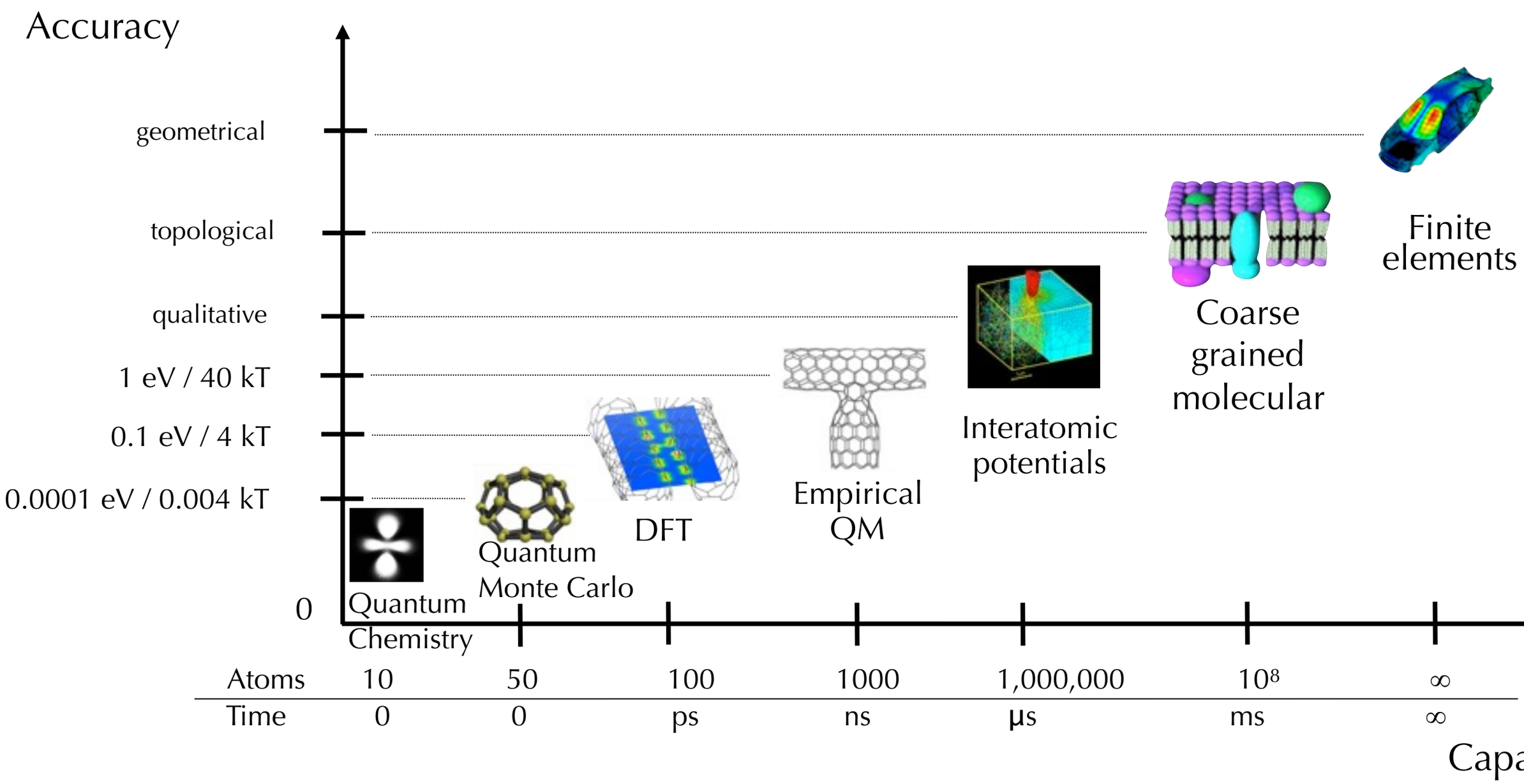
Representation and regression problems in molecular structure and dynamics

Gábor Csányi

Engineering Laboratory
UNIVERSITY OF CAMBRIDGE

Multiple scales of materials modelling



Multiple scales of materials modelling

Accuracy

geometrical

Quantum Mechanics

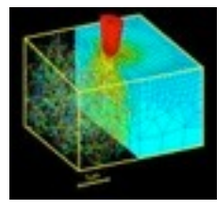
$$\hat{\mathscr{H}}\Psi = E\Psi$$

$\hat{\mathscr{H}}$ Nuclear positions, interaction of electrons and nuclei

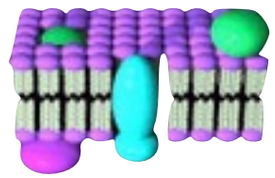
$\Psi(r_1 \ldots r_N)$ Wavefunction of electrons

Unknown 3N dimensional function, unique for each nuclear configuration

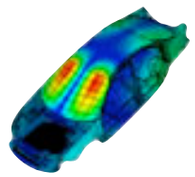
Empirical QM



Interatomic potentials



Coarse grained molecular

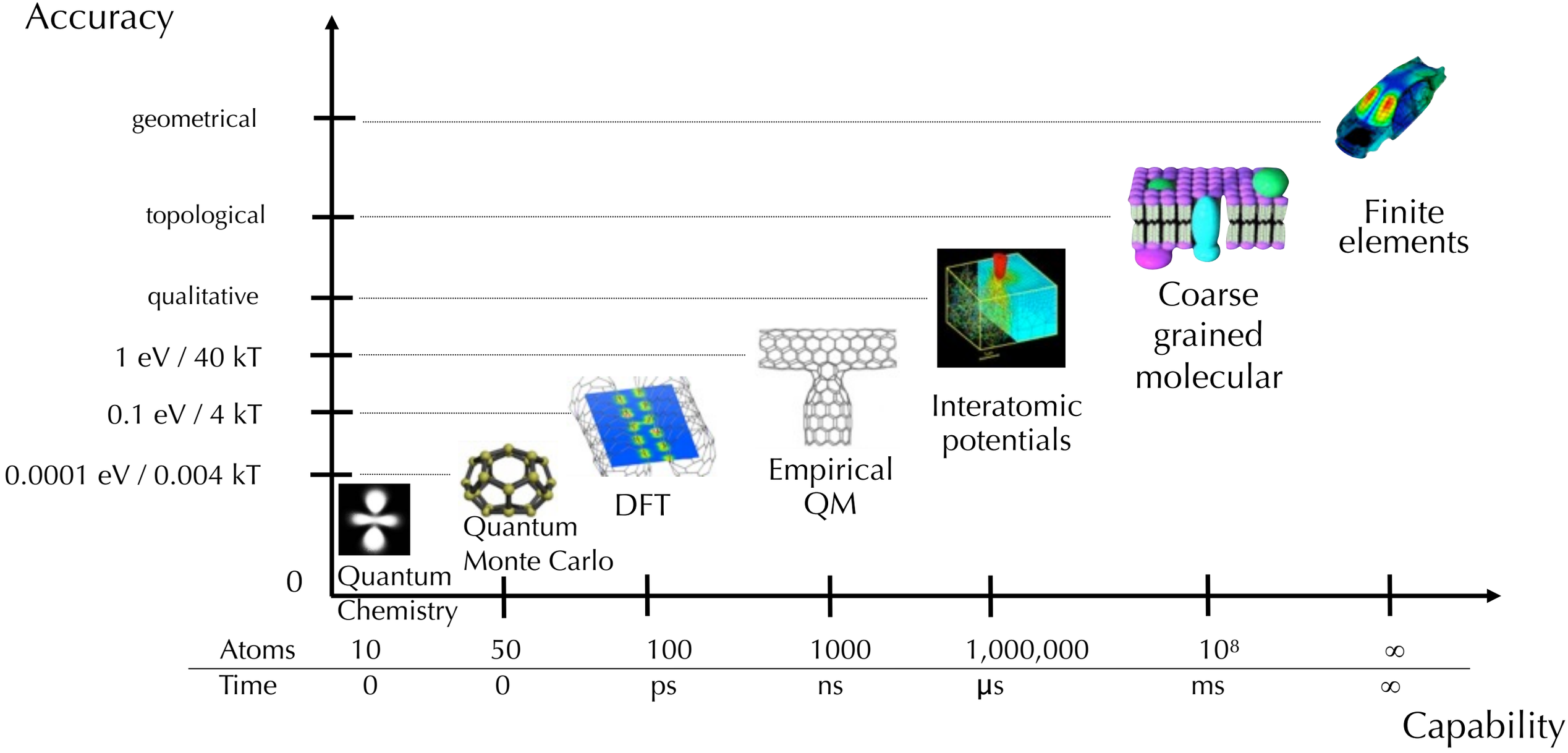


Finite elements

1000	1,000,000	10^8	∞
ns	µs	ms	∞

Capability

Multiple scales of materials modelling



Quantum Mechanics

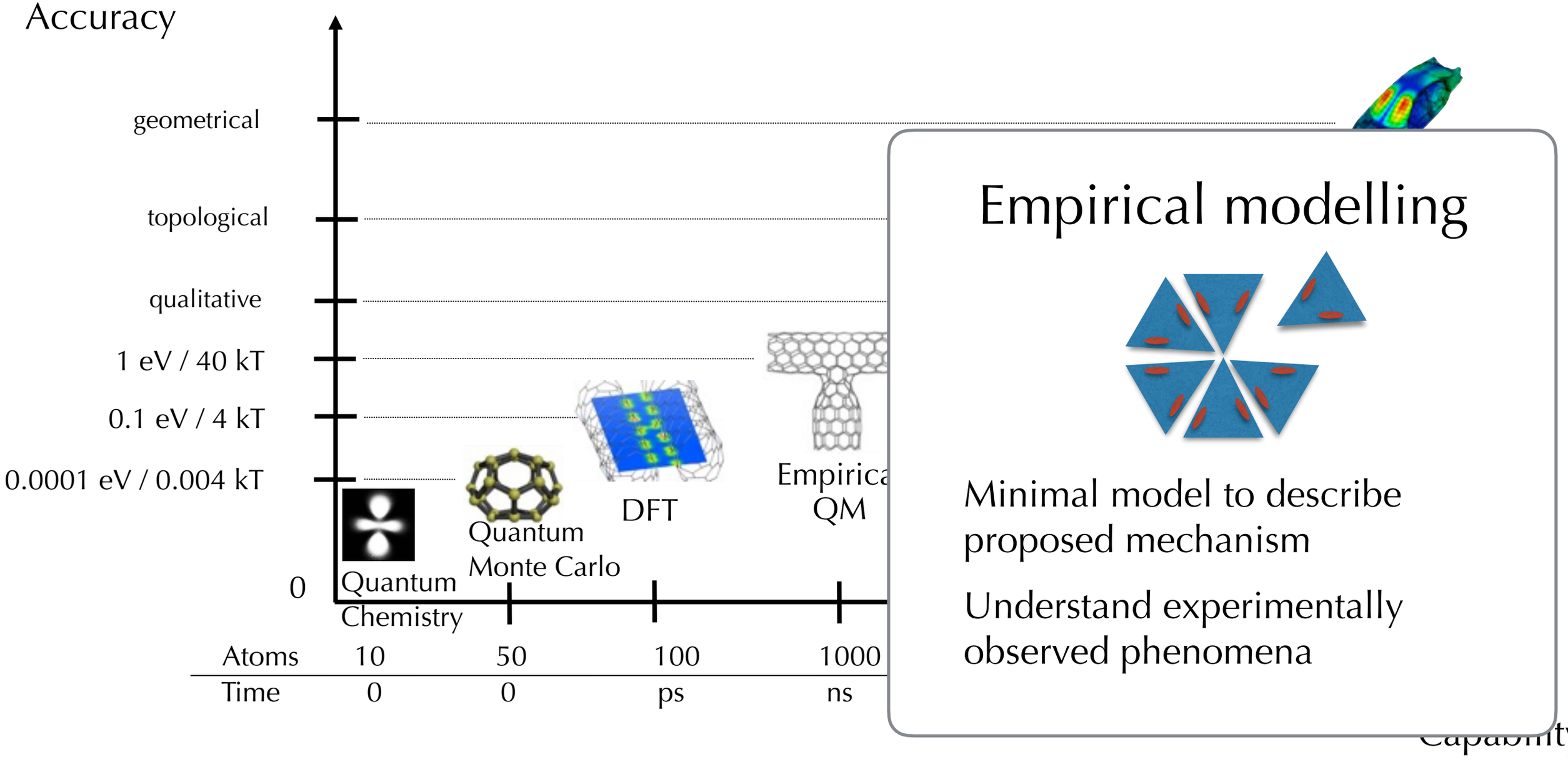
$\hat{\mathcal{H}}\Psi = E\Psi$

$\hat{\mathcal{H}}$ Nuclear positions, interaction of electrons and nuclei

$\Psi(r_1 \ldots r_N)$ Wavefunction of electrons

Unknown 3N dimensional function, unique for each nuclear configuration

Multiple scales of materials modelling



Empirical modelling

Minimal model to describe proposed mechanism

Understand experimentally observed phenomena

Quantum Mechanics

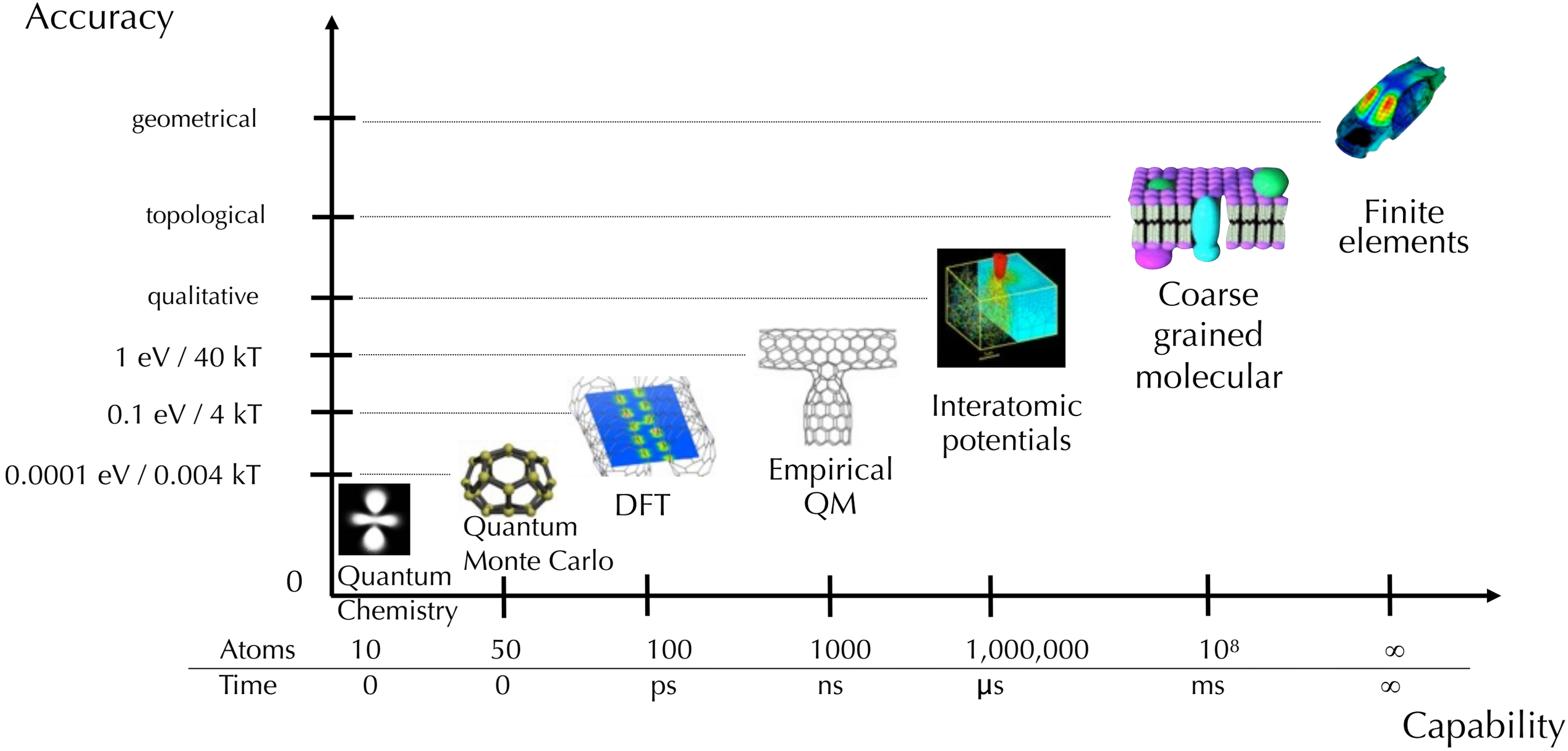
$$\hat{\mathcal{H}}\Psi = E\Psi$$

$\hat{\mathcal{H}}$ Nuclear positions, interaction of electrons and nuclei

$\Psi(r_1 \ldots r_N)$ Wavefunction of electrons

Unknown 3N dimensional function, unique for each nuclear configuration

Multiple scales of materials modelling



Quantum Mechanics

$\hat{\mathcal{H}}\Psi = E\Psi$

$\hat{\mathcal{H}}$ Nuclear positions, interaction of electrons and nuclei

$\Psi(r_1 \ldots r_N)$ Wavefunction of electrons

Unknown 3N dimensional function, unique for each nuclear configuration

Empirical modelling

Minimal model to describe proposed mechanism

Understand experimentally observed phenomena

Multiple scales of materials modelling

Accuracy

geometrical

topological

qualitative

1 eV / 40 kT

0.1 eV / 4 kT

0.0001 eV / 0.004 kT

0

Quantum Chemistry

Quantum Monte Ca

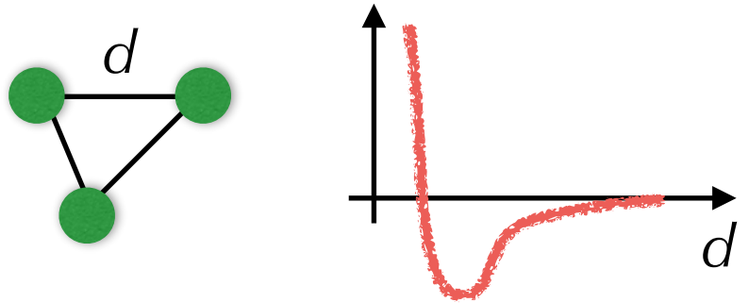
Finite elements

e

ed

ular

Force fields



3N positions of atoms → $E_{pot}(R)$

Newton's law → molecular dynamics

Structure stability and transformations

Atoms	10	50	∞
Time	0	0	∞

Capability

Quantum Mechanics

$\hat{\mathscr{H}}\Psi = E\Psi$

$\hat{\mathscr{H}}$ Nuclear positions, interaction of electrons and nuclei

$\Psi(r_1 \ldots r_N)$ Wavefunction of electrons

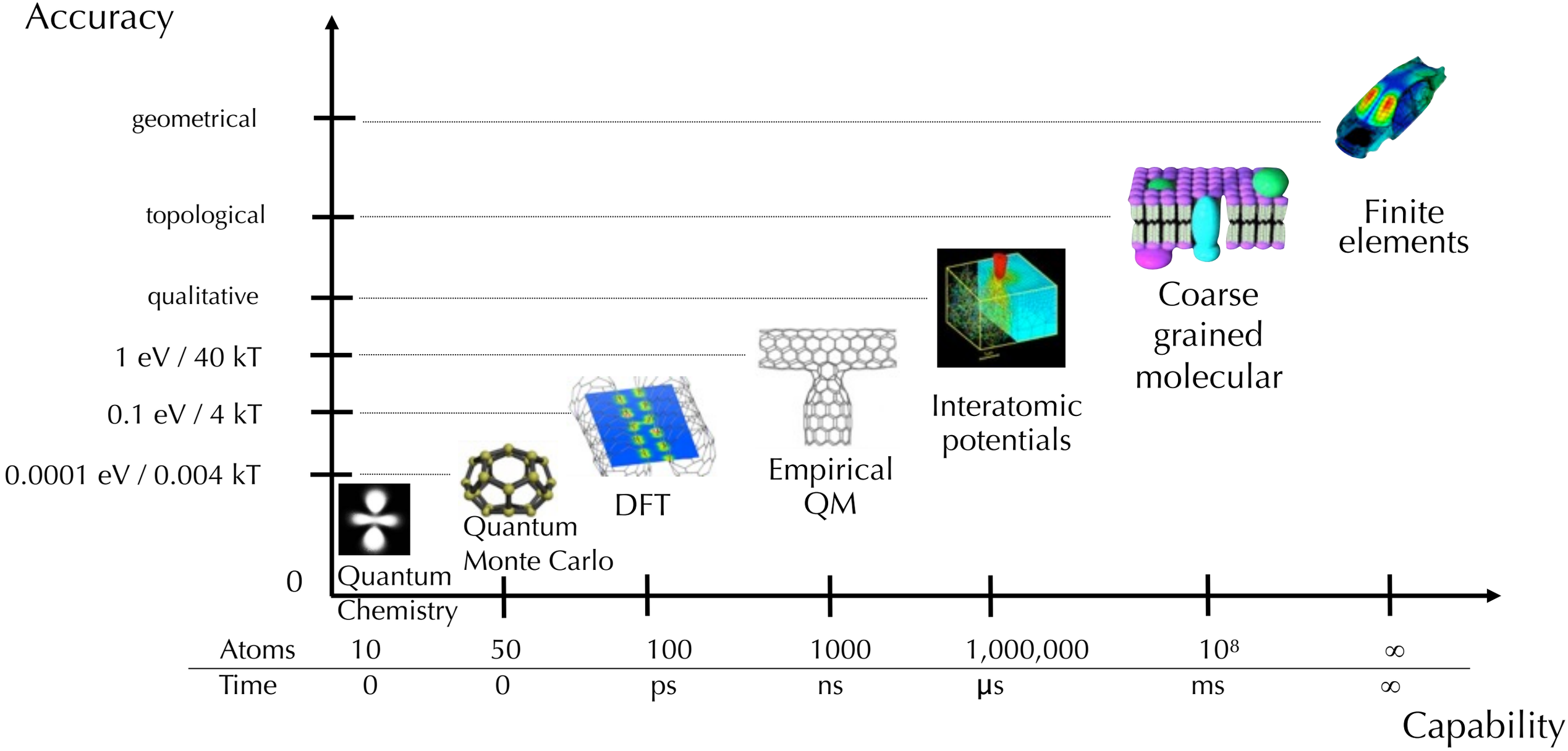
Unknown 3N dimensional function, unique for each nuclear configuration

Empirical modelling

Minimal model to describe proposed mechanism

Understand experimentally observed phenomena

Multiple scales of materials modelling



Quantum Mechanics

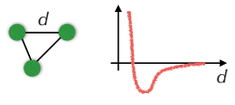
$\hat{\mathcal{H}}\Psi = E\Psi$

$\hat{\mathcal{H}}$ Nuclear positions, interaction of electrons and nuclei

$\Psi(r_1 \ldots r_N)$ Wavefunction of electrons

Unknown 3N dimensional function, unique for each nuclear configuration

Force fields



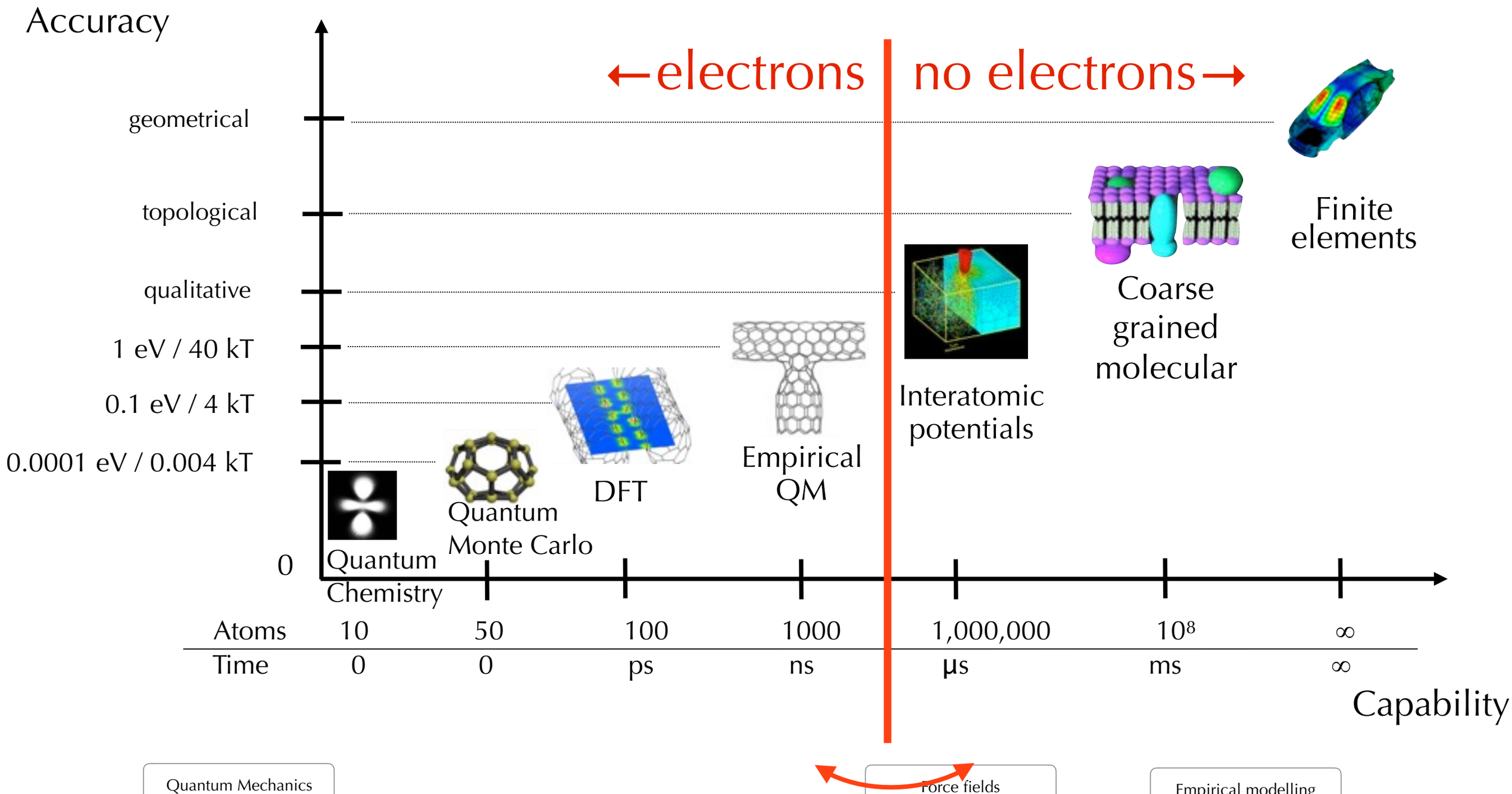
3N positions of atoms → $E_{pot}(R)$
Newton's law → molecular dynamics
Structure stability and transformations

Empirical modelling

Minimal model to describe proposed mechanism

Understand experimentally observed phenomena

Multiple scales of materials modelling



Quantum Mechanics

$\hat{\mathcal{H}}\Psi = E\Psi$

$\hat{\mathcal{H}}$ Nuclear positions, interaction of electrons and nuclei

$\Psi(r_1 \ldots r_N)$ Wavefunction of electrons

Unknown 3N dimensional function, unique for each nuclear configuration

GAP in accuracy and speed

Force fields

Newton's law → molecular dynamics

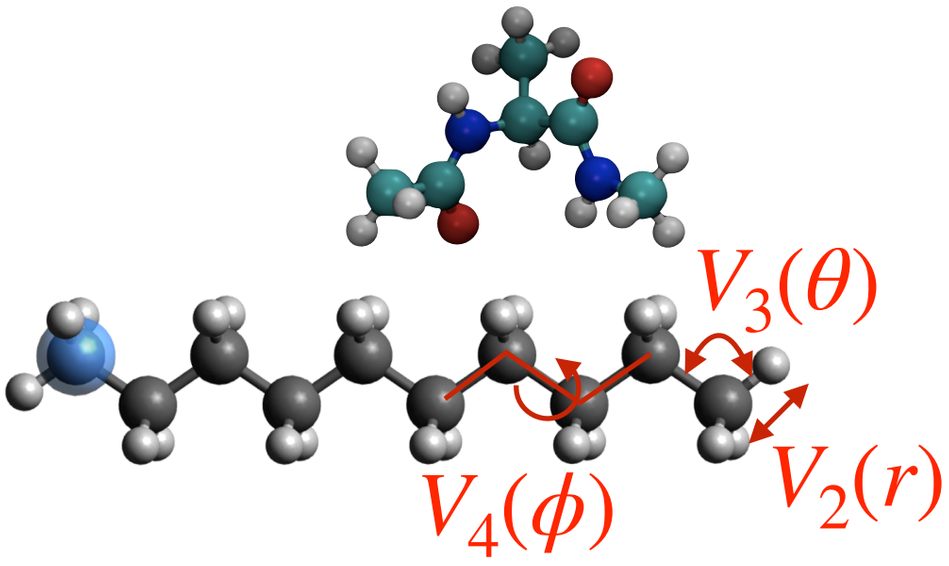
Structure stability and transformations

Empirical modelling

Minimal model to describe proposed mechanism

Understand experimentally observed phenomena

Interatomic potentials (force fields)



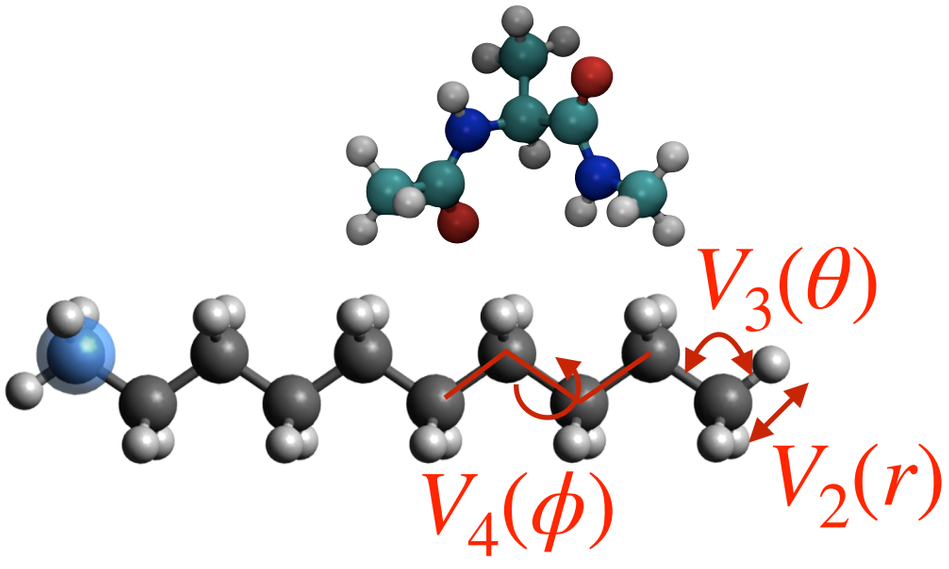
Transferability

organic force fields

- Accurate for small displacements
- Typically no reactions (some exceptions)
- Ultimate fit is to macroscopic experimental properties

AMBER
CHARMM
COMPASS
OPLS
...

Interatomic potentials (force fields)

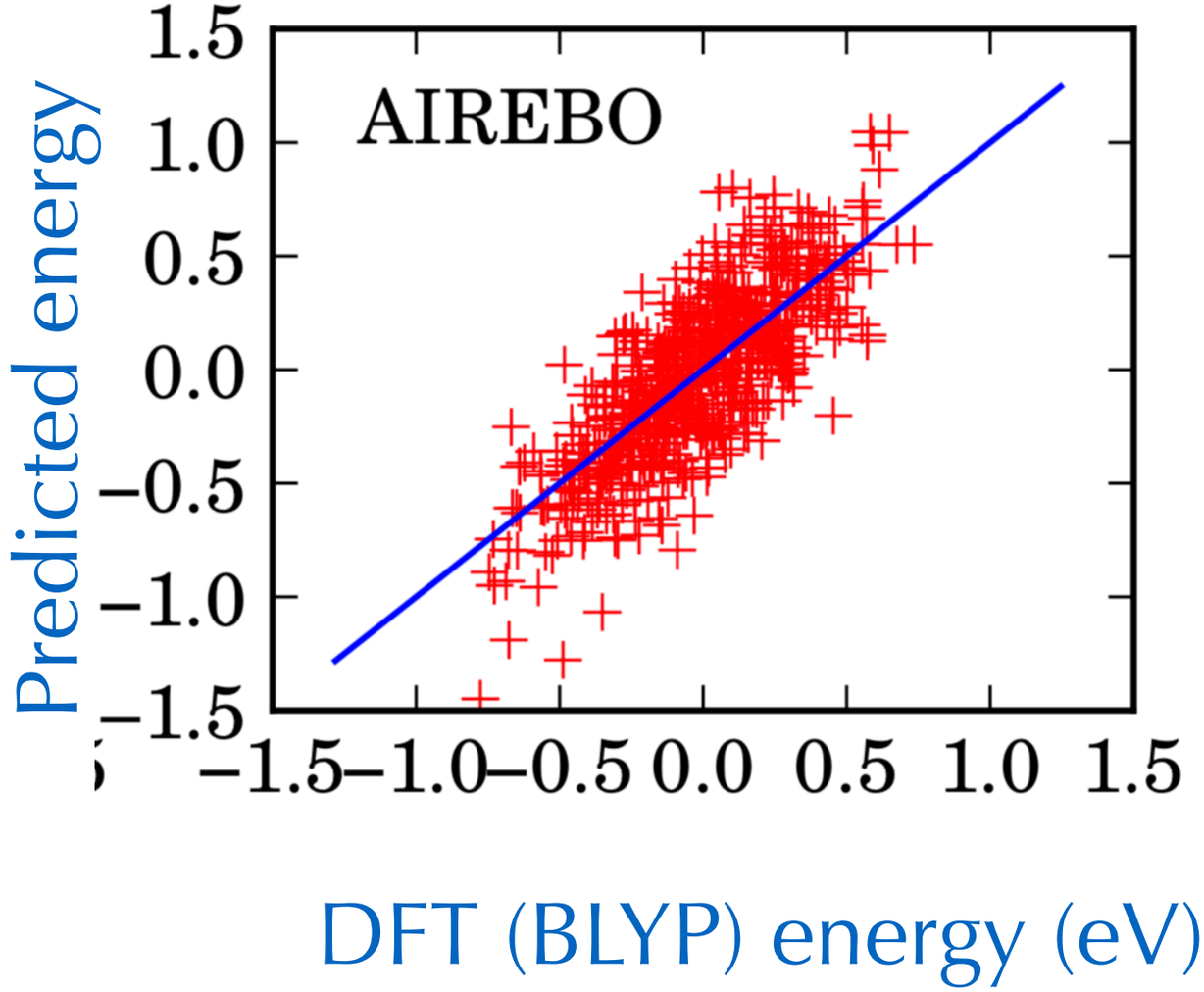


Transferability

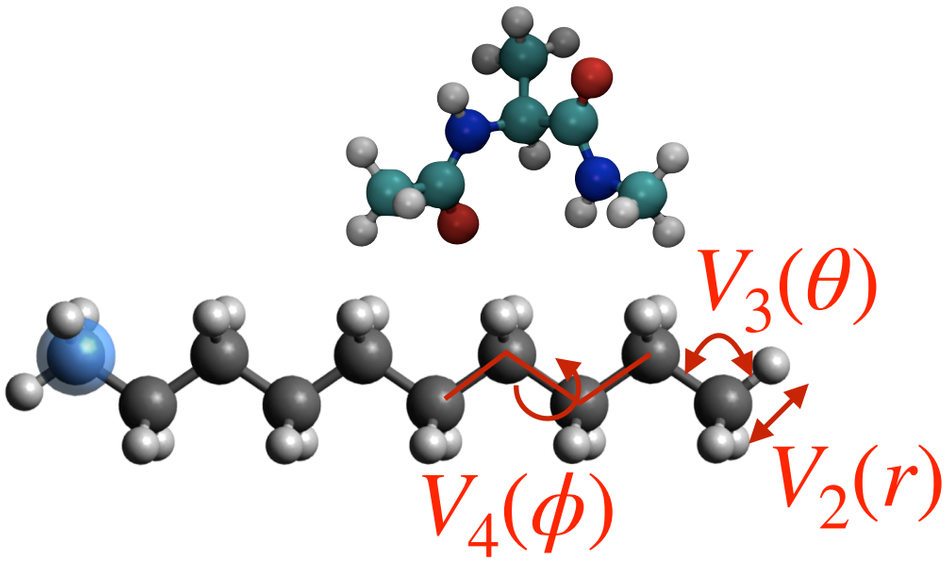
organic force fields

- Accurate for small displacements
- Typically no reactions (some exceptions)
- Ultimate fit is to macroscopic experimental properties

AMBER
CHARMM
COMPASS
OPLS
...



Interatomic potentials (force fields)



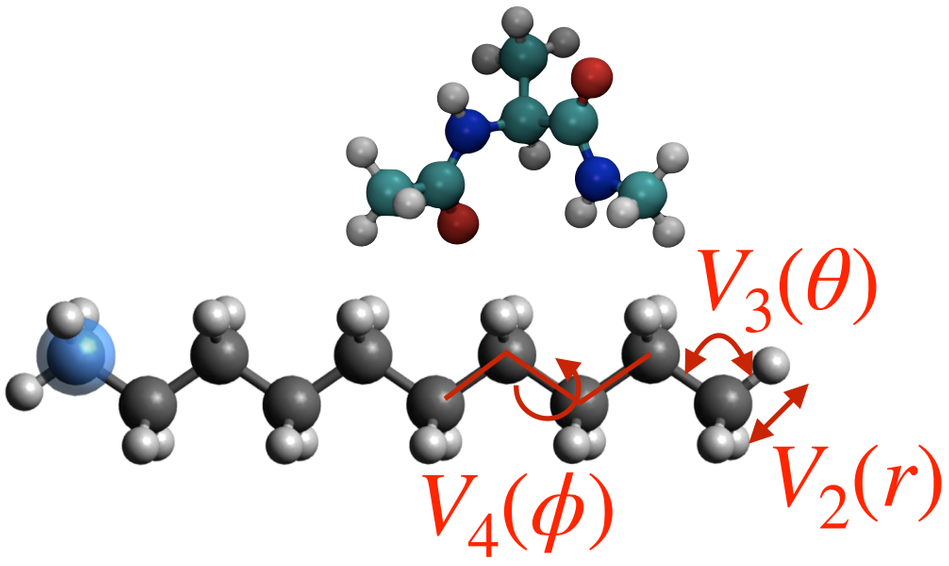
Transferability

organic force fields

- Accurate for small displacements
- Typically no reactions (some exceptions)
- Ultimate fit is to macroscopic experimental properties

AMBER
CHARMM
COMPASS
OPLS
...

Interatomic potentials (force fields)



Transferability

organic force fields

- Accurate for small displacements
- Typically no reactions (some exceptions)
- Ultimate fit is to macroscopic experimental properties

AMBER
CHARMM
COMPASS
OPLS
...

$V\left(\sum \rho(r)\right)$

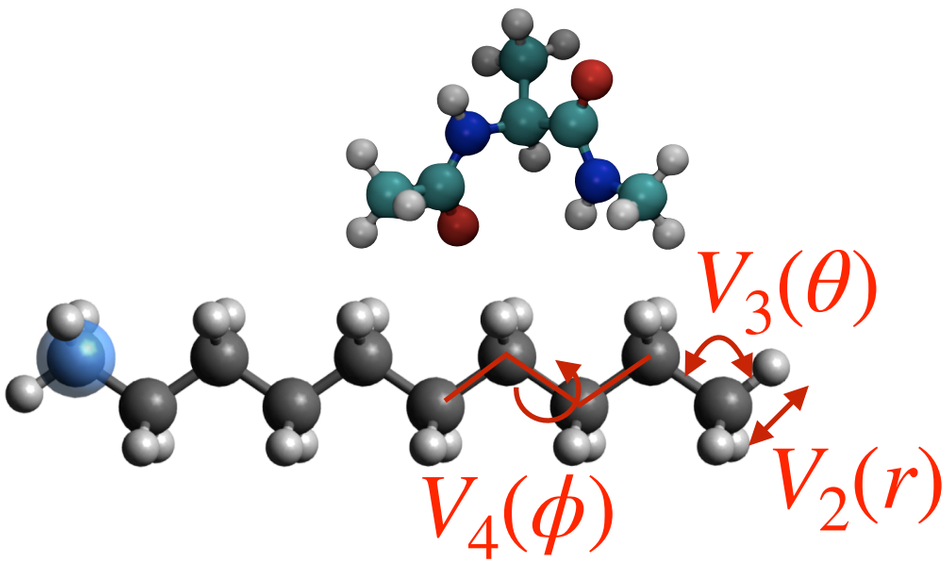
Reactive

solid state materials

- Physics-based functional forms
- Not systematic
- Simplest model to capture phenomena
- Mixture of fitting to experiment and ab initio data

Tersoff
EAM
BOP
...

Interatomic potentials (force fields)

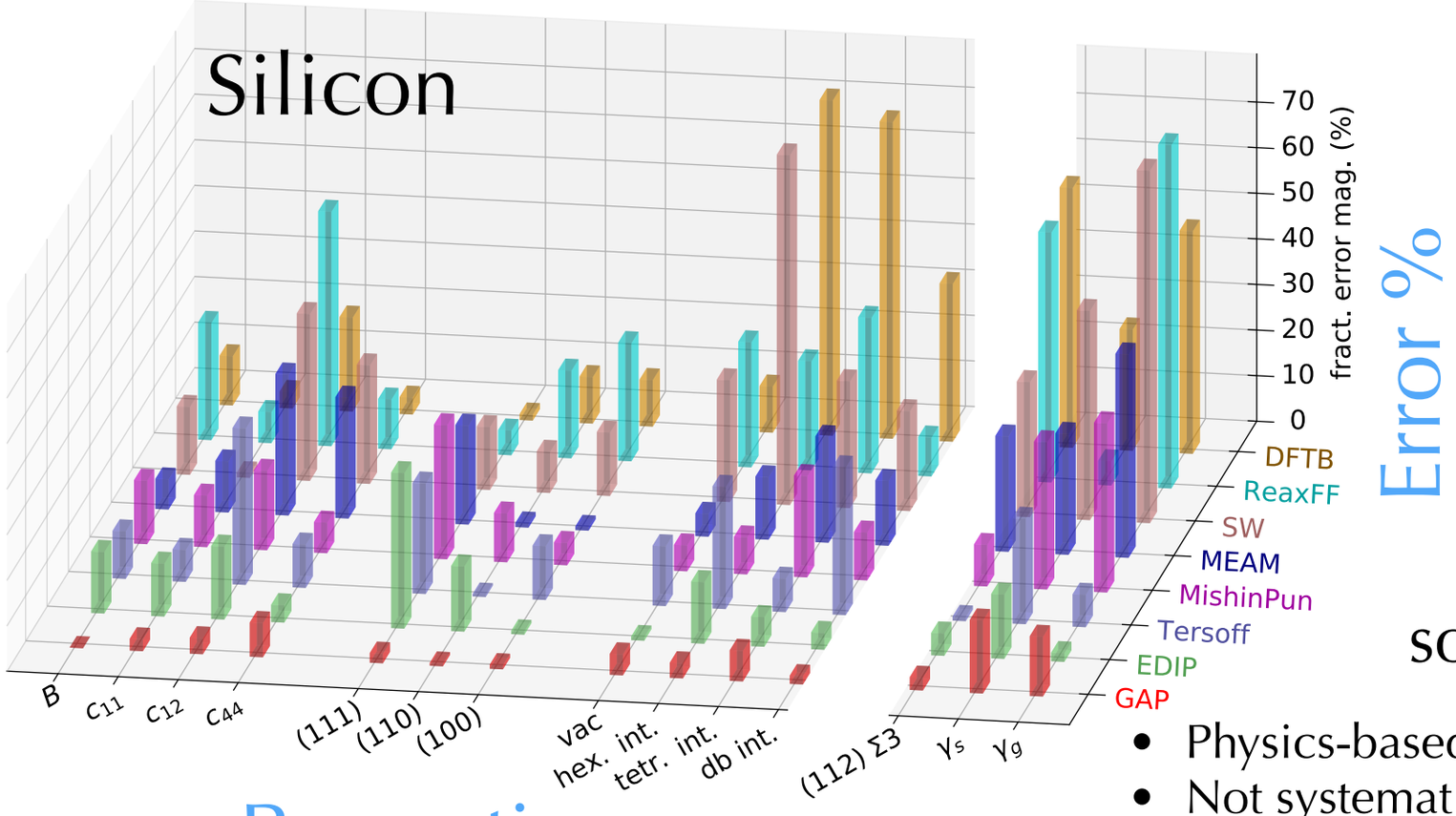


Transferability

organic force fields

- Accurate for small displacements
- Typically no reactions (some exceptions)
- Ultimate fit is to macroscopic experimental properties

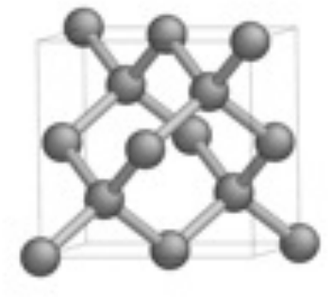
AMBER
CHARMM
COMPASS
OPLS
...



Reactive

solid state materials

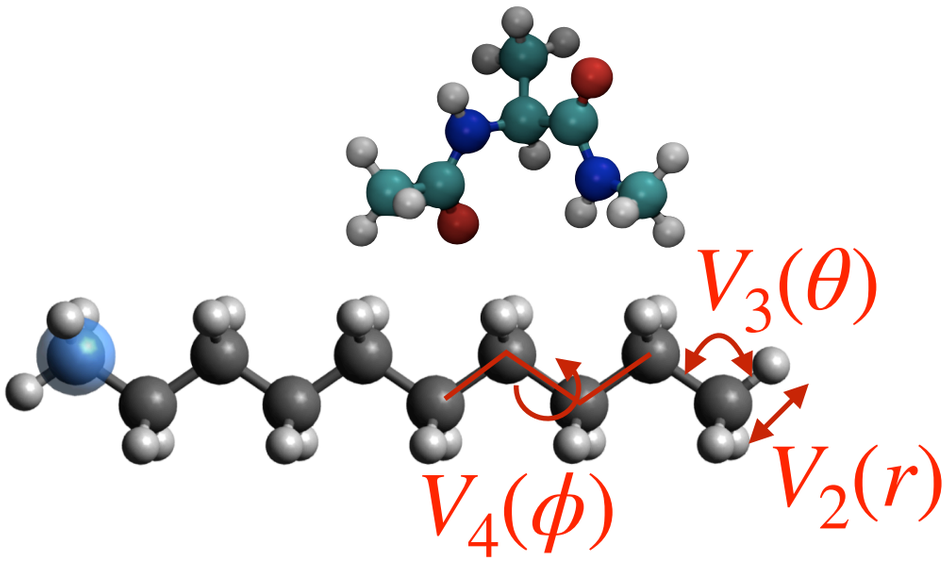
$$V\left(\sum \rho(r)\right)$$



- Physics-based functional forms
- Not systematic
- Simplest model to capture phenomena
- Mixture of fitting to experiment and ab initio data

Tersoff
EAM
BOP
...

Interatomic potentials (force fields)



Transferability

organic force fields

- Accurate for small displacements
- Typically no reactions (some exceptions)
- Ultimate fit is to macroscopic experimental properties

AMBER
CHARMM
COMPASS
OPLS
...

$V\left(\sum \rho(r)\right)$

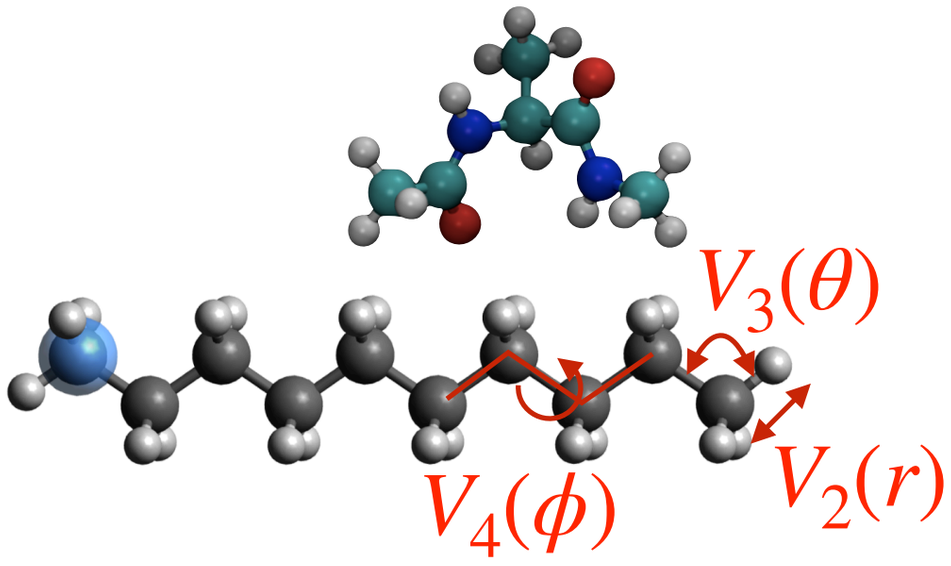
Reactive

solid state materials

- Physics-based functional forms
- Not systematic
- Simplest model to capture phenomena
- Mixture of fitting to experiment and ab initio data

Tersoff
EAM
BOP
...

Interatomic potentials (force fields)

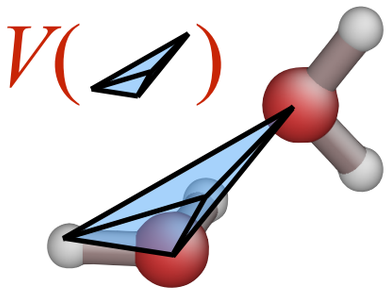


Transferability

organic force fields

- Accurate for small displacements
- Typically no reactions (some exceptions)
- Ultimate fit is to macroscopic experimental properties

AMBER
CHARMM
COMPASS
OPLS
…

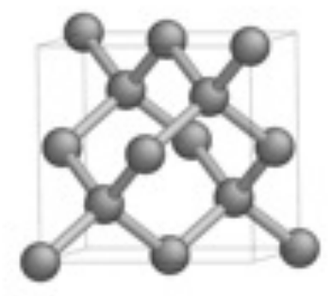


Accuracy

gas phase molecules

- Systematic polynomial basis
- Fit to ab initio data
- Exponential scaling in molecule size

Braams & Bowman
Szalewicz
Paesani
Thiel
…



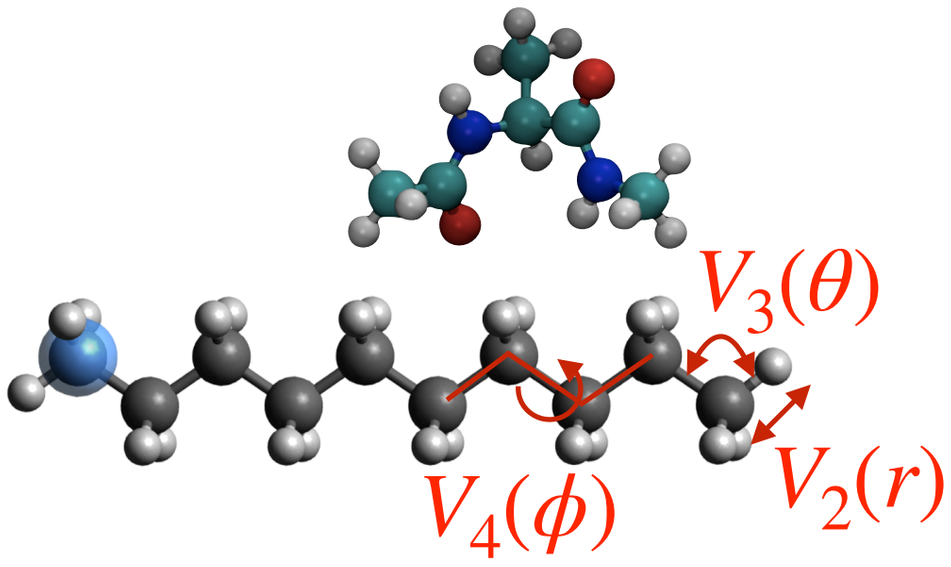
Reactive

solid state materials

- Physics-based functional forms
- Not systematic
- Simplest model to capture phenomena
- Mixture of fitting to experiment and ab initio data

Tersoff
EAM
BOP
…

Interatomic potentials (force fields)

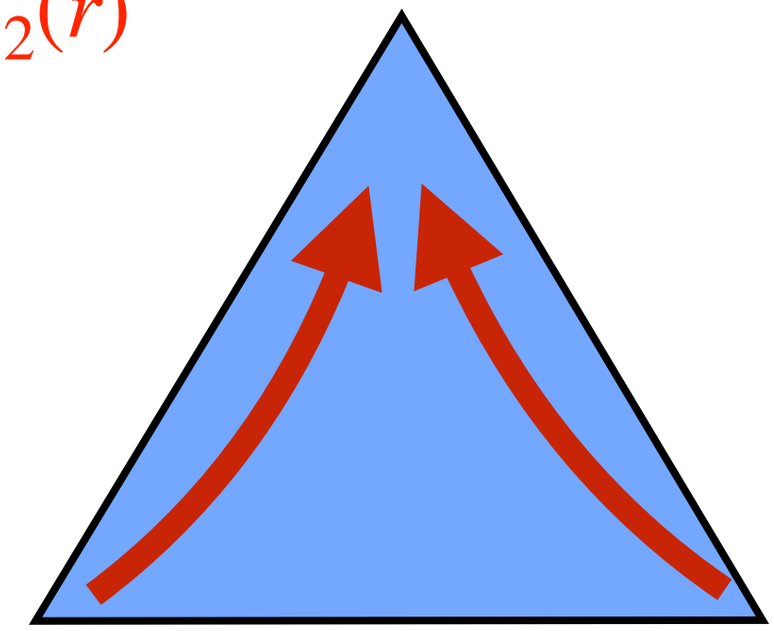


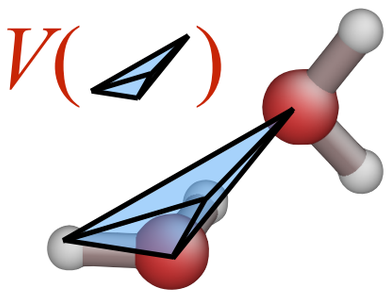
Transferability

organic force fields

- Accurate for small displacements
- Typically no reactions (some exceptions)
- Ultimate fit is to macroscopic experimental properties

AMBER
CHARMM
COMPASS
OPLS
…





Accuracy

gas phase molecules

- Systematic polynomial basis
- Fit to ab initio data
- Exponential scaling in molecule size

Braams & Bowman
Szalewicz
Paesani
Thiel
…

Reactive

solid state materials

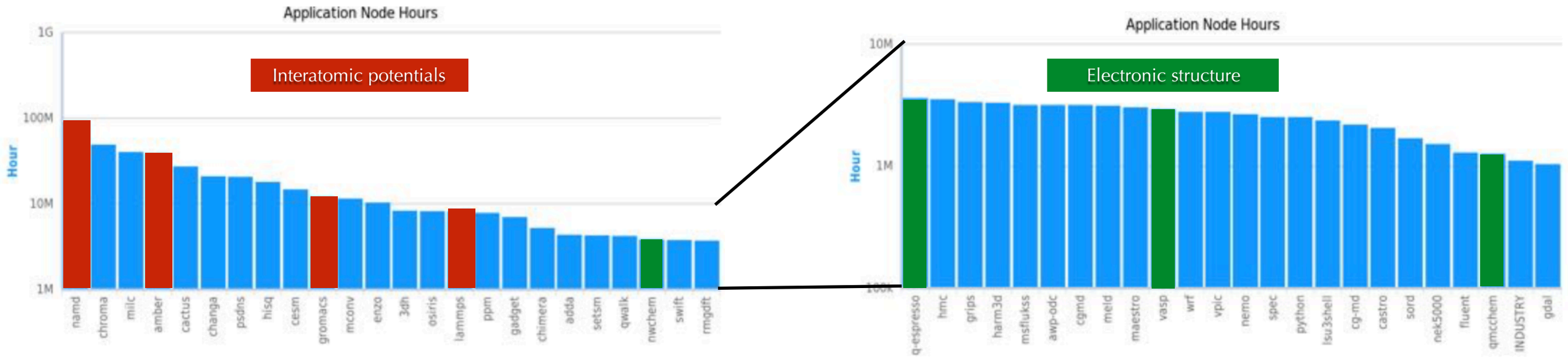
- Physics-based functional forms
- Not systematic
- Simplest model to capture phenomena
- Mixture of fitting to experiment and ab initio data

$V\left(\sum \rho(r)\right)$

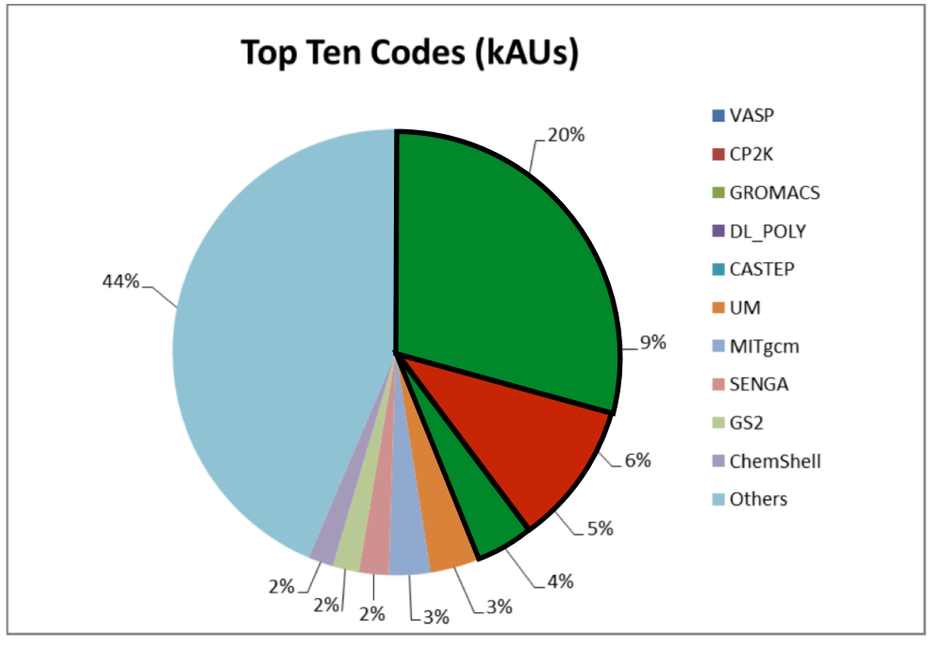
Tersoff
EAM
BOP
…

Significant HPC time

Blue Waters (USA) application workload analysis (2017 report)



UK national HPC (Hector) 2012



Is it *possible* to create an accurate, reactive interatomic potential for covalently bound materials?

Is it *possible* to create an accurate, reactive interatomic potential for covalently bound materials?

- Separation of short and long range interactions

Is it *possible* to create an accurate, reactive interatomic potential for covalently bound materials?

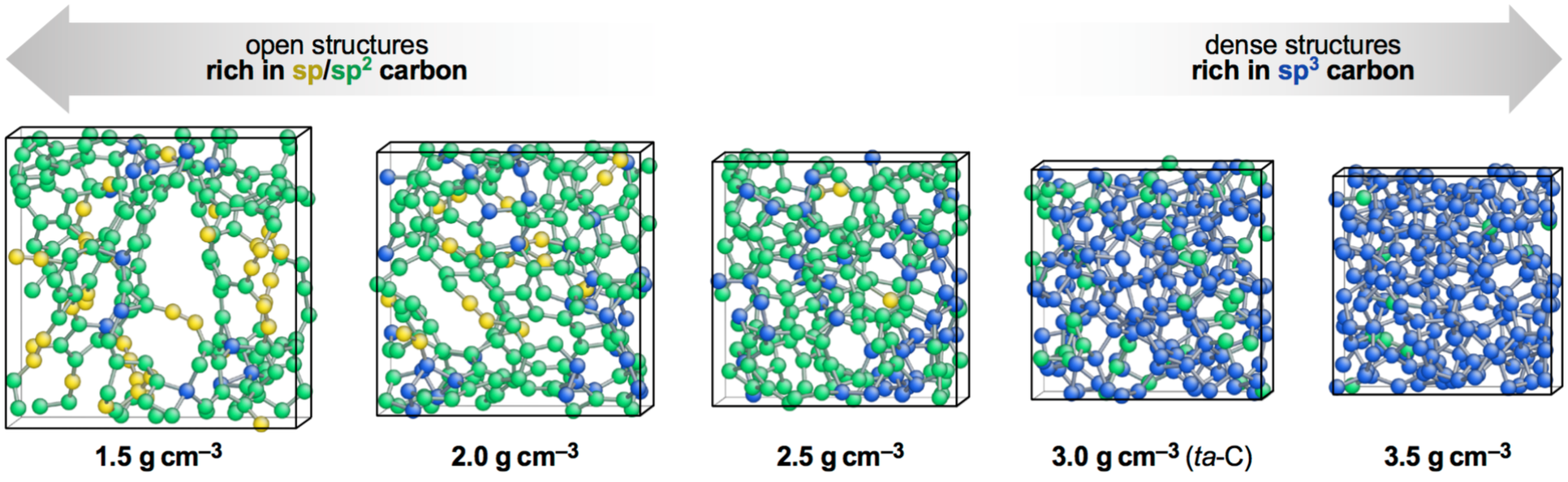
- Separation of short and long range interactions
- True quantum mechanics is many-body and long range: charge transfer, polarisable electrostatics, van der Waals dispersion

Is it *possible* to create an accurate, reactive interatomic potential for covalently bound materials?

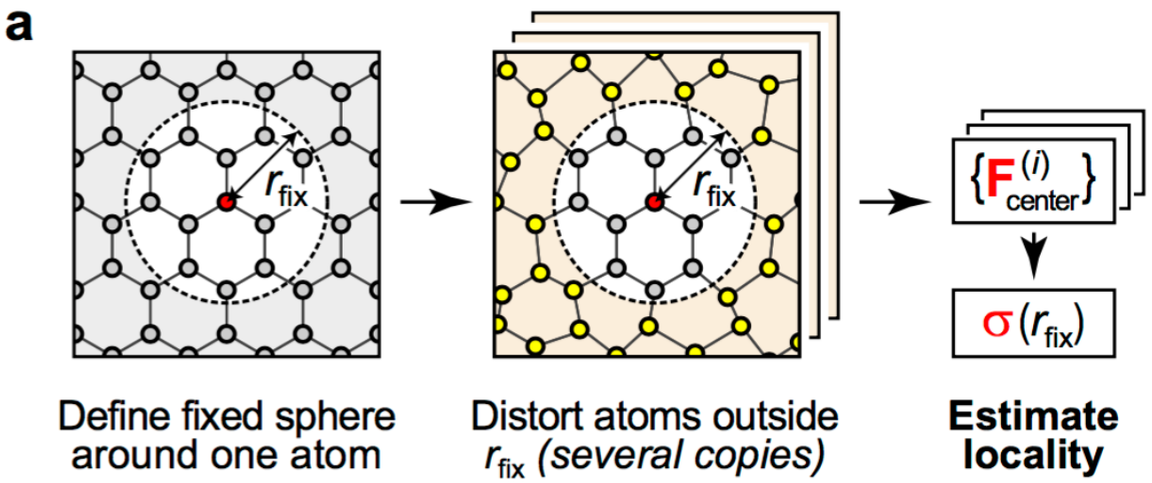
- Separation of short and long range interactions
- True quantum mechanics is many-body and long range: charge transfer, polarisable electrostatics, van der Waals dispersion
- Three possibilities to deal with long range interactions
  - They are small enough to be ignored (beyond a cutoff of e.g. 5-6-8 Å)
  - We know their analytical form, subtract it before fitting
  - We can try to fit parameters of the long range interactions

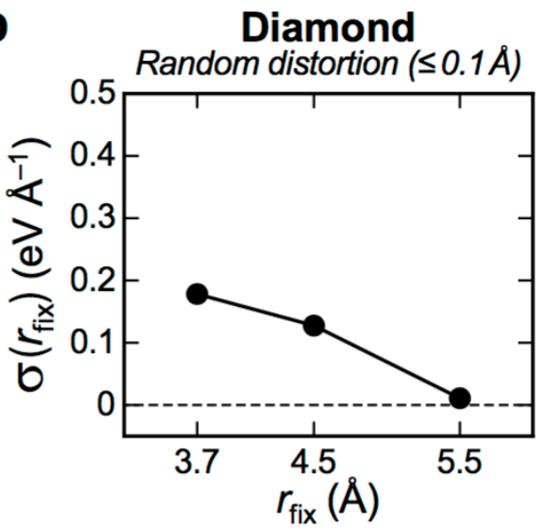
Is it *possible* to create an accurate, reactive interatomic potential for covalently bound materials?

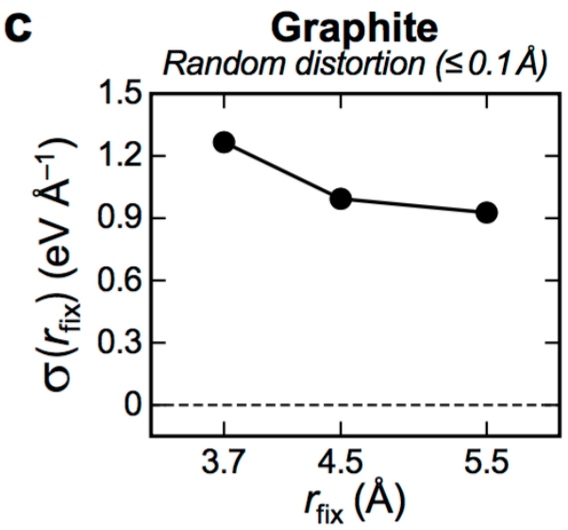
- Separation of short and long range interactions
- True quantum mechanics is many-body and long range: charge transfer, polarisable electrostatics, van der Waals dispersion
- Three possibilities to deal with long range interactions
  - They are small enough to be ignored (beyond a cutoff of e.g. 5-6-8 Å)
  - We know their analytical form, subtract it before fitting
  - We can try to fit parameters of the long range interactions
- How local is the remainder?

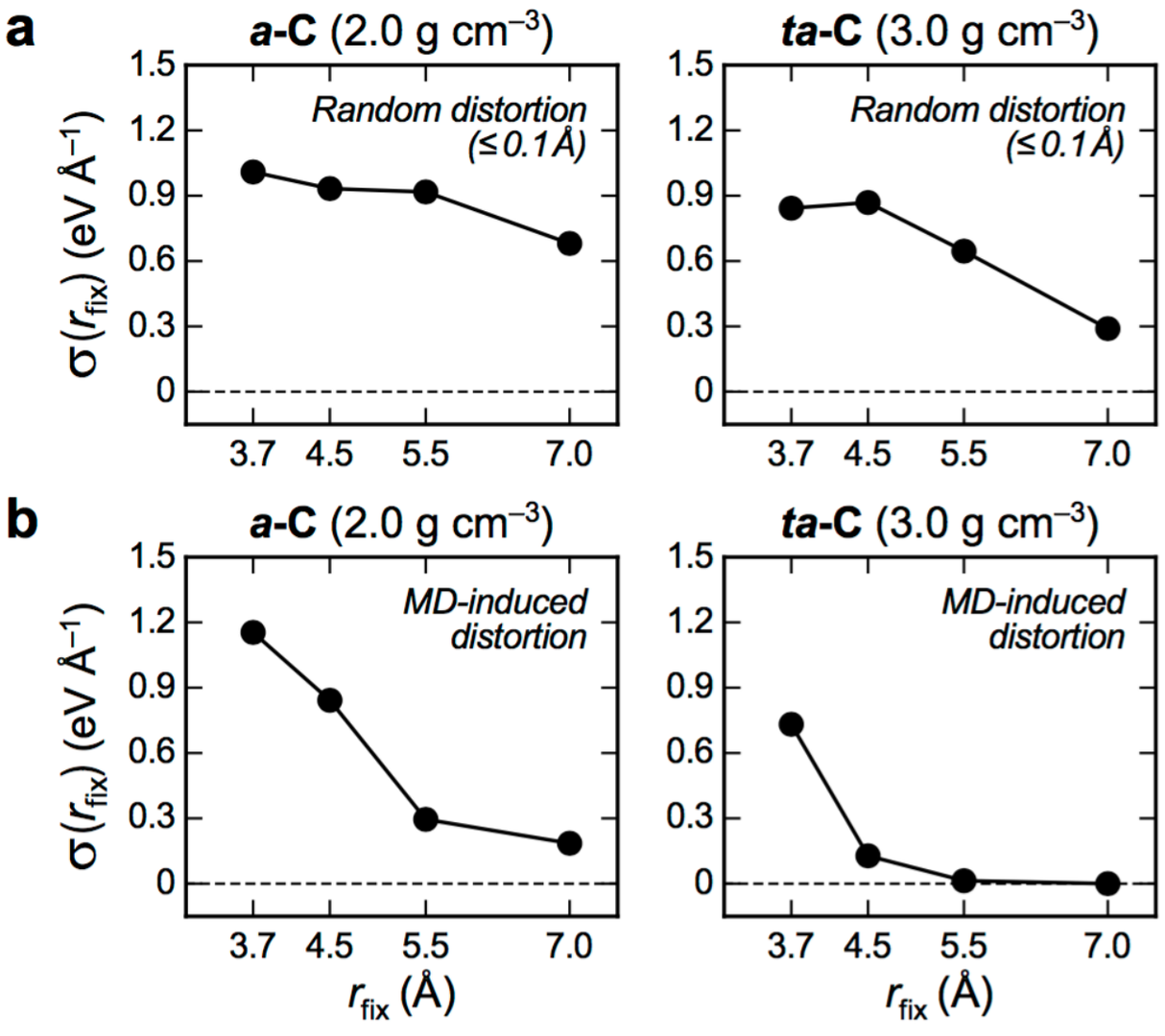


The "locality test"

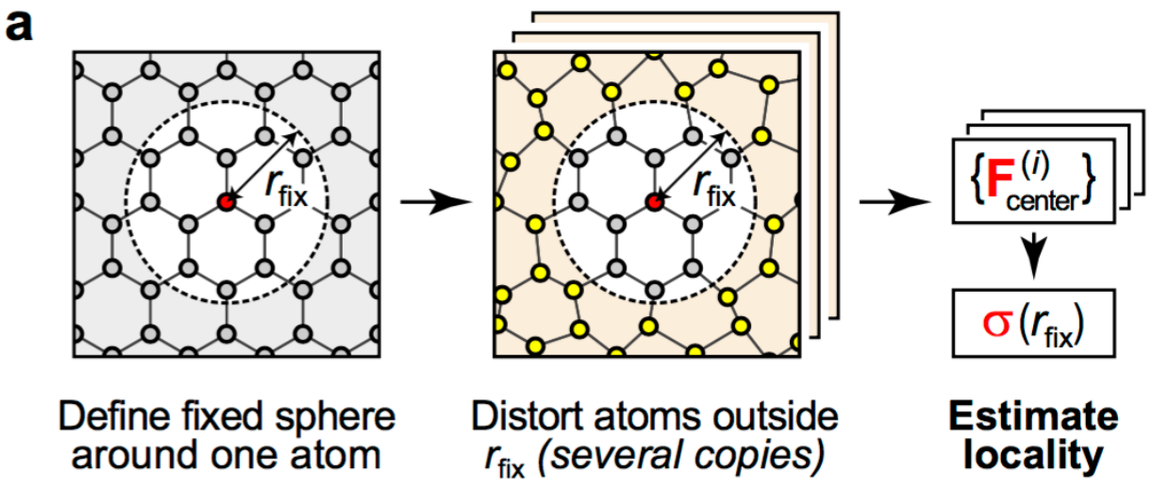


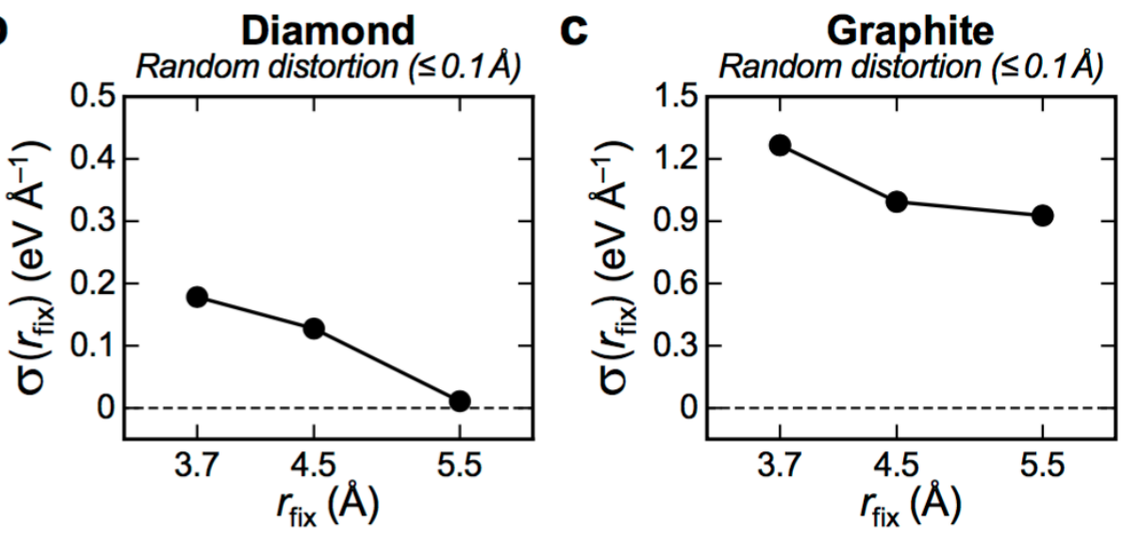


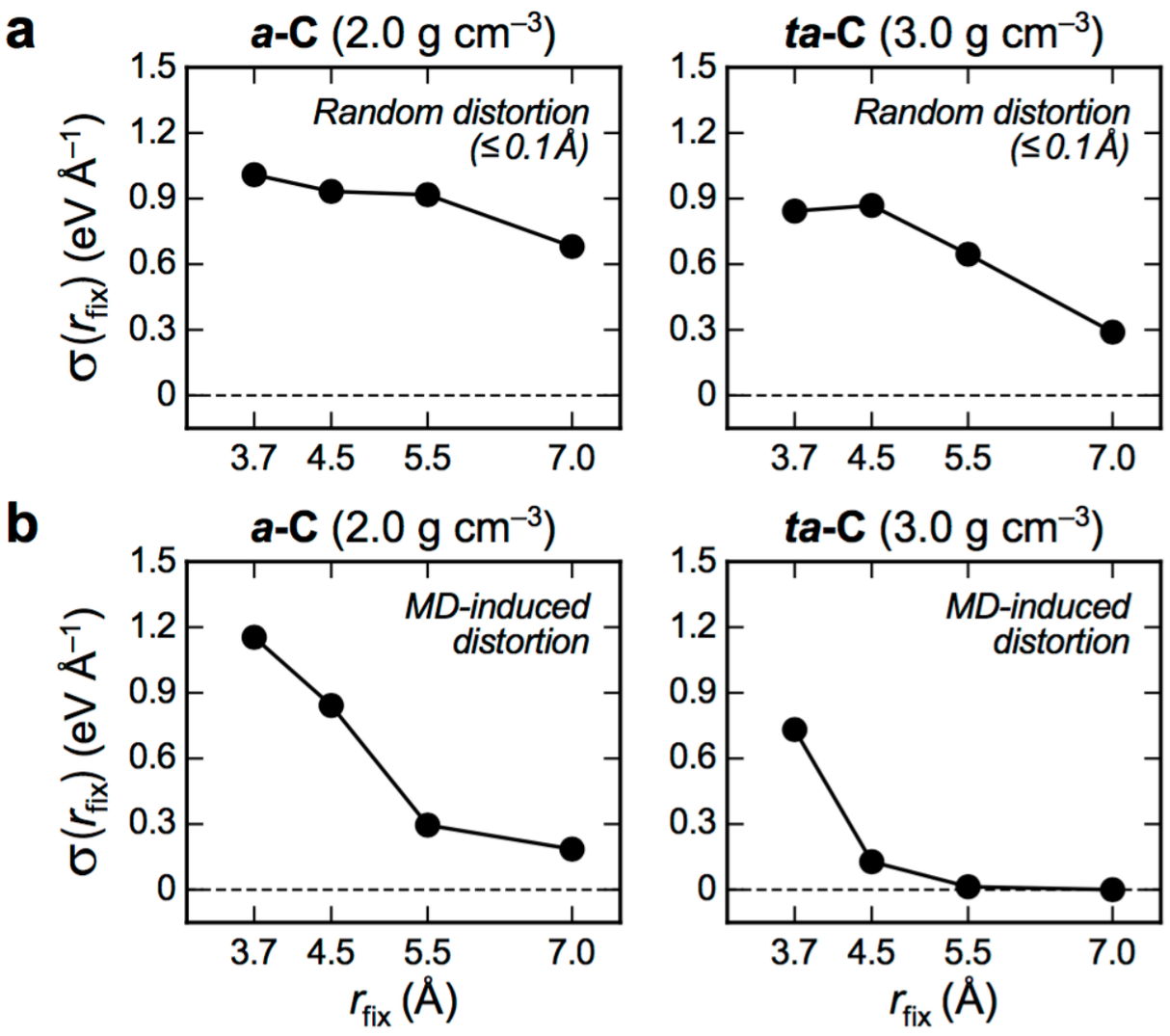




The "locality test"







Limit on force field accuracy: for a given cutoff, what force accuracy is achievable in principle?

A decade old promise. Where are we?

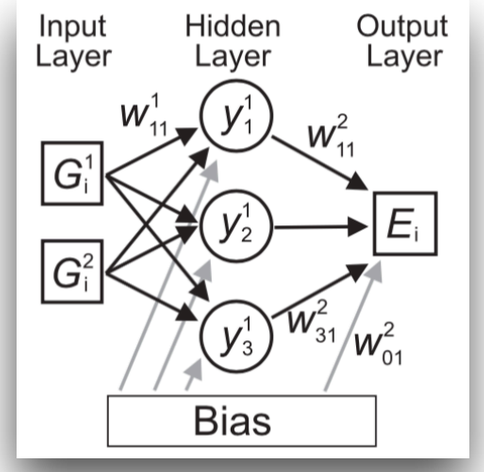
PRL **98,** 146401 (2007) PHYSICAL REVIEW LETTERS week ending 6 APRIL 2007

Generalized Neural-Network Representation of High-Dimensional Potential-Energy Surfaces

Jörg Behler and Michele Parrinello

Department of Chemistry and Applied Biosciences, ETH Zurich, USI-Campus, Via Giuseppe Buffi 13, CH-6900 Lugano, Switzerland

(Received 27 September 2006; published 2 April 2007)



Claim: short range QM regression problem is solved.

PRL **104,** 136403 (2010) PHYSICAL REVIEW LETTERS week ending 2 APRIL 2010

Gaussian Approximation Potentials: The Accuracy of Quantum Mechanics, without the Electrons

Albert P. Bartók and Mike C. Payne

Cavendish Laboratory, University of Cambridge, J J Thomson Avenue, Cambridge, CB3 0HE, United Kingdom

Risi Kondor

Center for the Mathematics of Information, California Institute of Technology, MC 305-16, Pasadena, California 91125, USA

Gábor Csányi

Engineering Laboratory, University of Cambridge, Trumpington Street, Cambridge, CB2 1PZ, United Kingdom

(Received 1 October 2009; published 1 April 2010)

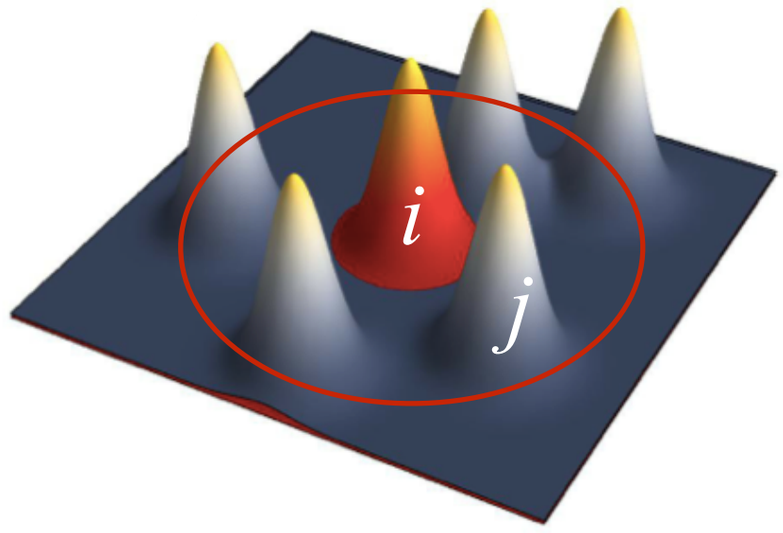
$$\mathbf{Q}_M = \mathbf{C}_M + \mathbf{C}_{MN}(\mathbf{\Lambda} + \sigma^2\mathbf{I})^{-1}\mathbf{C}_{NM},$$
$$\varepsilon_* = \mathbf{k}_*^T\mathbf{Q}_M^{-1}\mathbf{C}_{MN}(\mathbf{\Lambda} + \sigma^2\mathbf{I})^{-1}\mathbf{y},$$

Representation

Representation: atomic neighbour density

$$\rho^{(i)}(\mathbf{r}) = \sum_j G(\mathbf{r} - \mathbf{r}_{ij})\, f_{\text{cut}}(|\mathbf{r}_{ij}|)$$

Convolution kernel $\quad G(\mathbf{r}) = \begin{cases} \delta(\mathbf{r}) \\ e^{-|\mathbf{r}|^2/2\sigma^2} \end{cases}$



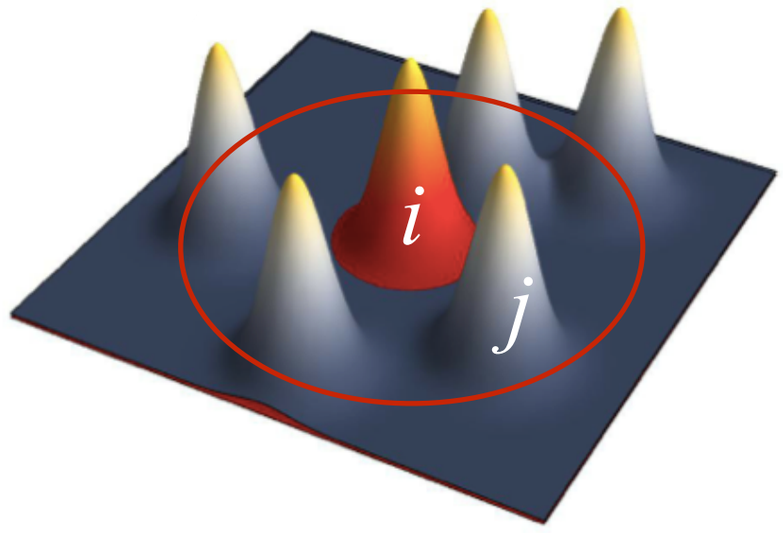
cutoff: compact support

ρ is permutational invariant, smooth and continuous

Representation: atomic neighbour density

$$\rho^{(i)}(\mathbf{r}) = \sum_j G(\mathbf{r} - \mathbf{r}_{ij})\, f_{\text{cut}}(|\mathbf{r}_{ij}|)$$

Convolution kernel $\quad G(\mathbf{r}) = \begin{cases} \delta(\mathbf{r}) \\ e^{-|\mathbf{r}|^2/2\sigma^2} \end{cases}$



cutoff: compact support

ρ is permutational invariant, smooth and continuous

Rotational invariance:

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

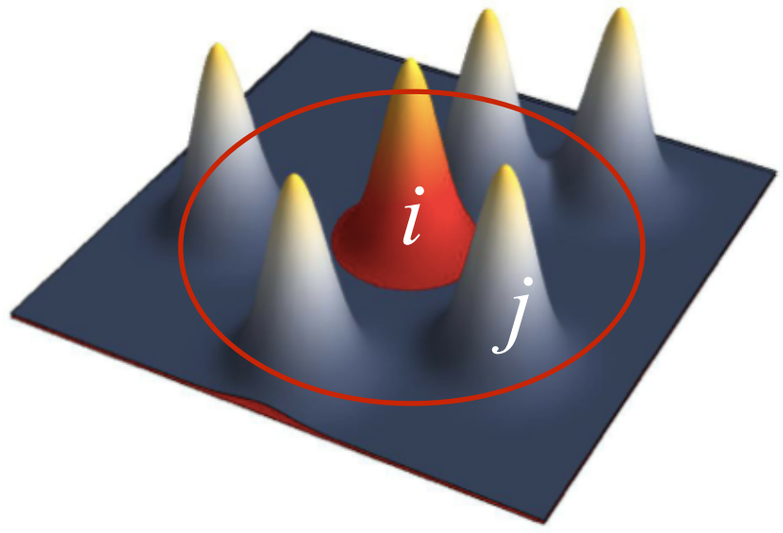
$$p_{nn'l} = \sum_m c^{\dagger}_{nlm} c_{n'lm}$$

is the power spectrum, rotationally invariant

Representation: atomic neighbour density

$$\rho^{(i)}(\mathbf{r}) = \sum_j G(\mathbf{r} - \mathbf{r}_{ij})\, f_{\text{cut}}(|\mathbf{r}_{ij}|)$$

Convolution kernel $\quad G(\mathbf{r}) = \begin{cases} \delta(\mathbf{r}) \\ e^{-|\mathbf{r}|^2/2\sigma^2} \end{cases}$



cutoff: compact support

ρ is permutational invariant, smooth and continuous

Rotational invariance:

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

$$p_{nn'l} = \sum_m c^{\dagger}_{nlm} c_{n'lm}$$

Smooth overlap of atomic positions (SOAP)

$$\mathbf{p} \cdot \mathbf{p}' = \int_{\hat{R} \in SO_3} d\hat{R} \left| \int d\mathbf{r} \rho(\mathbf{r}) \rho'(\hat{R}\mathbf{r}) \right|^2$$

SOAP kernel: $K(R, R') \equiv K(\rho, \rho') = |\mathbf{p} \cdot \mathbf{p}'|^{\zeta}$

is the power spectrum, rotationally invariant

Density projections

We can think of the soap vector $\mathbf{p}$ as the rotational invariant combination of projections of $\rho(\mathbf{r})$ onto basis functions:

$$b_\alpha(\mathbf{r}) = g_n(r) Y_{lm}(\hat{\mathbf{r}}) \qquad \alpha = [nlm]$$

(Ceriotti bra-ket notation: $\langle nlm | \chi^{(2)} \rangle$)

Many other representations: ACSF of Behler, MBTR of Rupp, von Lilienfeld's histograms, Mallat's scattering transforms, Bispectrum in SNAP (Thompson) are **all density projections** with some choice of basis and convolution kernel.

Which is the best basis to project onto?

Other nice properties of the Y_{lm} basis

Other nice properties of the Y_{lm} basis

- Natural extension to basis functions to expand non-scalar properties, vectors and tensors (Grisafi, Wilkins, GC, Ceriotti PRL 2018)

$$k^{\lambda}(\rho,\rho') = \int d\hat{R}\, \mathbf{D}^{\lambda}(\hat{R}) \left| \int d\mathbf{r} \rho(r) \rho'(\hat{R}\mathbf{r}) \right|^2$$

$\mathbf{D}$: Wigner matrix

Other nice properties of the Y_{lm} basis

- Natural extension to basis functions to expand non-scalar properties, vectors and tensors (Grisafi, Wilkins, GC, Ceriotti PRL 2018)

$$k^{\lambda}(\rho,\rho') = \int d\hat{R}\, \mathbf{D}^{\lambda}(\hat{R}) \left| \int d\mathbf{r}\rho(r)\rho'(\hat{R}\mathbf{r}) \right|^2$$

$\mathbf{D}$: Wigner matrix

- Further invariants (Ralf Drautz, 2019)

$$\sum \begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix} c_{n_1 l_1 m_1} c_{n_2 l_2 m_2} c_{n_3 l_3 m_3}$$

$$\sum \begin{pmatrix} l_1 & l_2 & l_3 & l_4 \\ m_1 & m_2 & m_3 & m_4 \end{pmatrix} c_{n_1 l_1 m_1} c_{n_2 l_2 m_2} c_{n_3 l_3 m_3} c_{n_4 l_4 m_4}$$

$$\vdots$$

Other nice properties of the Y_{lm} basis

- Natural extension to basis functions to expand non-scalar properties, vectors and tensors (Grisafi, Wilkins, GC, Ceriotti PRL 2018)

$$k^{\lambda}(\rho,\rho') = \int d\hat{R}\, \mathbf{D}^{\lambda}(\hat{R}) \left| \int d\mathbf{r}\rho(r)\rho'(\hat{R}\mathbf{r}) \right|^2$$

$\mathbf{D}$: Wigner matrix

- Further invariants (Ralf Drautz, 2019)

$$\sum \begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix} c_{n_1 l_1 m_1} c_{n_2 l_2 m_2} c_{n_3 l_3 m_3}$$

$$\sum \begin{pmatrix} l_1 & l_2 & l_3 & l_4 \\ m_1 & m_2 & m_3 & m_4 \end{pmatrix} c_{n_1 l_1 m_1} c_{n_2 l_2 m_2} c_{n_3 l_3 m_3} c_{n_4 l_4 m_4}$$

$$\vdots$$

- Variant to describe entire structures, rather than atomic neighbourhoods:

$$\bar{c}_{nlm} = \sum_i c^{(i)}_{nlm} \qquad \text{average over all atoms, no cutoff}$$

$$\bar{p}_{nn'l} = \sum_m \bar{c}^{\dagger}_{nlm} \bar{c}_{nlm} \qquad \text{or cutoff > periodic unit cell}$$

(Other ways to construct kernels for entire structures: De, Bartók, GC, Ceriotti PCCP 2016)

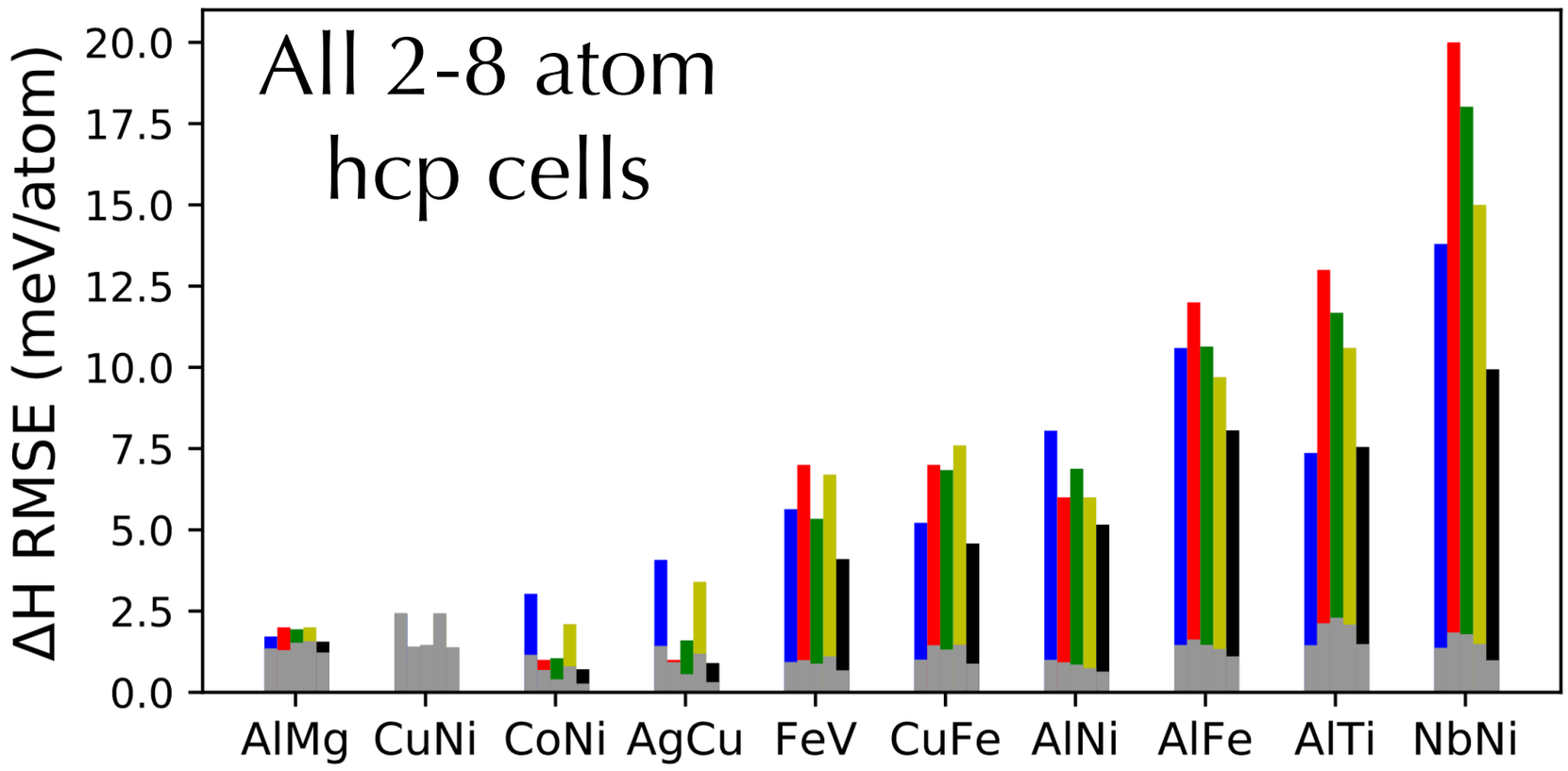
Method matrix

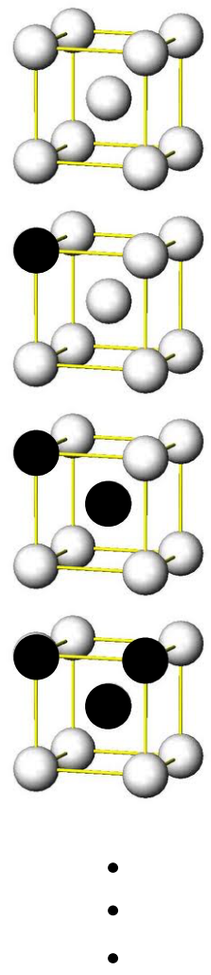
	SOAP/GAP	SNAP	BPNN	MTP	Drautz
Density	Smoothed	$\delta(\mathbf{r})$	$\delta(\mathbf{r})$	$\delta(\mathbf{r})$	$\delta(\mathbf{r})$
Basis set	Complete	Complete	Manual	Manual	Complete
Dimension	High	Medium	Medium	Low	Low
Regularised	Yes	No	Yes	No	Yes
Regression	Kernel	Linear	NN	Linear	Linear

Benchmarking

Metal alloys (with Gus Hart's group)

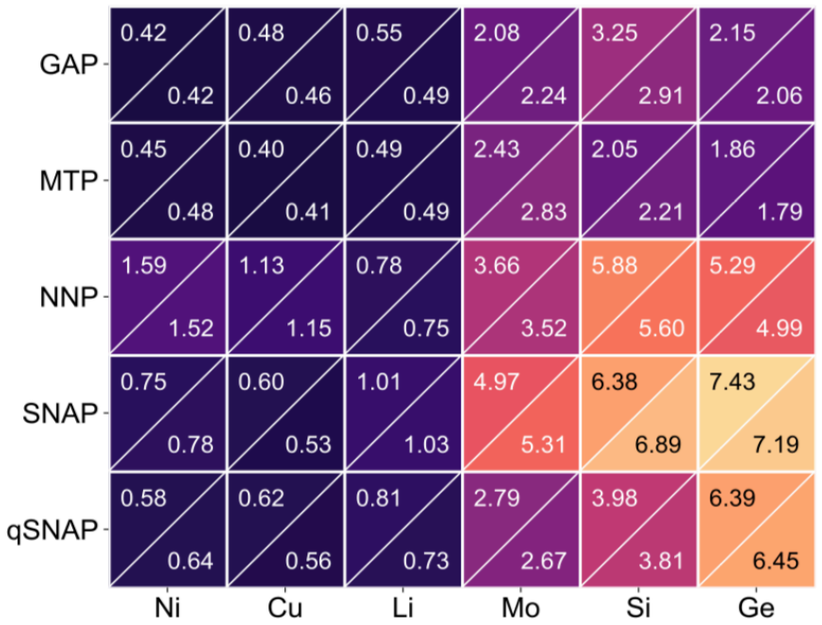
- For configurational prediction, Cluster Expansion is the successful method, meV accuracy
- Off-lattice is a problem, many components are difficult

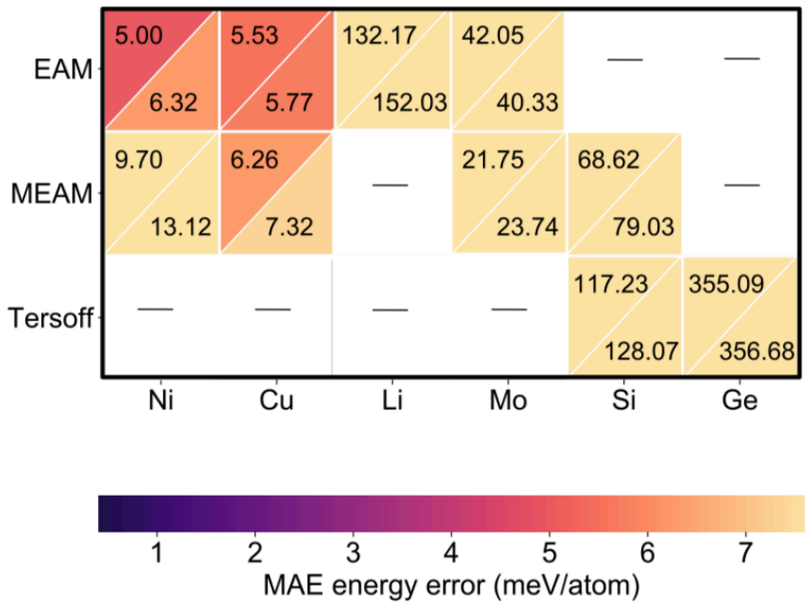




Cluster Expansion
MBTR+KRR
MBTR+DNN
SOAP+GP
MTP
δ_{RMSE}%

Alex Shapeev's Moment Tensor Potentials
(body order expansion+polynomial fit)





(a) Mean absolute errors in predicted energies

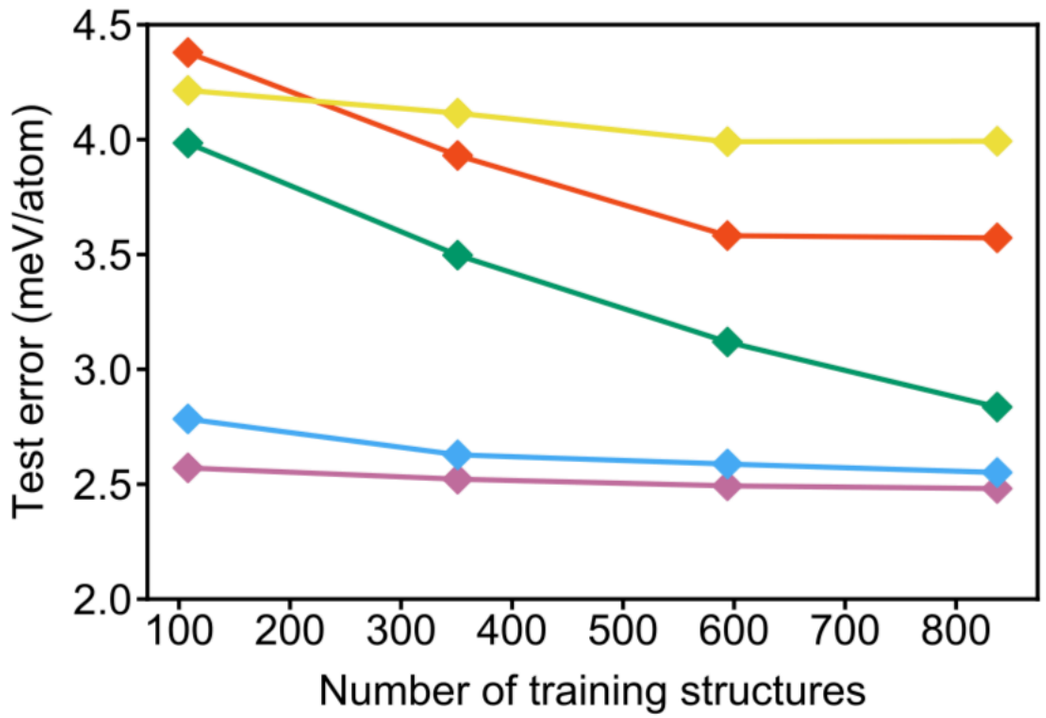


Figure 4: MAE in predicted (a) energies (

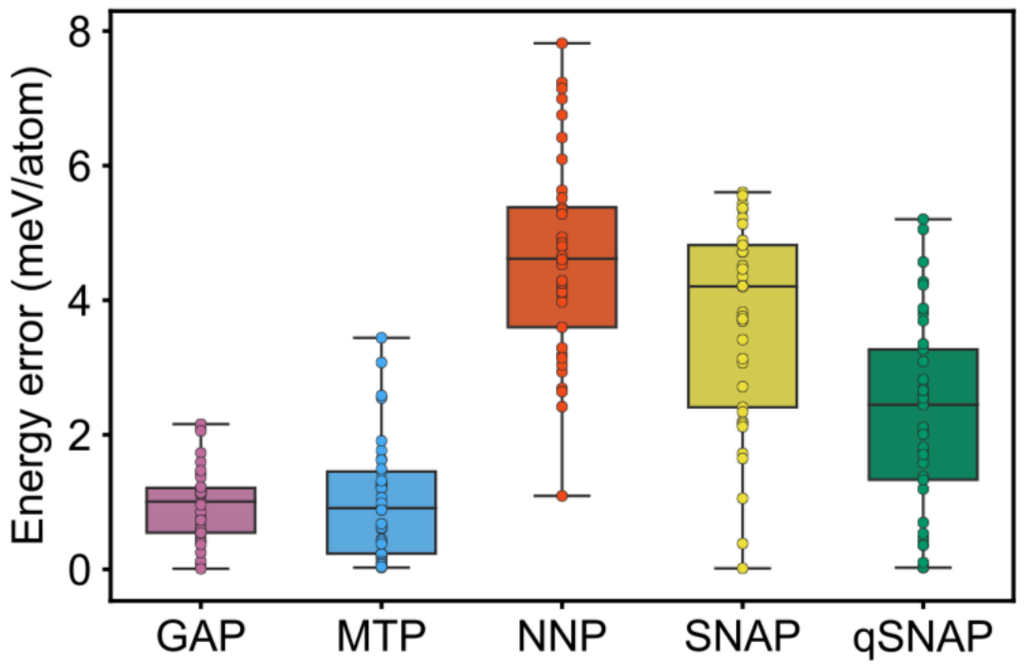
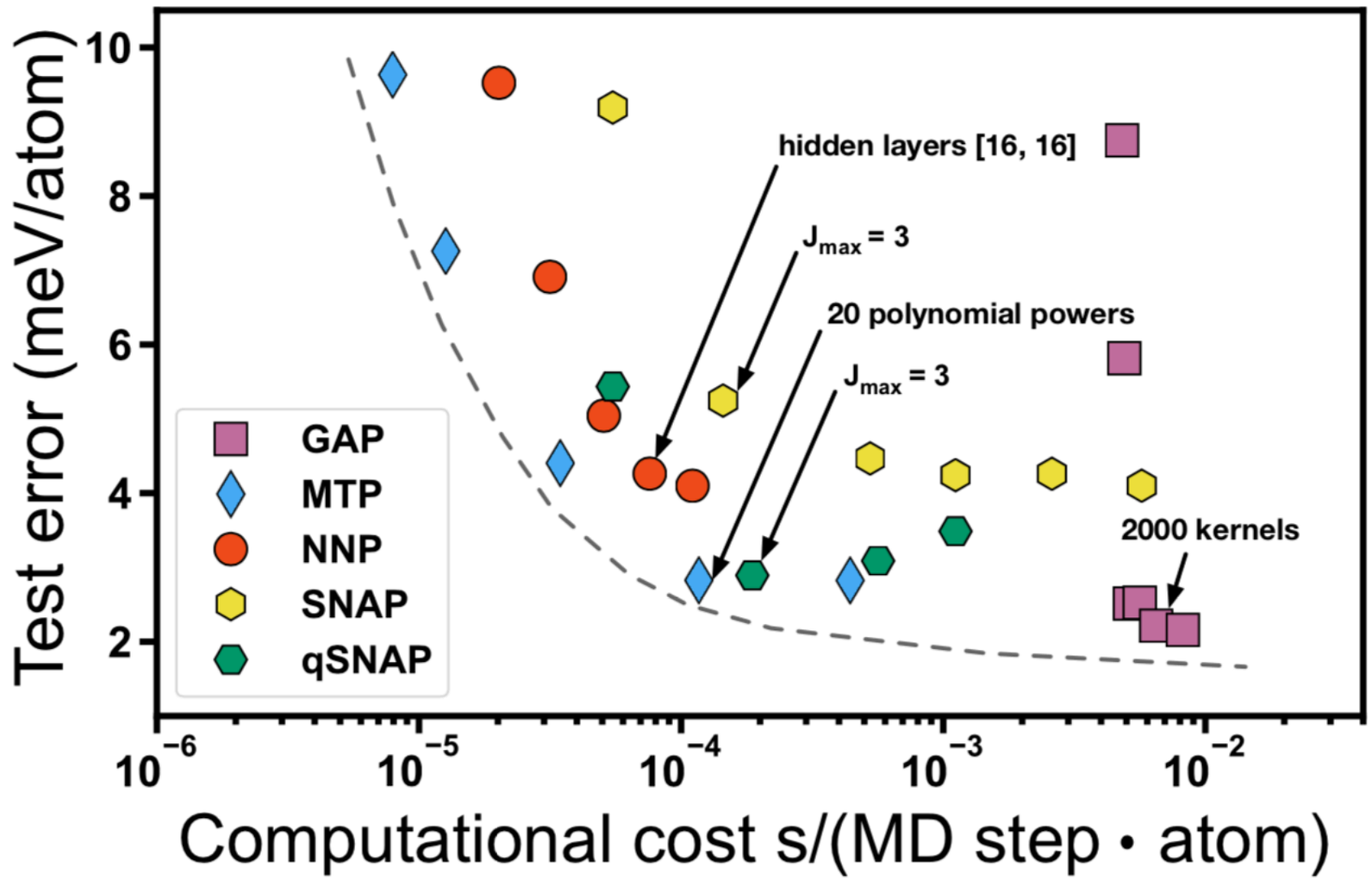


Figure 6: Error distributions in (a) predicted dynamics

Performance

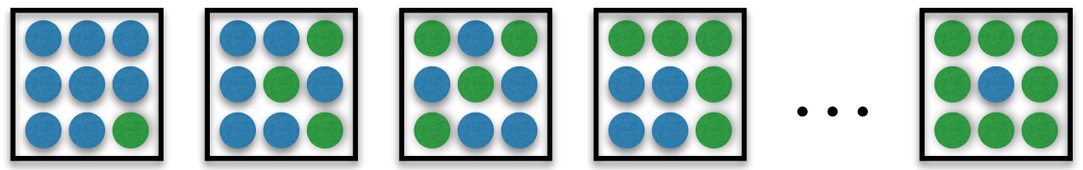
Several are in LAMMPS now :



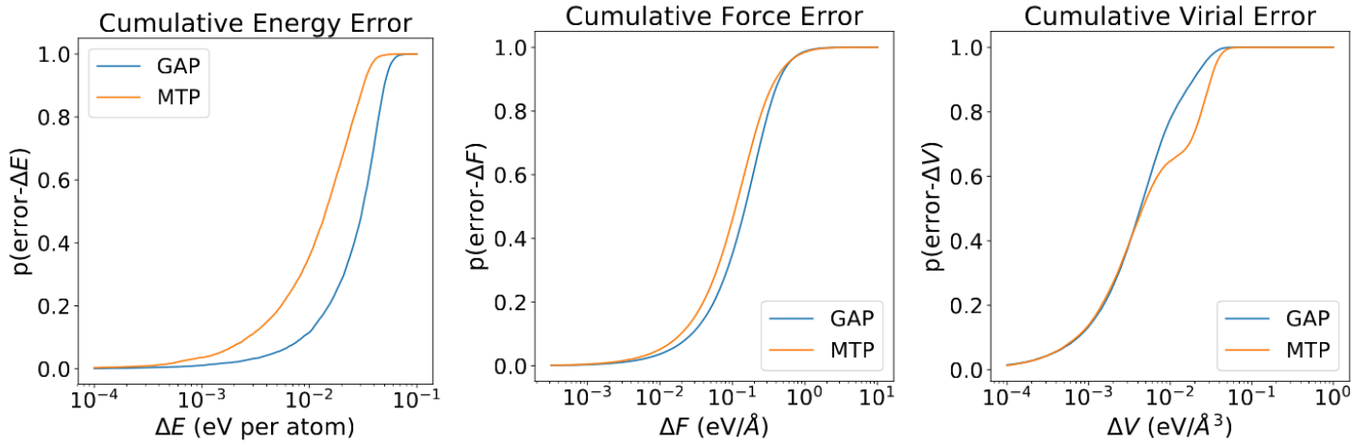
Mo deformations, surfaces, MD (arXiv, S Ping Ong et al)

Extrapolation: SOAP/GAP vs polynomials (MTP)

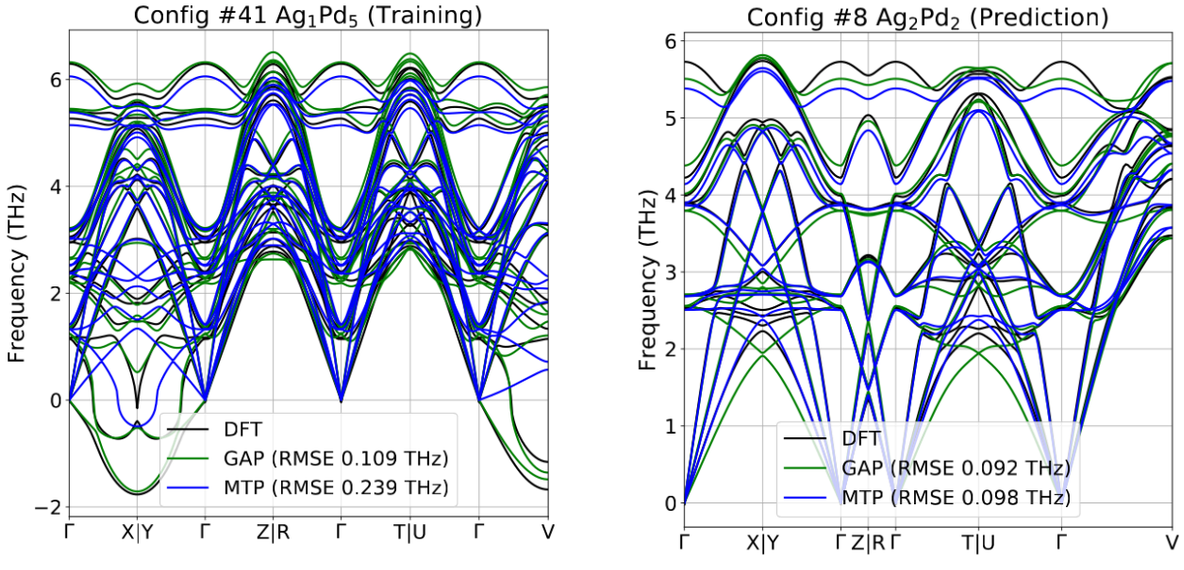
AgPd bcc crystal dataset (actively learned)



Liquid test

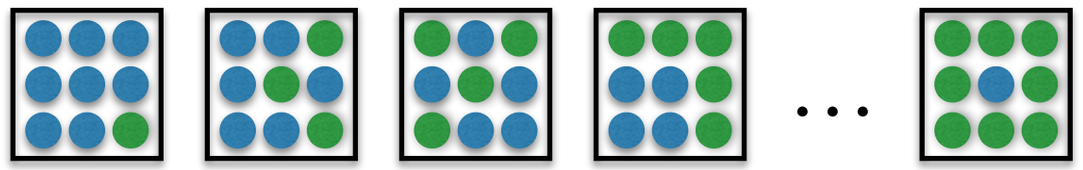


Phonon spectra

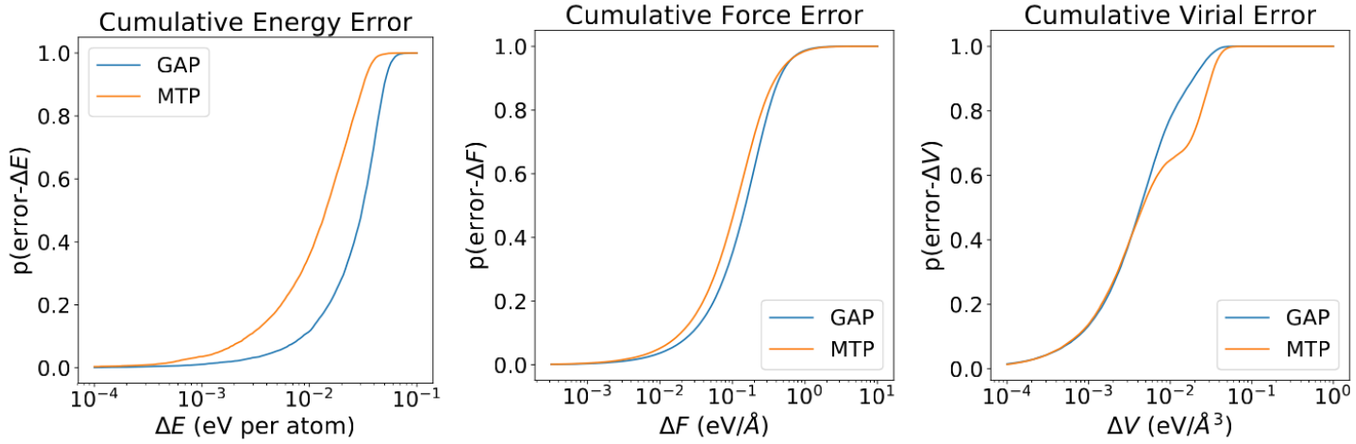


Extrapolation: SOAP/GAP vs polynomials (MTP)

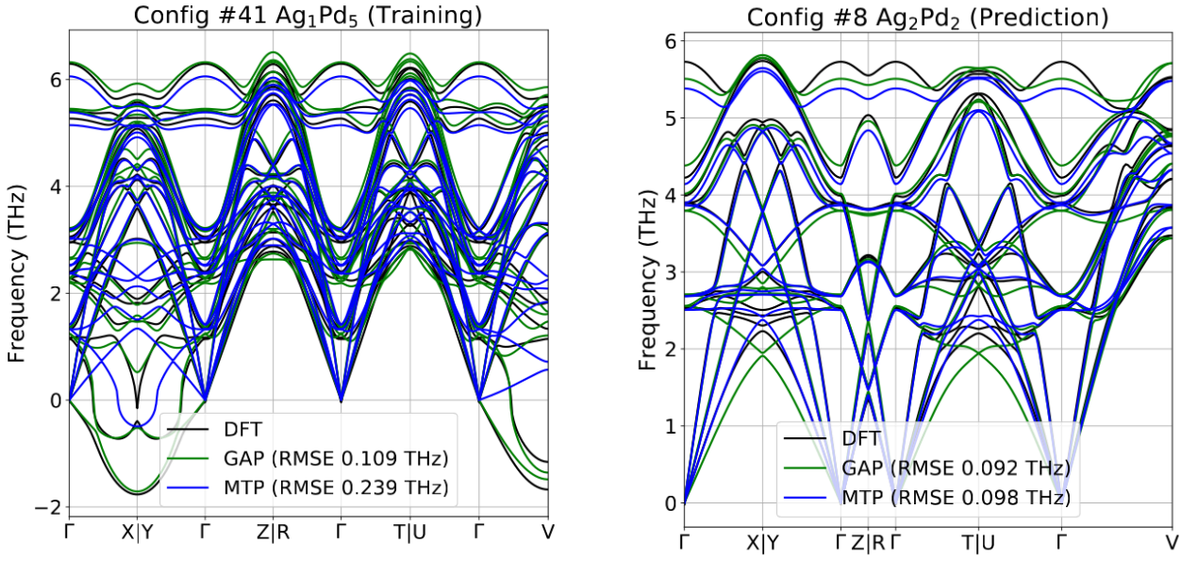
AgPd bcc crystal dataset (actively learned)



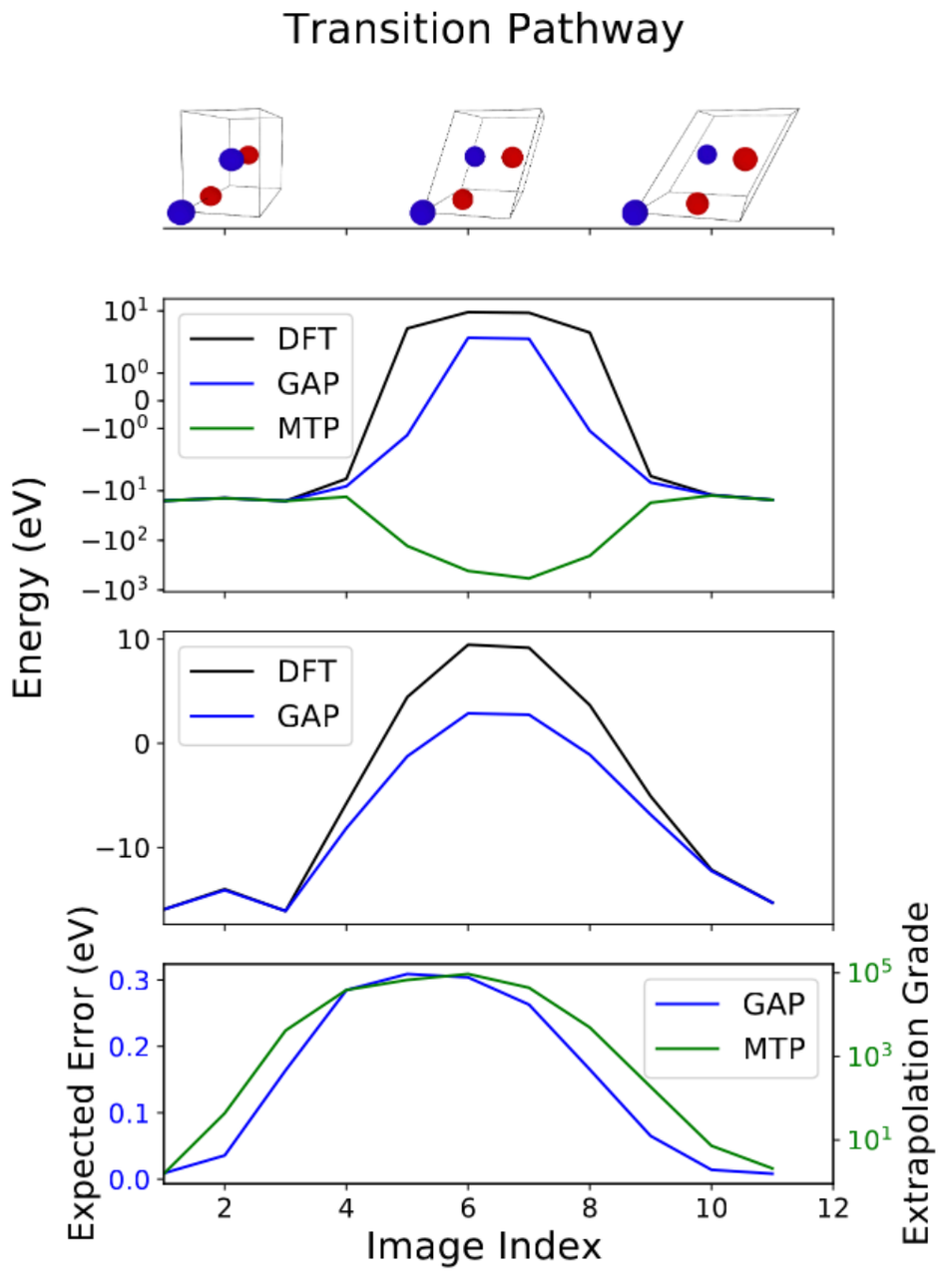
Liquid test



Phonon spectra

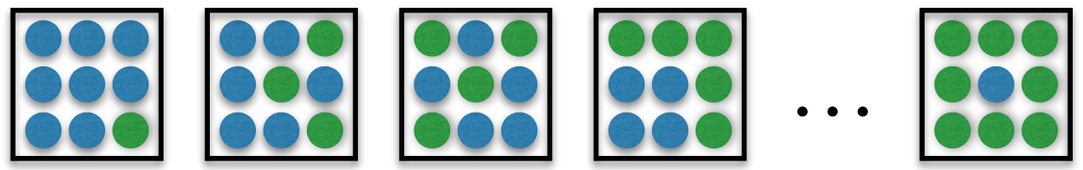


Transition Pathway

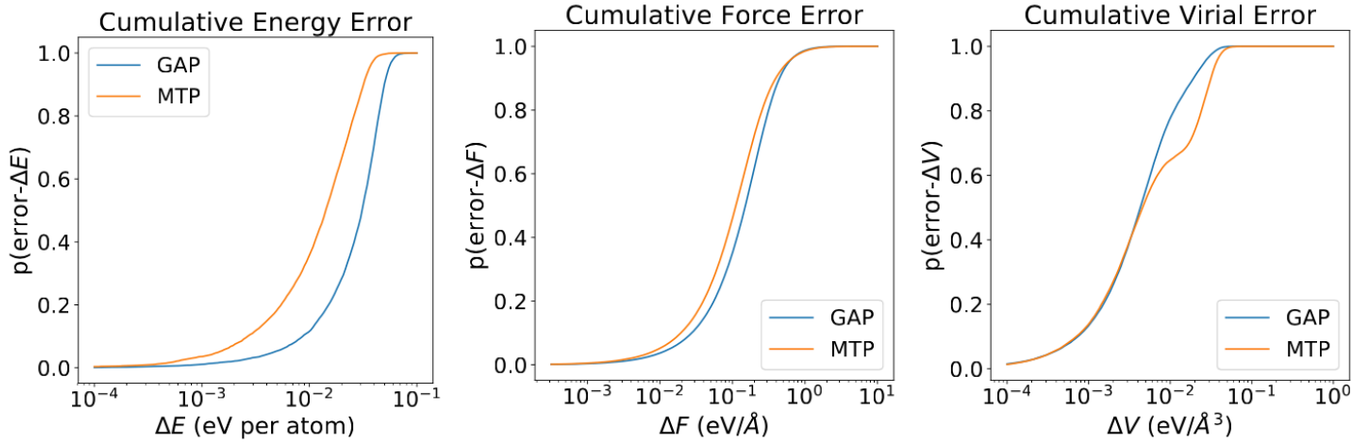


Extrapolation: SOAP/GAP vs polynomials (MTP)

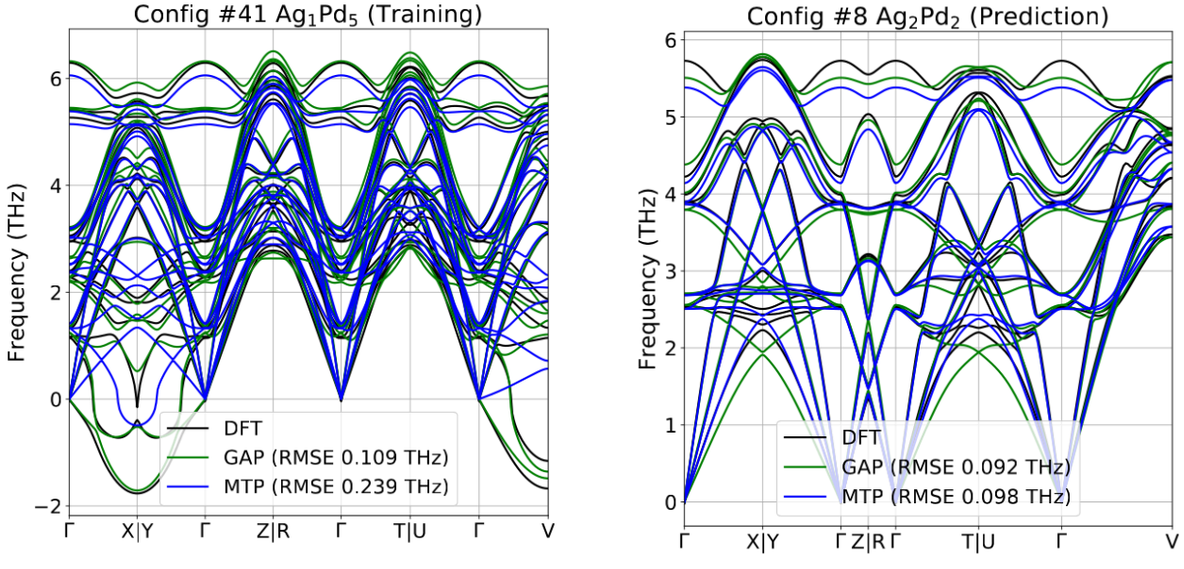
AgPd bcc crystal dataset (actively learned)



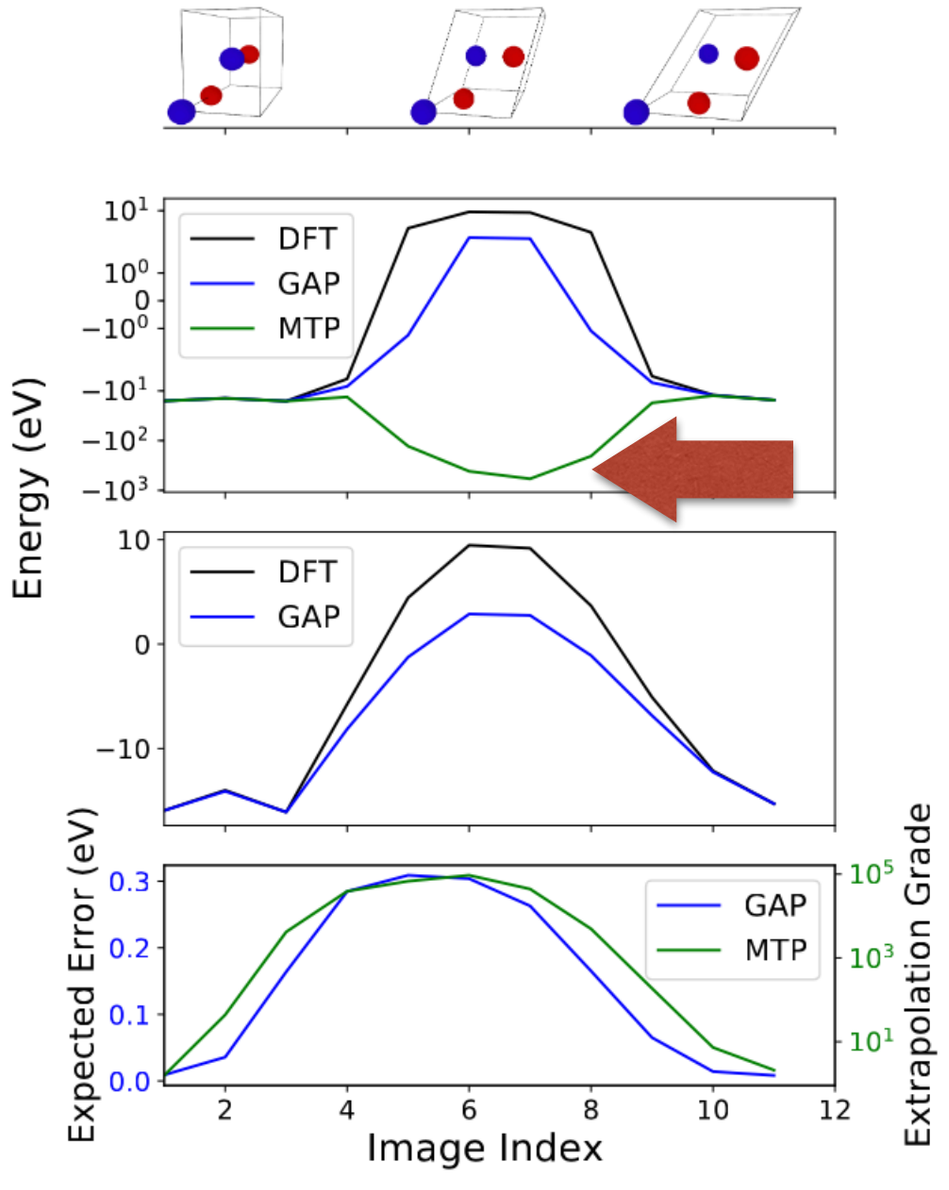
Liquid test



Phonon spectra



Transition Pathway



Applications

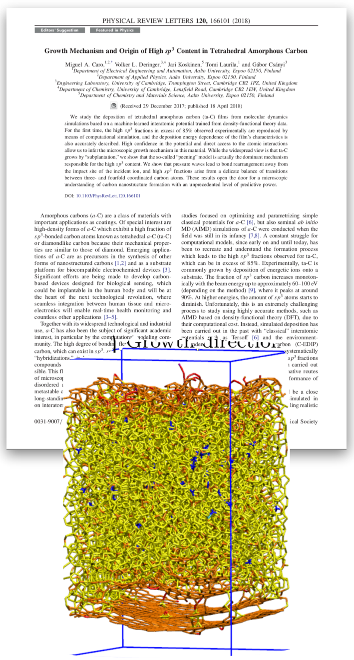
Successful *applications* to hard materials - 2018

With Volker Deringer
(now at Oxford)

Successful *applications* to hard materials - 2018

With Volker Deringer
(now at Oxford)

C


PHYSICAL REVIEW LETTERS **120,** 166101 (2018)

Editors' Suggestion | Featured in Physics

Growth Mechanism and Origin of High sp^3 Content in Tetrahedral Amorphous Carbon

Miguel A. Caro,[1,2,*] Volker L. Deringer,[3,4] Jari Koskinen,[5] Tomi Laurila,[1] and Gábor Csányi[3]
[1]*Department of Electrical Engineering and Automation, Aalto University, Espoo 02150, Finland*
[2]*Department of Applied Physics, Aalto University, Espoo 02150, Finland*
[3]*Engineering Laboratory, University of Cambridge, Trumpington Street, Cambridge CB2 1PZ, United Kingdom*
[4]*Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, United Kingdom*
[5]*Department of Chemistry and Materials Science, Aalto University, Espoo 02150, Finland*

(Received 29 December 2017; published 18 April 2018)

We study the deposition of tetrahedral amorphous carbon (ta-C) films from molecular dynamics simulations based on a machine-learned interatomic potential trained from density-functional theory data. For the first time, the high sp^3 fractions in excess of 85% observed experimentally are reproduced by means of computational simulation, and the deposition energy dependence of the film's characteristics is also accurately described. High confidence in the potential and direct access to the atomic interactions allow us to infer the microscopic growth mechanism in this material. While the widespread view is that ta-C grows by "subplantation," we show that the so-called "peening" model is actually the dominant mechanism responsible for the high sp^3 content. We show that pressure waves lead to bond rearrangement away from the impact site of the incident ion, and high sp^3 fractions arise from a delicate balance of transitions between three- and fourfold coordinated carbon atoms. These results open the door for a microscopic understanding of carbon nanostructure formation with an unprecedented level of predictive power.

DOI: 10.1103/PhysRevLett.120.166101

Amorphous carbons (*a*-C) are a class of materials with important applications as coatings. Of special interest are high-density forms of *a*-C which exhibit a high fraction of sp^3-bonded carbon atoms known as tetrahedral *a*-C (ta-C) or diamondlike carbon because their mechanical properties are similar to those of diamond. Emerging applications of *a*-C are as precursors in the synthesis of other forms of nanostructured carbons [1,2] and as a substrate platform for biocompatible electrochemical devices [3]. Significant efforts are being made to develop carbon-based devices designed for biological sensing, which could be implantable in the human body and will be at the heart of the next technological revolution, where seamless integration between human tissue and microelectronics will enable real-time health monitoring and countless other applications [3–5].

Together with its widespread technological and industrial use, *a*-C has also been the subject of significant academic interest, in particular by the computational modeling community. The high degree of bondin[…]

studies focused on optimizing and parametrizing simple classical potentials for *a*-C [6], but also seminal *ab initio* MD (AIMD) simulations of *a*-C were conducted when the field was still in its infancy [7,8]. A constant struggle for computational models, since early on and until today, has been to recreate and understand the formation process which leads to the high sp^3 fractions observed for ta-C, which can be in excess of 85%. Experimentally, ta-C is commonly grown by deposition of energetic ions onto a substrate. The fraction of sp^3 carbon increases monotonically with the beam energy up to approximately 60–100 eV (depending on the method) [9], where it peaks at around 90%. At higher energies, the amount of sp^3 atoms starts to diminish. Unfortunately, this is an extremely challenging process to study using highly accurate methods, such as AIMD based on density-functional theory (DFT), due to their computational cost. Instead, simulated deposition has been carried out in the past with "classical" interatomic […]

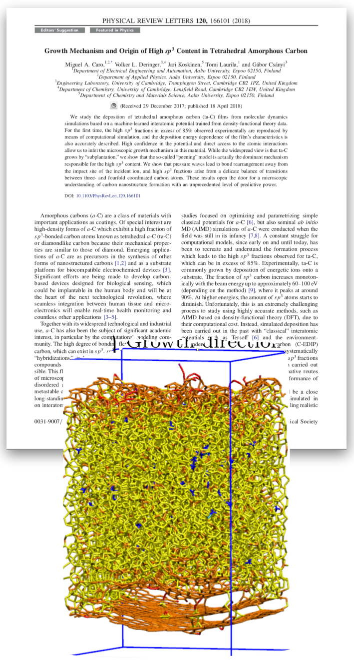
Structure and growth mechanism of amorphous carbon films

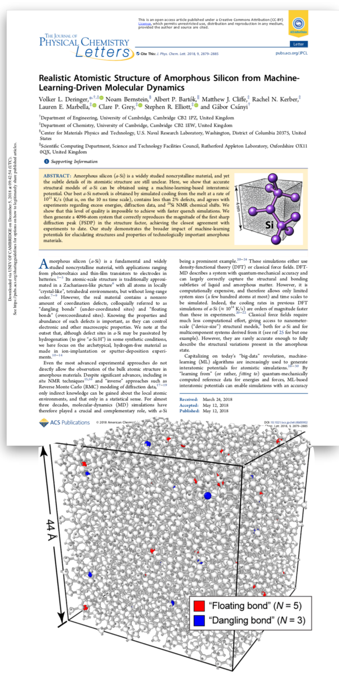
Successful *applications* to hard materials - 2018

With Volker Deringer
(now at Oxford)

C

Si





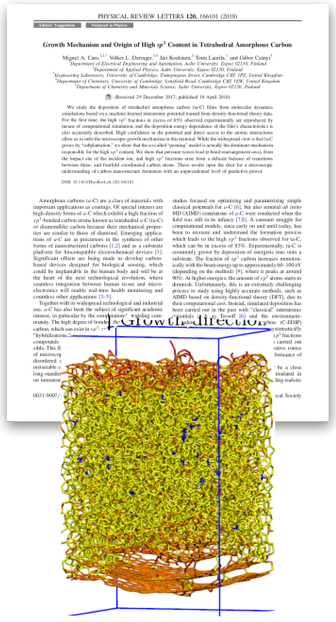
Structure and growth mechanism of amorphous carbon films

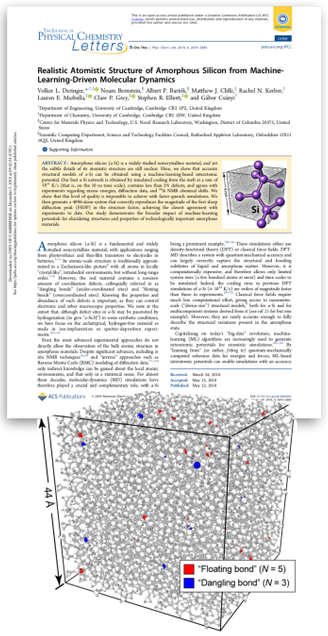
Realistic structure of slow-quenched amorphous silicon

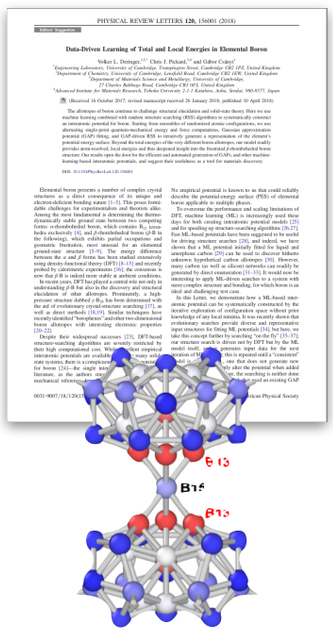
Successful *applications* to hard materials - 2018

With Volker Deringer
(now at Oxford)

C Si B







Structure and growth mechanism of amorphous carbon films

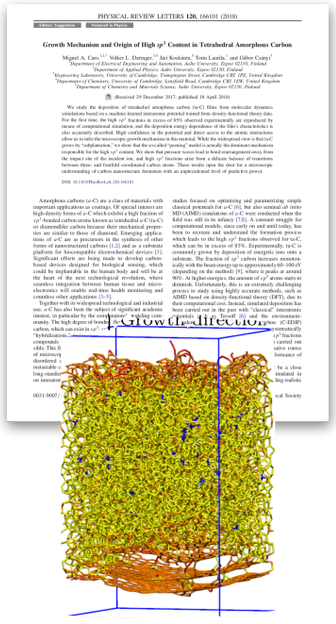
First potential for boron - the search for the elusive β phase

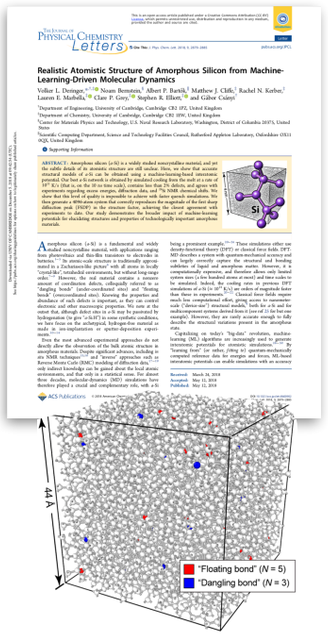
Realistic structure of slow-quenched amorphous silicon

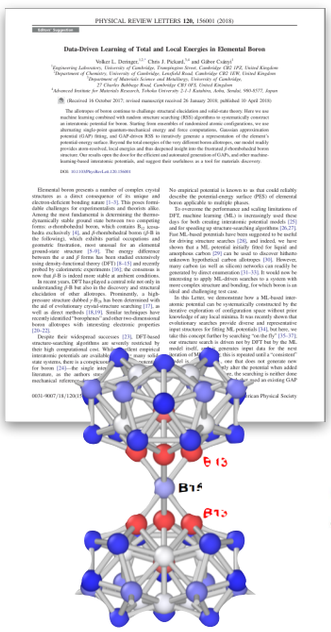
Successful *applications* to hard materials - 2018

With Volker Deringer
(now at Oxford)

C Si B P







Structure and growth mechanism of amorphous carbon films

First potential for boron - the search for the elusive β phase

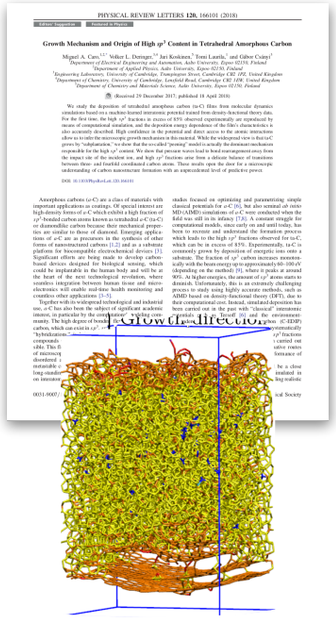
Realistic structure of slow-quenched amorphous silicon

Crystal structure search, fibrous phosphorous

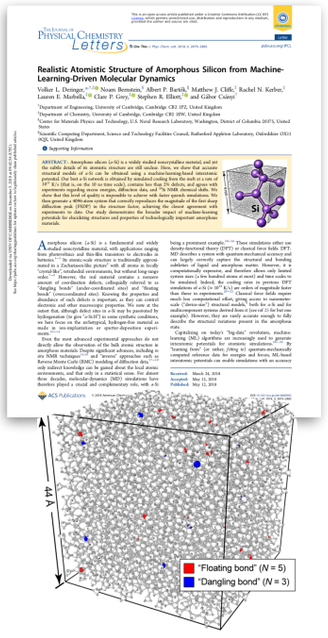
Successful *applications* to hard materials - 2018

With Volker Deringer
(now at Oxford)

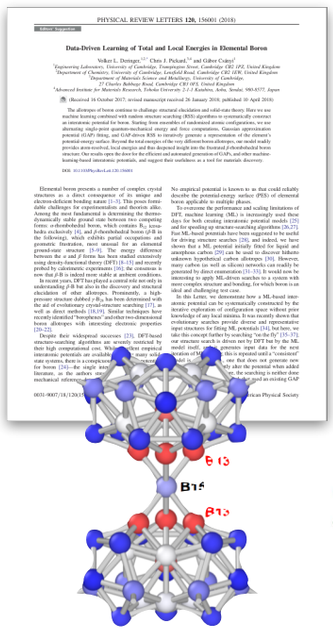
C Si B P Fe



Structure and growth mechanism of amorphous carbon films



Realistic structure of slow-quenched amorphous silicon



First potential for boron - the search for the elusive β phase

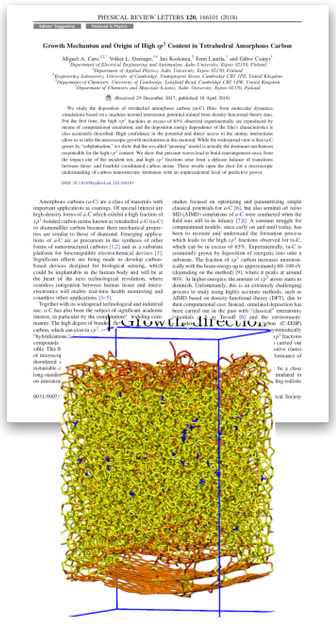
Crystal structure search, fibrous phosphorous

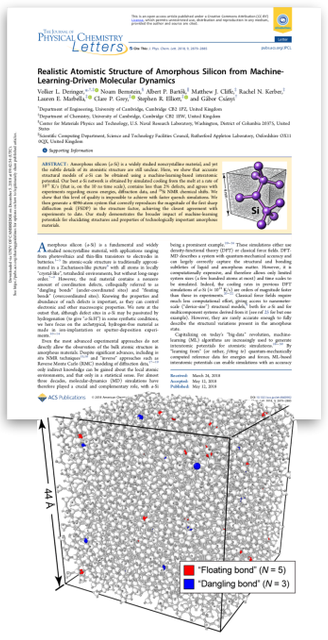
Screw dislocation glide in bcc iron

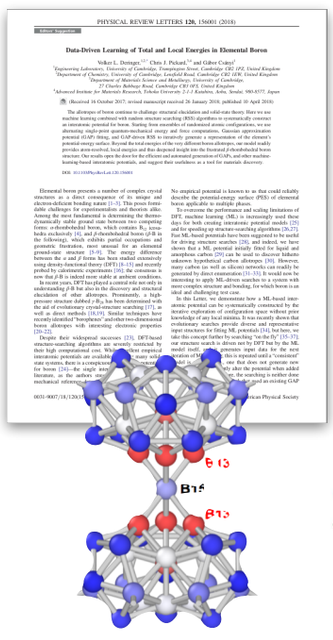
Successful *applications* to hard materials - 2018

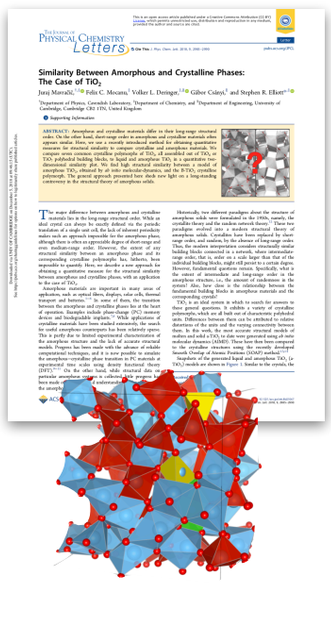
With Volker Deringer
(now at Oxford)

C Si B P Fe TiO_2









Structure and growth mechanism of amorphous carbon films

First potential for boron - the search for the elusive β phase

Screw dislocation glide in bcc iron

Realistic structure of slow-quenched amorphous silicon

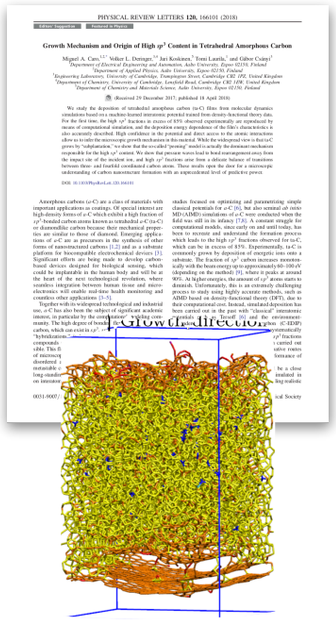
Crystal structure search, fibrous phosphorous

Crystalline and amorphous phases of titanium oxide

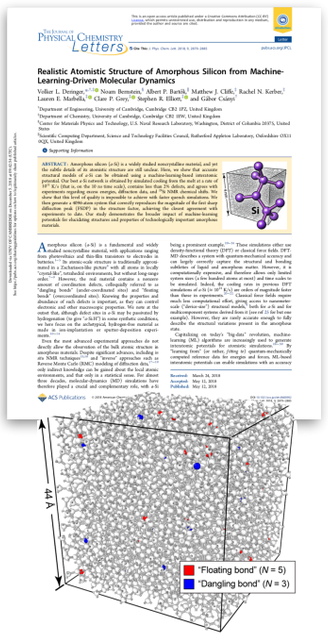
Successful *applications* to hard materials - 2018

With Volker Deringer
(now at Oxford)

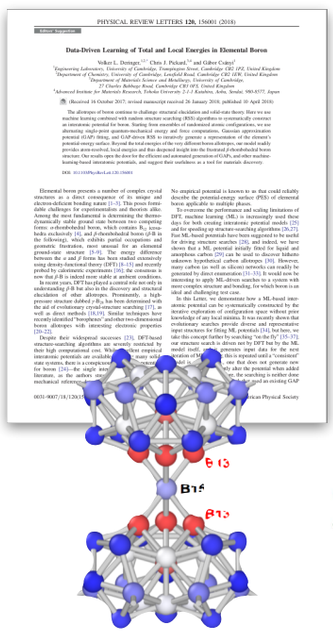
C Si B P Fe TiO_2 GeSbTe



Structure and growth mechanism of amorphous carbon films



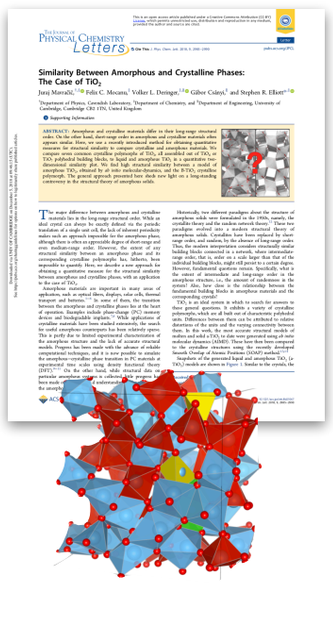
Realistic structure of slow-quenched amorphous silicon



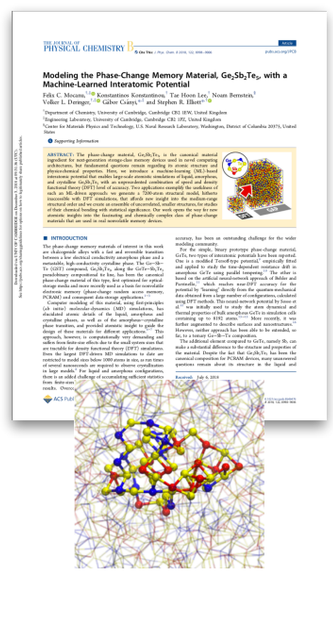
First potential for boron - the search for the elusive β phase

Crystal structure search, fibrous phosphorous

Screw dislocation glide in bcc iron



Crystalline and amorphous phases of titanium oxide



Medium-range order in the phase change material GST

The Silicon Challenge

PHYSICAL REVIEW X

Highlights Recent Subjects Accepted Collections Authors Referees Search

Open Access

Machine Learning a General-Purpose Interatomic Potential for Silicon

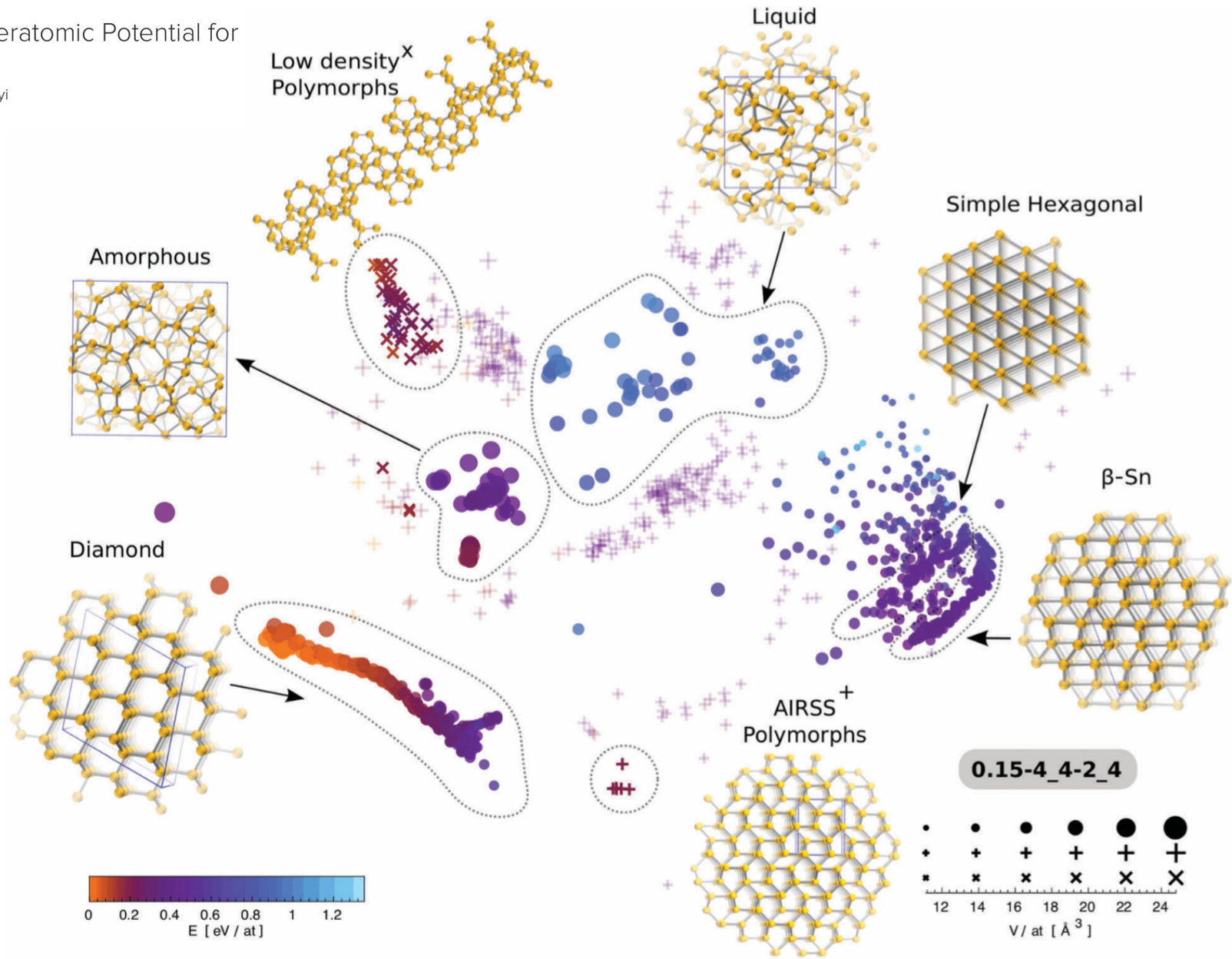
Albert P. Bartók, James Kermode, Noam Bernstein, and Gábor Csányi
Phys. Rev. X **8**, 041048 – Published 14 December 2018

CrossMark click for updates

Cite this: *Phys. Chem. Chem. Phys.*, 2016, **18**, 13754

Comparing molecules and solids across structural and alchemical space†

Sandip De,[ab] Albert P. Bartók,[c] Gábor Csányi[c] and Michele Ceriotti*[ab]

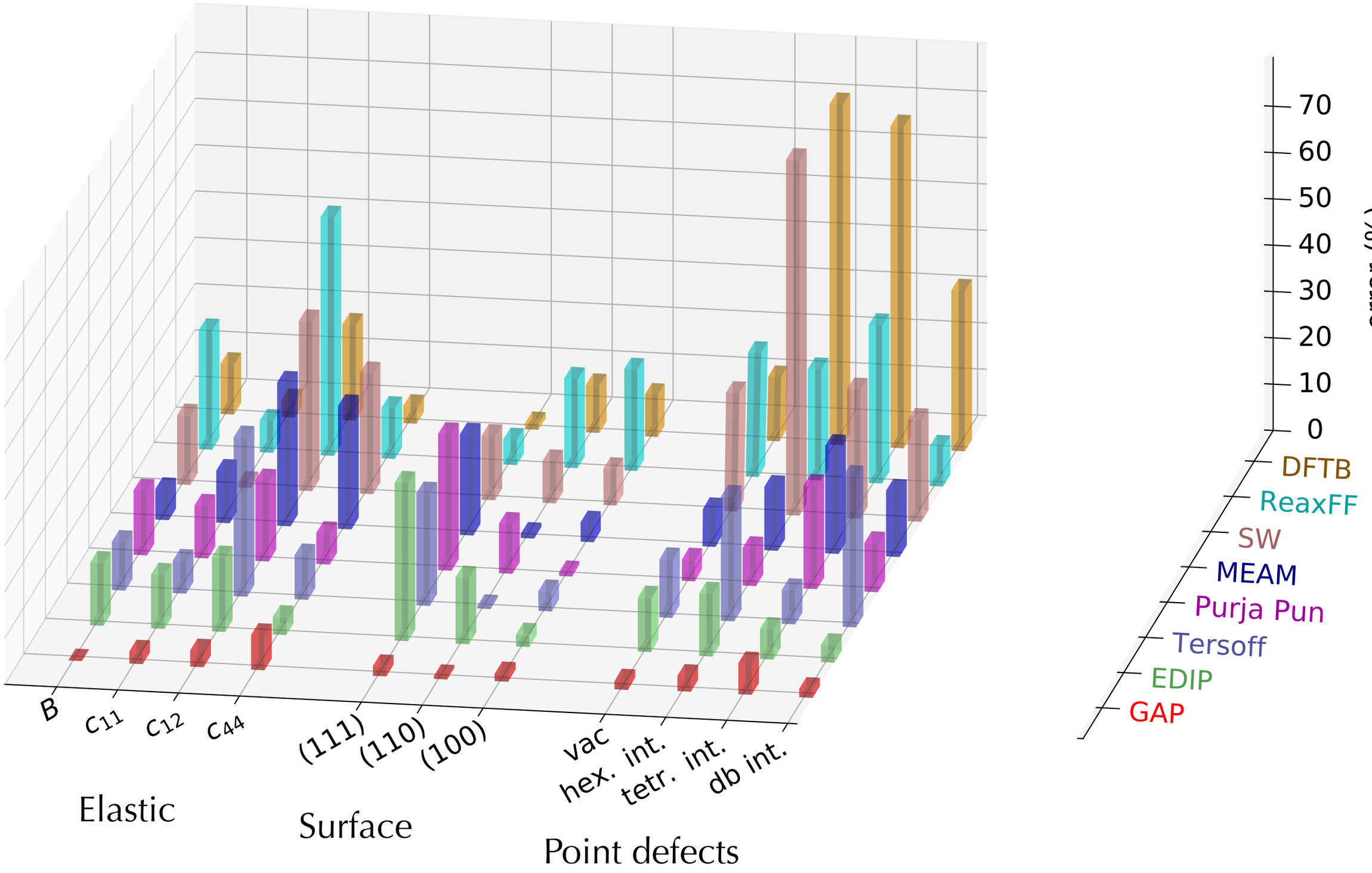


Can we cover **all** relevant configurations?

The monster database

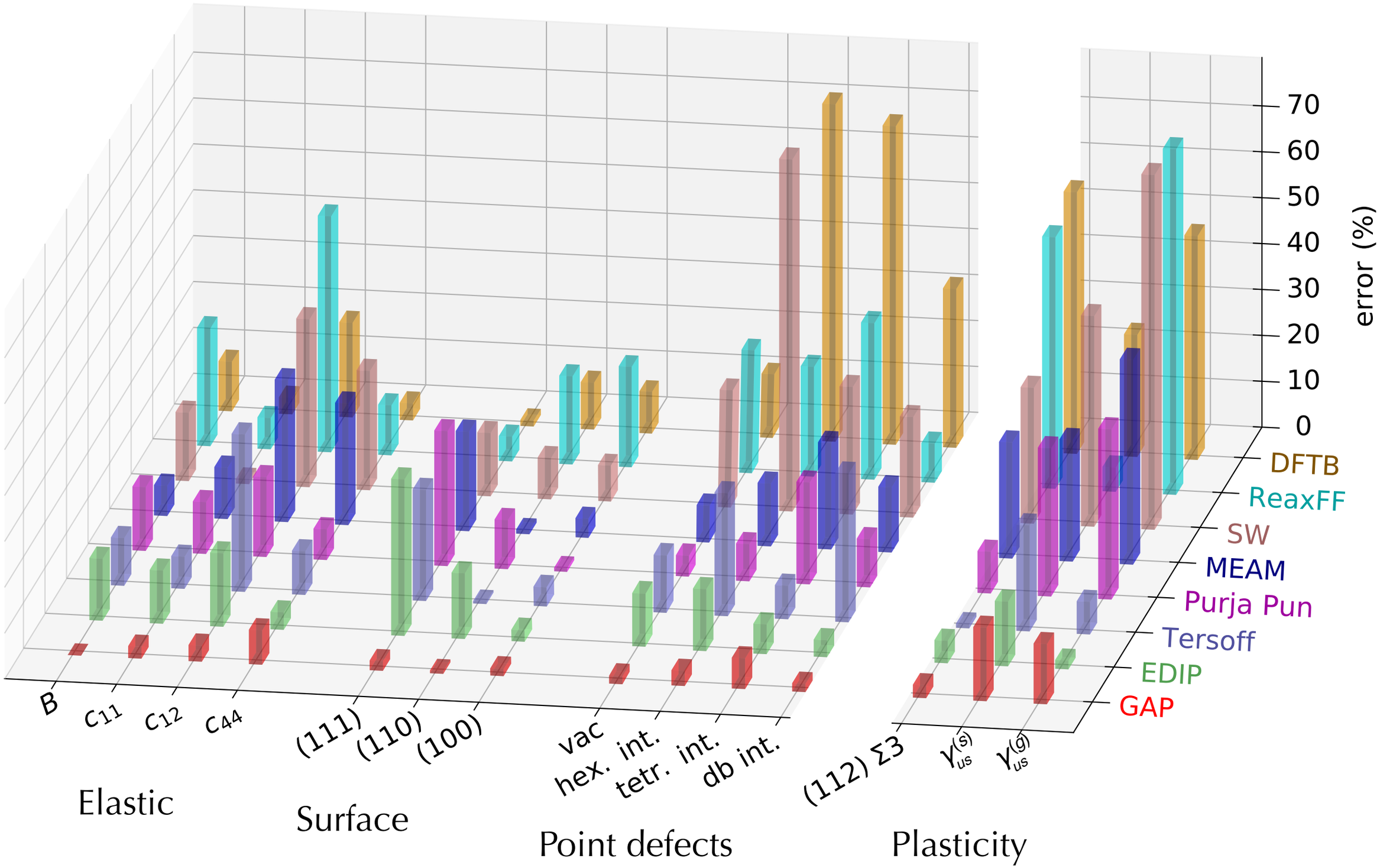
Structure type	# atoms	# structures	# environments	# representative pts	σ_{energy}	σ_{force}	σ_{virial}
					default values: 0.001	0.1	0.05
isolated atom	1	1	1	1			
diamond	2	104	208	6			
	16	220	3520	53			
	54	110	5940	58			
	128	55	7040	92			
β-Sn	2	60	120	32			
	16	220	3520	51			
	54	110	5940	66			
	128	55	7040	157			
simple hexagonal	1	110	110	13			
	8	30	240	15			
	27	30	810	42			
	64	53	3392	89			
bcc	2	49	392	40			
bc8	8	49	88	66			
fcc	4	49	196	46			
hcp	2	49	88	28			
st12	12	49	588	94			
liquid	64	69	4416	1114	0.003	0.15	0.2
	128	7	896	323			
amorphous	64	31	1984	231	0.01	0.2	0.4
	216	128	27648	1719			
diamond surface (001)	144	29	4176	514			
decohesion	32	11	352	28			
diamond surface (110)	108	26	2808	338			
decohesion	16	11	176	8			
diamond surface (111)							
decohesion	24	11	264	10			
unreconstructed	96	47	4512	573			
adatom	146	11	1606	62			
Pandey reconstruction	96	50	4800	632			
DAS 3x3 unrelaxed	52	1	52	6			
diamond vacancy	63	100	6300	168			
	215	111	23865	405			
diamond divacancy	214	78	16692	416			
diamond interstitial	217	115	24955	605			
small (110) crack tip	200	7	1400	130			
small (111) crack tip	192	10	1920	185			
screw dislocation core	144	19	2736	124			
sp^2 bonded	8	51	408	61			
sp bonded	4	100	400	392	0.01	0.2	0.4
Total							

Summary of material properties



Properties targeted by the training database

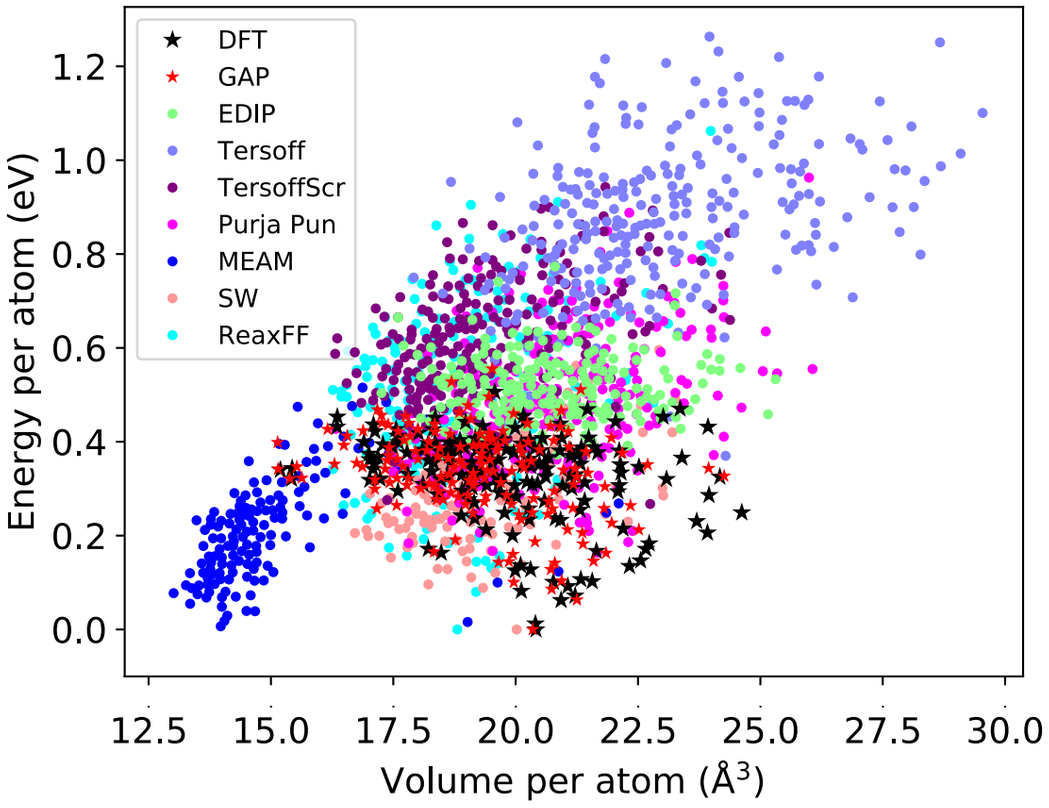
Summary of material properties



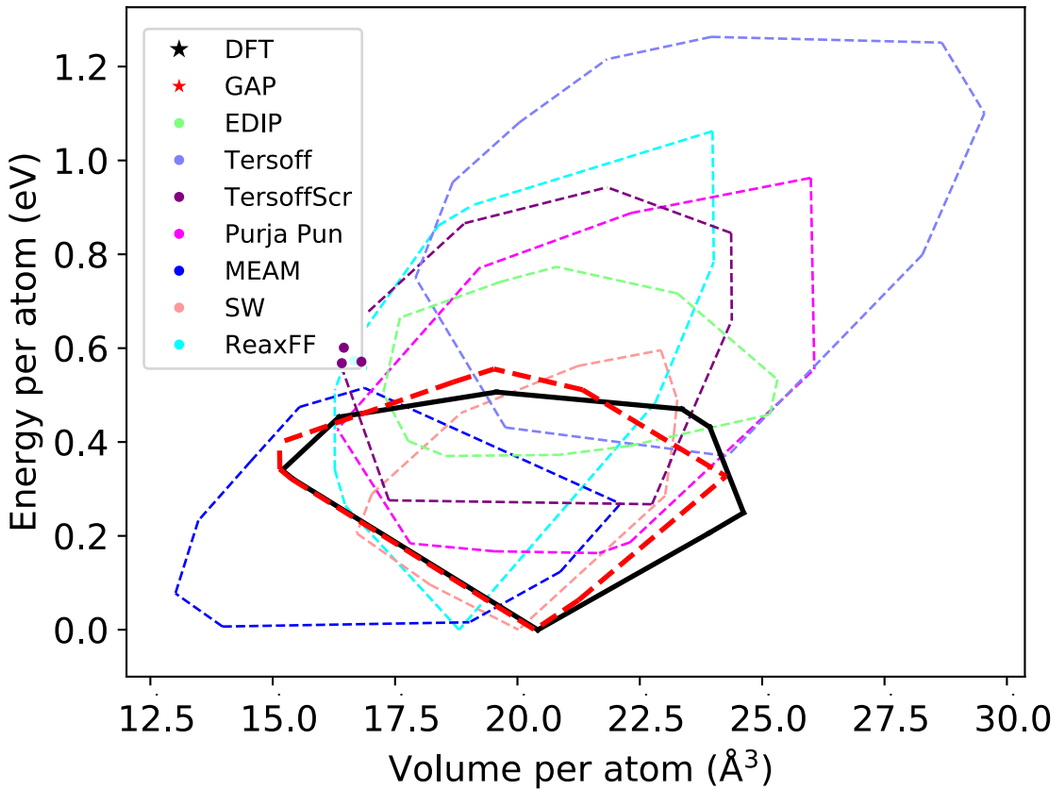
Properties targeted by the training database

Untargeted properties

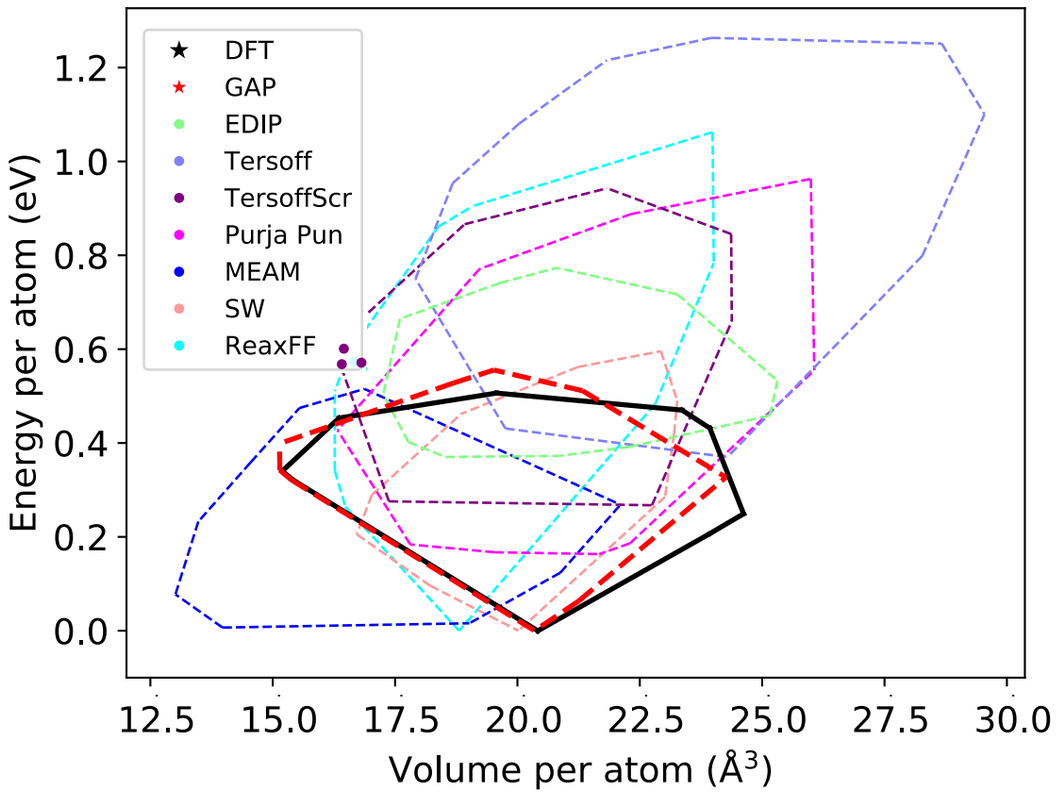
Stringent tests: crystal structure search and phases



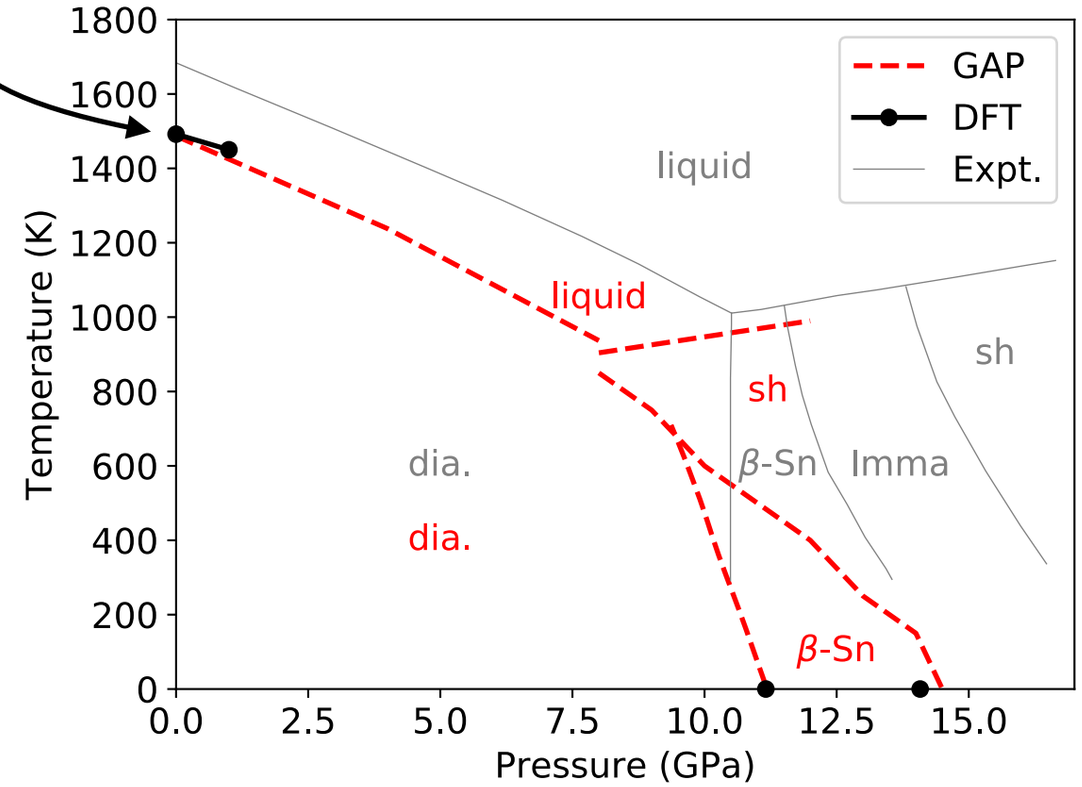
Stringent tests: crystal structure search and phases



Stringent tests: crystal structure search and phases

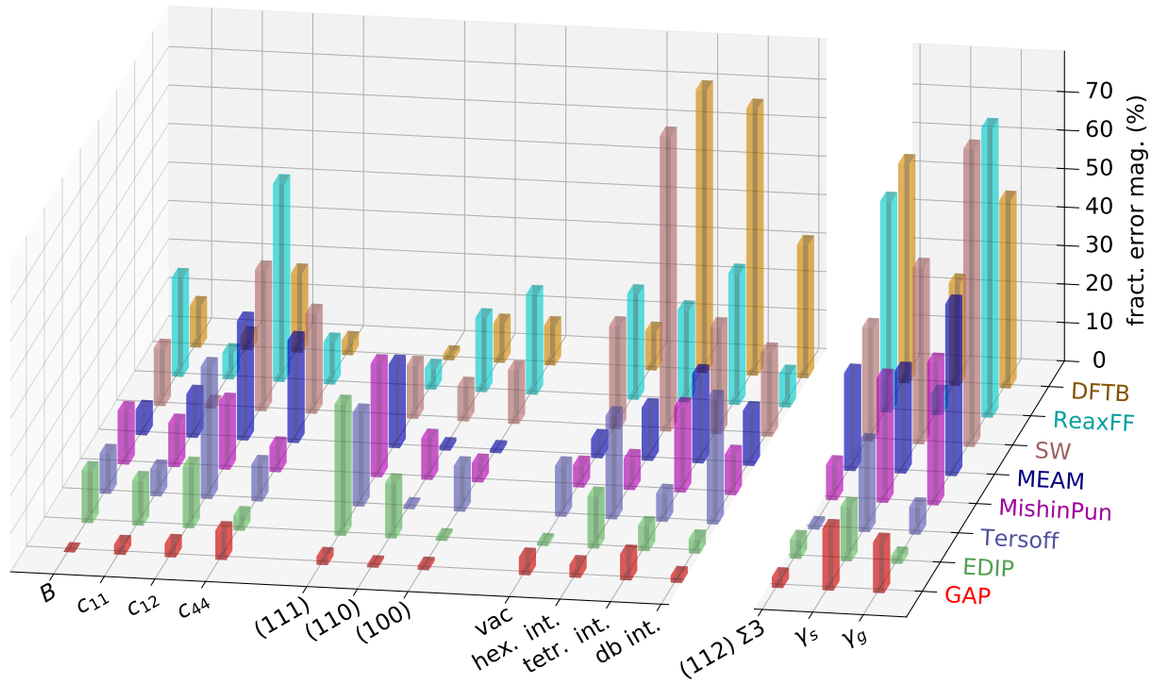


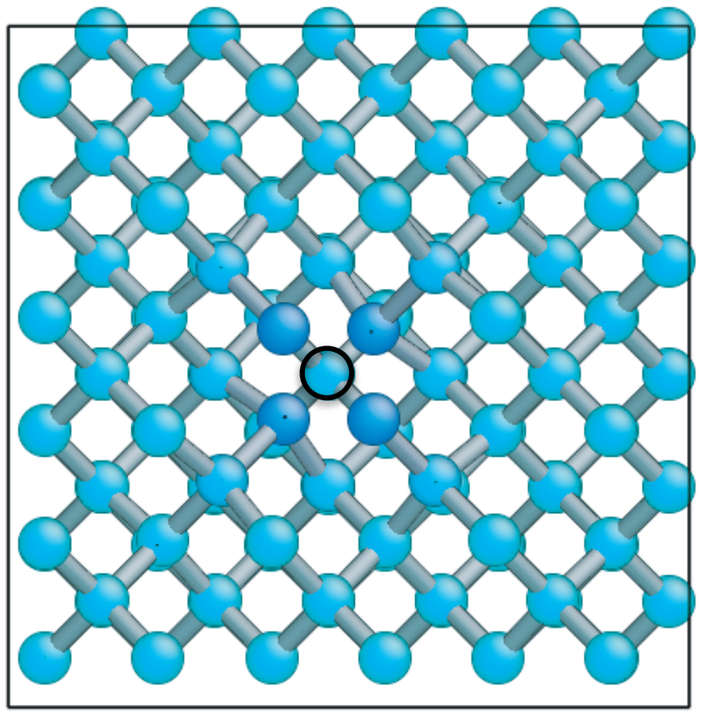
DFT : Many months on a supercomputer



Predicted errors from equivalent Gaussian Process

Predicted error^2 ~ Variance of GP



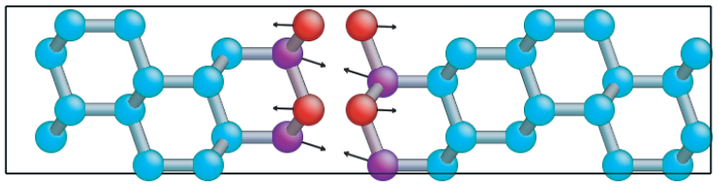


Vacancy

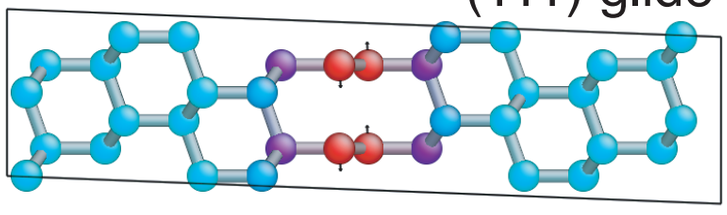
> 0.005 eV

< 0.001 eV

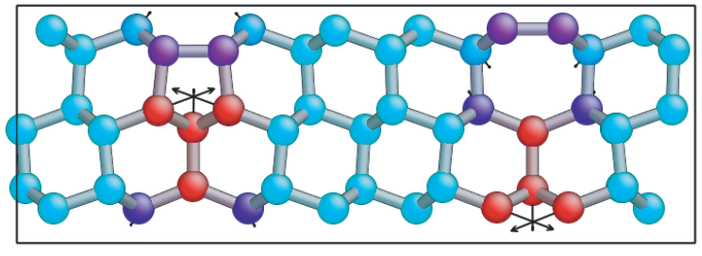
(111) shuffle



(111) glide



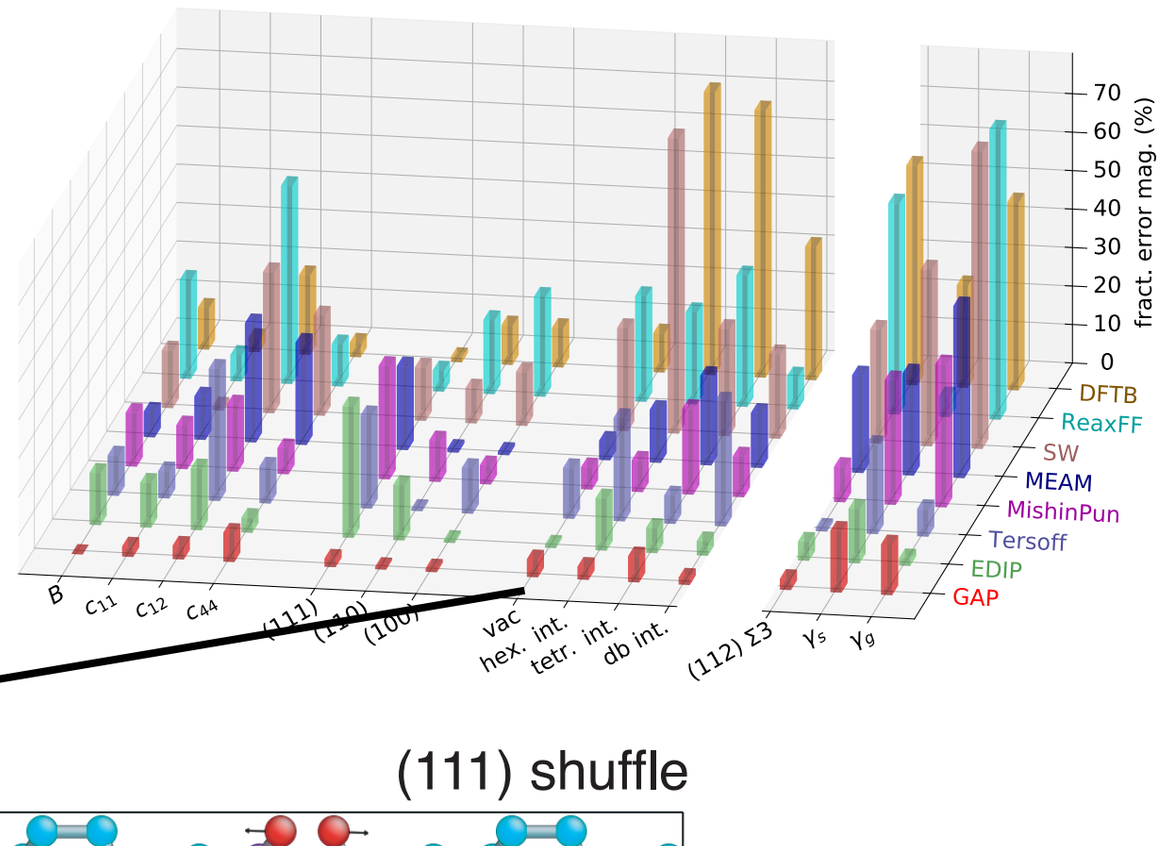
(112) Σ3

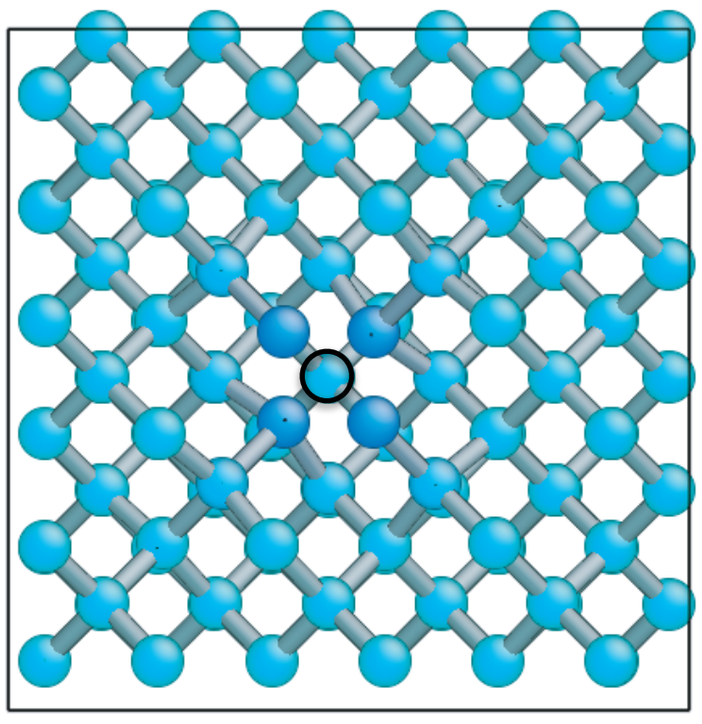


0.5 eV/Å ⟼

Predicted errors from equivalent Gaussian Process

Predicted error^2 ~ Variance of GP

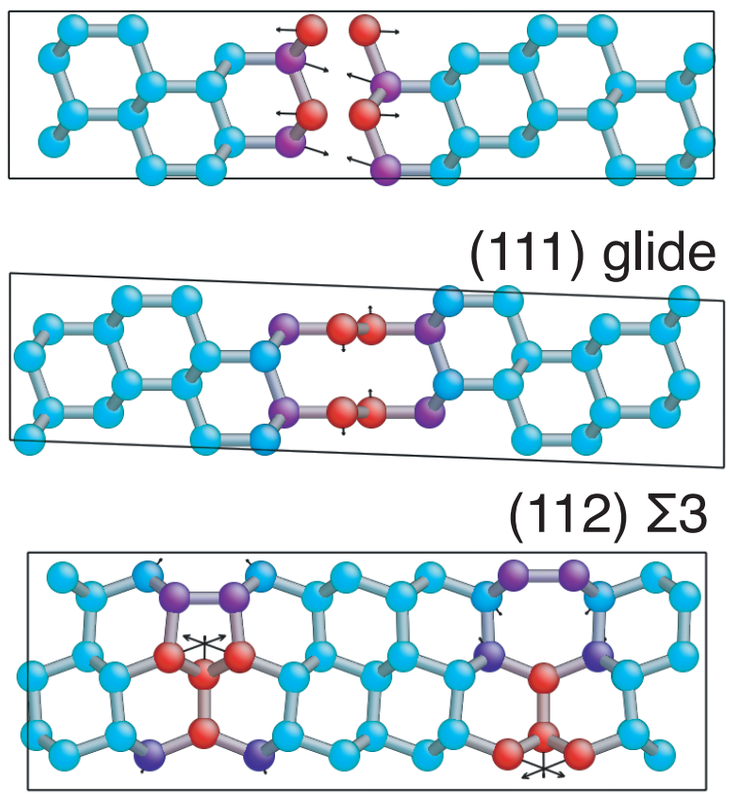




Vacancy

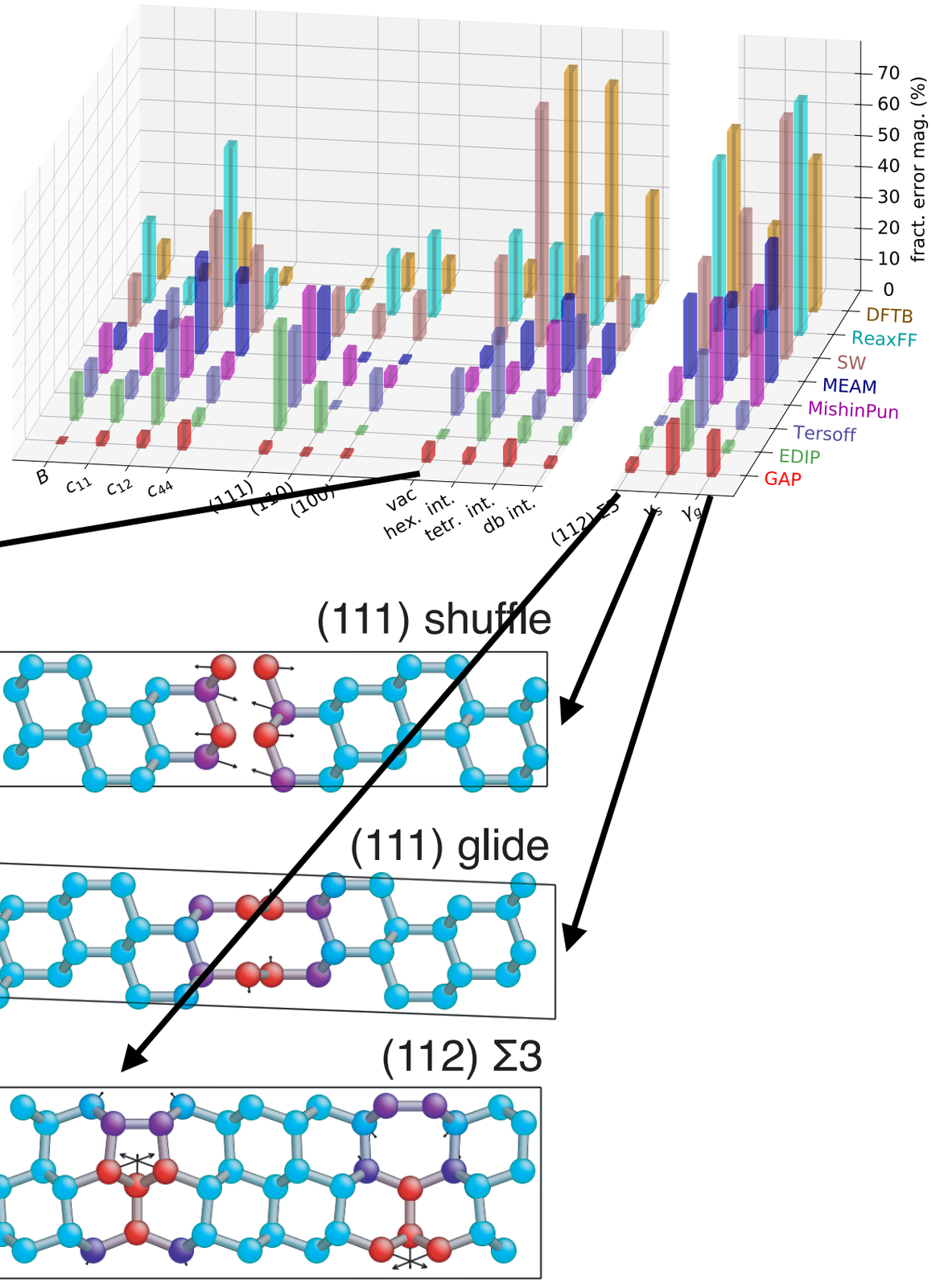
> 0.005 eV

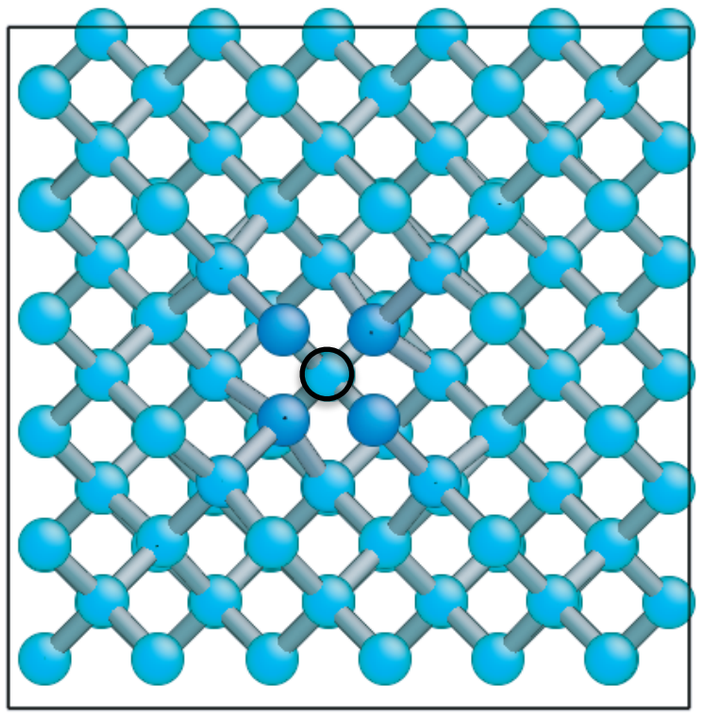
< 0.001 eV



Predicted errors from equivalent Gaussian Process

Predicted error^2 ~ Variance of GP





Vacancy

> 0.005 eV

< 0.001 eV

Application:

THE JOURNAL OF PHYSICAL CHEMISTRY Letters

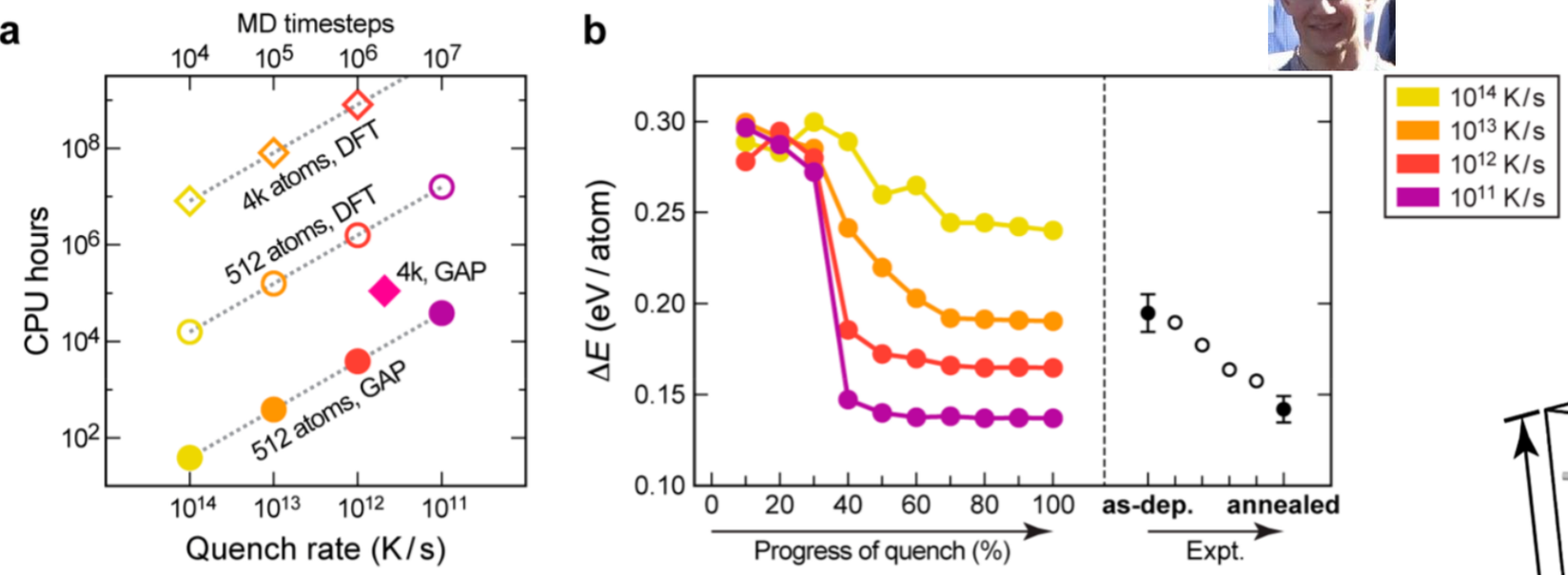
Letter

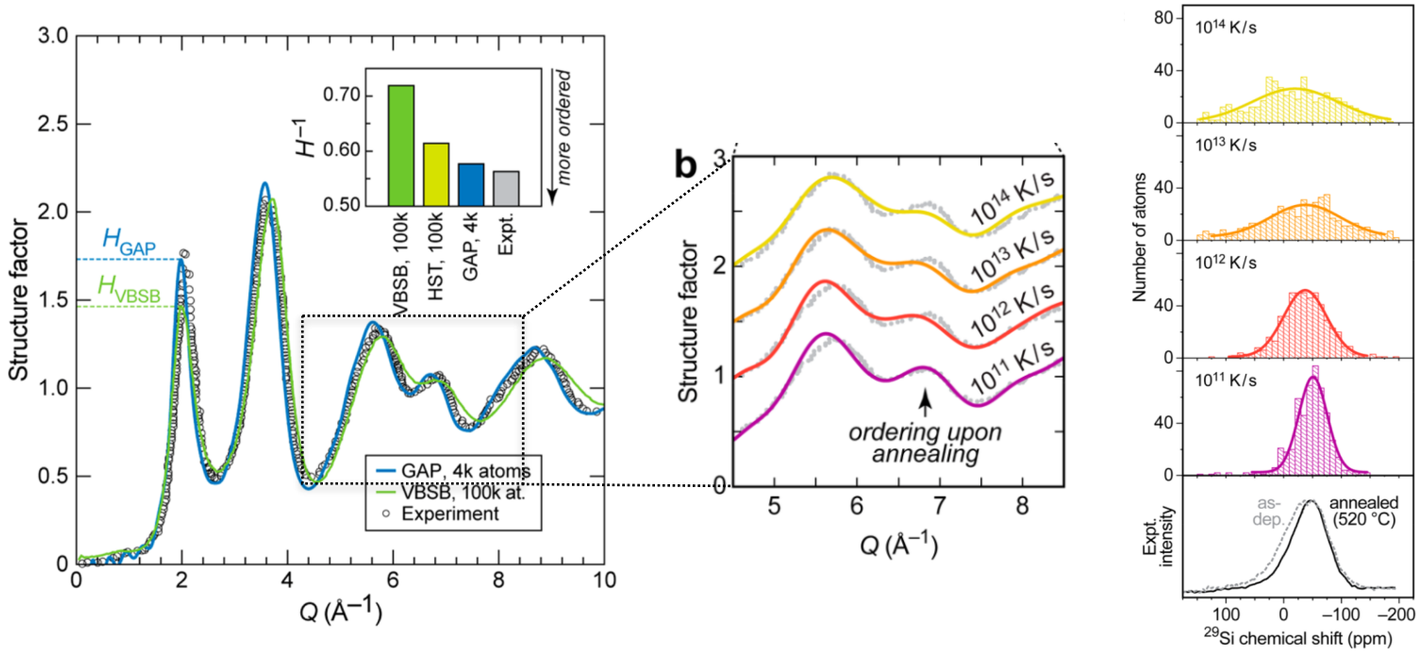
Cite This: J. Phys. Chem. Lett. 2018, 9, 2879−2885

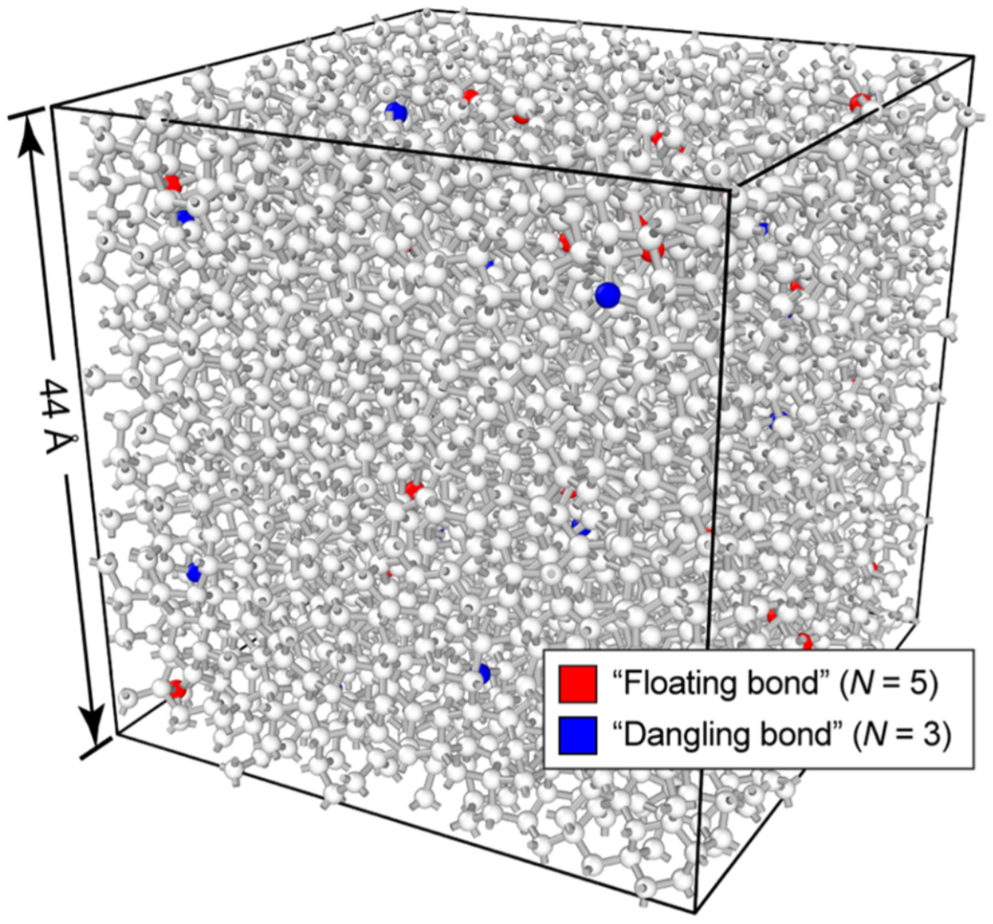
pubs.acs.org/JPCL

Realistic Atomistic Structure of Amorphous Silicon from Machine-Learning-Driven Molecular Dynamics

Volker L. Deringer,*[,†,‡] Noam Bernstein,[§] Albert P. Bartók,[∥] Matthew J. Cliffe,[‡] Rachel N. Kerber,[‡] Lauren E. Marbella,[‡] Clare P. Grey,[‡] Stephen R. Elliott,[‡] and Gábor Csányi[†]







Application:

THE JOURNAL OF PHYSICAL CHEMISTRY Letters

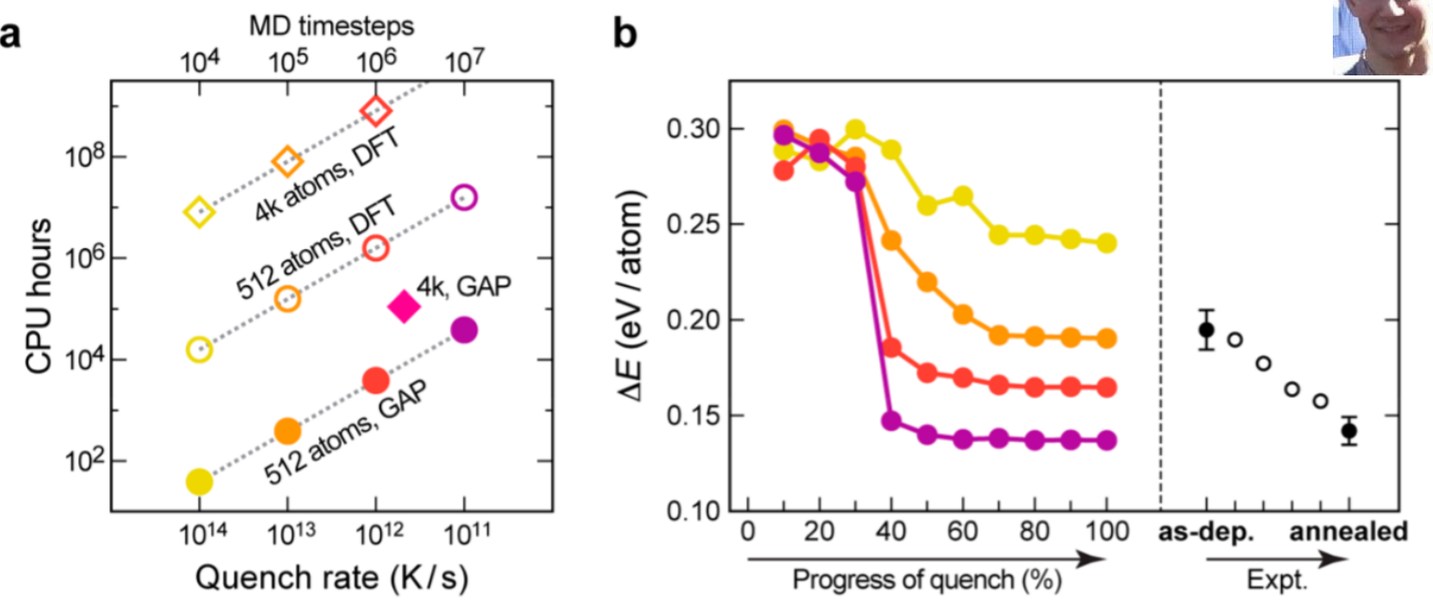
Letter

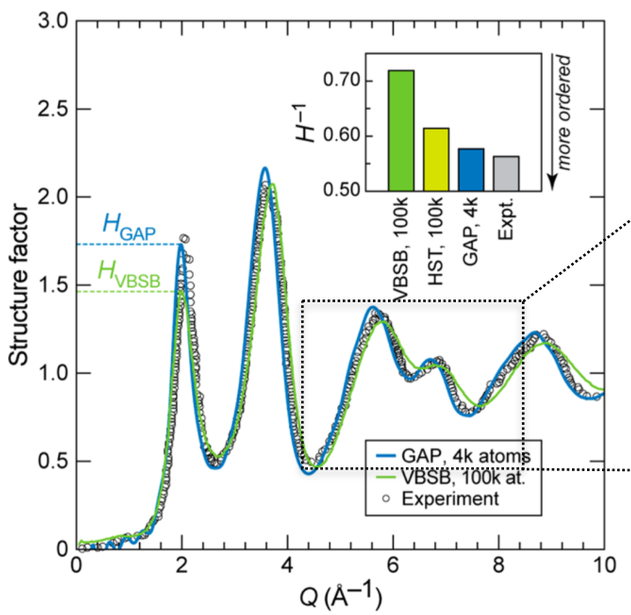
Cite This: J. Phys. Chem. Lett. 2018, 9, 2879–2885

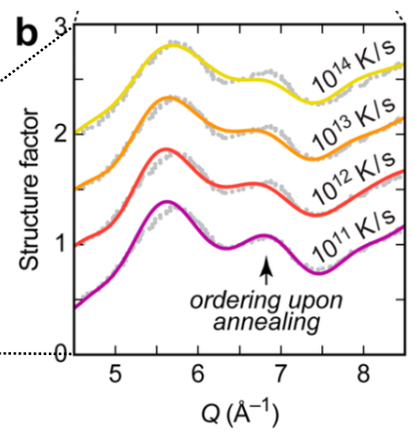
pubs.acs.org/JPCL

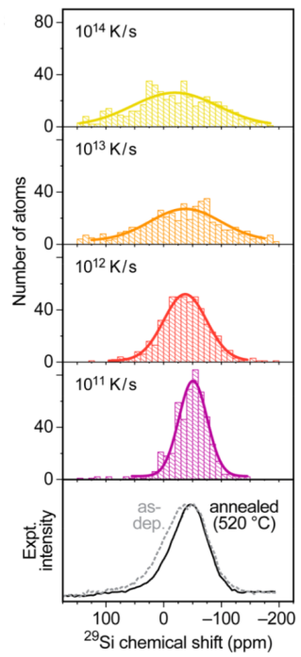
Realistic Atomistic Structure of Amorphous Silicon from Machine-Learning-Driven Molecular Dynamics

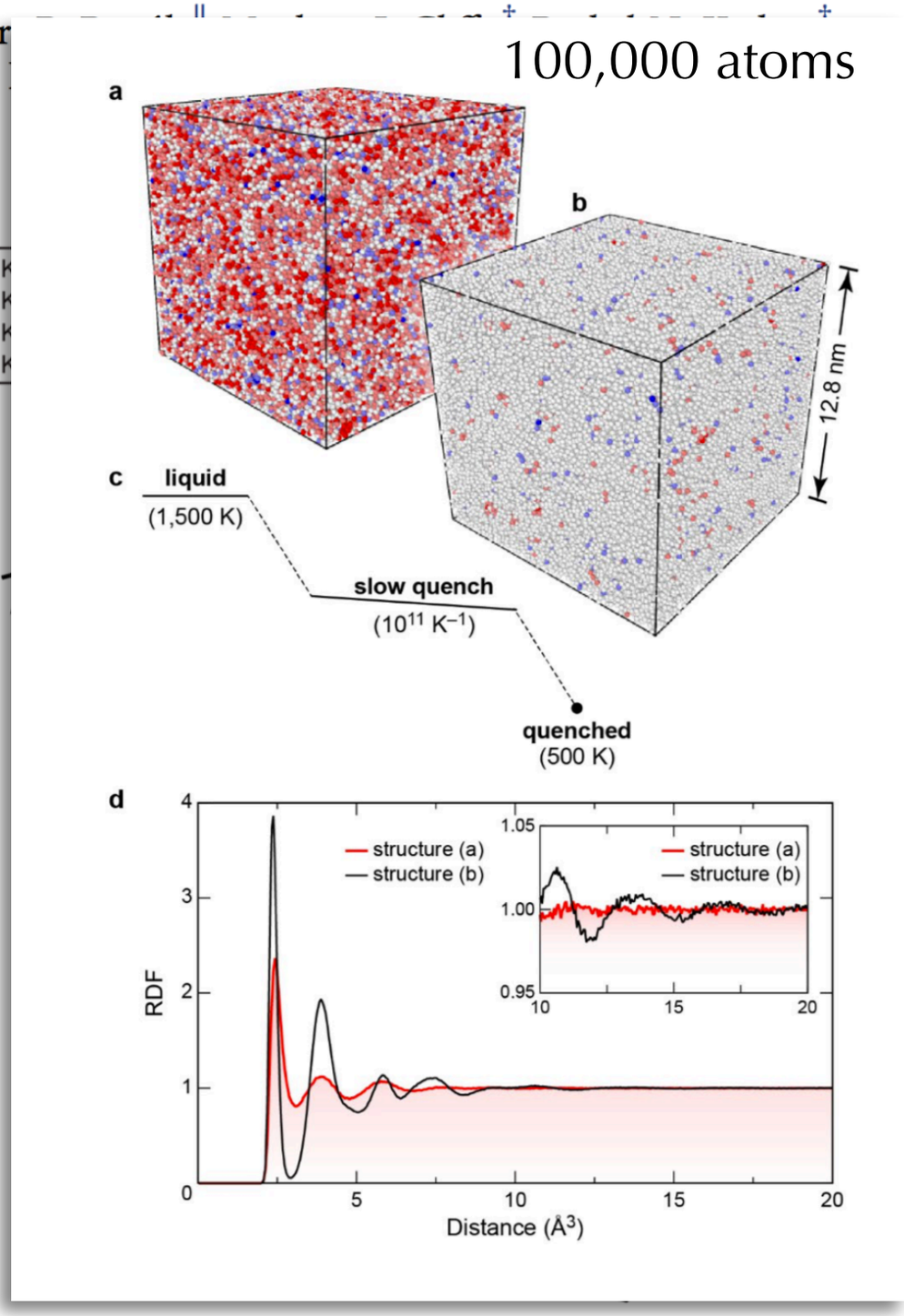
Volker L. Deringer,*,†,‡ Noam Bernstein,§ Alber
Lauren E. Marbella,‡ Clare P. Grey,‡ Stephen











Application:

THE JOURNAL OF PHYSICAL CHEMISTRY Letters

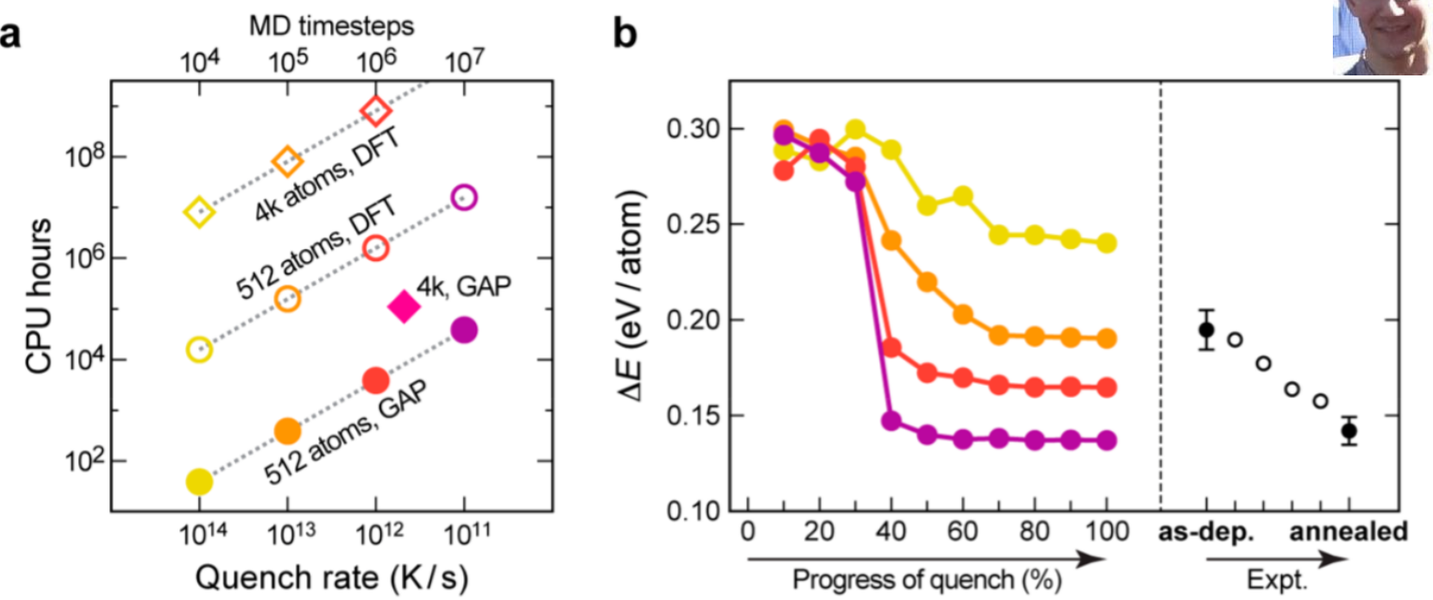
Letter

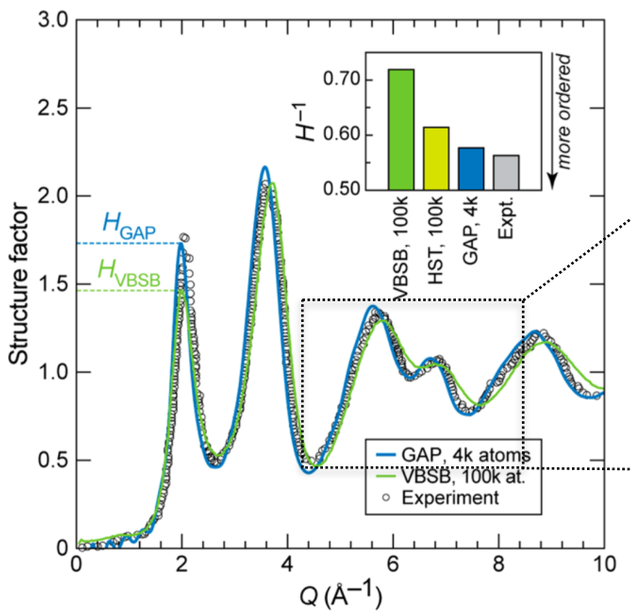
Cite This: J. Phys. Chem. Lett. 2018, 9, 2879–2885

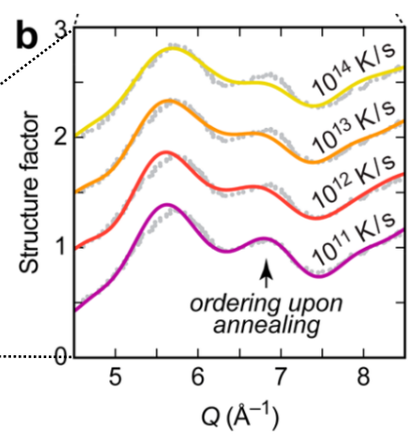
pubs.acs.org/JPCL

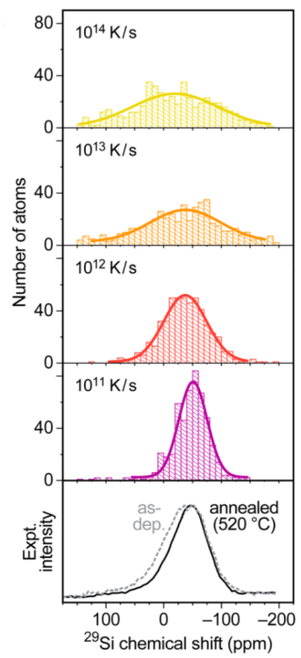
Realistic Atomistic Structure of Amorphous Silicon from Machine-Learning-Driven Molecular Dynamics

Volker L. Deringer,*,†,‡ Noam Bernstein,§ Alber
Lauren E. Marbella,‡ Clare P. Grey,‡ Stephen

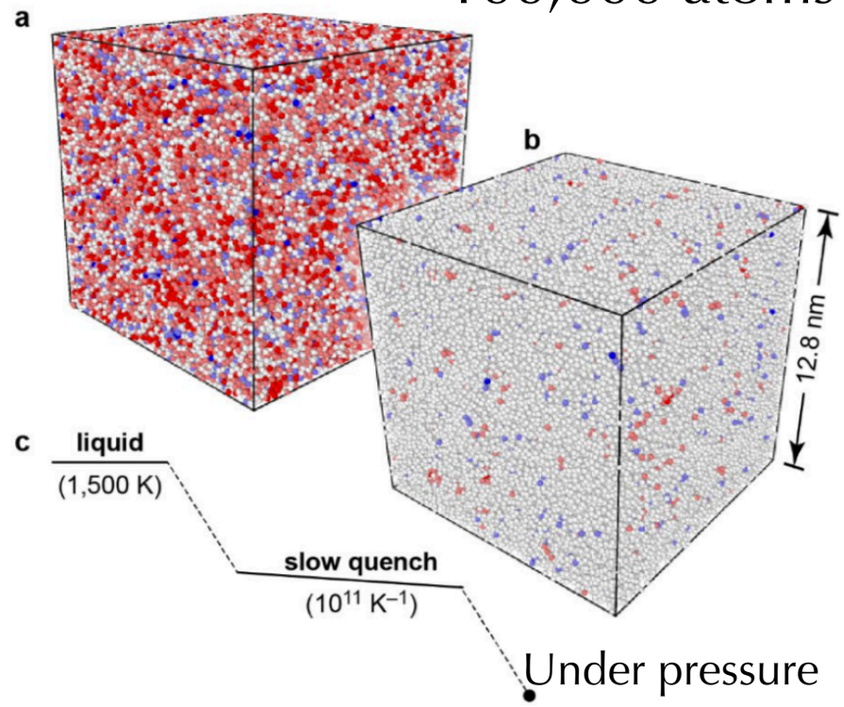




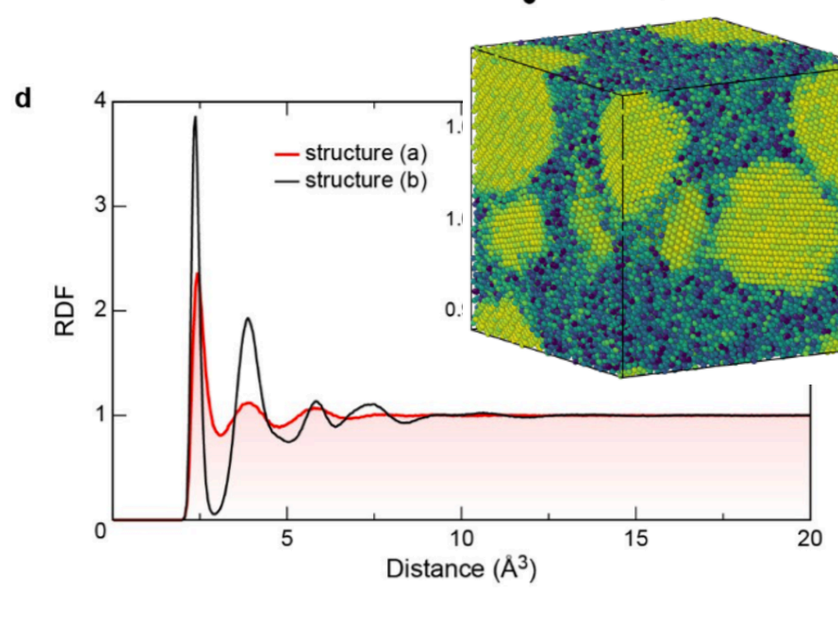




100,000 atoms



Under pressure



Fluid methane

- Dominated by weak dispersion interactions
- 2-body enough or need many-body effects?
- What level of quantum mechanics is required?

ACS Publications
Most Trusted. Most Cited. Most Read.

Search text, DOI, authors, etc.

RETURN TO ISSUE | < PREV **ARTICLE** NEXT >

Equation of State of Fluid Methane from First Principles with Machine Learning Potentials

Max Veit*, Sandeep Kumar Jain, Satyanarayana Bonakala, Indranil Rudra, Detlef Hohl and Gábor Csányi

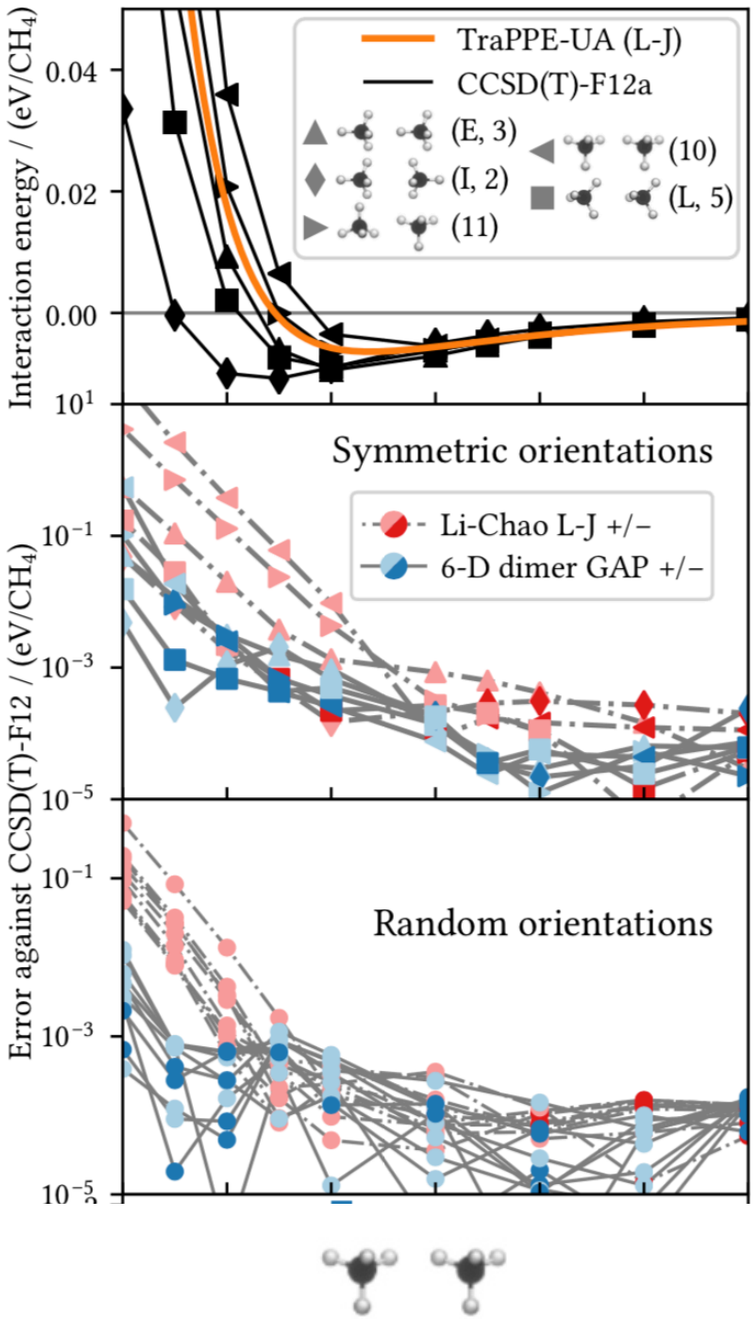
Cite This: *J. Chem. Theory Comput.* 2019, 15, 4, 2574-2586
Publication Date: February 22, 2019
https://doi.org/10.1021/acs.jctc.8b01242
Copyright © 2019 American Chemical Society
RIGHTS & PERMISSIONS

Article Views	Altmetric	Citations	Share	Add to
393	7	1		

LEARN ABOUT THESE METRICS

Fluid methane

- Dominated by weak dispersion interactions
- 2-body enough or need many-body effects?
- What level of quantum mechanics is required?



ACS Publications
Most Trusted. Most Cited. Most Read.

Search text, DOI, authors, etc.

RETURN TO ISSUE | < PREV **ARTICLE** NEXT >

Equation of State of Fluid Methane from First Principles with Machine Learning Potentials

Max Veit*, Sandeep Kumar Jain, Satyanarayana Bonakala, Indranil Rudra, Detlef Hohl and Gábor Csányi

Cite This: *J. Chem. Theory Comput.* 2019, 15, 4, 2574-2586

Publication Date: February 22, 2019

https://doi.org/10.1021/acs.jctc.8b01242

Copyright © 2019 American Chemical Society

RIGHTS & PERMISSIONS

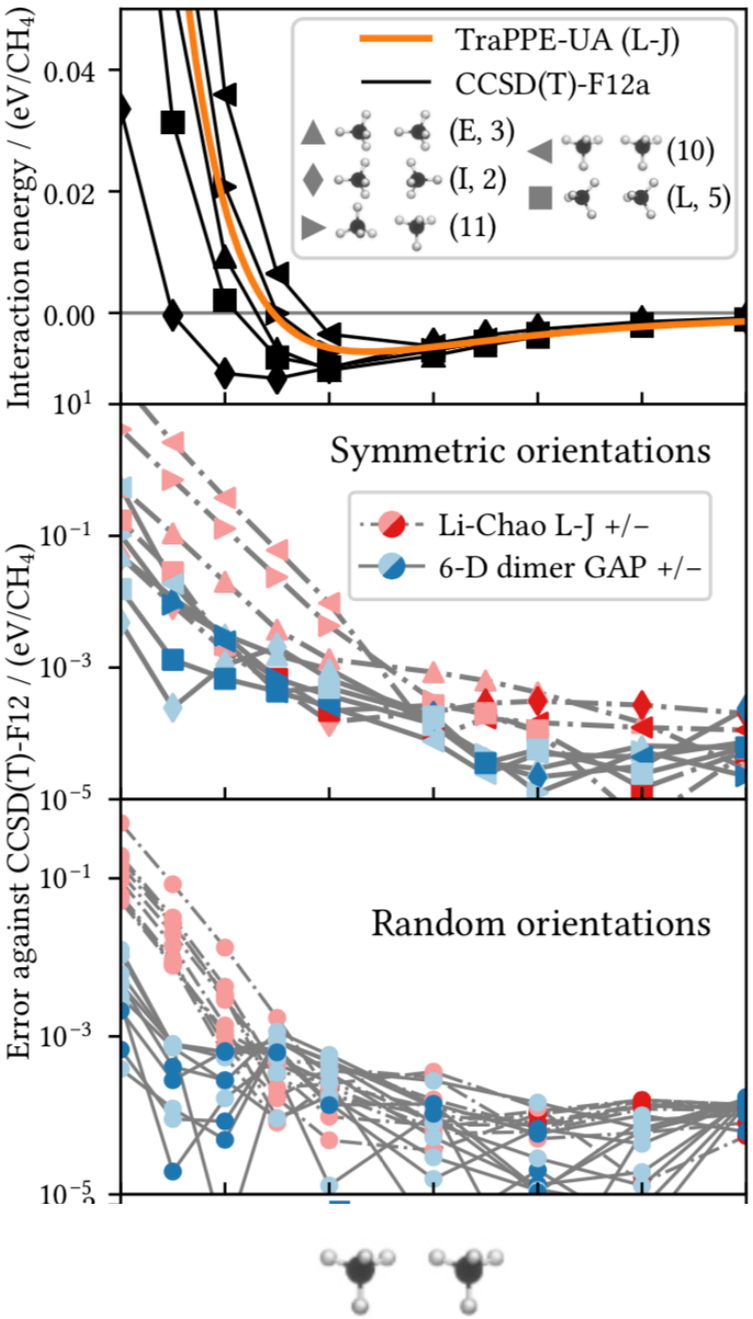
Article Views	Altmetric	Citations
393	7	1

LEARN ABOUT THESE METRICS

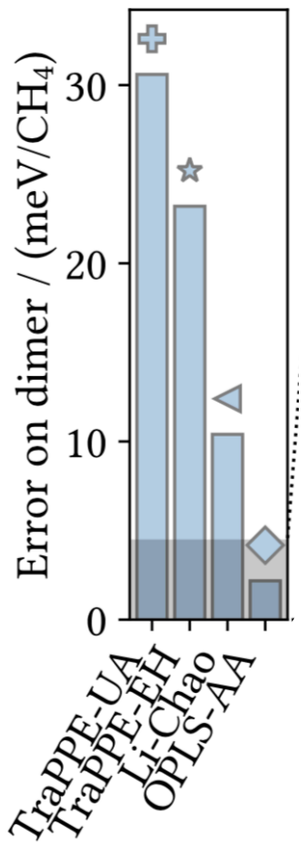
Share Add to

Fluid methane

- Dominated by weak dispersion interactions
- 2-body enough or need many-body effects?
- What level of quantum mechanics is required?



Empirical models



ACS Publications
Most Trusted. Most Cited. Most Read.

Search text, DOI, authors, etc.

RETURN TO ISSUE | < PREV **ARTICLE** NEXT >

Equation of State of Fluid Methane from First Principles with Machine Learning Potentials

Max Veit*, Sandeep Kumar Jain, Satyanarayana Bonakala, Indranil Rudra, Detlef Hohl and Gábor Csányi

Cite This: *J. Chem. Theory Comput.* 2019, 15, 4, 2574-2586

Publication Date: February 22, 2019

https://doi.org/10.1021/acs.jctc.8b01242

Copyright © 2019 American Chemical Society

RIGHTS & PERMISSIONS

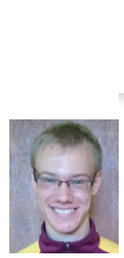
Article Views	Altmetric	Citations
393	7	1

LEARN ABOUT THESE METRICS

Share Add to

Fluid methane

- Dominated by weak dispersion interactions
- 2-body enough or need many-body effects?
- What level of quantum mechanics is required?



ACS Publications
Most Trusted. Most Cited. Most Read.

Search text, DOI, authors, etc.

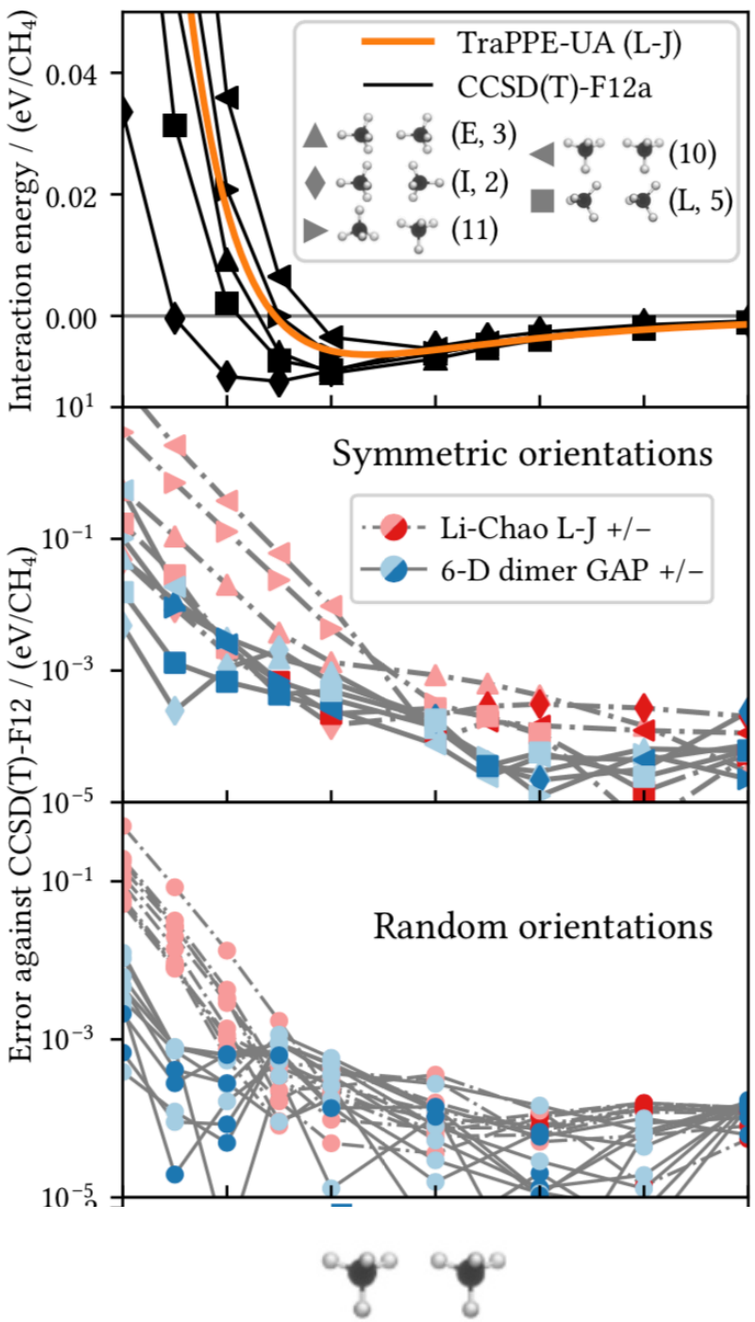
RETURN TO ISSUE | < PREV ARTICLE NEXT >

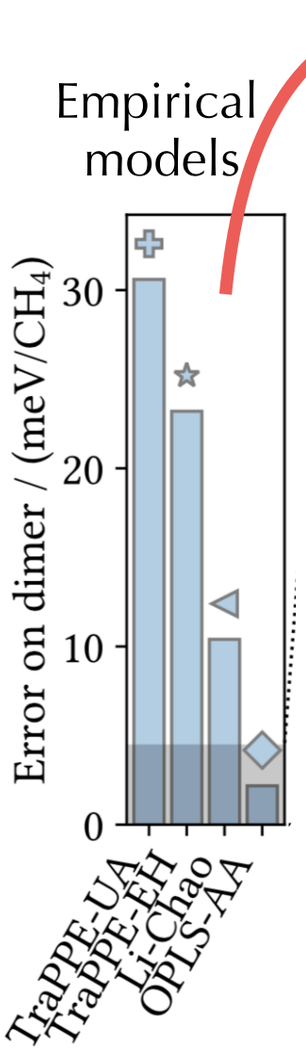
Equation of State of Fluid Methane from First Principles with Machine Learning Potentials

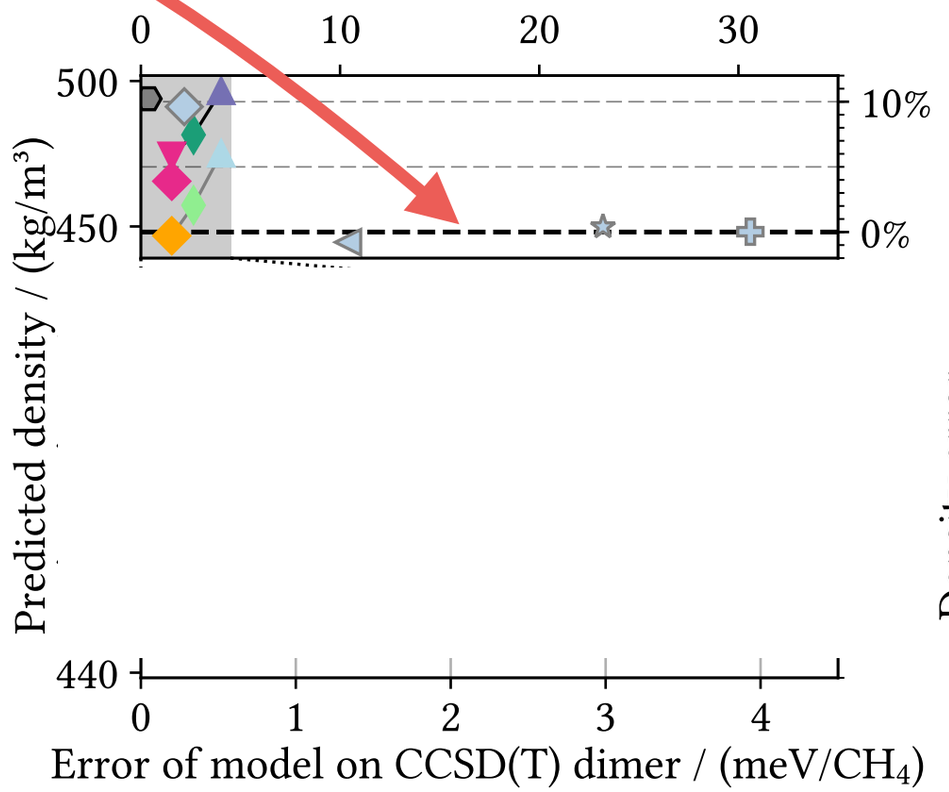
Max Veit*, Sandeep Kumar Jain, Satyanarayana Bonakala, Indranil Rudra, Detlef Hohl and Gábor Csányi

Cite This: *J. Chem. Theory Comput.* 2019, 15, 4, 2574-2586
Publication Date: February 22, 2019
https://doi.org/10.1021/acs.jctc.8b01242

Article Views 393 | Altmetric 7 | Citations 1 | Share | Add to
LEARN ABOUT THESE METRICS

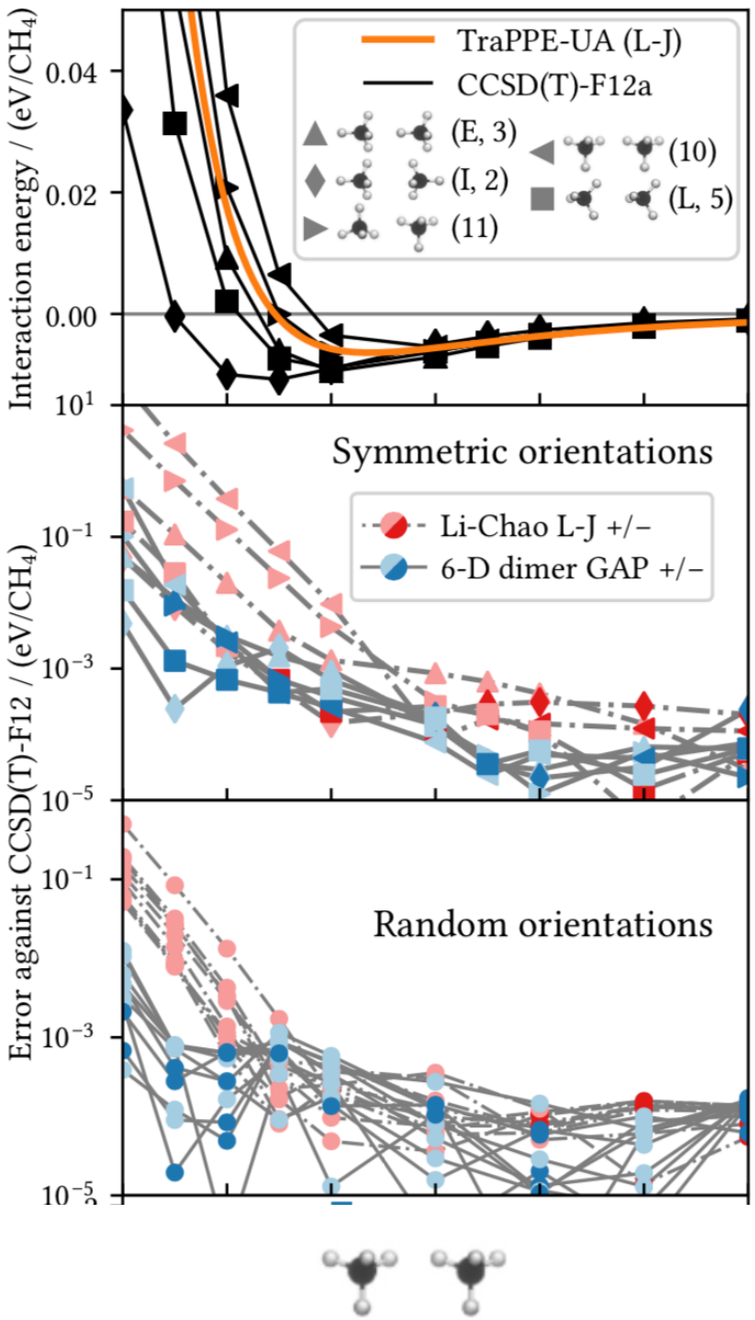


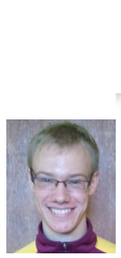




Fluid methane

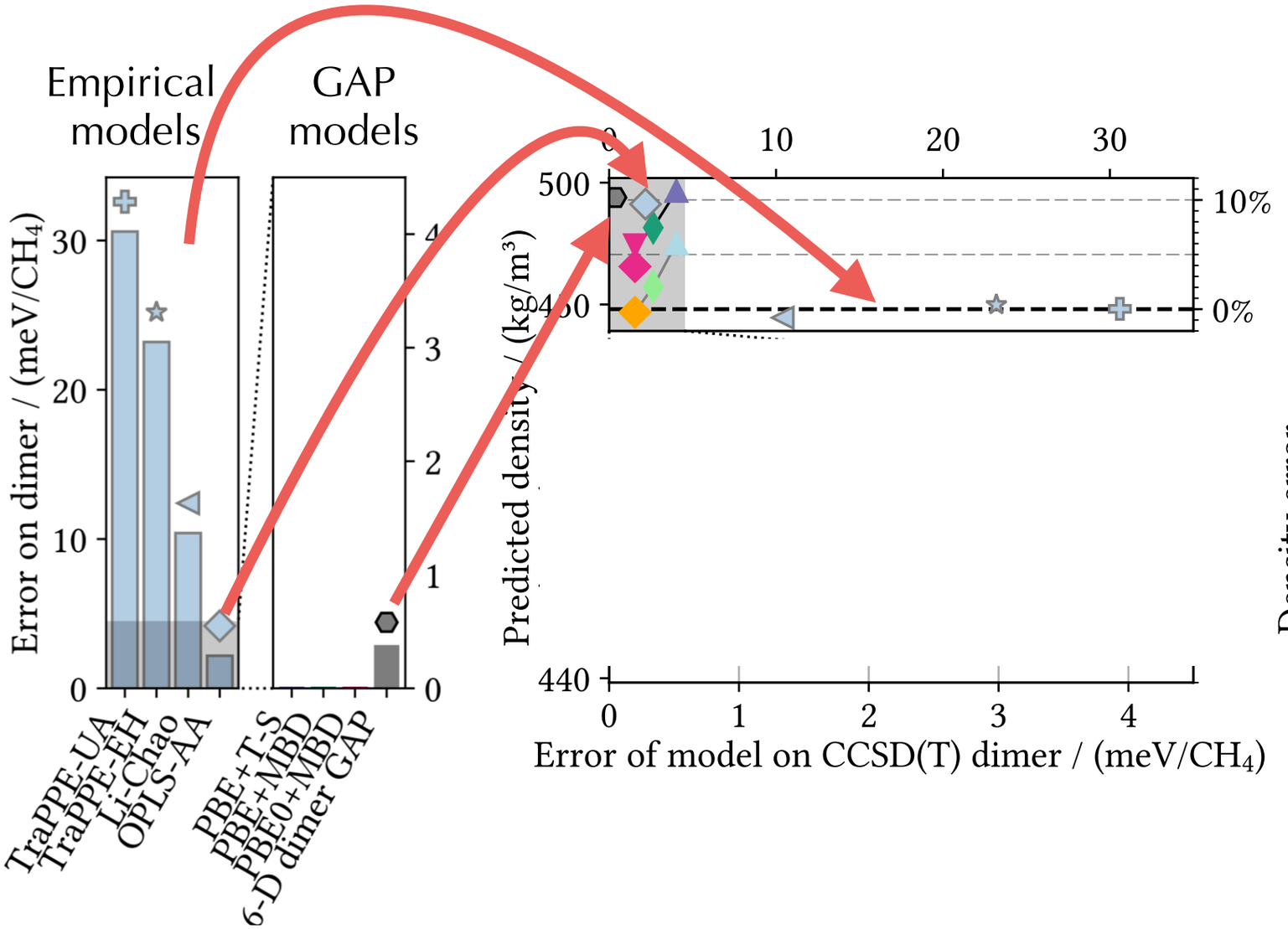
- Dominated by weak dispersion interactions
- 2-body enough or need many-body effects?
- What level of quantum mechanics is required?





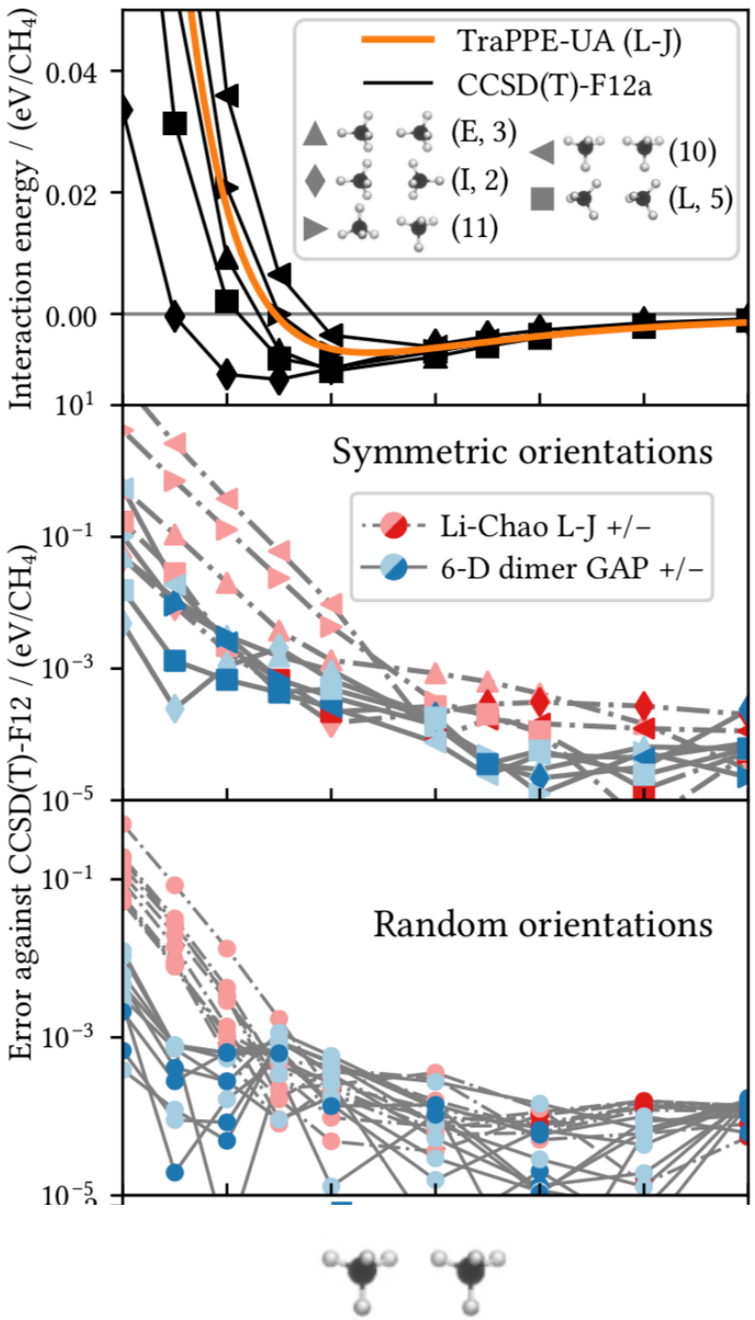
ACS Publications
Most Trusted. Most Cited. Most Read.
Search text, DOI, authors, etc.
RETURN TO ISSUE | < PREV ARTICLE NEXT >
Equation of State of Fluid Methane from First Principles with Machine Learning Potentials
Max Veit*, Sandeep Kumar Jain, Satyanarayana Bonakala, Indranil Rudra, Detlef Hohl and Gábor Csányi
Cite This: *J. Chem. Theory Comput.* 2019, 15, 4, 2574-2586
Publication Date: February 22, 2019
https://doi.org/10.1021/acs.jctc.8b01242

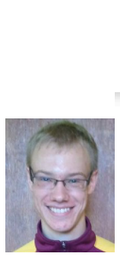
Article Views 393 | Altmetric 7 | Citations 1
LEARN ABOUT THESE METRICS
Share Add to

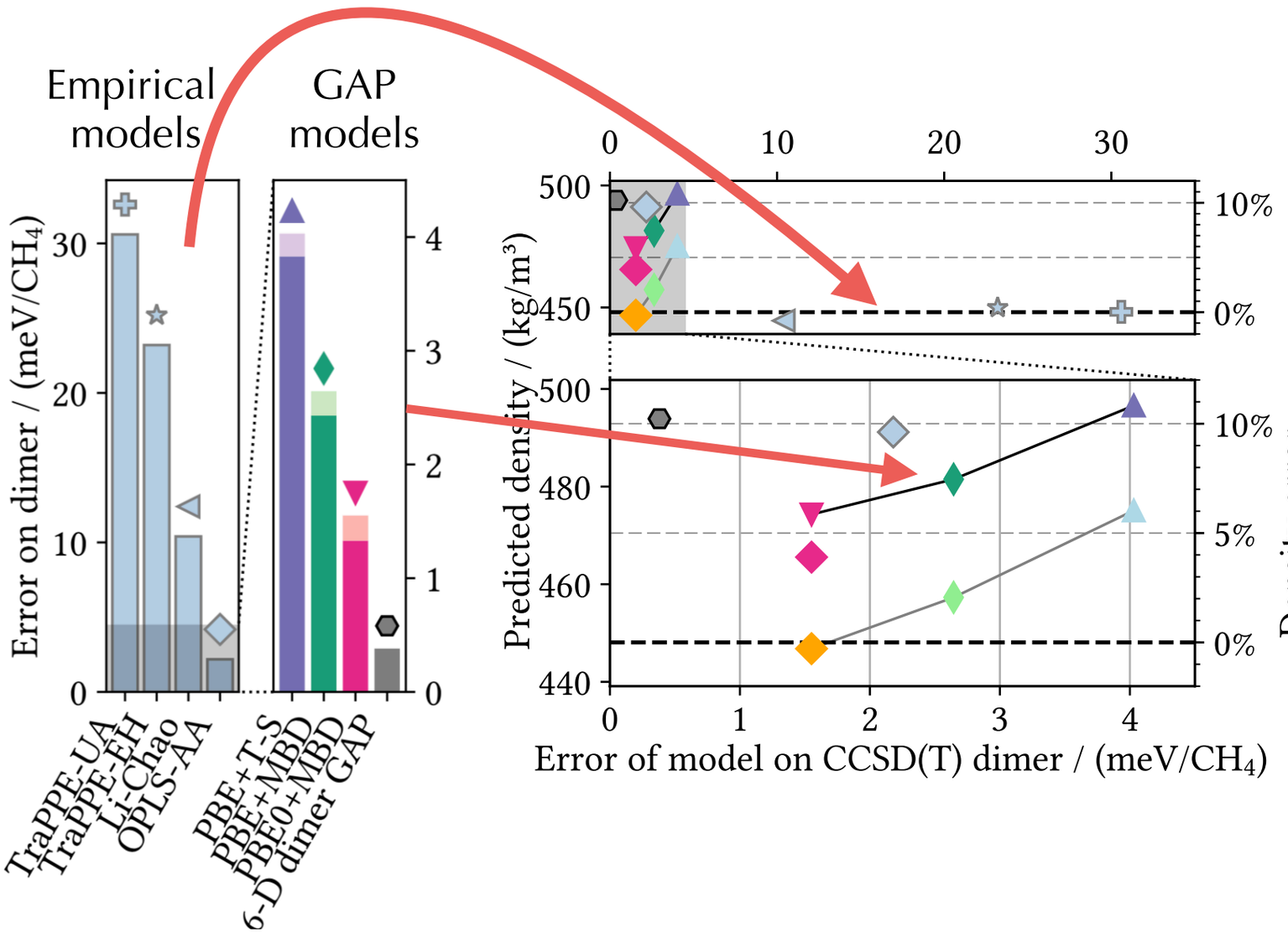


Fluid methane

- Dominated by weak dispersion interactions
- 2-body enough or need many-body effects?
- What level of quantum mechanics is required?

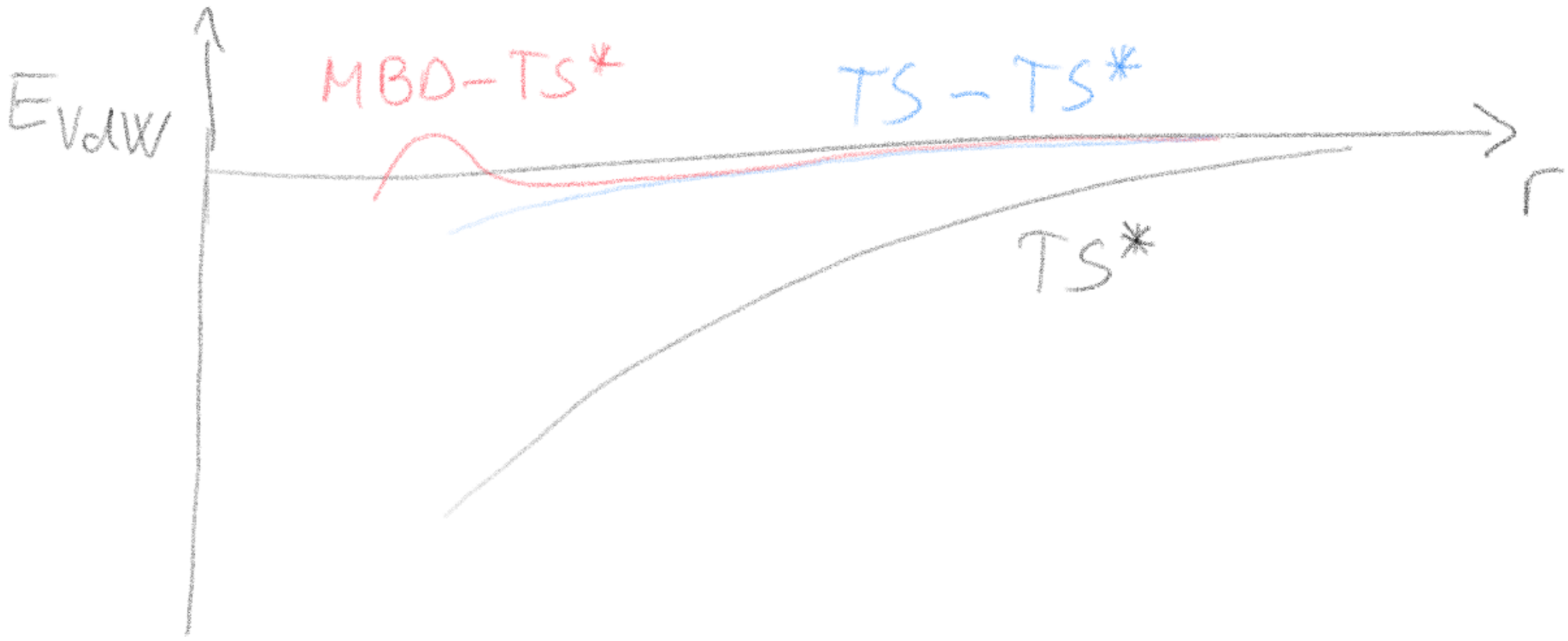






Medium-range physics

- Best many-body QM we can do: Hybrid DFT + VdW
- VdW is long range : $\sim 1/r^6$
- Hierarchy of models:
  - Tkatchenko-Scheffler pair potential with fixed coefficient: TS*
  - TS correction with on-the-fly-dft-computed coefficients
  - MBD (many-body dispersion also by Tkatchenko et al), also needs DFT-computed coefficients

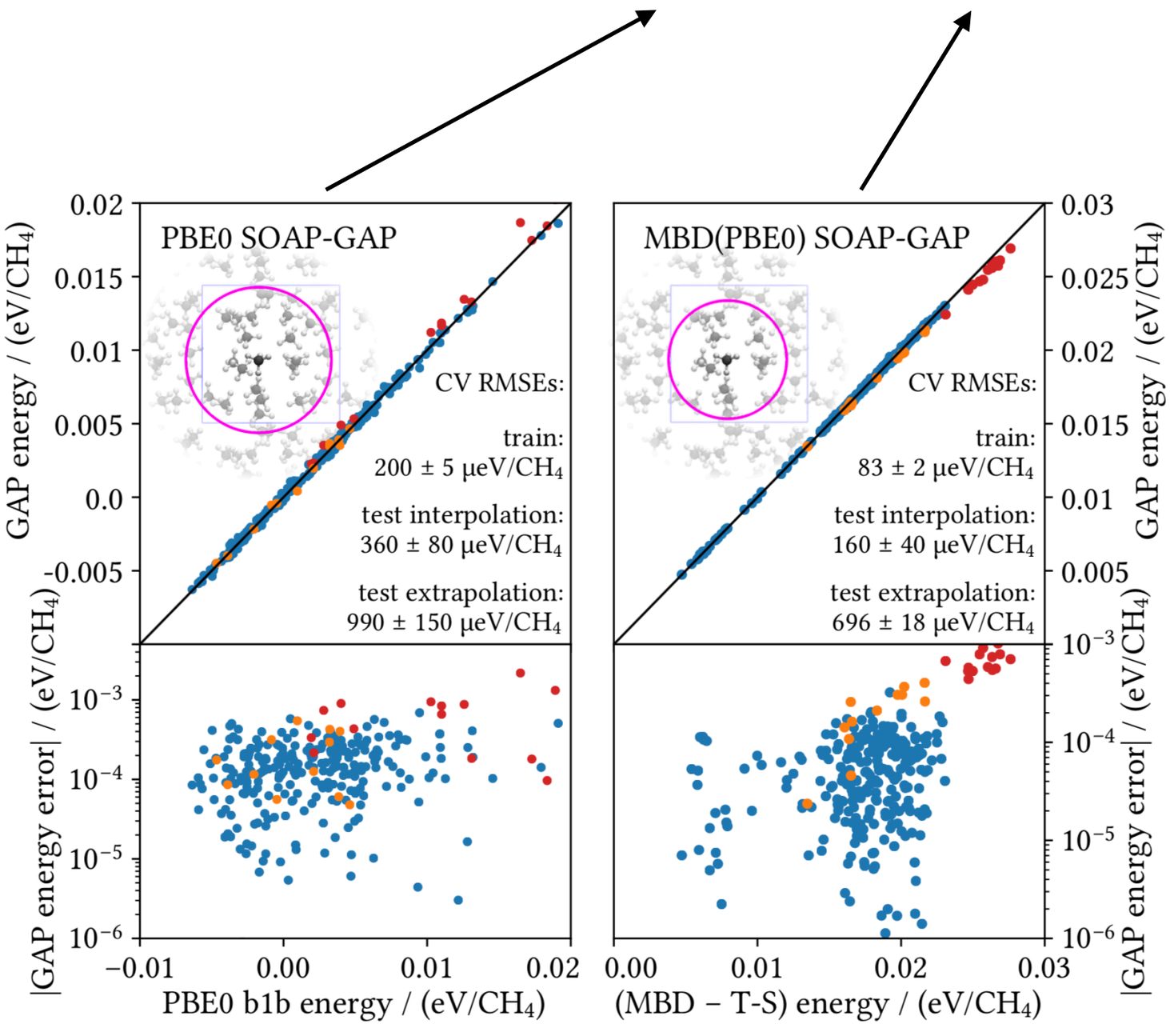


Many-body SOAP-GAP model

$$E_{\text{potential}} = N E_{\text{CH}_4} + E_{\text{short-range}} + E_{\text{VdW}}(\text{MBD-TS*}) + E_{\text{VdW}}(\text{TS*})$$

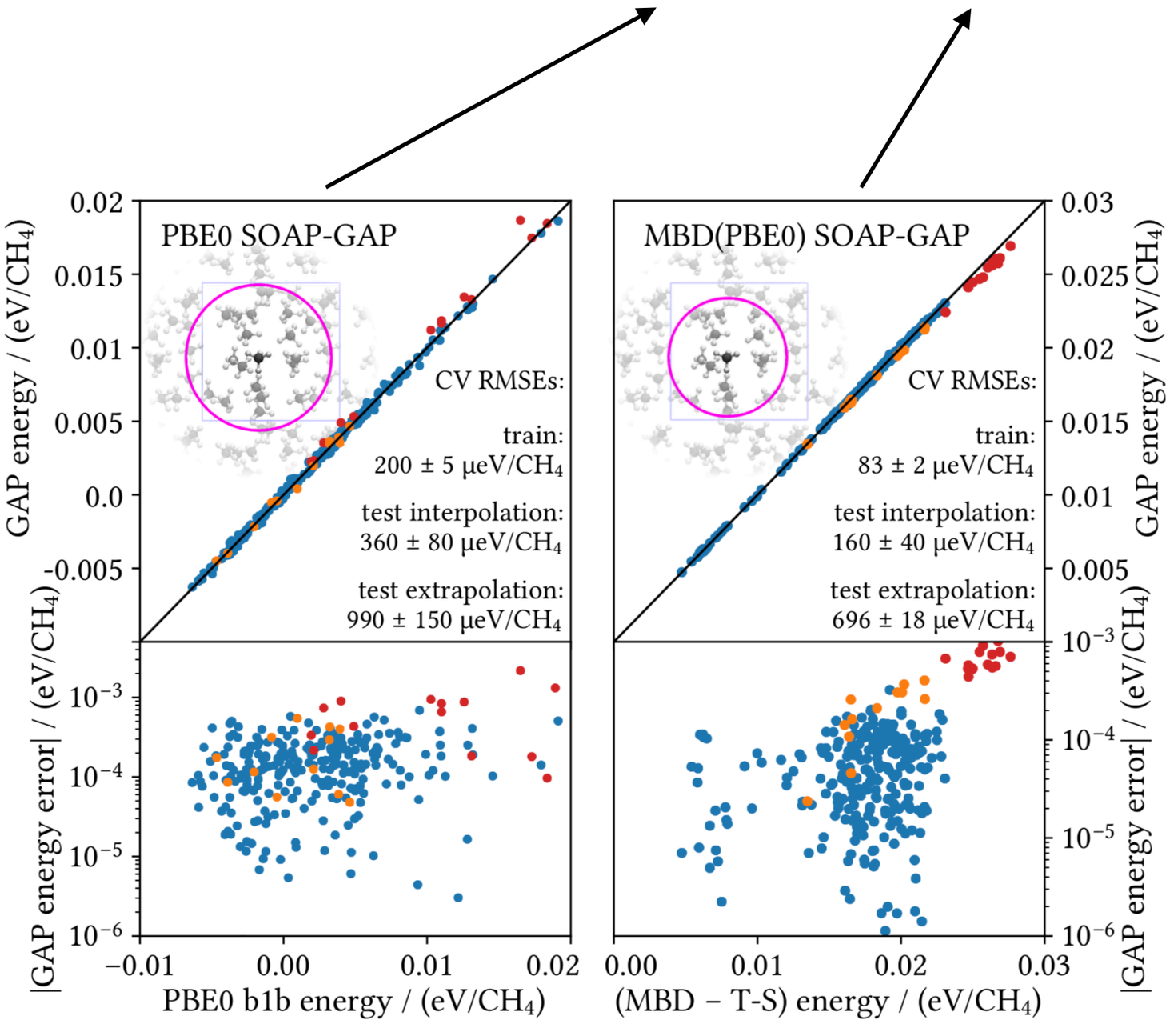
Many-body SOAP-GAP model

$$E_{\text{potential}} = N E_{\text{CH}_4} + E_{\text{short-range}} + E_{\text{VdW}}(\text{MBD-TS*}) + E_{\text{VdW}}(\text{TS*})$$

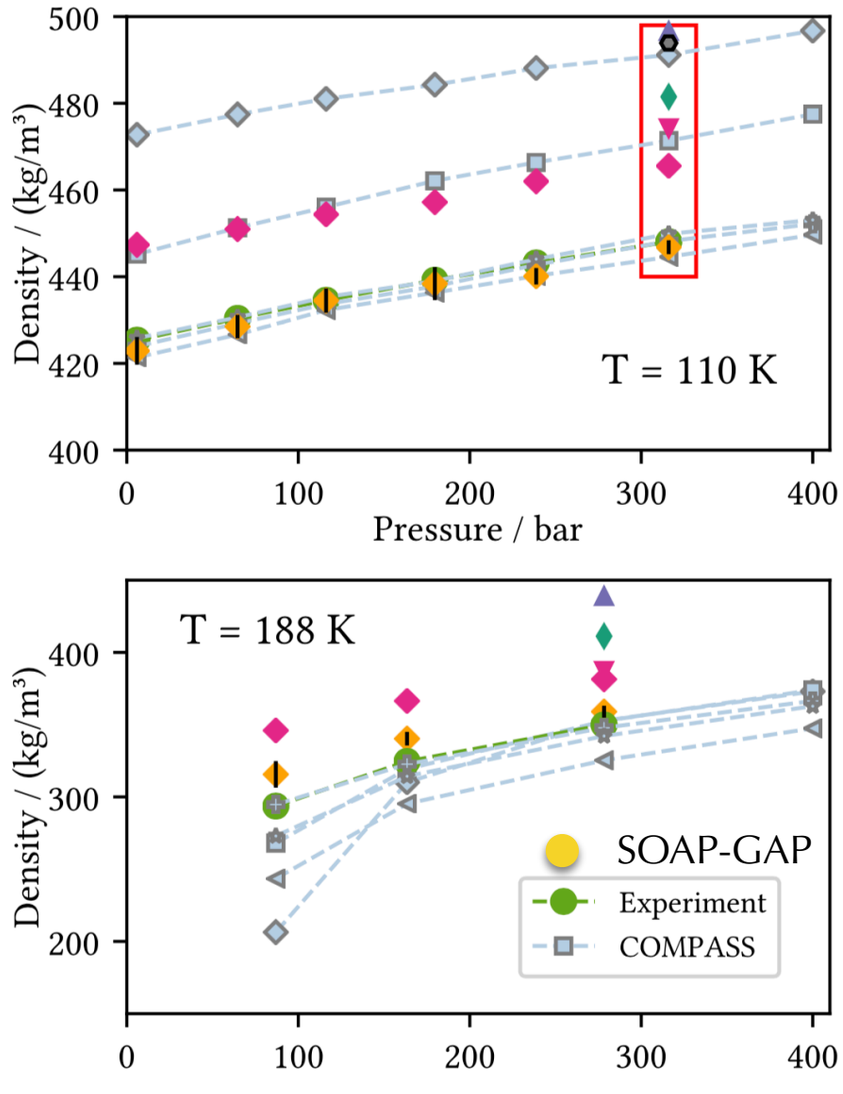


Many-body SOAP-GAP model

$$E_{\text{potential}} = N E_{\text{CH}_4} + E_{\text{short-range}} + E_{\text{VdW}}(\text{MBD-TS*}) + E_{\text{VdW}}(\text{TS*})$$



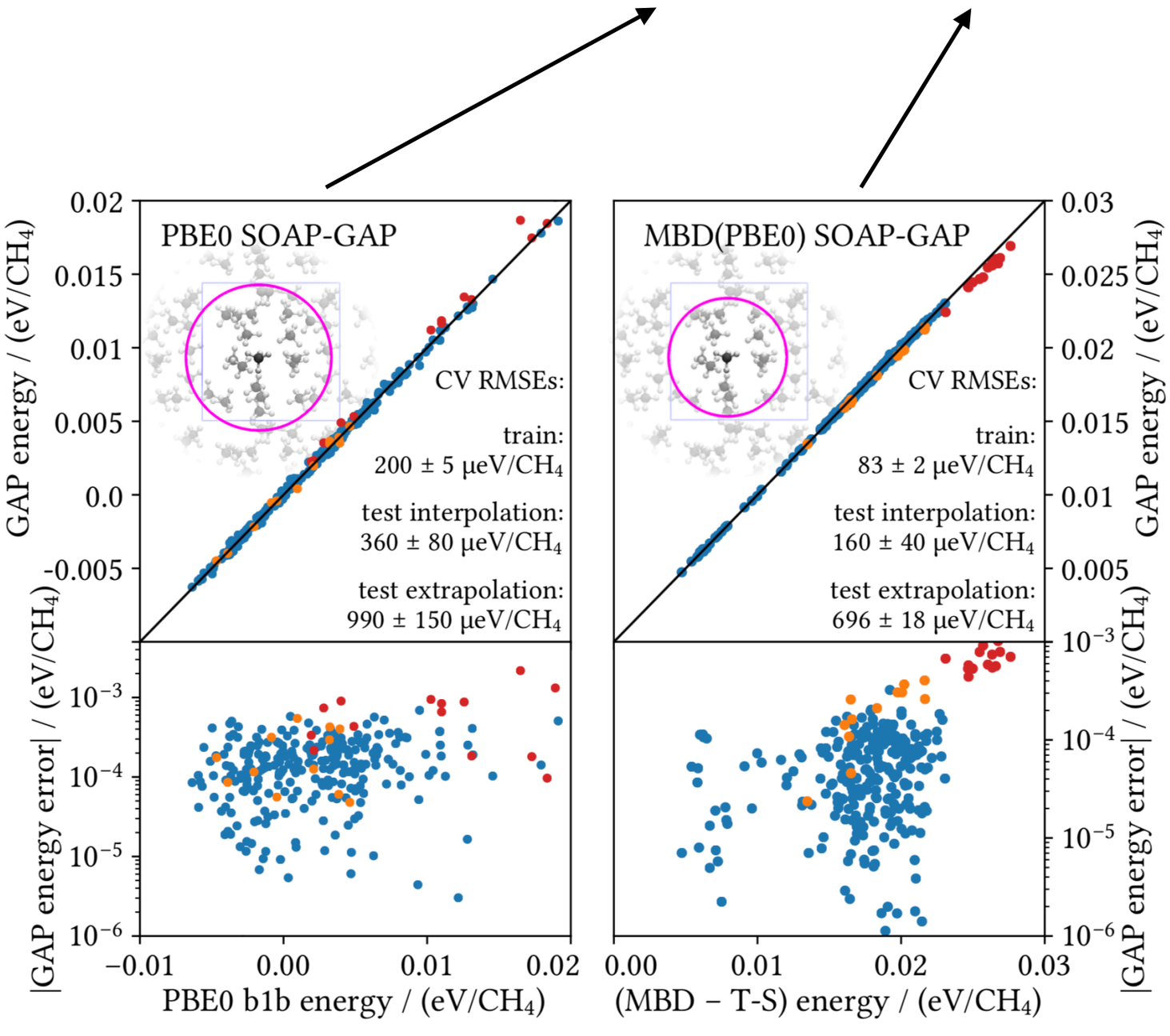
Density prediction



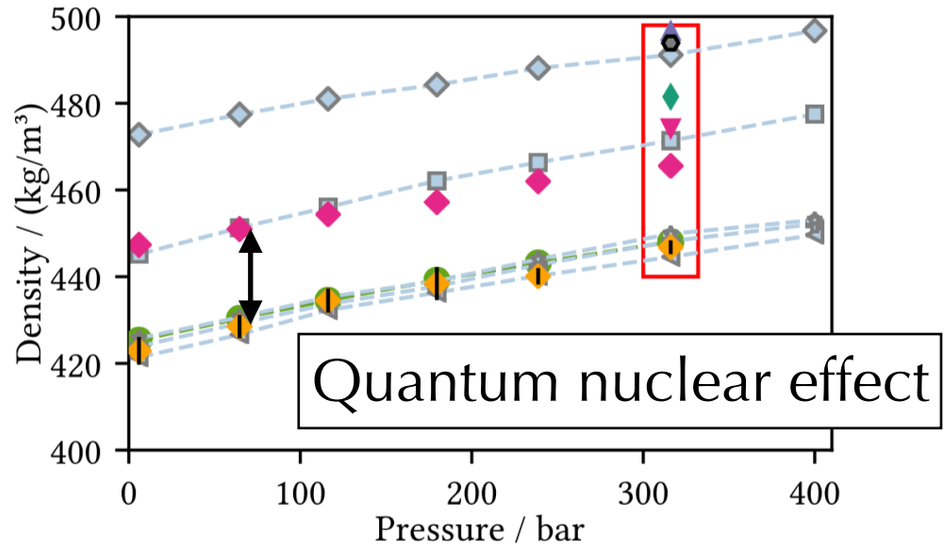
Pressure / bar

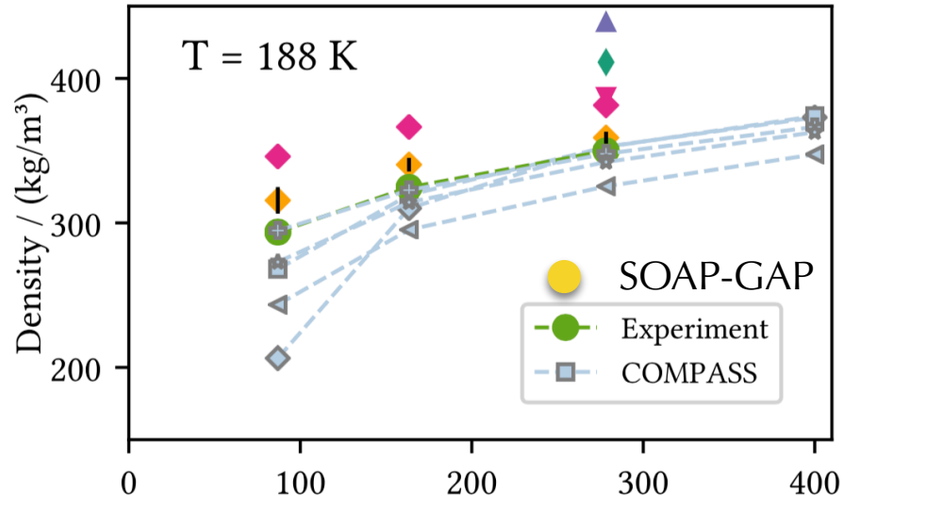
Many-body SOAP-GAP model

$$E_{\text{potential}} = N E_{\text{CH}_4} + E_{\text{short-range}} + E_{\text{VdW}}(\text{MBD-TS*}) + E_{\text{VdW}}(\text{TS*})$$



Density prediction

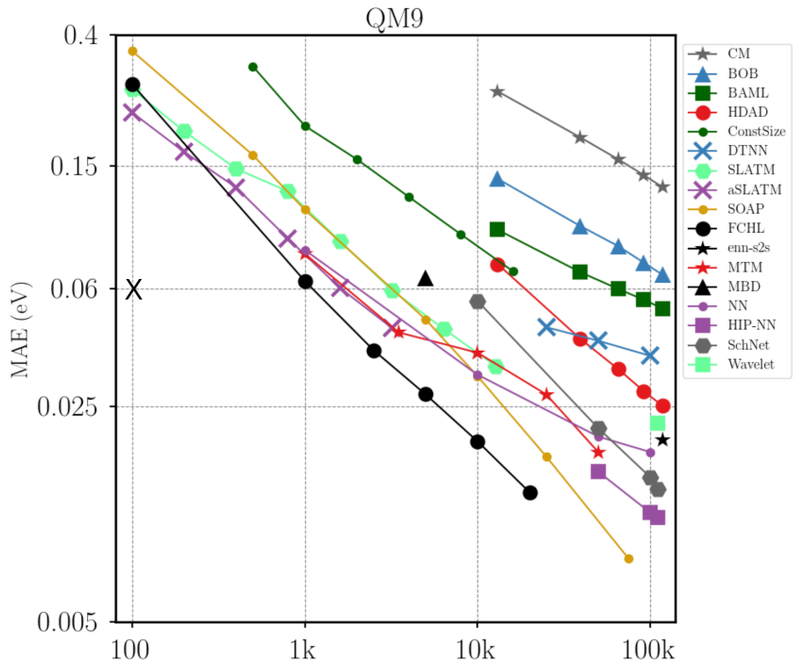




Pressure / bar

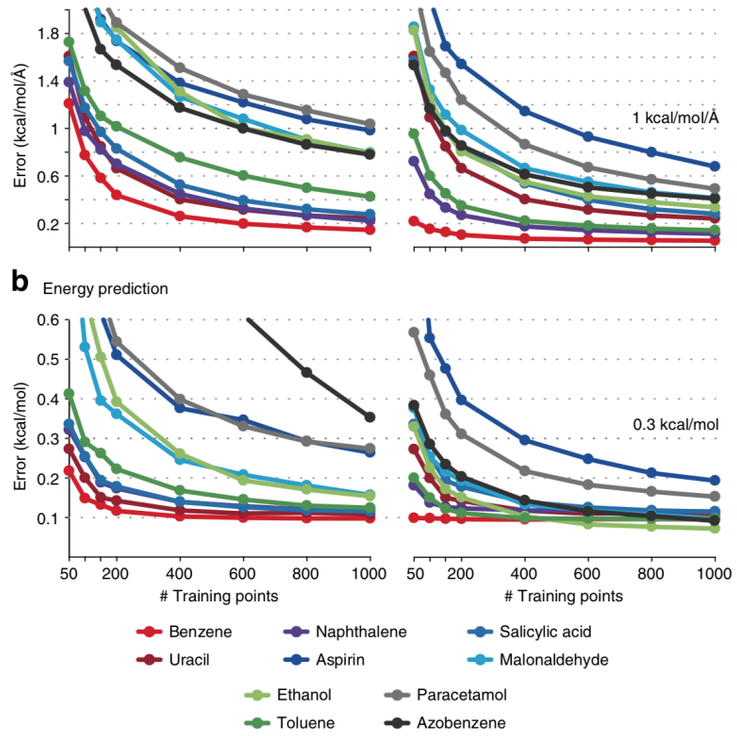
Molecular successes (of others, mostly)

QM9 (local minima)

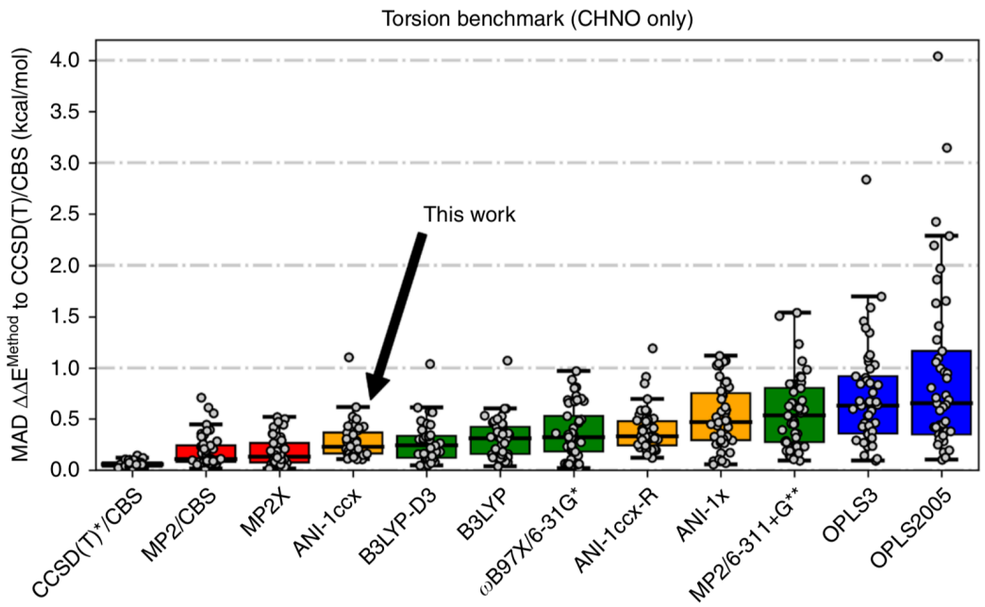


(Courtesy of O. A von Lilienfeld)

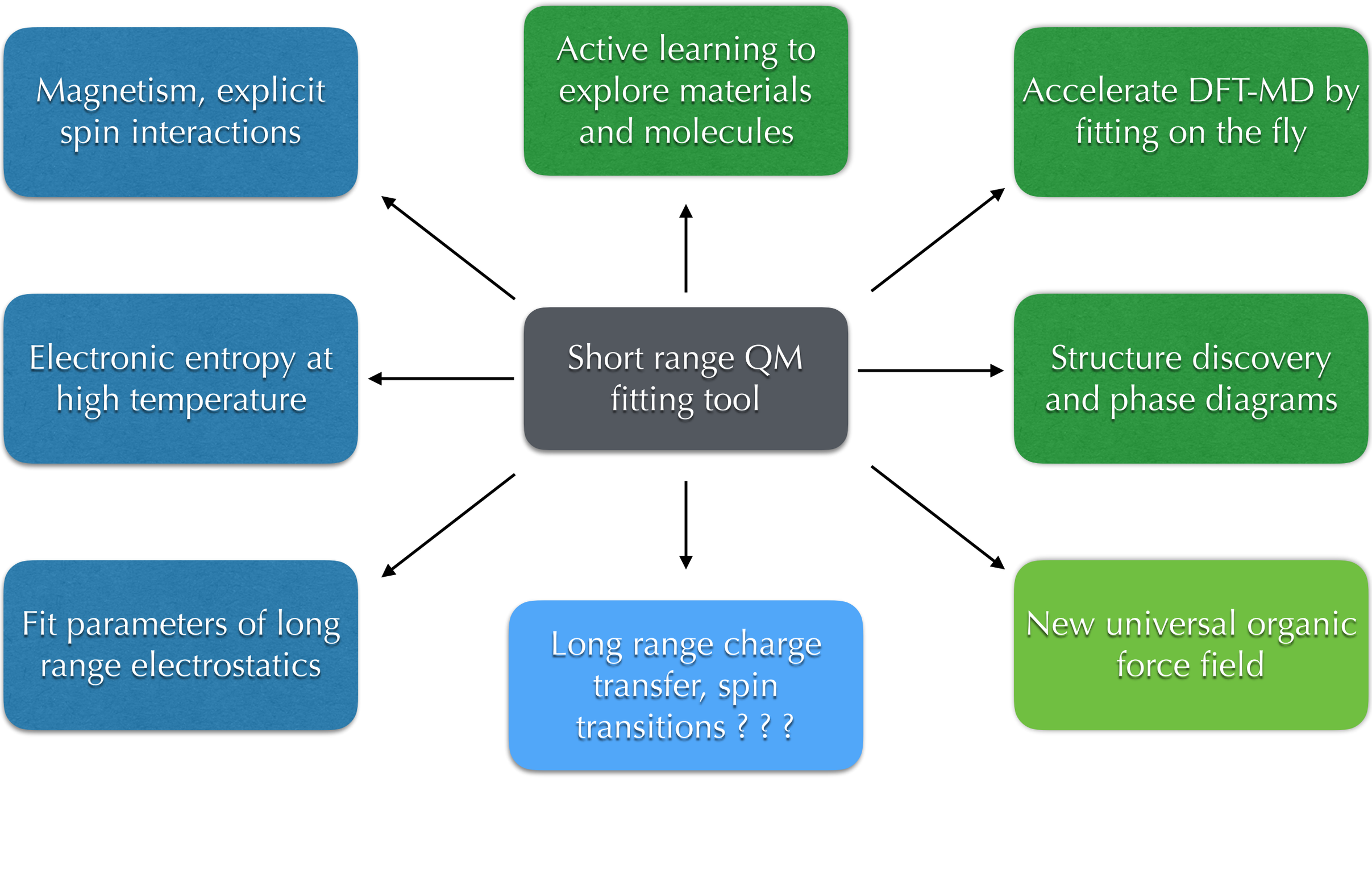
(s)GDML custom force fields of Tkatchenko et al. (KRR model on interatomic distances)



ANI-CC model of Isayev, Roitberg and coworkers for arbitrary molecules (2-3b descriptors, 3-layer ANN)



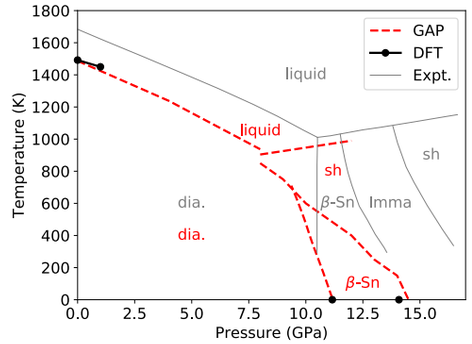
Go back to doing physics and materials modelling

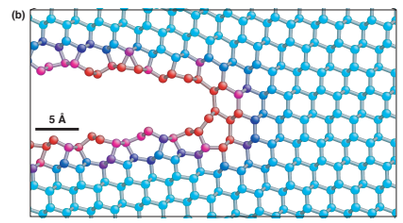


Challenge: rigorous approximation

- How do we characterise relationship between fitting database and accuracy of predicted results?
- How do we give **global** accuracy guarantees?
- Inverse problem: what fitting database do we need to predict a given property ?

Structure type	# atoms	# structures	# environments	# representative pts	σ_{energy}	σ_{force}	σ_{virial}
					default values: 0.001	0.1	0.05
isolated atom	1	1	1	1			
diamond	2	104	208	6			
	16	220	3520	53			
	54	110	5940	58			
	128	55	7040	92			
β-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			
bcc	2	49	392	40			
bc8	8	49	88	66			
fcc	4	49	196	46			
hcp	2	49	88	28			
st12	12	49	588	94			
liquid	64	69	4416	1114	0.003	0.15	0.2
	128	7	896	323			
amorphous	64	31	1984	231	0.01	0.2	0.4
	216	128	27648	1719			
diamond surface (001)	144	29	4176	514			
decohesion	32	11	352	28			
diamond surface (110)	108	26	2808	338			
decohesion	16	11	176	8			
diamond surface (111)							
decohesion	24	11	264	10			
unreconstructed	96	47	4512	573			
adatom	146	11	1606	62			
Pandey reconstruction	96	50	4800	632			
DAS 3x3 unrelaxed	52	1	52	6			
diamond vacancy	63	100	6300	168			
	215	111	23865	405			
diamond divacancy	214	78	16692	416			
diamond interstitial	217	115	24955	605			
small (110) crack tip	200	7	1400	130			
small (111) crack tip	192	10	1920	185			
screw dislocation core	144	19	2736	124			
sp^2 bonded	8	51	408	61			
sp bonded	4	100	400	392	0.01	0.2	0.4
Total							





???