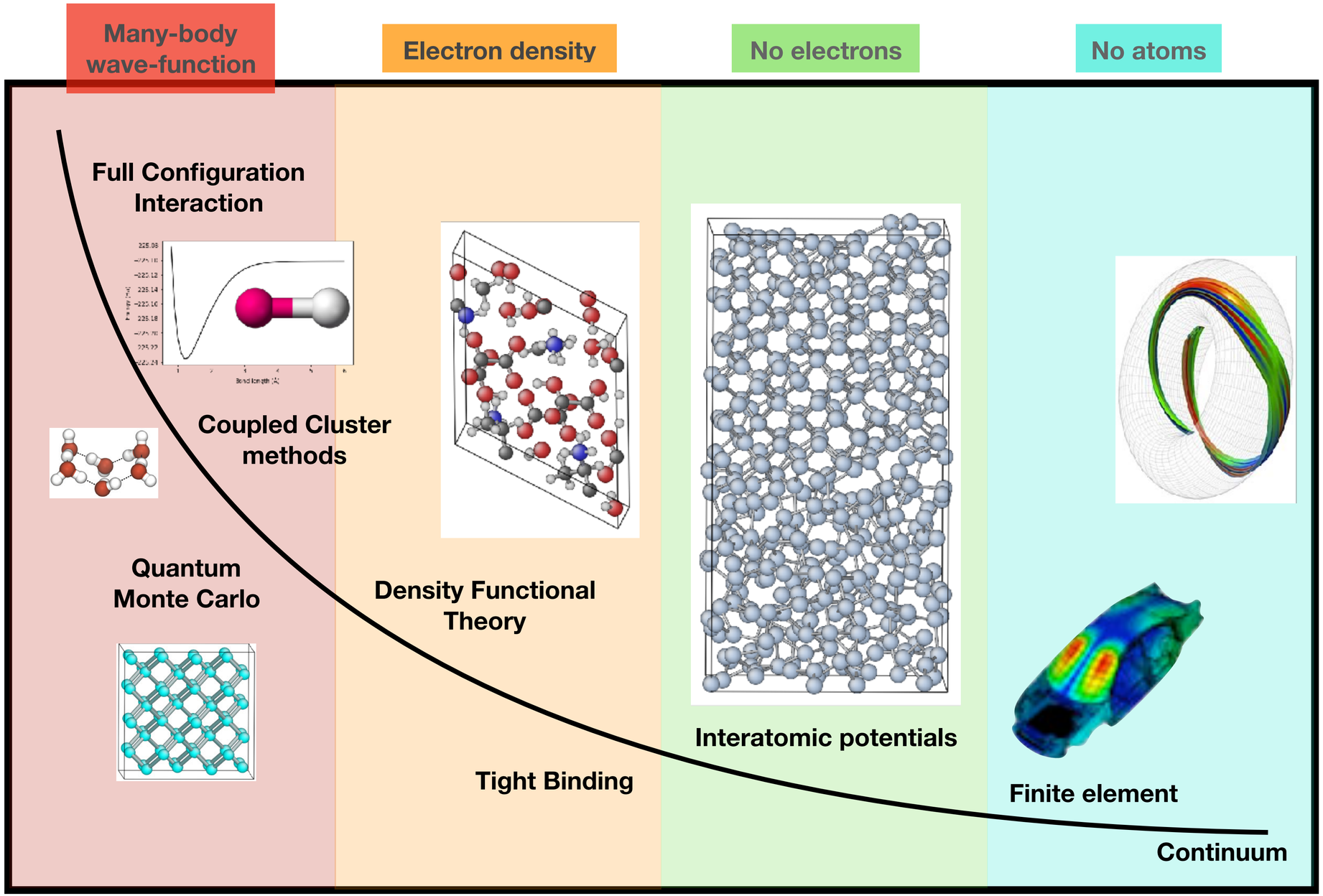
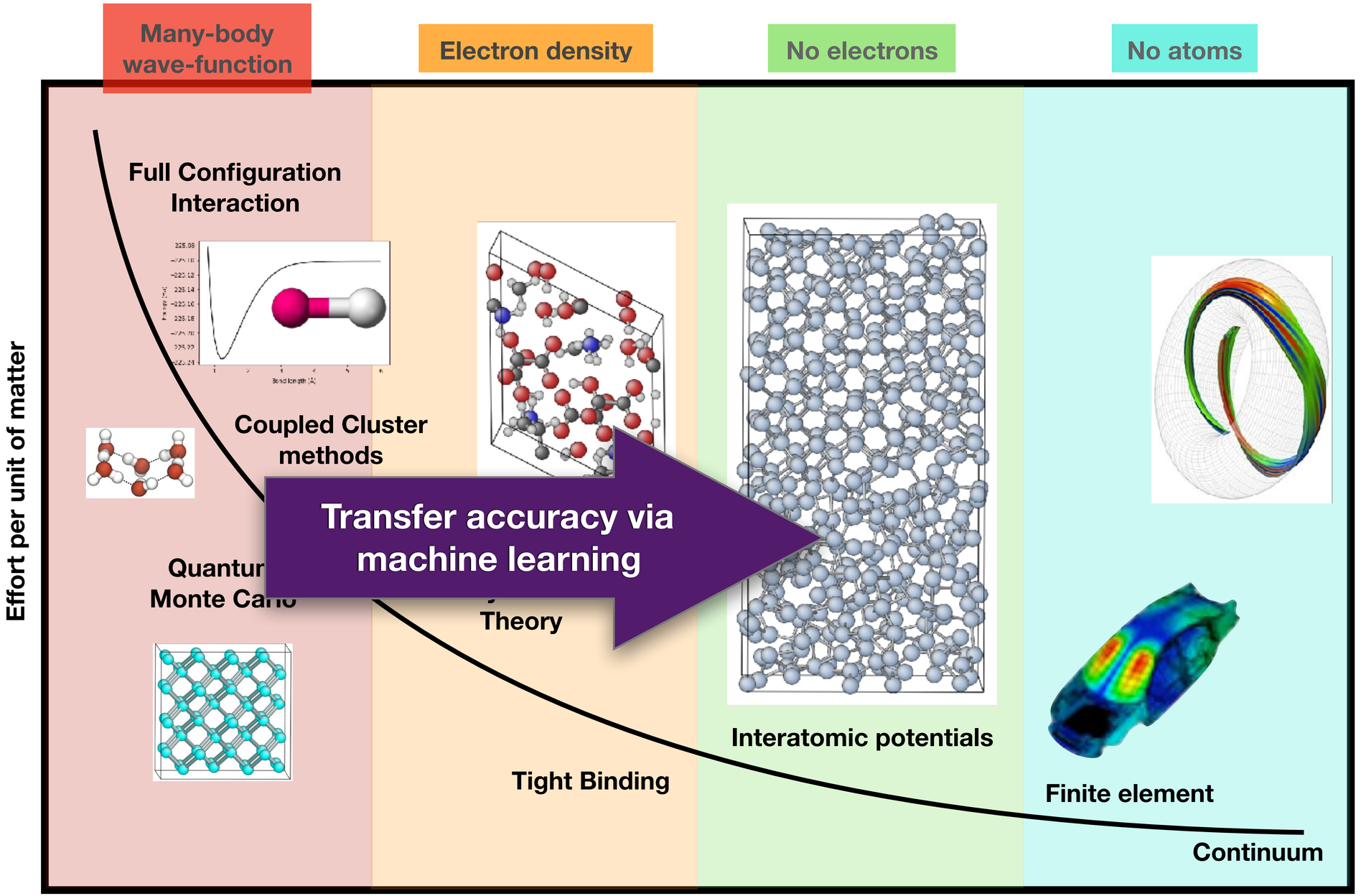
# Learning atomic and electronic interactions from microscopic observables

Albert Bartók-Pártay

Department of Physics
School of Engineering

WARWICK
THE UNIVERSITY OF WARWICK


Many-body wave-function
Electron density
No electrons
No atoms
Full Configuration Interaction
Coupled Cluster methods
Quantum Monte Carlo
Density Functional Theory
Tight Binding
Interatomic potentials
Finite element
Continuum
Effort per unit of matter
Level of coarse graining

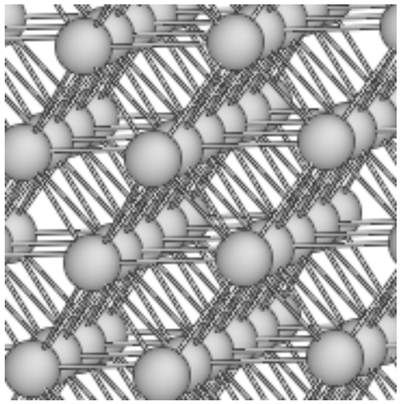

Many-body wave-function
Electron density
No electrons
No atoms
Effort per unit of matter
Full Configuration Interaction
Coupled Cluster methods
Quantur
Monte Carlo
Theory
Transfer accuracy via machine learning
Tight Binding
Interatomic potentials
Finite element
Continuum
Level of coarse graining

# Interatomic potentials

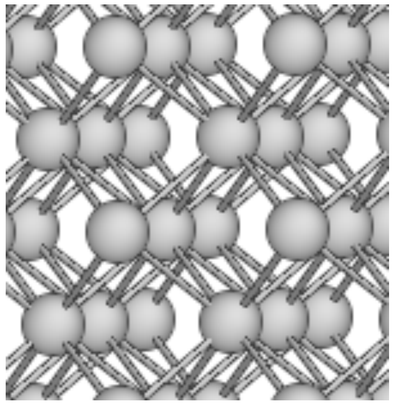
$$\varepsilon_{ij} = 4\epsilon\left[\left(\frac{\sigma}{r_{ij}}\right)^{12} - \left(\frac{\sigma}{r_{ij}}\right)^{6}\right]$$

$$\varepsilon_i = f\left[\sum_j \phi(r_{ij})\right]$$

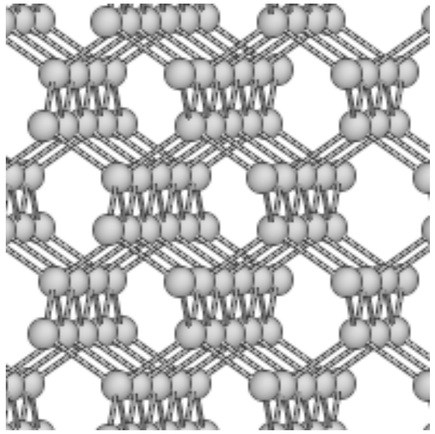
$$\varepsilon_i = \sum_j f_2(r_{ij}) + \sum_{jk} f_3(r_{ij}, r_{jk}, \theta_{ijk})$$



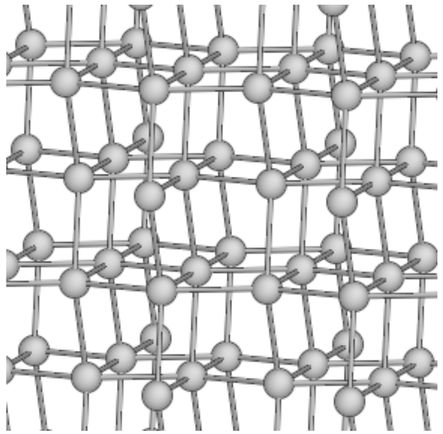
fcc



bcc



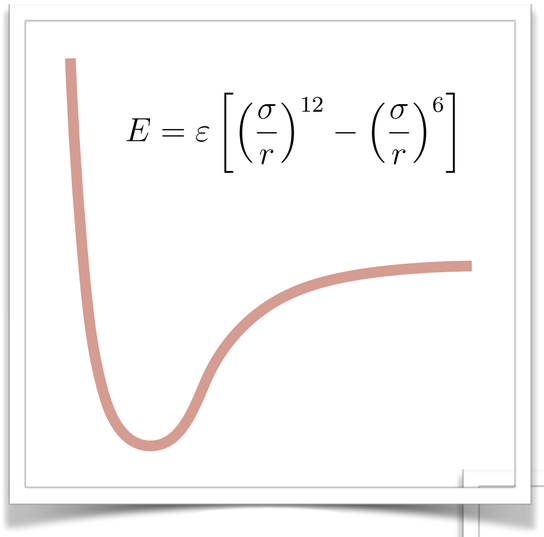
diamond

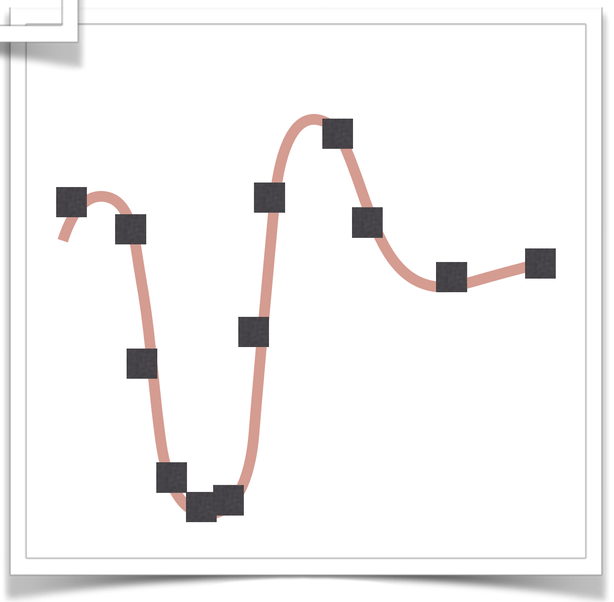


β-tin

# Interatomic potentials

- Analytic potentials
  - fixed functional formula
  - based on physical understanding
  - few fixed parameters
  - fit to experimental and/or computational data
- Machine learning potentials
  - flexible functional form
  - no physical motivation
  - data driven - mostly computational



# Potential Energy Surface fitting

$$E_{\text{total}} = \sum_{i}^{\text{atoms}} \varepsilon(\text{neighbourhood}_i) + \text{long range}$$

$$E_{\text{total}} = \sum_{i}^{\text{atoms}} \varepsilon(\{\mathbf{r}_{ij}\}_{j}^{\text{neighbours}_i})$$

Fit to electronic structure data!

Representations:

$$\{\mathbf{r}_{ij}\}_{j}^{\text{neighbours}_i}$$

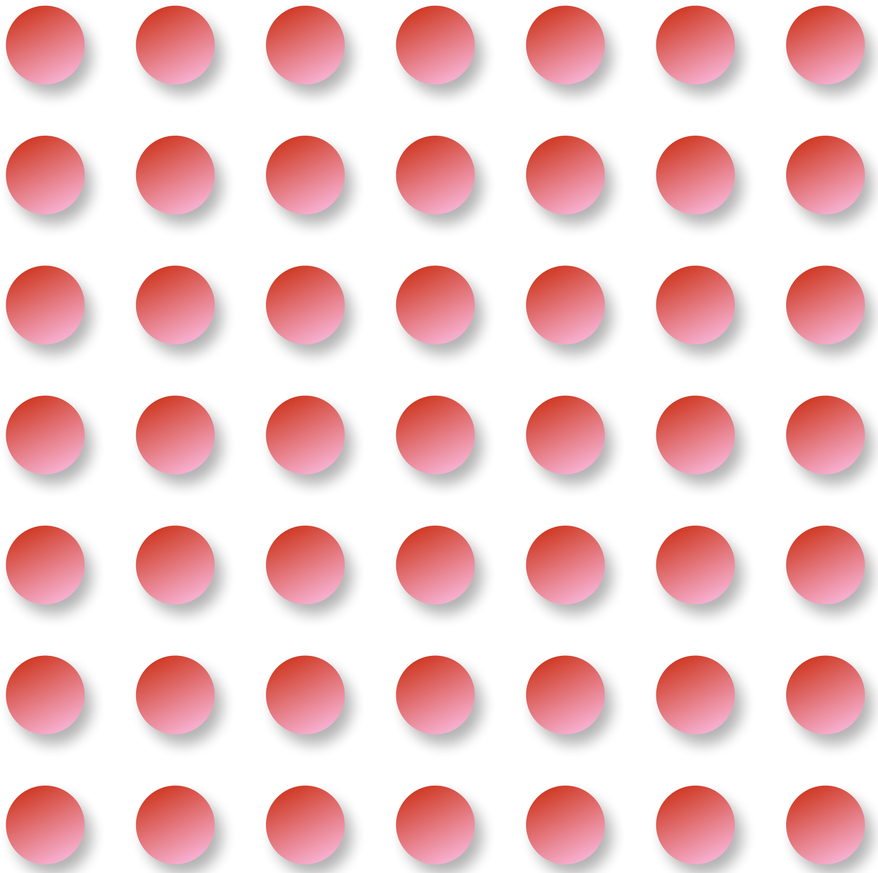
- Parrinello-Behler symmetry functions
- SO(4) bispectrum
- Coulomb matrix
- Smooth Overlap of Atomic Positions

Regression methods:

- Linear regression $\varepsilon()$
  - SNAP
- Neural networks
  - Parrinello-Behler
  - ANI
- Gaussian Process
  - (s)GDML
  - GAP

Database

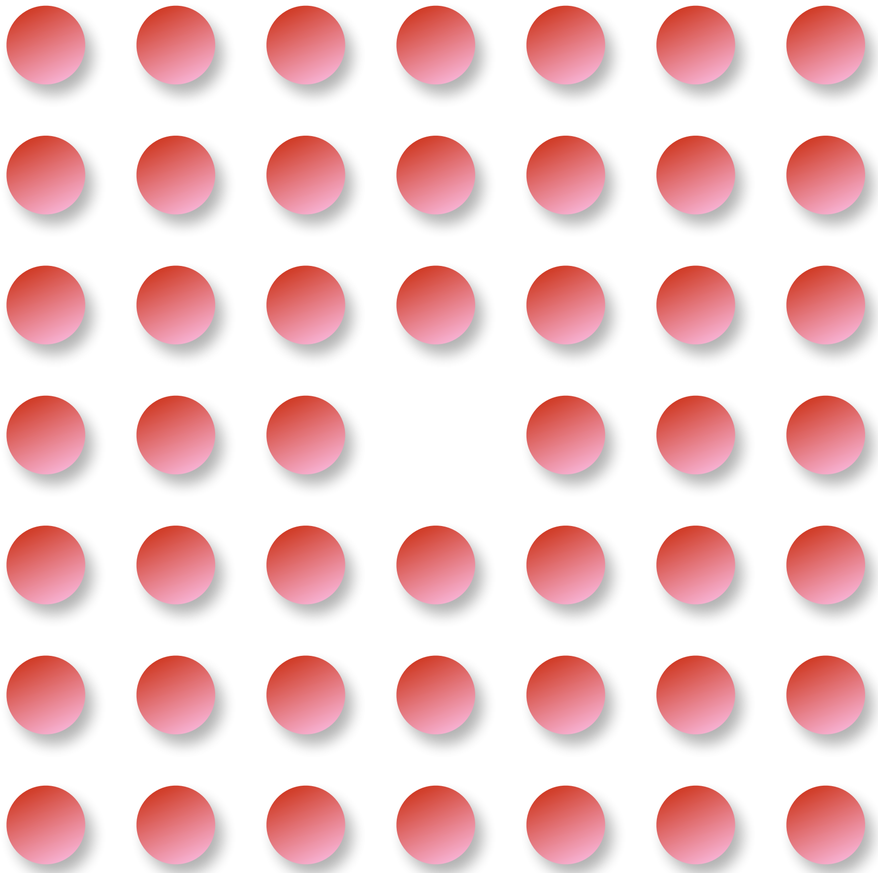
# ~~How~~ Why do ML potentials work?



**Toy example:**
**vacancy formation**

- **start with bulk**

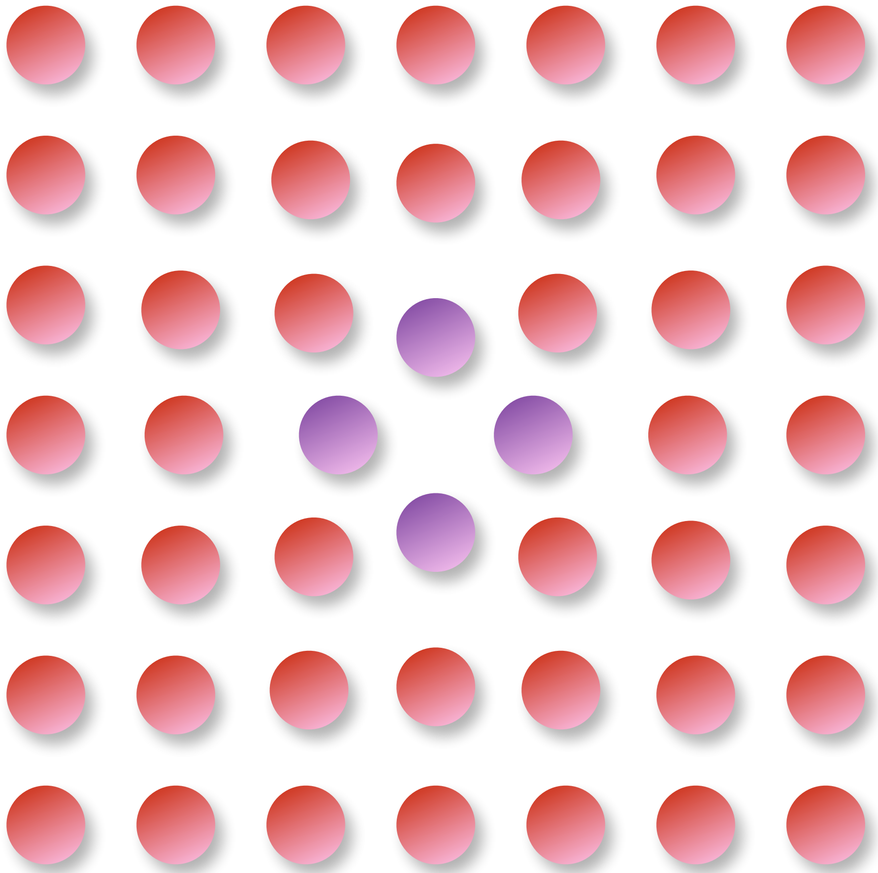
# ~~How~~ Why do ML potentials work?



**Toy example:**
**vacancy formation**

- **start with bulk**
- **remove an atom**

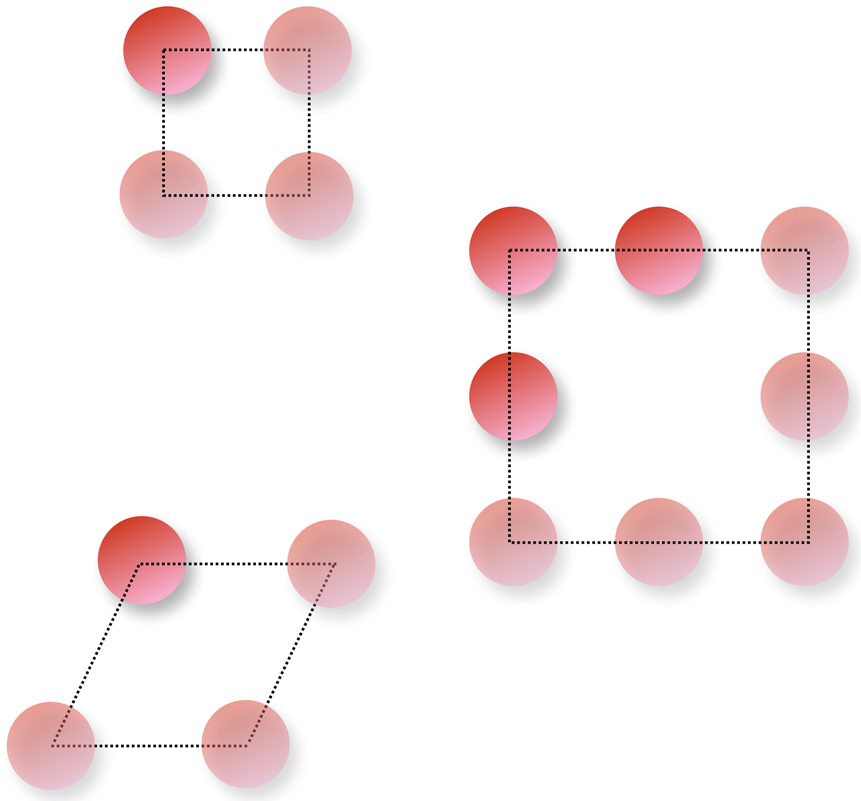
# ~~How~~ Why do ML potentials work?



**Toy example:**
**vacancy formation**

- **start with bulk**
- **remove an atom**
- **relax**

# ~~How~~ Why do ML potentials work?



**Toy example: vacancy formation**

- start with bulk
- remove an atom
- relax

↓

**Unified, atomic level model of vacancy formation**

# Gaussian Process Regression

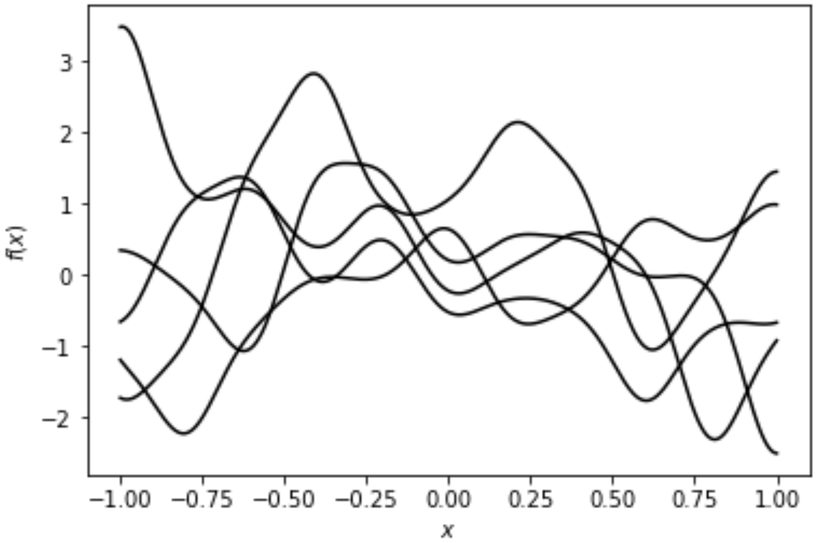
WARWICK
THE UNIVERSITY OF WARWICK

**Gaussian Process:**

$$f_i = f(\mathbf{d}_i, \mathbf{w}) = \sum_h w_h \phi_h(\mathbf{d}_i)$$

$$P(\mathbf{w}) = \text{Normal}(\mathbf{w}; \mathbf{0}, \sigma_w \mathbf{I})$$

**Samples from a Gaussian Process:**




**Covariance of Gaussian Processes at *i* and *j***

$$\langle f_i \, f_j \rangle = \left\langle \sum_{hh'} w_h w_{h'} \phi_h(\mathbf{d}_i) \phi_{h'}(\mathbf{d}_j) \right\rangle = \sigma_w^2 \sum_h \phi_h(\mathbf{d}_i) \phi_h(\mathbf{d}_j)$$

$$C(\mathbf{d}, \mathbf{d}') \equiv \sigma_w^2 \sum_h \phi_h(\mathbf{d}) \phi_h(\mathbf{d}')$$

**We don't need the basis functions any more!**

# Gaussian Process Regression

**Gaussian Process with noisy observations:**

$$f_i = f(\mathbf{d}_i, \mathbf{w}) + \varepsilon_i = \sum_h w_h \phi_h(\mathbf{d}_i) + \varepsilon_i$$

$$P(\boldsymbol{\varepsilon}) = \text{Normal}(\boldsymbol{\varepsilon}; \mathbf{0}, \lambda^2 \mathbf{I})$$

↓ **for *N* observations**

$$P(\mathbf{f}) = \text{Normal}(\mathbf{f}; \mathbf{0}, \mathbf{C} + \lambda^2 \mathbf{I}) \propto \exp\left(-\frac{1}{2}\mathbf{f}^T \mathbf{K}^{-1} \mathbf{f}\right) \quad \textbf{with} \quad \mathbf{K} = \mathbf{C} + \lambda^2 \mathbf{I}$$

$$C_{ij} = C(\mathbf{d}_i, \mathbf{d}_j)$$

**How to use this for regression?**

$$P(f|\mathbf{f}) \propto \exp\left(-\frac{1}{2}\mathbf{f}_{N+1}^T \mathbf{K}_{N+1}^{-1} \mathbf{f}_{N+1}\right) \quad \textbf{with} \quad \mathbf{K}_{N+1} = \begin{bmatrix} \mathbf{C}_N + \lambda^2 \mathbf{I} & \mathbf{c} \\ \mathbf{c}^T & c + \lambda^2 \end{bmatrix}$$

$$c_i = C(\mathbf{d}_i, \mathbf{d}_{N+1})$$

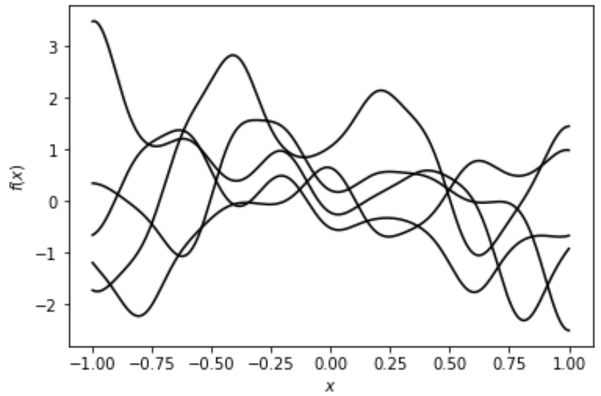
$$c = C(\mathbf{d}_{N+1}, \mathbf{d}_{N+1})$$

**Mean and variance of *f* calculated analytically:**

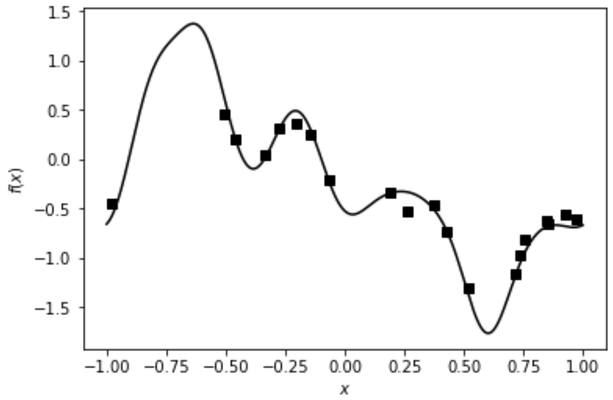
$$\hat{f} = \mathbf{c}^T \mathbf{K}^{-1} \mathbf{f}$$

$$\text{var}(f) = c + \lambda^2 - \mathbf{c}^T \mathbf{K}^{-1} \mathbf{c}$$

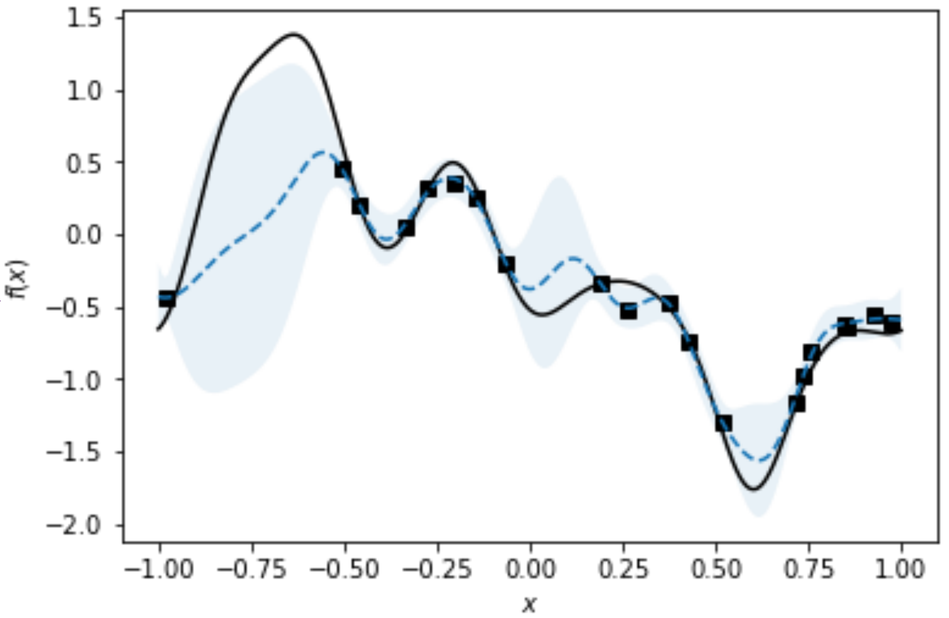
# Gaussian Process Regression




prior (set of functions)




data with noise




prediction with error estimates

# Interatomic Potentials from Machine Learning

WARWICK
THE UNIVERSITY OF WARWICK

- model: total energy is a sum of atomic many-body energies
- result: covariance of total energies is sum up atomic covariance functions
- we have just defined a covariance function for total energies!

**Gaussian Process model**

$$\varepsilon_i = \varepsilon(\mathbf{d}_i, \mathbf{w}) = \sum_h w_h \phi_h(\mathbf{d}_i)$$

$$E = \sum_i \varepsilon_i$$

+ long range

**Quantum Mechanical observable (total energy)**

$$\langle E_N E_M \rangle = \left\langle \sum_{i\in N} \varepsilon(\mathbf{d}_i) \sum_{j\in M} \varepsilon(\mathbf{d}_j) \right\rangle = \left\langle \sum_{i\in N} \sum_{j\in M} \sum_{hh'} w_h w_{h'} \phi_h(\mathbf{d}_i) \phi_{h'}(\mathbf{d}_j) \right\rangle =$$

$$\sum_{i\in N} \sum_{j\in M} \sum_{hh'} \langle w_h w_{h'} \rangle \phi_h(\mathbf{d}_i) \phi_{h'}(\mathbf{d}_j) = \sigma_w^2 \sum_{i\in N} \sum_{j\in M} \sum_{h} \phi_h(\mathbf{d}_i) \phi_h(\mathbf{d}_j) = \sigma_w^2 \sum_{i\in N} \sum_{j\in M} C(\mathbf{d}_i, \mathbf{d}_j)$$

# Interatomic Potentials from Machine Learning

WARWICK
THE UNIVERSITY OF WARWICK

- compute the covariance between energies and forces/ virials
- or compute the covariances between forces/virials
- new covariance functions of the same total energy model!

$$\left\langle \frac{\partial E_N}{\partial \xi_k} E_M \right\rangle = \frac{\partial \langle E_N E_M \rangle}{\partial \xi_k} = \sigma_w^2 \sum_{i \in N} \sum_{j \in M} \nabla_{\mathbf{d}_i} C(\mathbf{d}_i, \mathbf{d}_j) \cdot \frac{\partial \mathbf{d}_i}{\partial \xi_k}$$

**Cartesian coordinate or cell deformation**

**Further quantum mechanical observables: forces, stresses**

$$\left\langle \frac{\partial E_N}{\partial \xi_k} \frac{\partial E_M}{\partial \chi_l} \right\rangle = \frac{\partial^2 \langle E_N E_M \rangle}{\partial \xi_k \partial \chi_l} = \sigma_w^2 \sum_{i \in N} \sum_{j \in M} \frac{\partial \mathbf{d}_i^\top}{\partial \xi_k} (\nabla_{\mathbf{d}_i} C(\mathbf{d}_i, \mathbf{d}_j) \nabla_{\mathbf{d}_j}^\top) \frac{\partial \mathbf{d}_j}{\partial \chi_l}$$

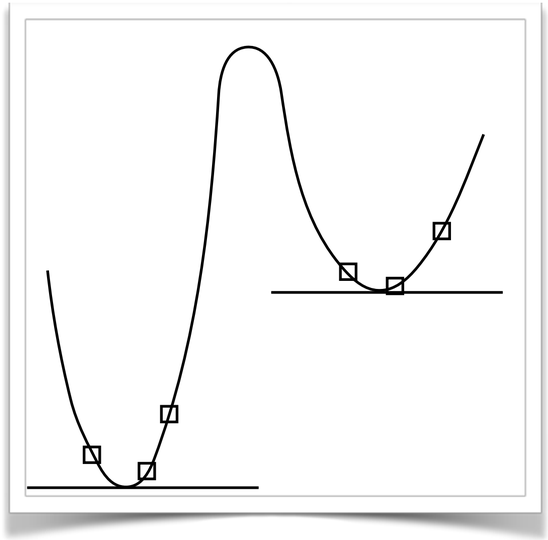
E Solak *et al*, NIPS **15**, 529 (2003)

# Interatomic Potentials from Machine Learning

- Generate database of relevant configurations
- Compute QM observables (energy, force, stress)
- forces provide rich information on PES
- energies set scale and connect minima correctly
- viral stresses capture the really soft response to deformation

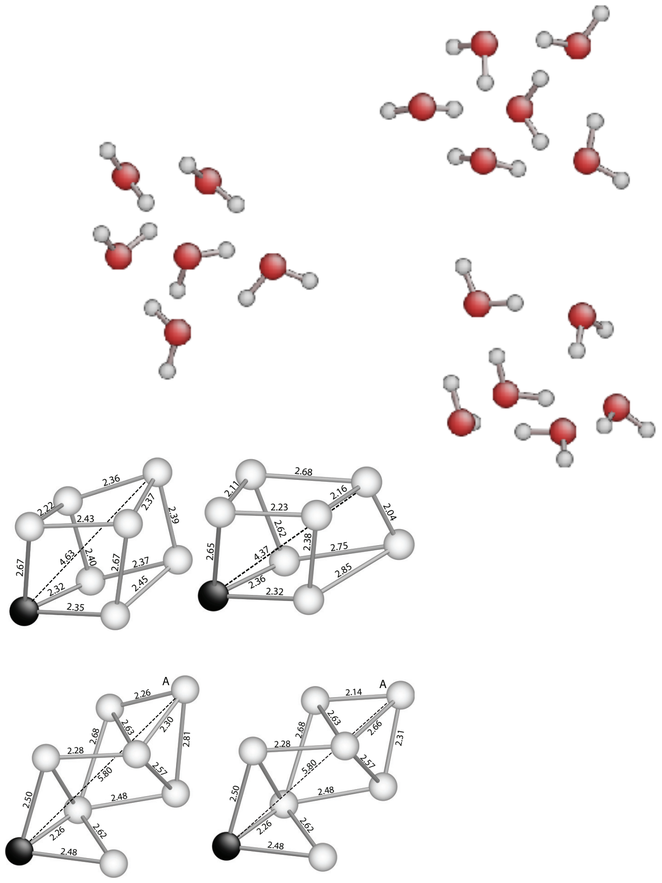
| Structure type | No. atoms | No. structures | No. environments | No. representative atoms |
|---|---|---|---|---|
| Isolated atom | 1 | 1 | 1 | 1 |
| Diamond | 2 | 104 | 208 | 6 |
| | 16 | 220 | 3520 | 53 |
| | 54 | 110 | 5940 | 58 |
| | 128 | 55 | 7040 | 92 |
| $\beta$-Sn | 2 | 60 | 120 | 32 |
| | 16 | 220 | 3520 | 51 |
| | 54 | 110 | 5940 | 66 |
| | 128 | 55 | 7040 | 157 |
| Simple hexagonal | 1 | 110 | 110 | 13 |
| | 8 | 30 | 240 | 15 |
| | 27 | 30 | 810 | 42 |
| | 64 | 53 | 3392 | 89 |
| Hexagonal diamond | 4 | 49 | 196 | 7 |
| bcc | 2 | 49 | 98 | 40 |
| bc8 | 8 | 49 | 392 | 66 |
| fcc | 4 | 49 | 196 | 46 |
| hcp | 2 | 49 | 98 | 28 |
| st12 | 12 | 49 | 588 | 94 |
| Liquid | 64 | 69 | 4416 | 1114 |
| | 128 | 7 | 896 | 323 |

AP Bartók *et al*, PRX **8**, 041048 (2018)



# Representing atomic environments

$$\{\mathbf{r}_{ij}\}_j^{\text{neighbours}_i} \longrightarrow (d_1, d_2, \ldots, d_k)$$

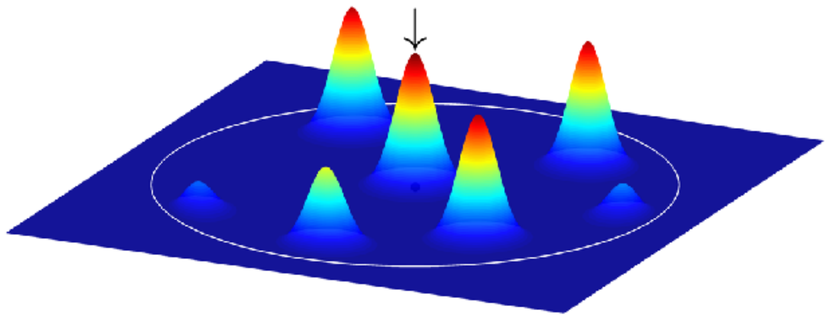


Represent a local neighbourhood configuration: “descriptor”, “fingerprint”, “feature vector”, “symmetry function”

- Rotational, reflectional, translational and permutational invariance
- Faithfulness - no two different configuration give the same representation
- Continuous, differentiable and smooth

# Smooth Overlap of Atomic Positions

- represent atomic environment by a sum of Gaussians
- similarity is overlap of environment densities
- permuational invariance satisfied (it's a sum!)
- rotational invariance obtained by integrating over all rotations



$$\rho_i(\mathbf{r}) \equiv \sum_j^{\text{neigh.}} \exp\left(-\frac{|\mathbf{r}-\mathbf{r}_{ij}|^2}{2\sigma_{\text{atom}}^2}\right) \rightarrow S(\rho,\rho') = \int \rho(\mathbf{r})\rho'(\mathbf{r})d\mathbf{r} \rightarrow k(\rho,\rho') = \int |S(\rho,\hat{R}\rho')|^n d\hat{R}$$

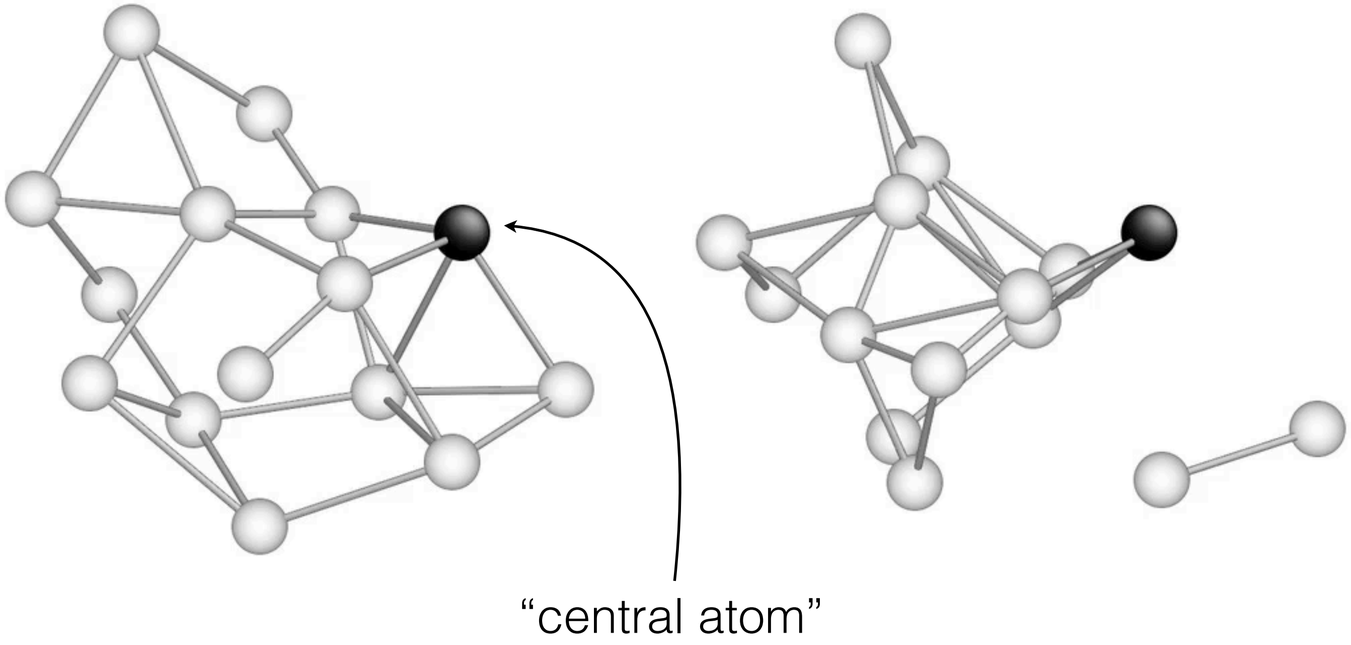
$$= \int d\hat{R} \left| \int \rho(\mathbf{r})\rho'(\hat{R}\mathbf{r})d\mathbf{r} \right|^n$$

$$\rho_i(\mathbf{r}) = \sum c_{nlm} g_n(r) Y_{lm}(\hat{\mathbf{r}}) \rightarrow S(\rho,\hat{R}\rho') = \sum_{l,m,m'} I^l_{mm'}[\rho,\rho'] D^l_{mm'}(\hat{R})$$

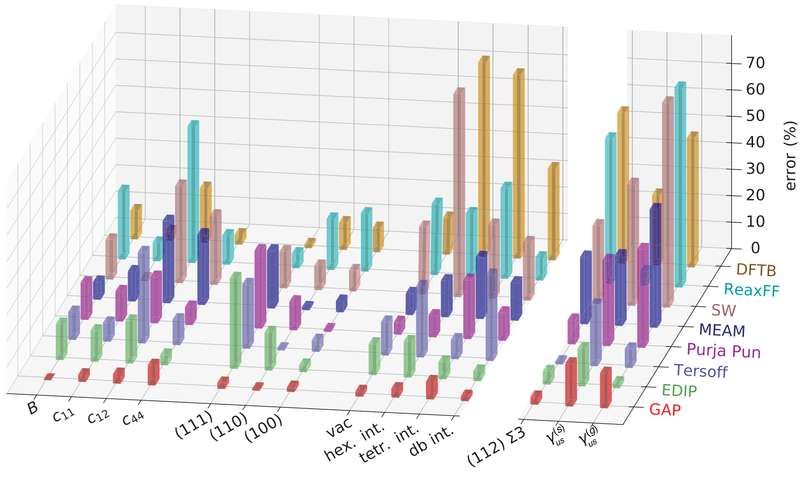
- basis set expansion
- overlap transforms with the Wigner D-matrix
- Wigner D-matrices orthogonal functions
- rotational integral becomes *analytic*!
- rotational invariant similarity obtained as a *dot-product*

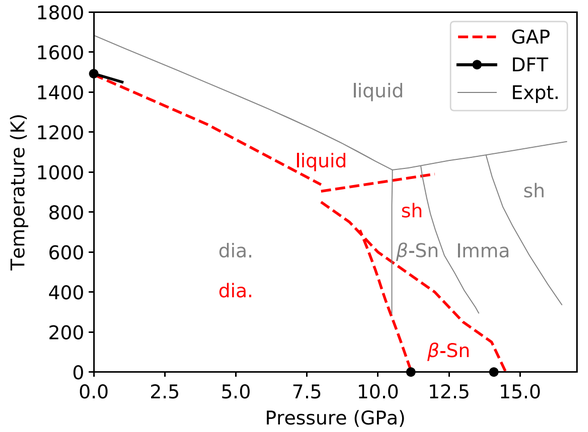
$$k(\rho,\rho') = \sum_{l,n,n'} p_{lnn'} p'_{lnn'}$$

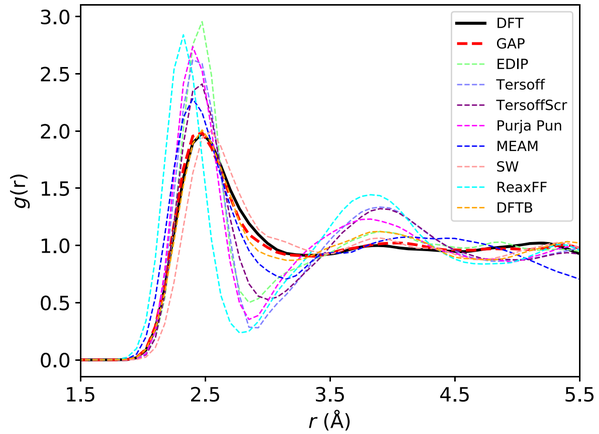
# Smooth Overlap of Atomic Positions

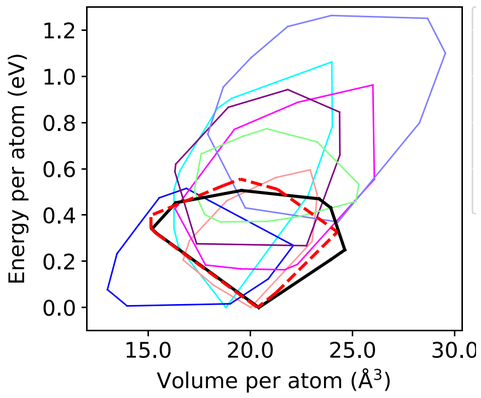


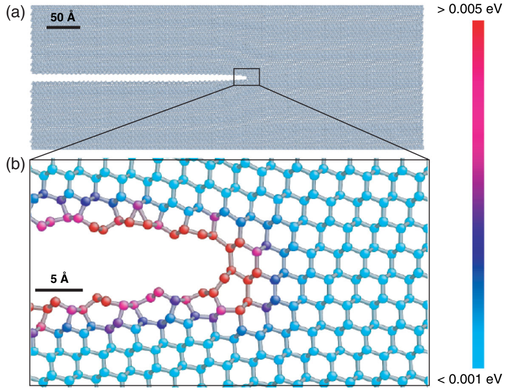
# Gaussian Approximation Potentials: Silicon

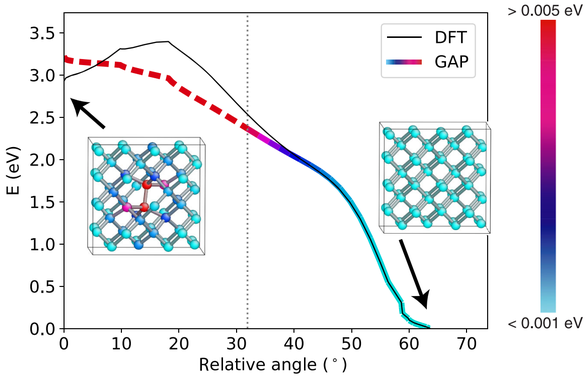


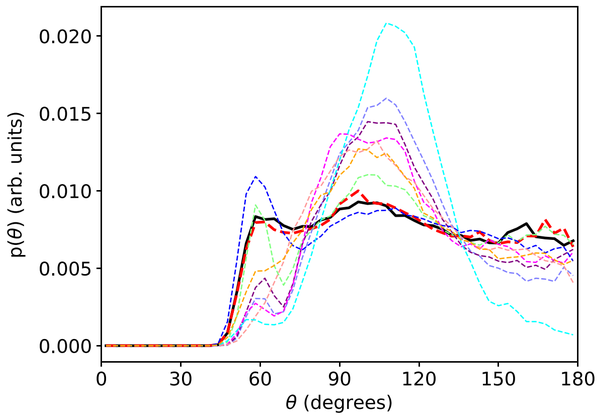



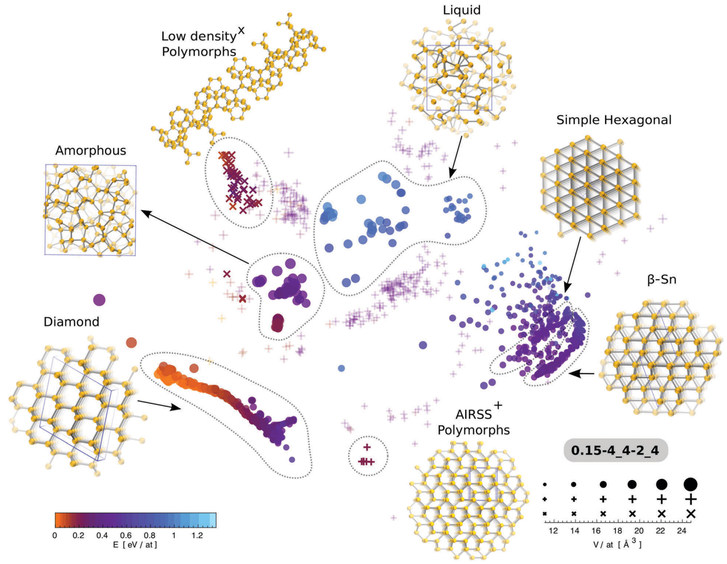



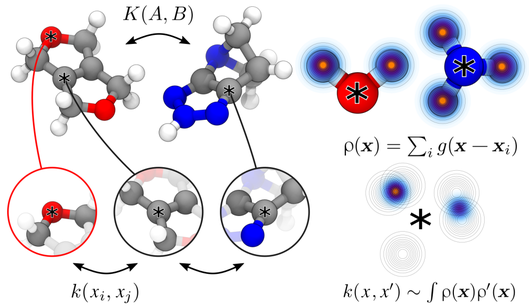



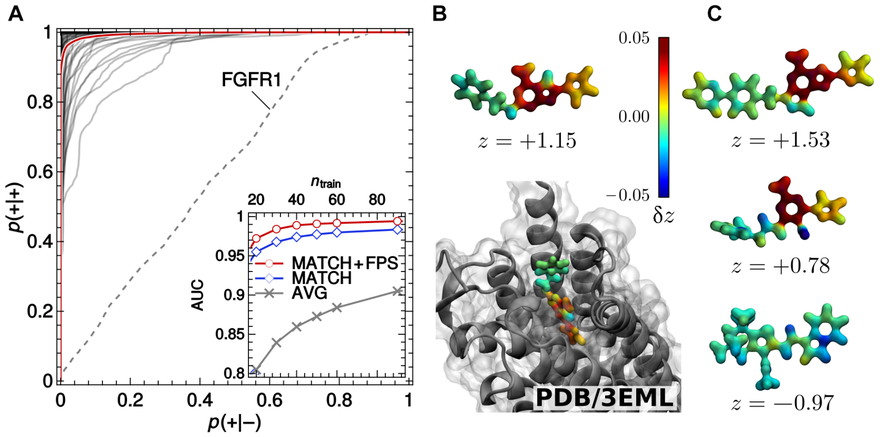

AP Bartók *et al*, PRX **8**, 041048 (2018)

# Similarity of atomic structures

- in collaboration with Michele Ceriotti
- SOAP: atomic similarity
- match atomic environments across molecules or structures
- compare complete structures





AP Bartók *et al*, Sci. Adv. **3**, e1701816 (2017)
S De *et al*, PCCP **18**, 13754 (2016)

# Conclusions

WARWICK
THE UNIVERSITY OF WARWICK

- Inform models based on higher level theory
- Interatomic potentials based on QM data
- Machine learning glues it together