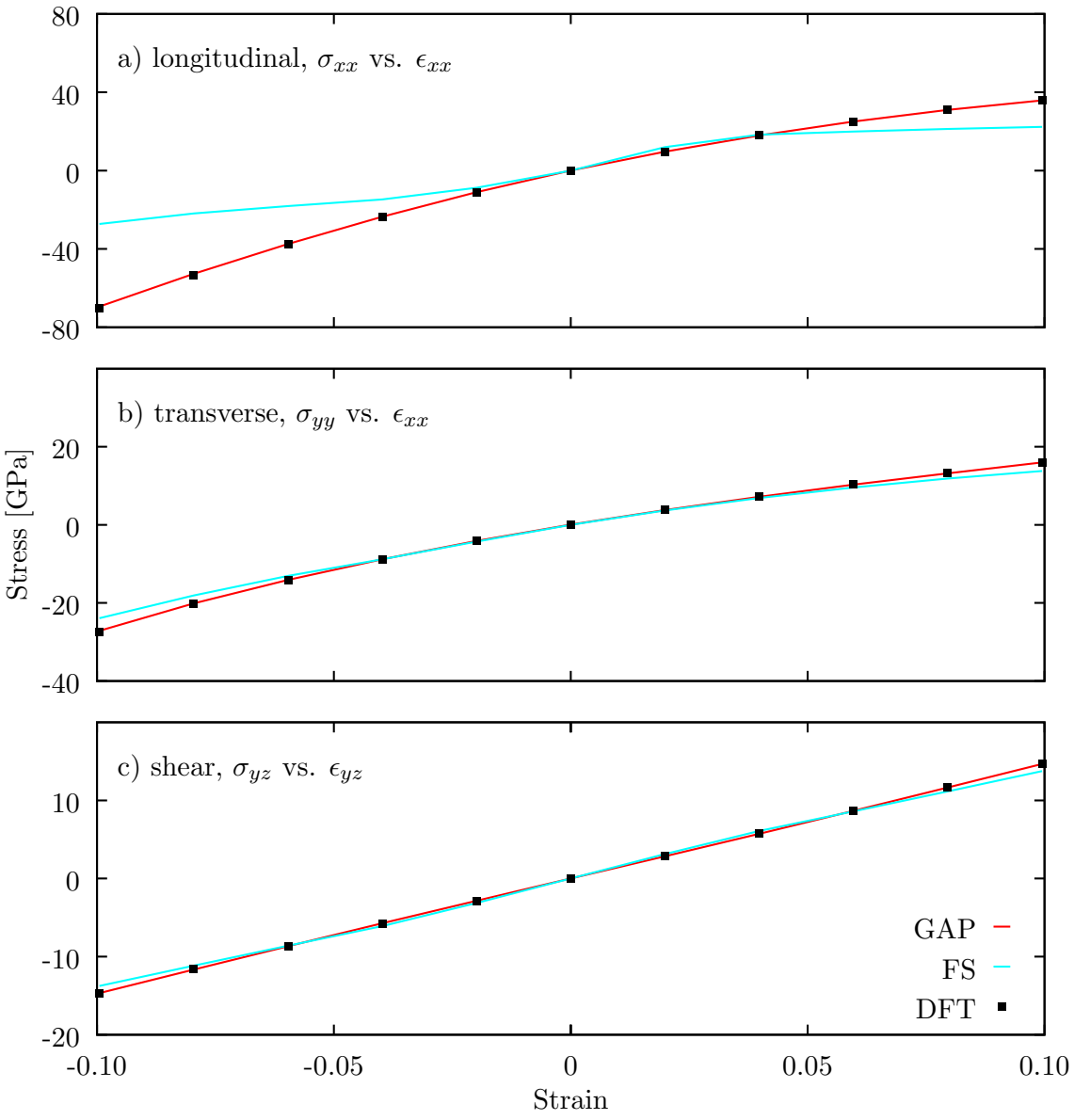
Elastic Constants & Lattice Defects



Phonon Spectrum

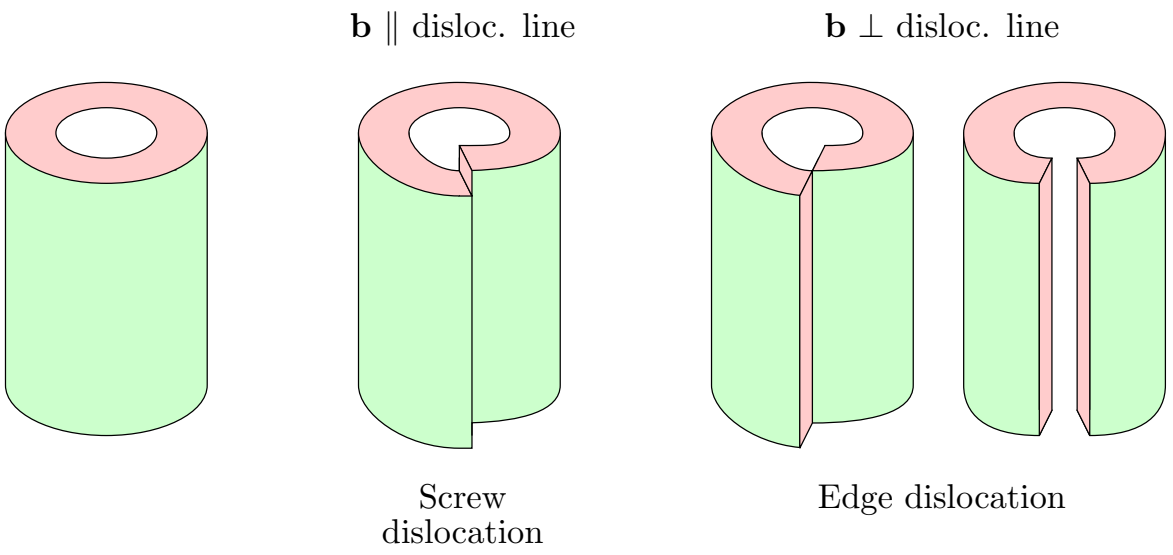


Stress-Strain Curves



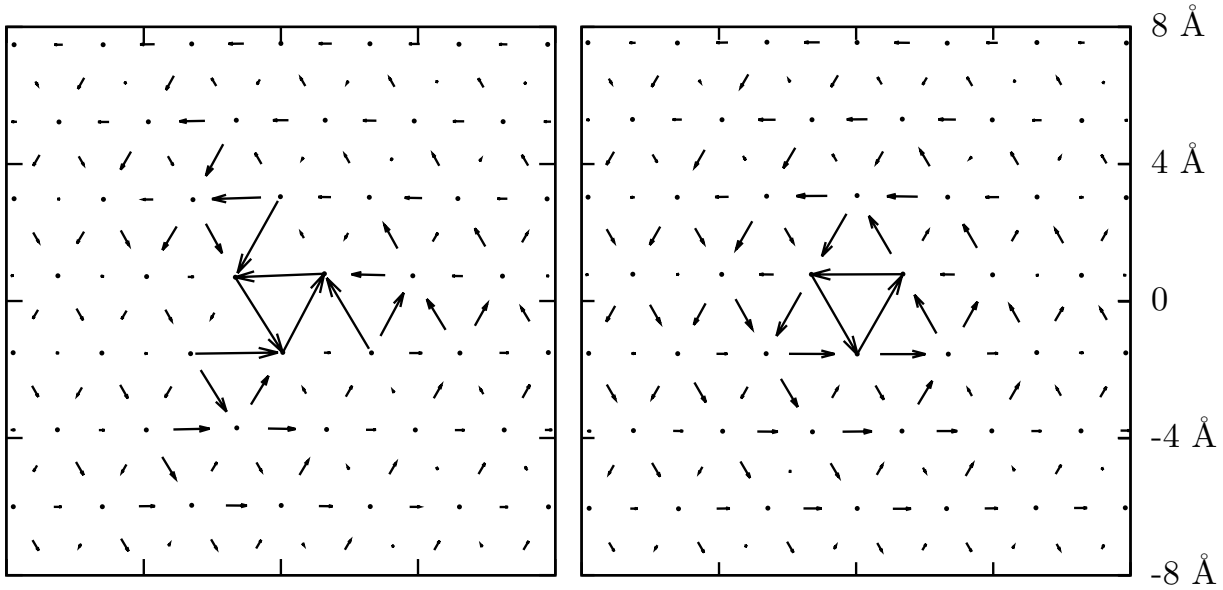
Towards Description of Plasticity

- Plasticity behaviour largely attributed to lattice crystallography
- Dominant dislocation type in bcc metals is $\frac{1}{2}\langle 111\rangle$ screw
- $\langle 110\rangle$ dislocations observed, but believed to be product of the dominant $\frac{1}{2}\langle 111\rangle$ screw dislocations



Screw Dislocation Core Structure

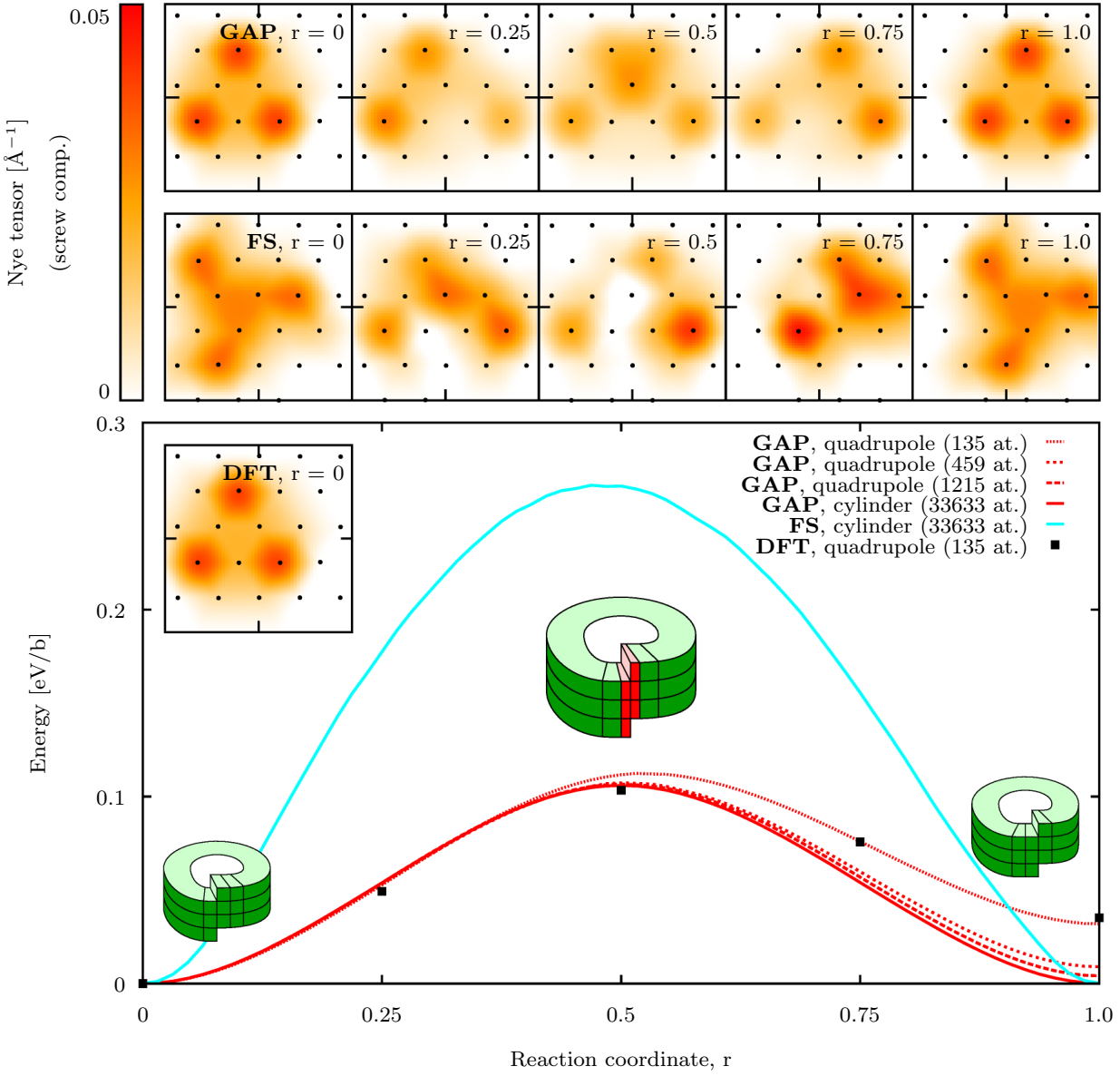
- Screw dislocation core structure determined by the properties of the interatomic potential and the boundary conditions
- Dislocation mobility is a direct consequence of the core structure energetics



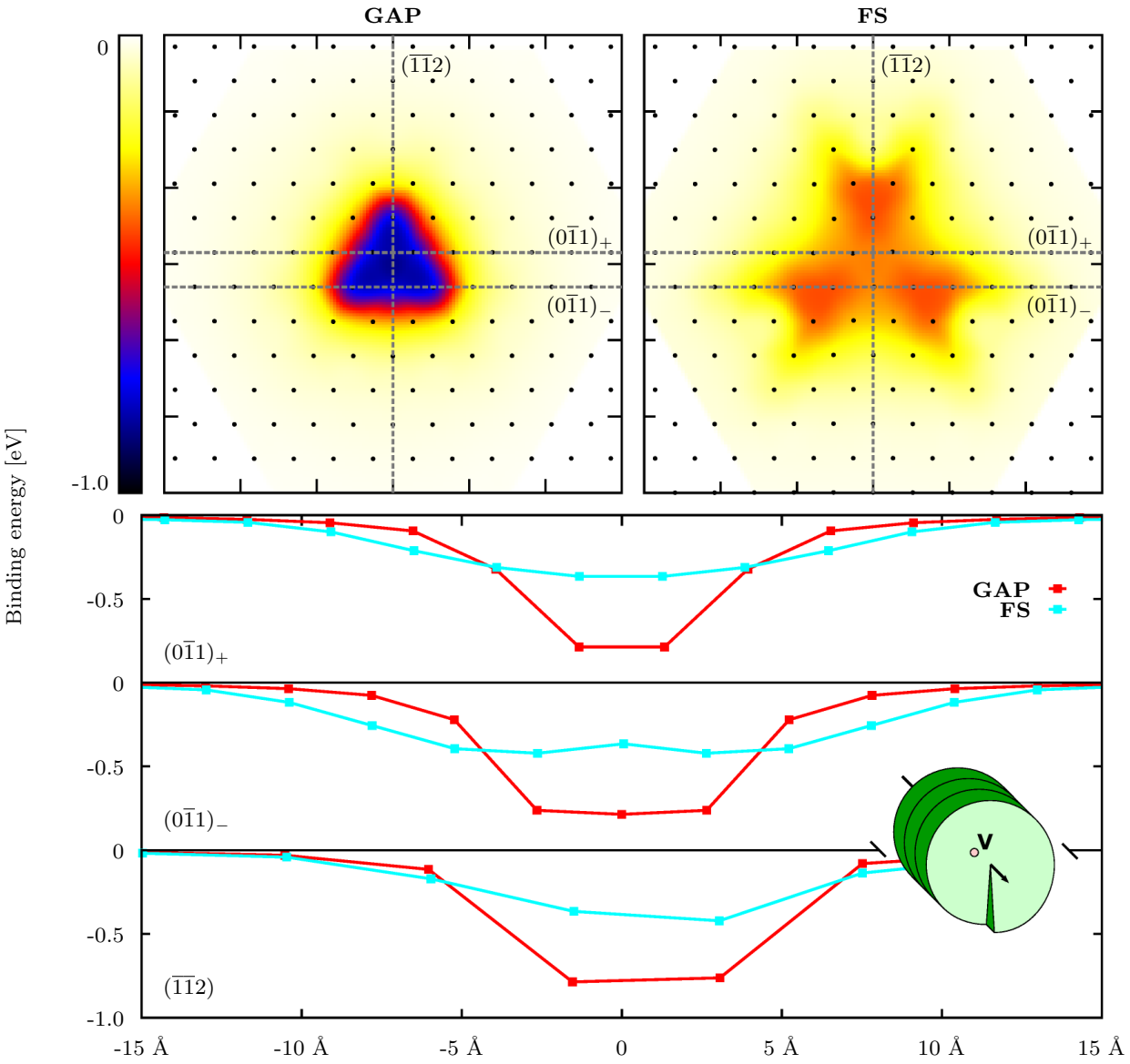
Non-symmetric core

Symmetric core

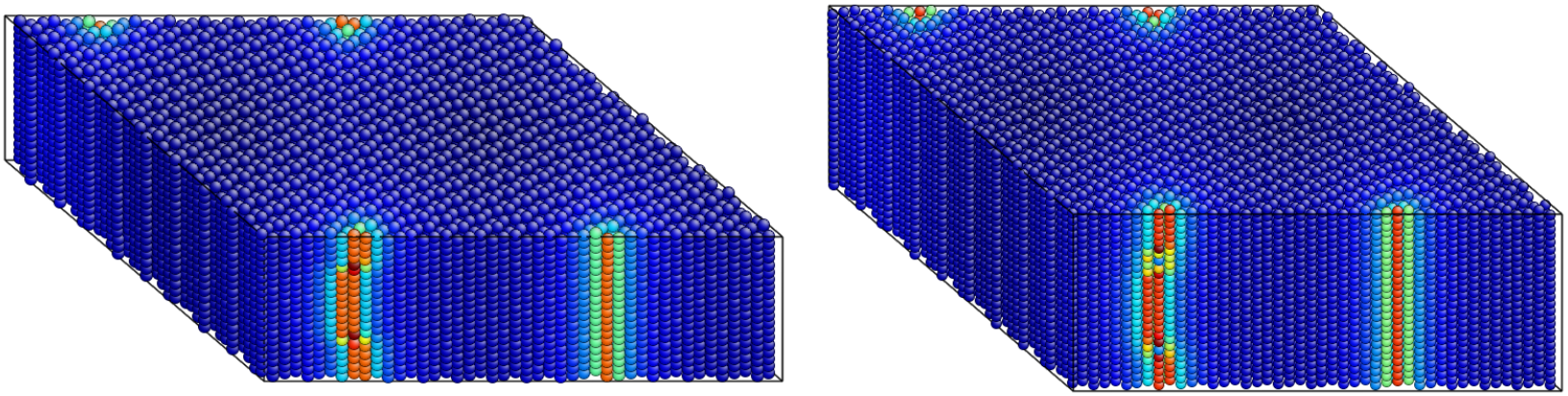
Screw Dislocation Peierls Barrier



Dislocation-Vacancy Binding Energy



Further Work



- Screw dislocation kink in tungsten
- Other defects/phases are also on the way

- Automatic generation of GAP databases for other elements

- Code and data available at: `www.libatoms.org`

Thank you for your attention!

More details in:
W. J. Szlachta, A. P. Bartók, and G. Csányi,
Phys. Rev. B 90, 104108 (2014)

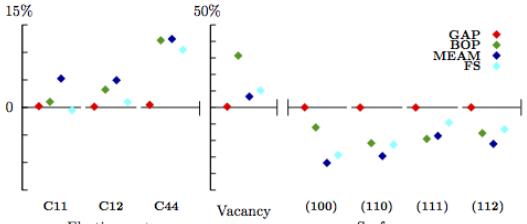
Accuracy and transferability of GAP models for tungsten

Wojciech J. Szlachta, Albert P. Bartók, and Gábor Csányi
Engineering Laboratory, University of Cambridge, Trumpington Street, Cambridge, CB2 1PZ, UK

We introduce interatomic potentials for tungsten in the bcc crystal phase and its defects within the Gaussian Approximation Potential (GAP) framework, fitted to a database of first principles density functional calculations. We investigate the performance of a number of models based on a series of databases of increasing coverage in configuration space and showcase our strategy of choosing representative small unit cells to train models that predict properties only observable using thousands of atoms. The most comprehensive model is then used to calculate properties of the screw dislocation, including its structure, the Peierls barrier and the energetics of the vacancy-dislocation interaction. All software and data are available at www.libatoms.org.

PACS numbers: 65.40.De,71.15.Nc,31.50.-x,34.20.Cf

Tungsten is a hard, refractory metal with the highest melting point (3695 K) among metals, and its alloys are utilised in numerous technological applications. The details of the atomistic processes behind the plastic behaviour of tungsten have been investigated for a long time and many interatomic potentials exist in the literature reflecting an evolution, over the past three decades, in their level of sophistication, starting with the Finnis Sinclair (FS) potential[1], embedded atom model



Appendix

- Compute many-body atomic energy function

$$\epsilon_i = \epsilon(\{\mathbf{x}_j - \mathbf{x}_i\}_j^N) = \epsilon(\mathbf{q}_i)$$

- Fit to arbitrary precision QM data

- **Need rotationally and permutationally invariant description of atomic environment!**

Decriptor vs. covariance function symmetries

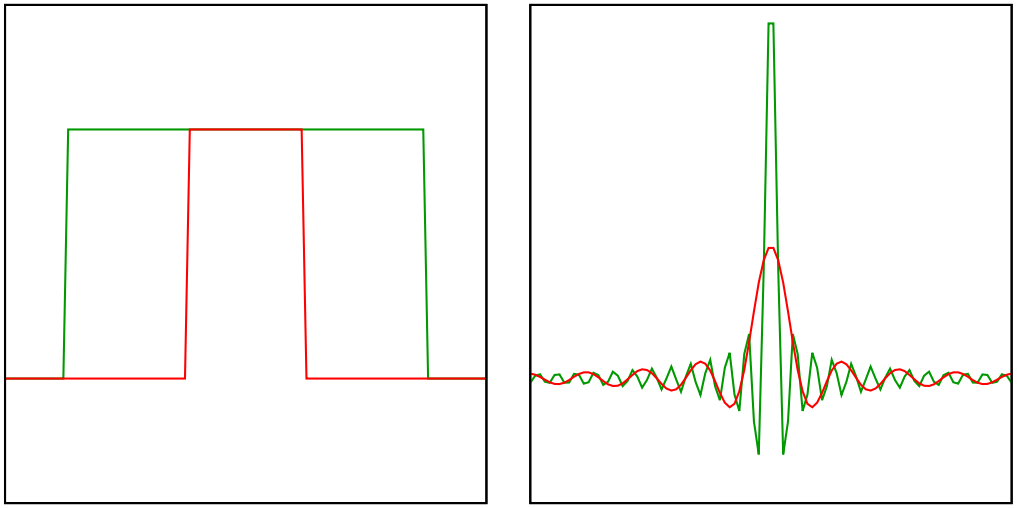
$$\begin{array}{c} k(\mathbf{q}_j, \mathbf{q}_i) \\ + \\ \{\mathbf{x}_{k'} - \mathbf{x}_i\}_{k'} \to \mathbf{q}_i \end{array} \longleftrightarrow \begin{array}{c} k^*(\{\mathbf{x}_{k''} - \mathbf{x}_j\}_{k''}, \\ \{\mathbf{x}_{k'} - \mathbf{x}_i\}_{k'}) \end{array}$$

Appendix

Bispectrum	**Smooth Overlap of Atomic Positions**
expand atomic density using 4d spherical harmonics basis	expand atomic density using 3d spherical harmonics and radial basis
atoms represented by Dirac δ function	atoms represented by Gaussian function
square exponential covariance	dot product covariance

Appendix

- Both bispectrum and SOAP are based on expansion in spherical harmonics that needs to be truncated:



Fourier / bispectrum space Real space

- Mapping from $\{\mathbf{x}_j - \mathbf{x}_i\}_j$ to $\mathbf{q}_i$ should be an invertible function such that both the function and its inverse are smooth!

Appendix

- Calculation of α (and β) requires inversion of the covariance matrix $\mathbf{K}$:

$$\mathbf{K}_{ij} = k(\mathbf{q}_i, \mathbf{q}_j)$$

$$\alpha = [\mathbf{K}_{MM} + \mathbf{K}_{MN}\mathbf{L}\mathbf{\Lambda}^{-1}\mathbf{L}^T\mathbf{K}_{NM}]^{-1}\mathbf{K}_{MN}\mathbf{L}\mathbf{\Lambda}^{-1}\mathbf{y}$$

- Where:

$$\mathbf{K}_{DD} = \mathbf{L}^T\mathbf{K}_{NN}\mathbf{L}$$

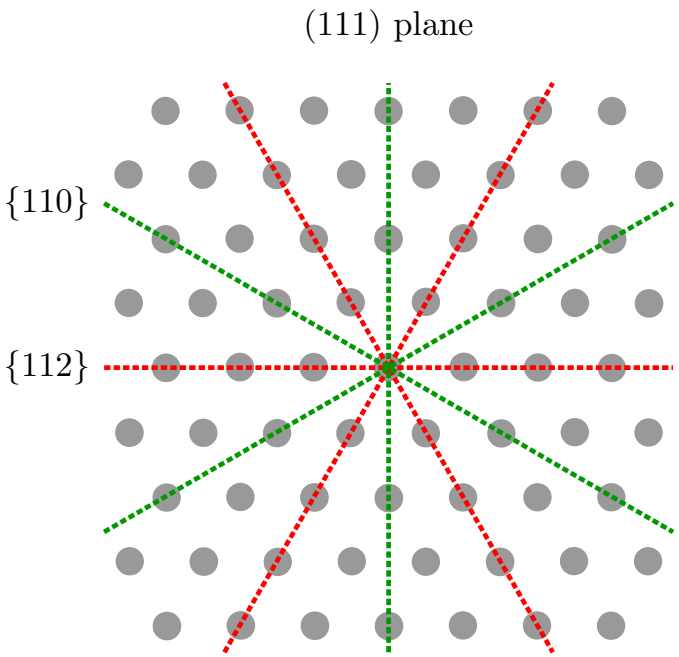
$$\mathbf{\Lambda} = \sigma_\nu^2\mathbf{I}$$

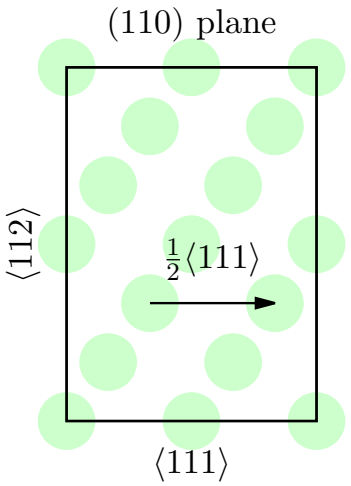
Appendix

	Database:						
	1	2	3	4	5	6	Total
GAP_1	2000						2000
GAP_2	814	3186					4000
GAP_3	366	1378	4256				6000
GAP_4	187	617	1890	6306			9000
GAP_5	158	492	1604	5331	2415		10000
GAP_6	140	450	1500	4874	2211	825	10000

Appendix

- Slip can occur along nearest neighbour direction $\frac{1}{2}\langle 111\rangle$ (also the shortest Burgers vector)
- The most densely packed planes of the $\langle 111\rangle$ zone are the $\{110\}$ planes

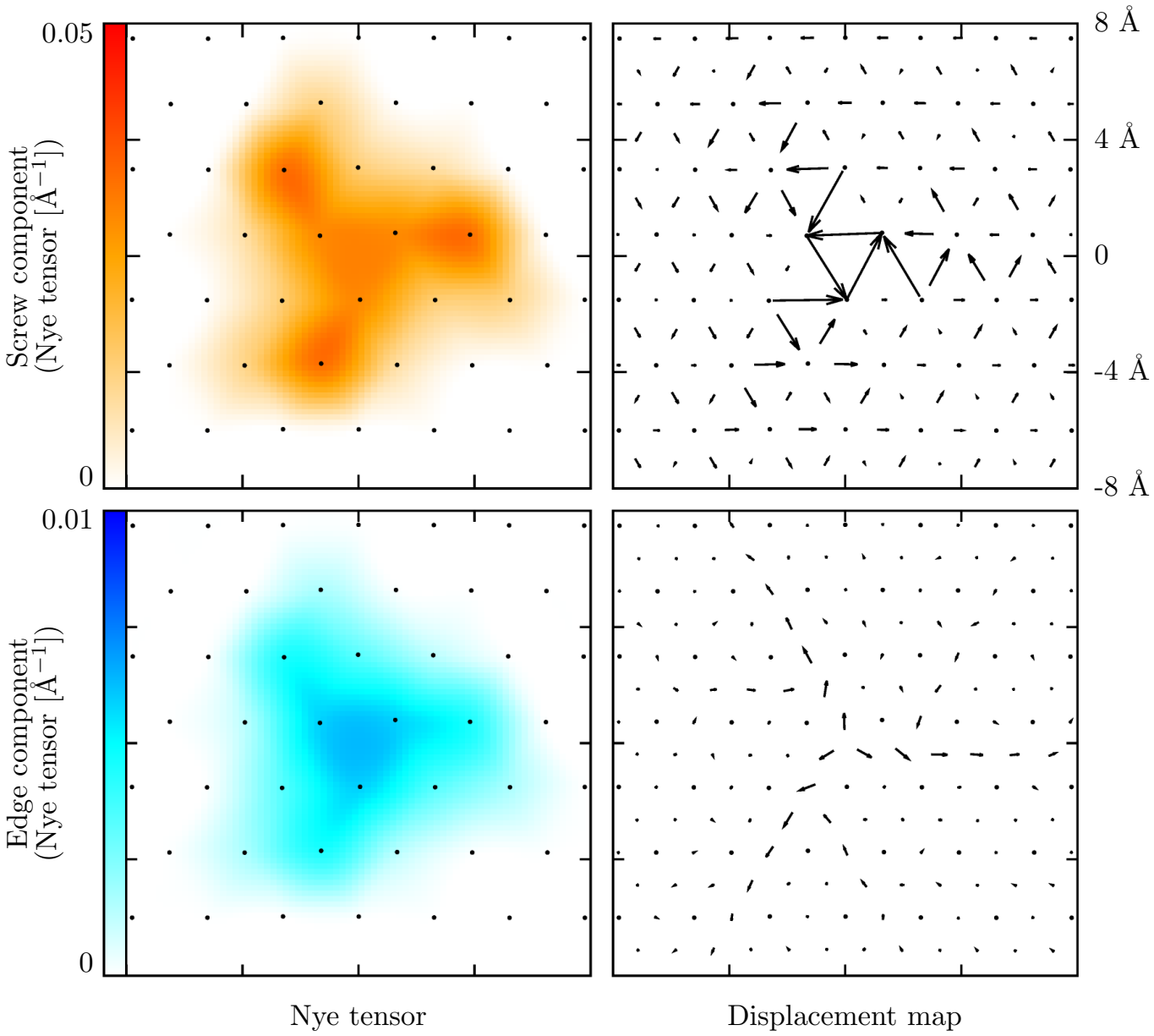




Appendix

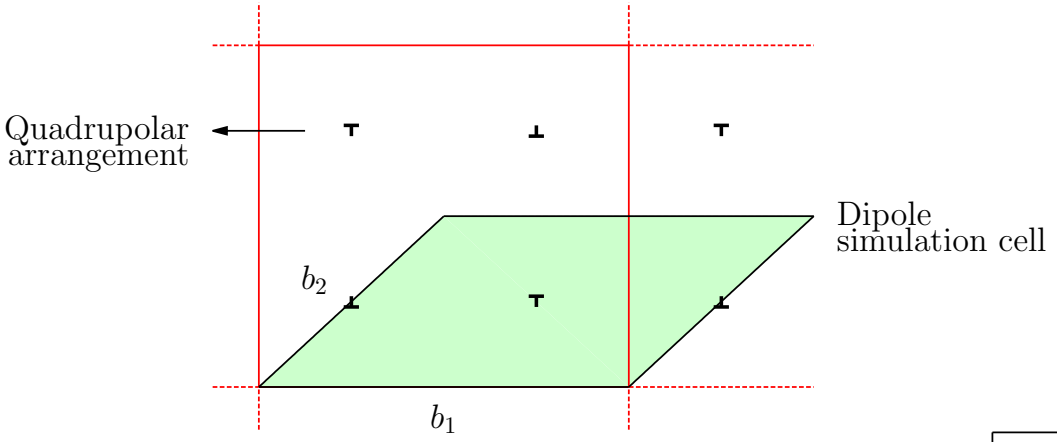
Slip plane separation distance		
$\langle 111 \rangle \{110\}$	$\rightarrow$	$\frac{1}{\sqrt{2}}\, a$
$\langle 111 \rangle \{112\}$	$\rightarrow$	$\frac{1}{\sqrt{6}}\, a$
$\langle 111 \rangle \{123\}$	$\rightarrow$	$\frac{1}{\sqrt{14}}\, a$
$\langle 111 \rangle \{134\}$	$\rightarrow$	$\frac{1}{\sqrt{26}}\, a$

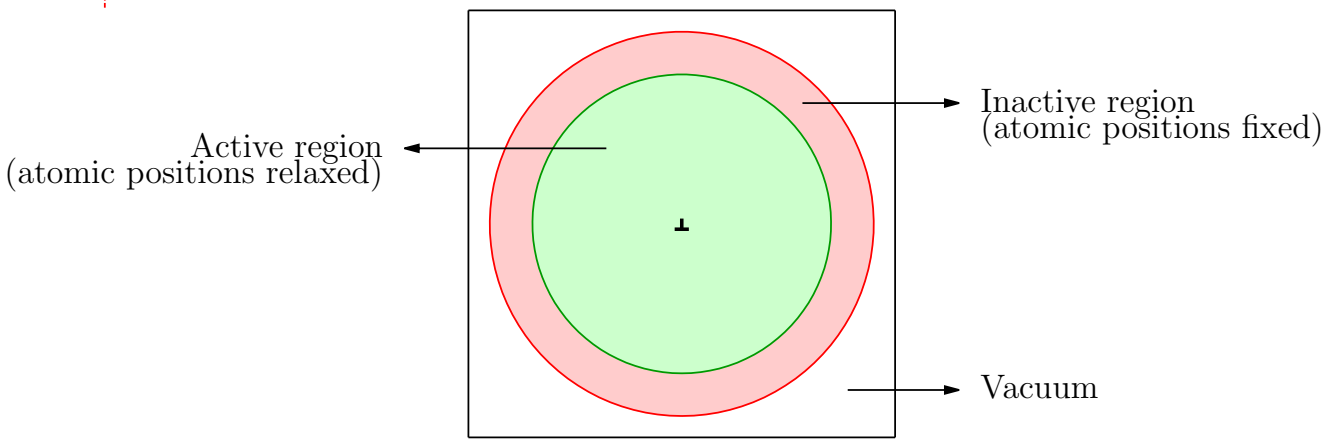
Appendix


0.05
0
Screw component
(Nye tensor [Å^{-1}])
0.01
0
Edge component
(Nye tensor [Å^{-1}])
8 Å
4 Å
0
-4 Å
-8 Å
Nye tensor
Displacement map

Appendix

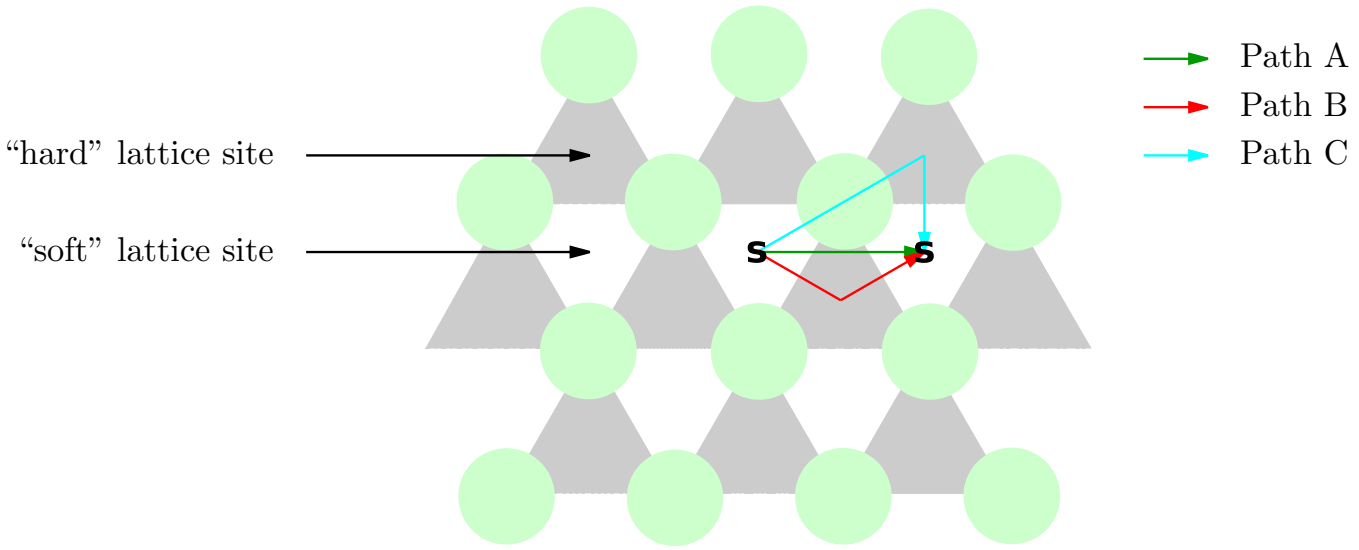
Dislocation Quadrupole



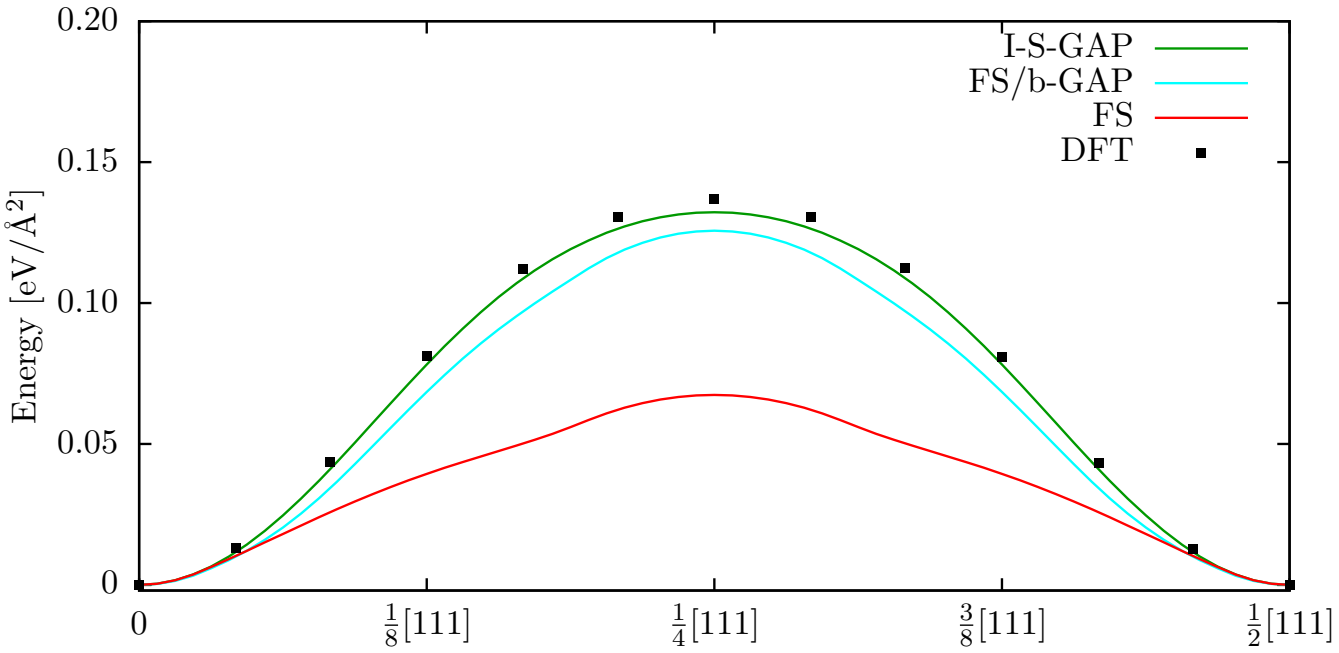


Isolated Dislocation

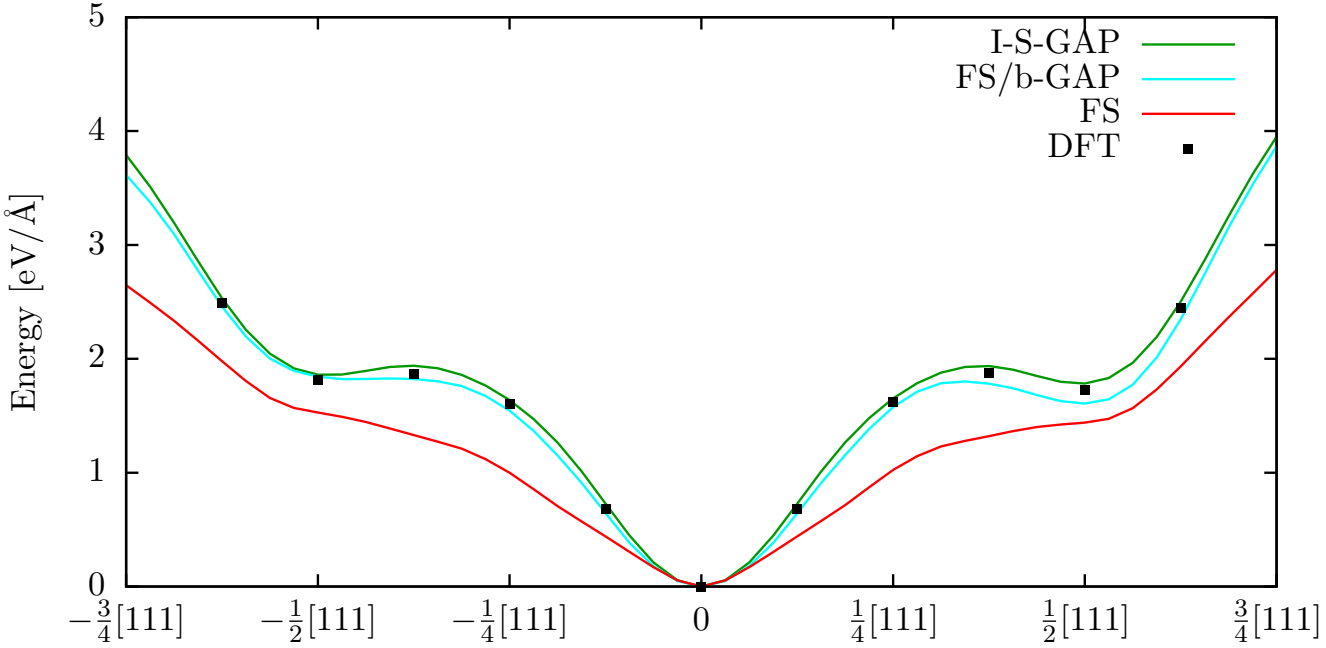
Appendix



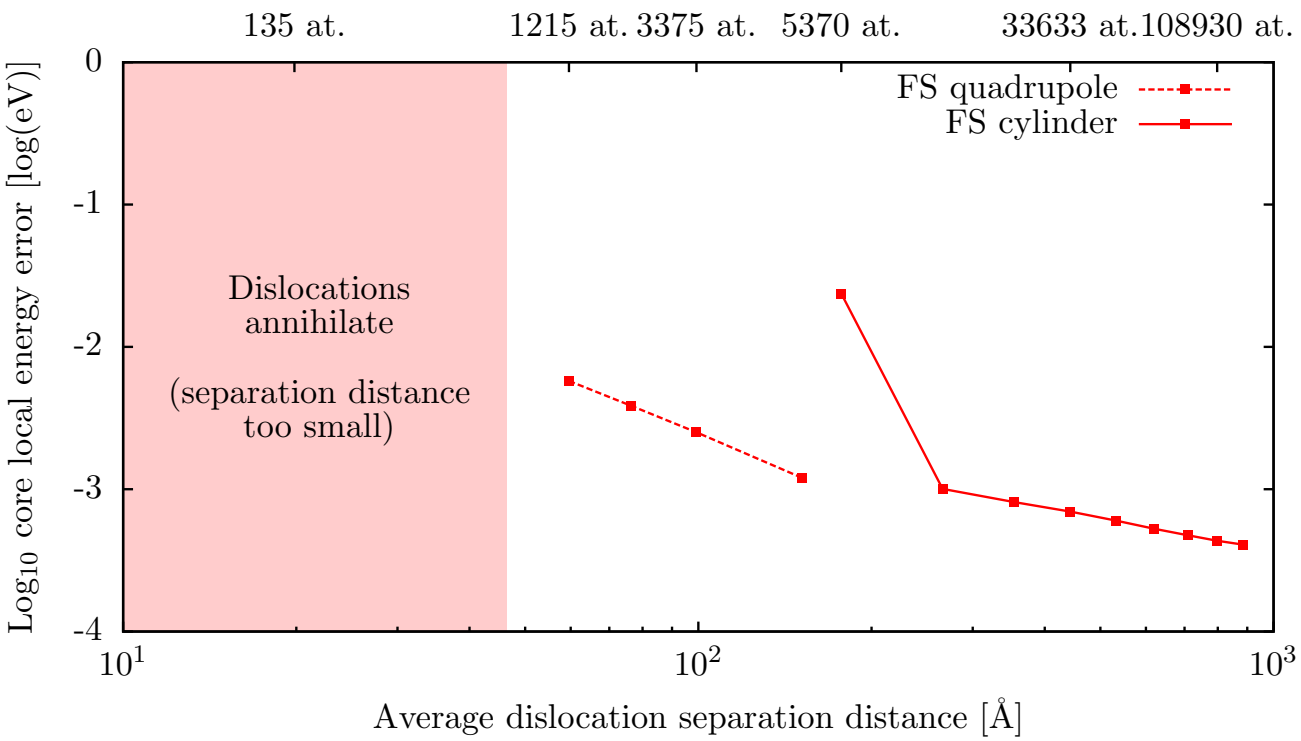
Appendix


0.20
0.15
0.10
0.05
0
Energy [eV/Å²]
I-S-GAP
FS/b-GAP
FS
DFT
0
1/8[111]
1/4[111]
3/8[111]
1/2[111]

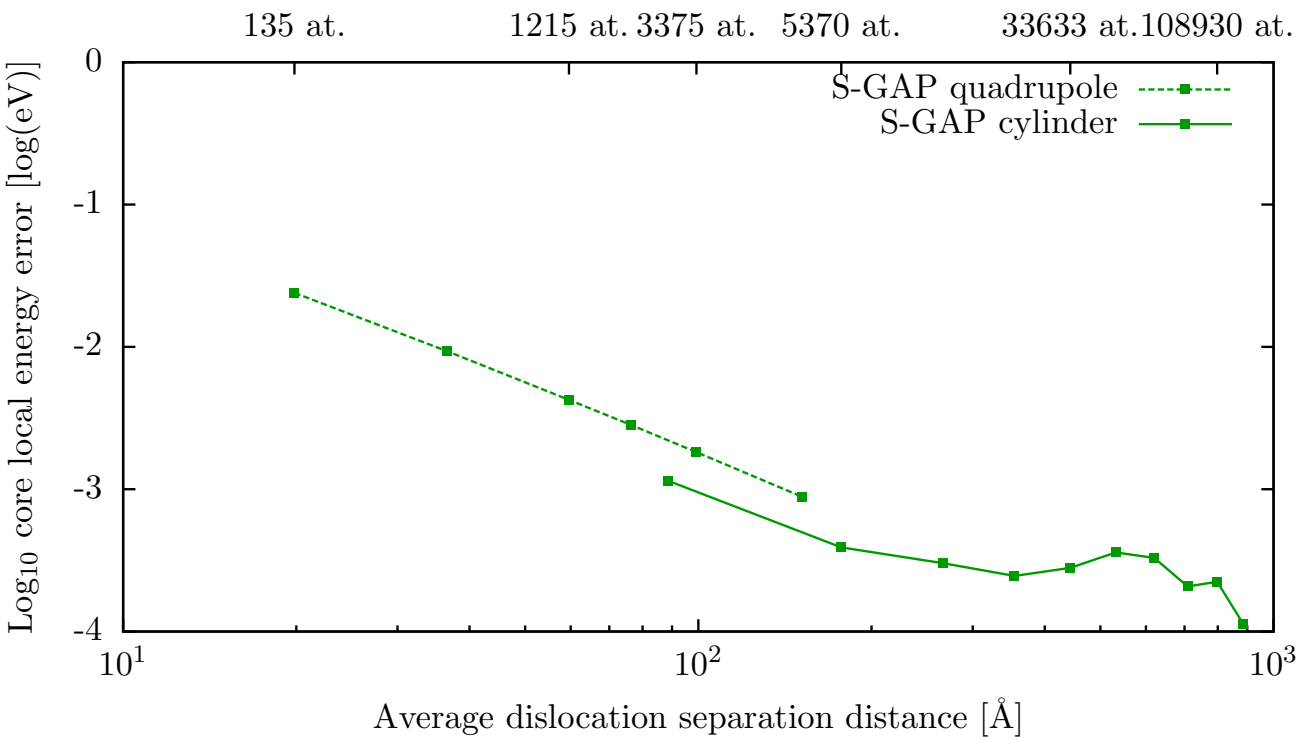
Appendix


I-S-GAP
FS/b-GAP
FS
DFT
Energy [eV/Å]
0
1
2
3
4
5
$-\frac{3}{4}[111]$
$-\frac{1}{2}[111]$
$-\frac{1}{4}[111]$
0
$\frac{1}{4}[111]$
$\frac{1}{2}[111]$
$\frac{3}{4}[111]$

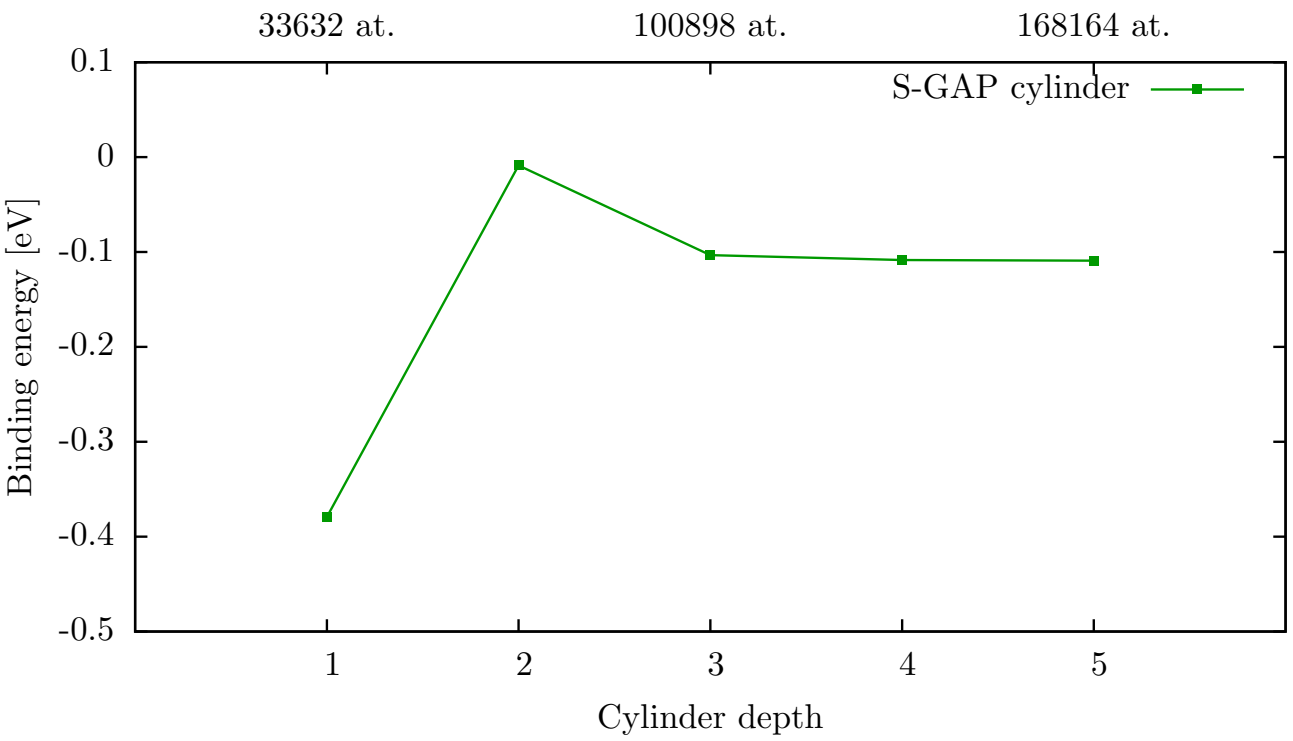
Appendix


135 at.
1215 at.
3375 at.
5370 at.
33633 at.
108930 at.
FS quadrupole
FS cylinder
Dislocations annihilate
(separation distance too small)
Log10 core local energy error [log(eV)]
0
-1
-2
-3
-4
10^1
10^2
10^3
Average dislocation separation distance [Å]

Appendix


135 at.
1215 at.
3375 at.
5370 at.
33633 at.
108930 at.
S-GAP quadrupole
S-GAP cylinder
0
-1
-2
-3
-4
10^1
10^2
10^3
Log_{10} core local energy error [log(eV)]
Average dislocation separation distance [Å]

Appendix


33632 at.
100898 at.
168164 at.
S-GAP cylinder
0.1
0
-0.1
-0.2
-0.3
-0.4
-0.5
Binding energy [eV]
1
2
3
4
5
Cylinder depth