

An overview of how descriptors are used for ML tasks in Physics, Chemistry and Materials

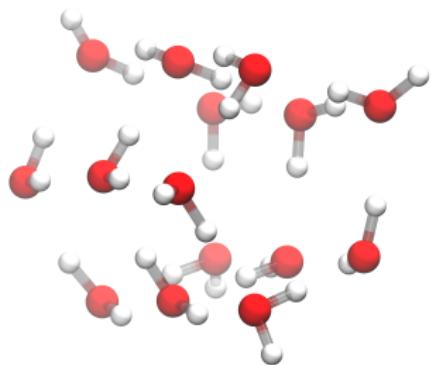
Bingqing Cheng

Trinity College, the University of Cambridge

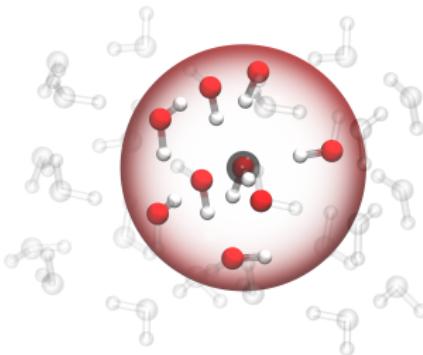
What we will talk about:

- Machine learning for visualization, classification and analysis of structures.
 - Similarity measurement
 - Applications
- Machine learning potentials
 - How does it work?
 - An example on water.
 - Limitations.

Representing atomistic environments



A



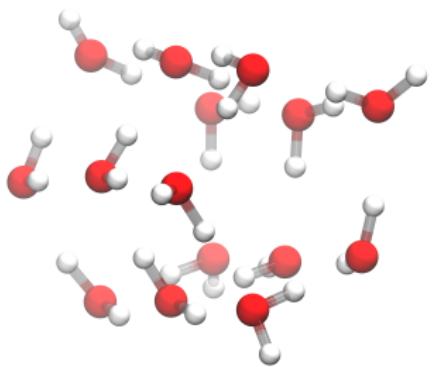
x_i

The first step is to divide the system into a set of atomic environments.
So the task becomes representing atomic environments.

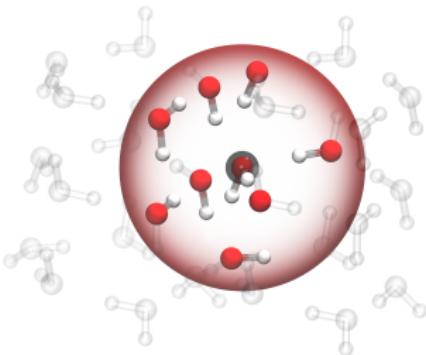
Popular representations (Invariant with respect to translation, rotation and permutation.):

- Smooth overlap of atomic positions (SOAP) [Bartók, Kondor & Csányi PRB 2013]
- Behler-Parrinello symmetry functions [Behler & Parrinello PRL 2008]

Representing atomistic environments



$$\Phi \leftarrow \{\Psi(\mathcal{X}_i)\}$$



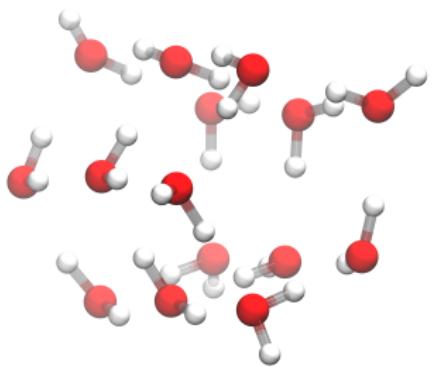
$$\Psi(\mathcal{X}_i)$$

The first step is to divide the system into a set of atomic environments.
So the task becomes representing atomic environments.

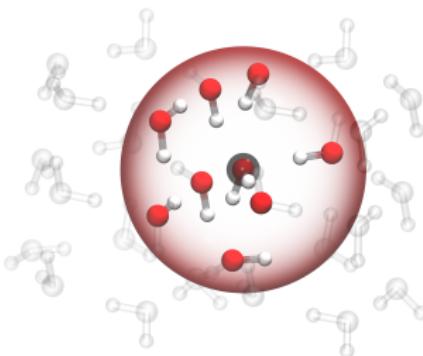
Popular representations (Invariant with respect to translation, rotation and permutation.):

- Smooth overlap of atomic positions (SOAP) [Bartók, Kondor & Csányi PRB 2013]
- Behler-Parrinello symmetry functions [Behler & Parrinello PRL 2008]

Representing atomistic environments



$$O(A) = F(\Phi(A))$$



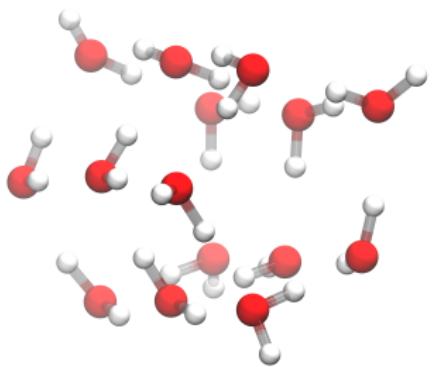
$$O_i = f(\Psi(\mathcal{X}_i))$$

The first step is to divide the system into a set of atomic environments.
So the task becomes representing atomic environments.

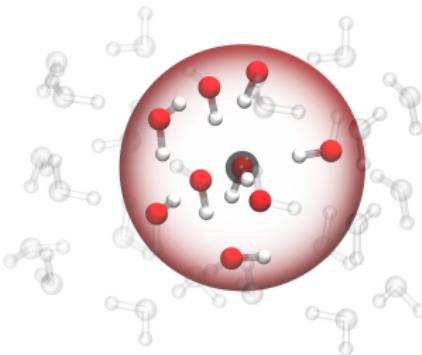
Popular representations (Invariant with respect to translation, rotation and permutation.):

- Smooth overlap of atomic positions (SOAP) [Bartók, Kondor & Csányi PRB 2013]
- Behler-Parrinello symmetry functions [Behler & Parrinello PRL 2008]

Representing atomistic environments



$$E(A) = \sum E_i$$



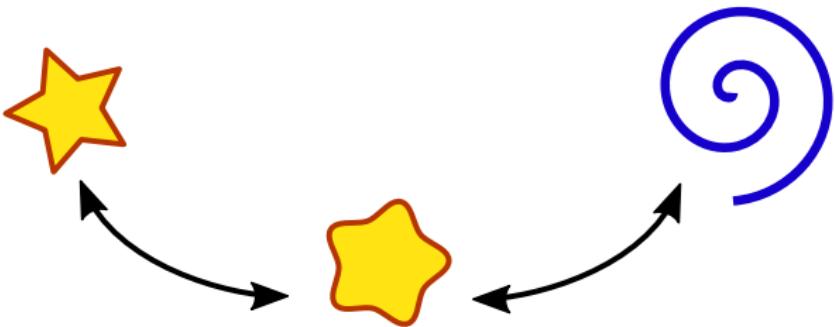
$$E_i = e(\Psi(\mathcal{X}_i))$$

The first step is to divide the system into a set of atomic environments.
So the task becomes representing atomic environments.

Popular representations (Invariant with respect to translation, rotation and permutation.):

- Smooth overlap of atomic positions (SOAP) [Bartók, Kondor & Csányi PRB 2013]
- Behler-Parrinello symmetry functions [Behler & Parrinello PRL 2008]

[Bartók, Kondor & Csányi PRB 2013]



$$K(\mathcal{X}, \mathcal{X}') \approx 1$$

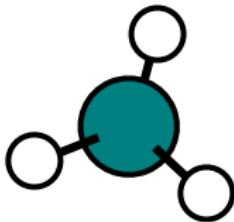
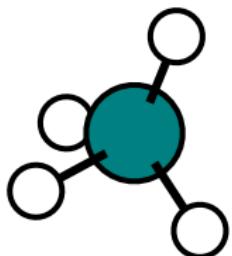
$$D(\mathcal{X}, \mathcal{X}') \approx 0$$

$$K(\mathcal{X}, \mathcal{X}') \approx 0$$

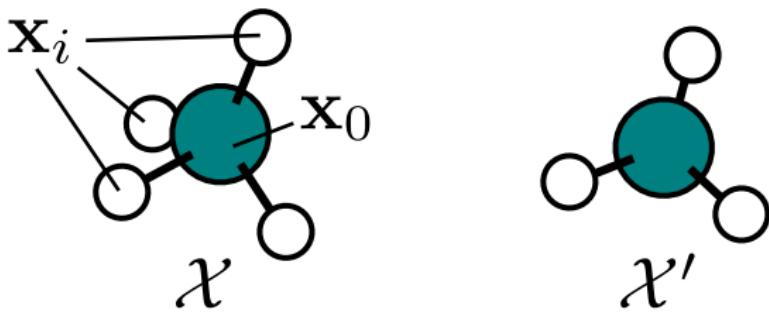
$$D(\mathcal{X}, \mathcal{X}') \gg 0$$

Representing atomic environments

[Bartók, Kondor & Csányi PRB 2013]

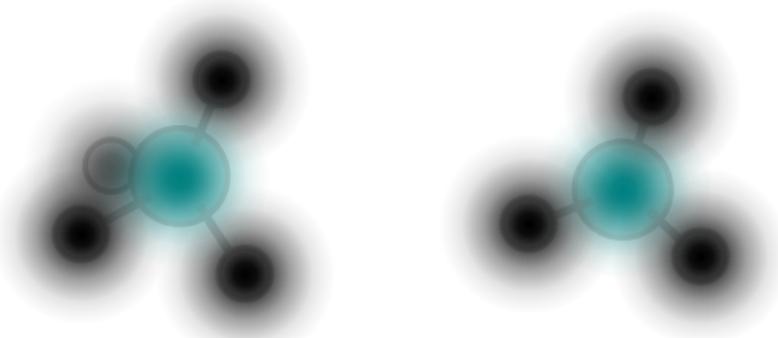


[Bartók, Kondor & Csányi PRB 2013]



$$\{\mathbf{x}_i - \mathbf{x}_0\} \leftrightarrow \mathcal{X}$$

[Bartók, Kondor & Csányi PRB 2013]



$$\rho_\alpha(\mathbf{x}) = \sum_{i \in \alpha} g(\mathbf{x} - \mathbf{x}_i)$$

[Bartók, Kondor & Csányi PRB 2013]

$$k(\mathcal{X}, \mathcal{X}') = \int \rho(\mathbf{x}) \rho'(\mathbf{x})$$

[Bartók, Kondor & Csányi PRB 2013]

$$k(\mathcal{X}, \mathcal{X}') = \int d\hat{R} \left| \int \rho(\mathbf{x}) \rho'(\hat{R}\mathbf{x}) \right|^2$$

[Bartók, Kondor & Csányi PRB 2013]

$$k(\mathcal{X}, \mathcal{X}') = \int d\hat{R} \left| \int \rho(\mathbf{x}) \rho'(\hat{R}\mathbf{x}) \right|^2$$

[Bartók, Kondor & Csányi PRB 2013]

$$\kappa(\mathcal{X}, \mathcal{X}') = \int d\hat{R} |\rho(\mathbf{x})\rho'(\hat{R}\mathbf{x})|^2$$

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

$$k_{nn'l}(\mathcal{X}) = \pi \sqrt{\frac{8}{2l+1}} \sum_m (c_{nlm})^* c_{n'l m}$$

The list of vector $\{k_{nn'l}(\mathcal{X})\}$ is the descriptor of the atomic environment \mathcal{X} .

- The kernel matrix $\{K\}$ records the similarity measurement for each pair of structures in the data set.
- The kernel function $K(A, B)$ for structure A and B is

$$K(A, B) = \Phi(A)^T \Phi(B) = \sum_{i=1}^M \phi_i(A) \phi_i(B)$$

- Global features are constructed from local features by taking the average:

$$\Phi(A) = \frac{1}{N_A} \sum_{n=1}^{N_A} \Psi(\mathcal{X}_n^A)$$

- We use SOAP vectors as local features. [Bartók et al. PRB 2013]
We use DScribe Python library [Himanen et al. arxiv 2019]

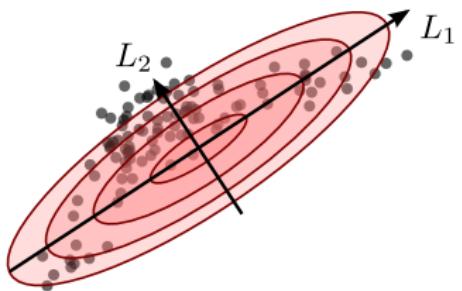
$$k_{nn'l}(\mathcal{X}) = \pi \sqrt{\frac{8}{2l+1}} \sum_m (c_{nlm})^* c_{n'lm}$$

- Build two-dimensional map using kernel principal component analysis (KPCA)



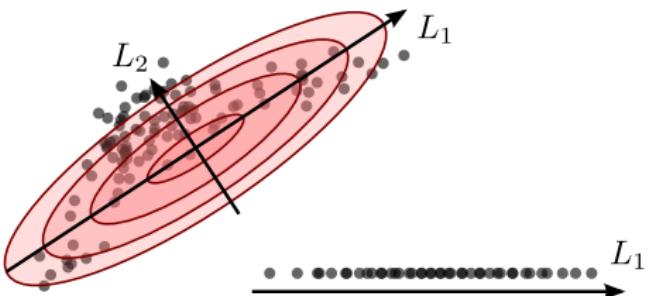
- Sparsity the train set using farthest point sampling, CUR or uniform sampling
- Clustering
- Regression (Kernel ridge regression (KRR), neural networks)

- Build two-dimensional map using kernel principal component analysis (KPCA)



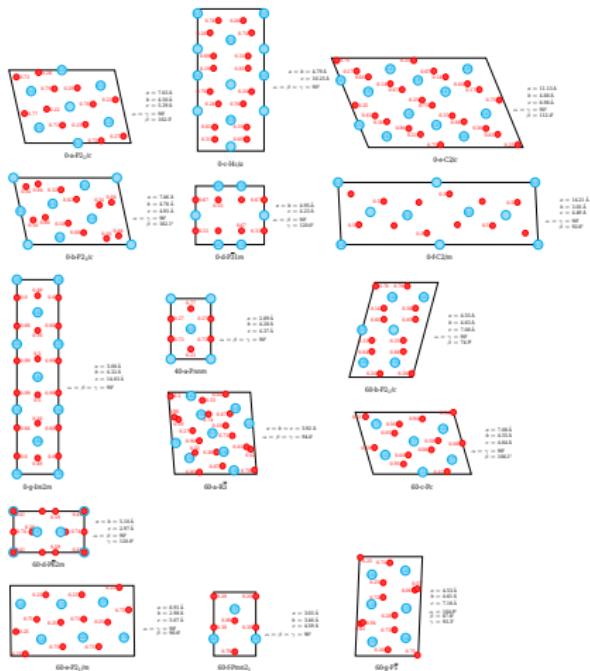
- Sparsity the train set using farthest point sampling, CUR or uniform sampling
- Clustering
- Regression (Kernel ridge regression (KRR), neural networks)

- Build two-dimensional map using kernel principal component analysis (KPCA)



- Sparsity the train set using farthest point sampling, CUR or uniform sampling
- Clustering
- Regression (Kernel ridge regression (KRR), neural networks)

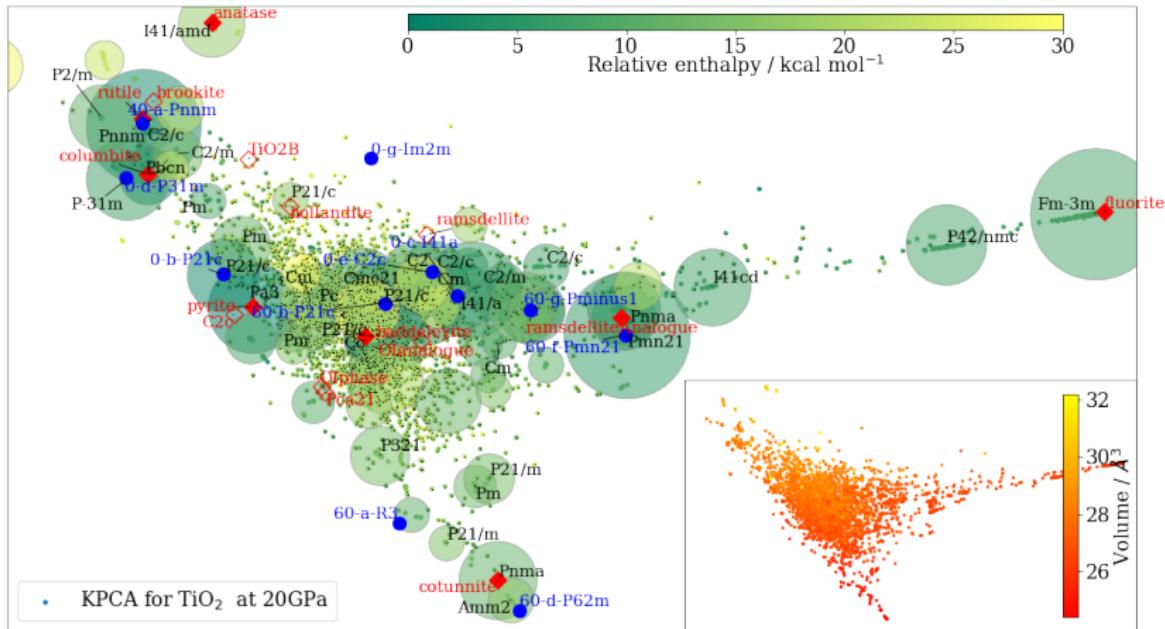
Random structure search of TiO₂



Matsui–Akaogi (MA) empirical pair potential
Random Structure Search: [Pickard and Needs JPCM 2011]

Random structure search of TiO₂

[Reinhardt, Pickard & Cheng arXiv 2019]

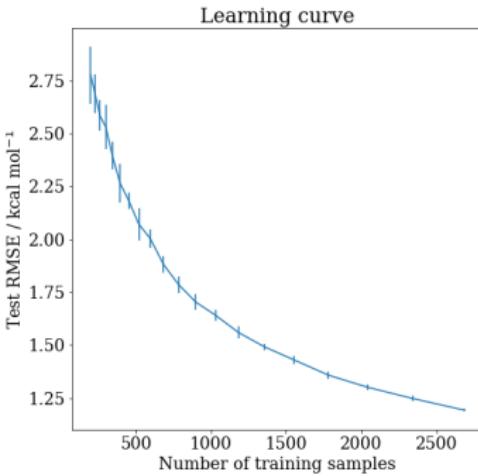
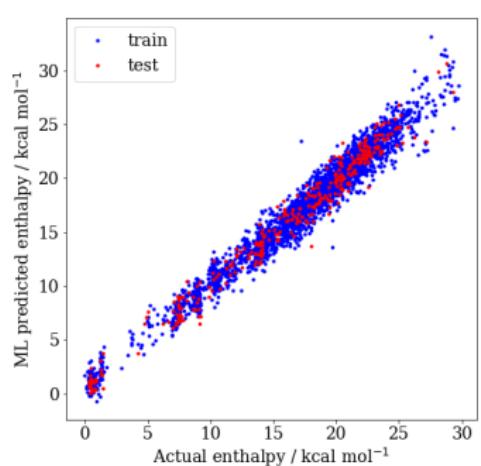


Kernel ridge regression

[Reinhardt, Pickard & Cheng arXiv 2019]

$$H^{\text{ML}}(A) = \sum_{B \in M} w(B) K(A, B)$$

$$w_M = (K_{MM} + K_{MN}\Lambda^{-1}K_{MN}^T)^{-1} K_{MN}\Lambda^{-1} H_N$$



What we will talk about:

- Machine learning for visualization, classification and analysis of structures.
 - Similarity measurement
 - Applications
- Machine learning potentials
 - How does it work?
 - An example on water.
 - Limitations.

- We use Behler-Parrinello neural network [Behler Parrinello PRL 2008]

n2p2 - The neural network potential package

DOI: [10.5281/zenodo.1344447](https://doi.org/10.5281/zenodo.1344447)

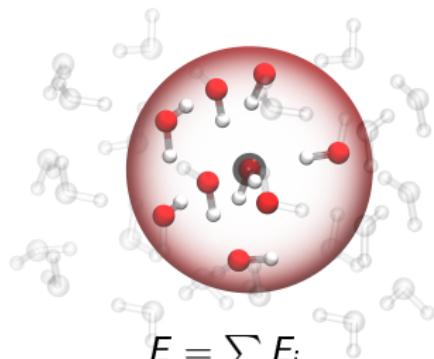
Authors

- Andreas Singraber (University of Vienna)
- Gaussian Approximation Potentials (GAPs)

QUIP - QUantum mechanics and Interatomic Potentials

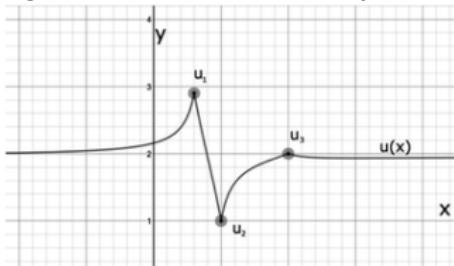
[docs](#) [public](#) [build](#) [persons](#) [docker-pulls](#)  Docker size

Step 1: Collect environments.



$$E = \sum E_i$$

Step 2: The rest is interpolation.

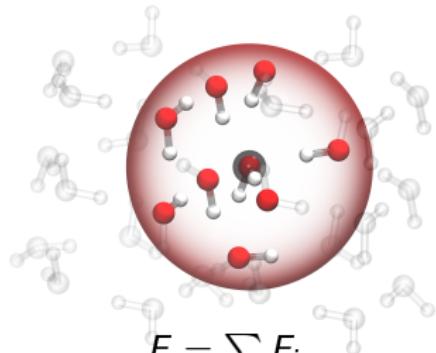


Construct ML potentials

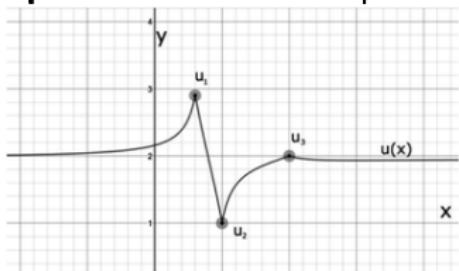
Ways to collect atomic environments:

- Enumerating possible structures
- Random displacement
- Stretching and compression
- Molecular dynamics (MD) and PIMD
- On-the-fly learning [Li, Kermode & Vita PRL 2015]
- Random searches [Deringer, Pickard & Csányi PRL 2018]
- Active learning [Podryabinkin & Shapeev Com. Mat. Sci. 2017]

Step 1: Collect environments.

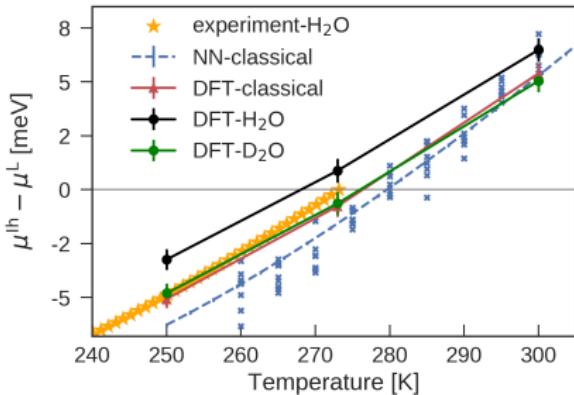
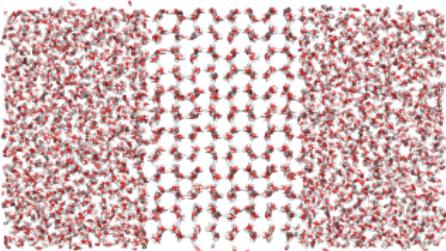
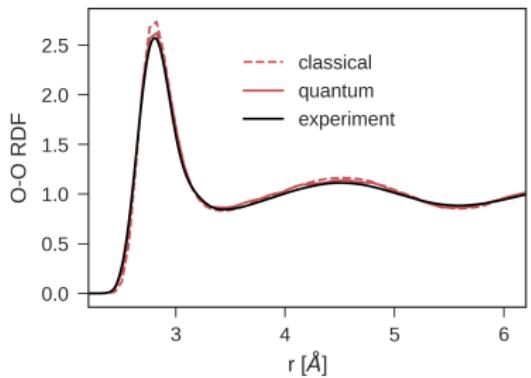
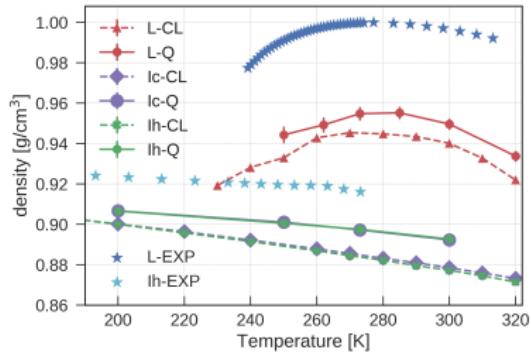


Step 2: The rest is interpolation.



Neural network potential for water

[Cheng, Engel, Behler, Dellago & Ceriotti PNAS 2019]



DFT with revPBE0+D3 functional

- AIMD and PIMD simulations [Marsalek & Markland JPCL 2017]
- Benchmarks with CCSD(T) and DMC [Brandenburg 2019]

25/04/2019

<https://archive.materialscloud.org/2018.0020/v1>

(<https://www.nature.com/scientificdata/policies/repositories#materials>)



FAIRsharing.org

DOI: 10.5281/zenodo.1250000

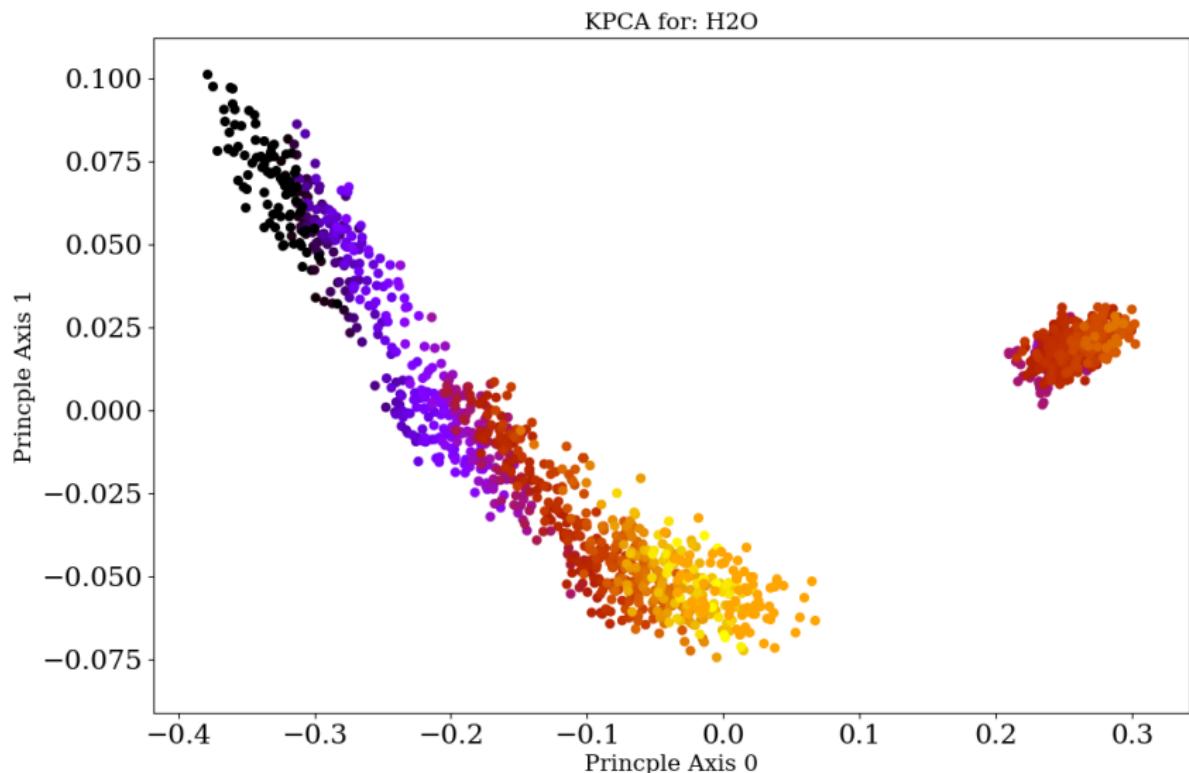
materialscloud:2018.0020 (/2018.0020/v1)

Ab initio thermodynamics of liquid and solid water: supplemental materials

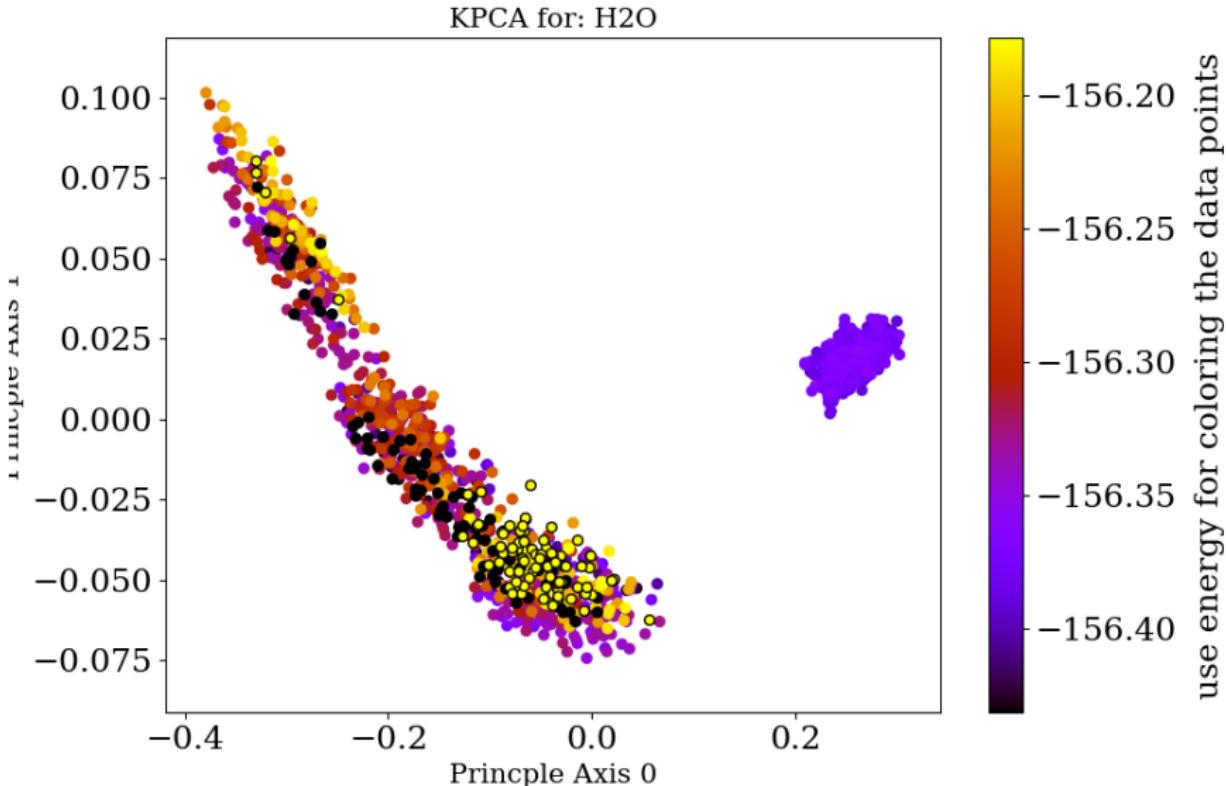
- 1593 bulk liquid 64 water molecules (energy + forces)
- 1000 structures from quenching at a wide range of densities.
- 593 structures from PIMD.

	WATER27		ICE		2D-ICE	
	MD	MAD	MD	MAD	MD	MAD
LDA	-201	209	-376	376	-286	286
PBE	-9	17	26	60	-1	21
PBE-D2	-49	52	-105	105	-70	70
PBE-D3	-38	42	-80	80	-54	54
PBE-D4	-37	41			-52	52
PBE-TS			-93	93	-50	50
PBE-MBD			-93	93	-58	58
PBE-VV10	-51	55	-137	137	-78	78
revPBE	86	86	233	233	145	145
revPBE-D3	21	25	23	33	32	32
revPBE-D4	11	17			19	19
SCAN	-130	133	-87	87	-72	72
optB86b-vdW			-109	109	-34	34
rev-vdW-DF2					72	72
BLYP-D3	-4	14	-46	46	-22	22
BLYP-D4	-8	14			-26	26
revPBE0-D3	11	17	13	13	25	25
revPBE0-D4	12	18			28	28
B3LYP-D4	-15	17				
PW6B95-DISP	5	10				

Training set for bulk water

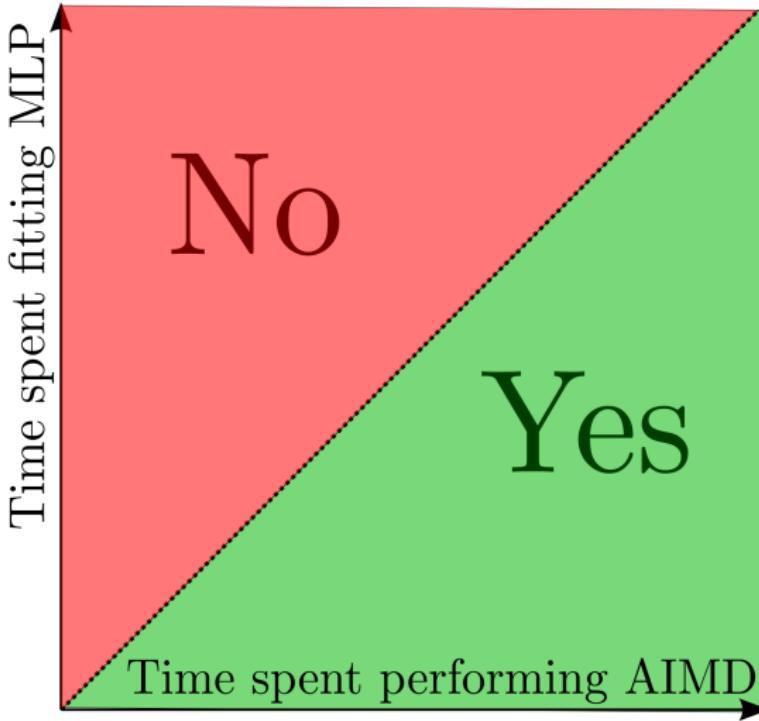


Training set for bulk water

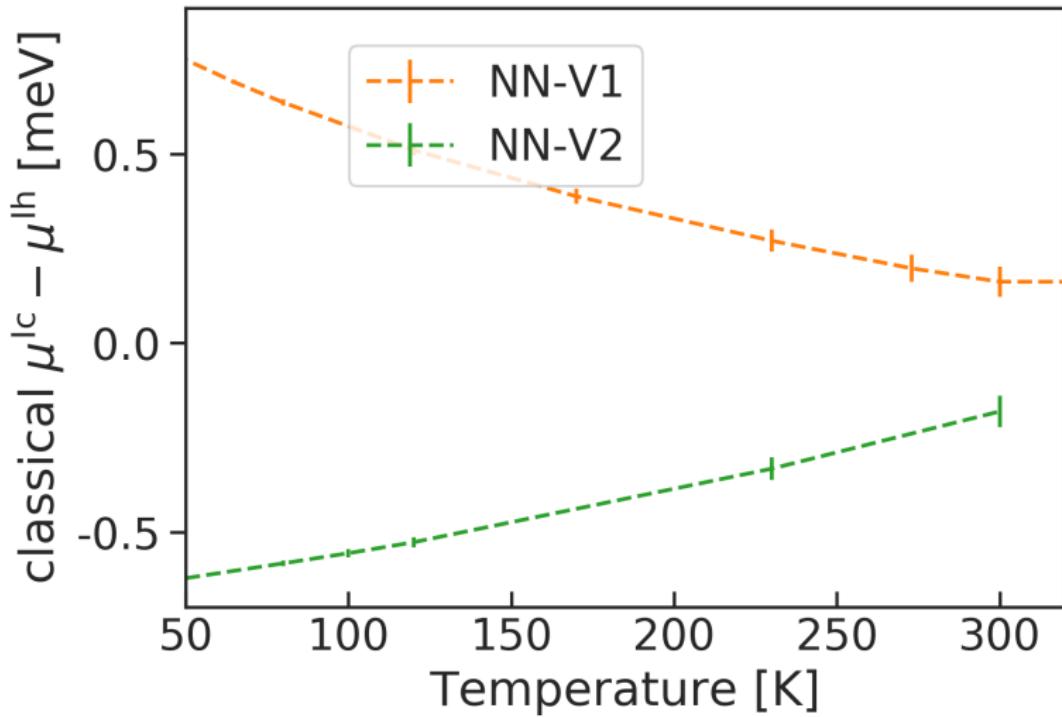


Shortcomings of machine-learning potentials (MLPs)

- Can be tedious to fit a MLP.
- Residual errors (fitting, sparse training set, incomplete representations, ...).
- The lack of long-range interactions.



Comparison between two NN potentials



The Gibbs free energy of the system described by the DFT:

$$G = -k_B T \ln \int d\mathbf{q} \exp \left[-\frac{U(\mathbf{q}) + PV}{k_B T} \right]$$

The Gibbs free energy of the system described by the ML potential:

$$G_{ML} = -k_B T \ln \int d\mathbf{q} \exp \left[-\frac{U_{ML}(\mathbf{q}) + PV}{k_B T} \right]$$

$$G - G_{ML} = -k_B T \ln \frac{\int d\mathbf{q} \exp \left[-\frac{U_{ML}(\mathbf{q}) + PV}{k_B T} \frac{U(\mathbf{q}) - U_{ML}(\mathbf{q})}{k_B T} \right]}{\int d\mathbf{q} \exp \left[-\frac{U_{ML}(\mathbf{q}) + PV}{k_B T} \right]}$$

$$G - G_{ML} = -k_B T \ln \left\langle \exp \left[\frac{U(\mathbf{q}) - U_{ML}(\mathbf{q})}{k_B T} \right] \right\rangle_{ML}$$

Free energy perturbation method

The Gibbs free energy of the system described by the DFT:

$$G = -k_B T \ln \int d\mathbf{q} \exp \left[-\frac{U(\mathbf{q}) + PV}{k_B T} \right]$$

The Gibbs free energy of the system described by the ML potential:

$$G_{ML} = -k_B T \ln \int d\mathbf{q} \exp \left[-\frac{U_{ML}(\mathbf{q}) + PV}{k_B T} \right]$$

$$G - G_{ML} = -k_B T \ln \frac{\int d\mathbf{q} \exp \left[-\frac{U_{ML}(\mathbf{q}) + PV}{k_B T} \frac{U(\mathbf{q}) - U_{ML}(\mathbf{q})}{k_B T} \right]}{\int d\mathbf{q} \exp \left[-\frac{U_{ML}(\mathbf{q}) + PV}{k_B T} \right]}$$

$$G - G_{ML} = -k_B T \ln \left\langle \exp \left[\frac{U(\mathbf{q}) - U_{ML}(\mathbf{q})}{k_B T} \right] \right\rangle_{ML}$$

Free energy perturbation method

The Gibbs free energy of the system described by the DFT:

$$G = -k_B T \ln \int d\mathbf{q} \exp \left[-\frac{U(\mathbf{q}) + PV}{k_B T} \right]$$

The Gibbs free energy of the system described by the ML potential:

$$G_{ML} = -k_B T \ln \int d\mathbf{q} \exp \left[-\frac{U_{ML}(\mathbf{q}) + PV}{k_B T} \right]$$

$$G - G_{ML} = -k_B T \ln \frac{\int d\mathbf{q} \exp \left[-\frac{U_{ML}(\mathbf{q}) + PV}{k_B T} \frac{U(\mathbf{q}) - U_{ML}(\mathbf{q})}{k_B T} \right]}{\int d\mathbf{q} \exp \left[-\frac{U_{ML}(\mathbf{q}) + PV}{k_B T} \right]}$$

$$G - G_{ML} = -k_B T \ln \left\langle \exp \left[\frac{U(\mathbf{q}) - U_{ML}(\mathbf{q})}{k_B T} \right] \right\rangle_{ML}$$

Free energy perturbation method

The Gibbs free energy of the system described by the DFT:

$$G = -k_B T \ln \int d\mathbf{q} \exp \left[-\frac{U(\mathbf{q}) + PV}{k_B T} \right]$$

The Gibbs free energy of the system described by the ML potential:

$$G_{ML} = -k_B T \ln \int d\mathbf{q} \exp \left[-\frac{U_{ML}(\mathbf{q}) + PV}{k_B T} \right]$$

$$G - G_{ML} = -k_B T \ln \frac{\int d\mathbf{q} \exp \left[-\frac{U_{ML}(\mathbf{q}) + PV}{k_B T} \frac{U(\mathbf{q}) - U_{ML}(\mathbf{q})}{k_B T} \right]}{\int d\mathbf{q} \exp \left[-\frac{U_{ML}(\mathbf{q}) + PV}{k_B T} \right]}$$

$$G - G_{ML} = -k_B T \ln \left\langle \exp \left[\frac{U(\mathbf{q}) - U_{ML}(\mathbf{q})}{k_B T} \right] \right\rangle_{ML}$$

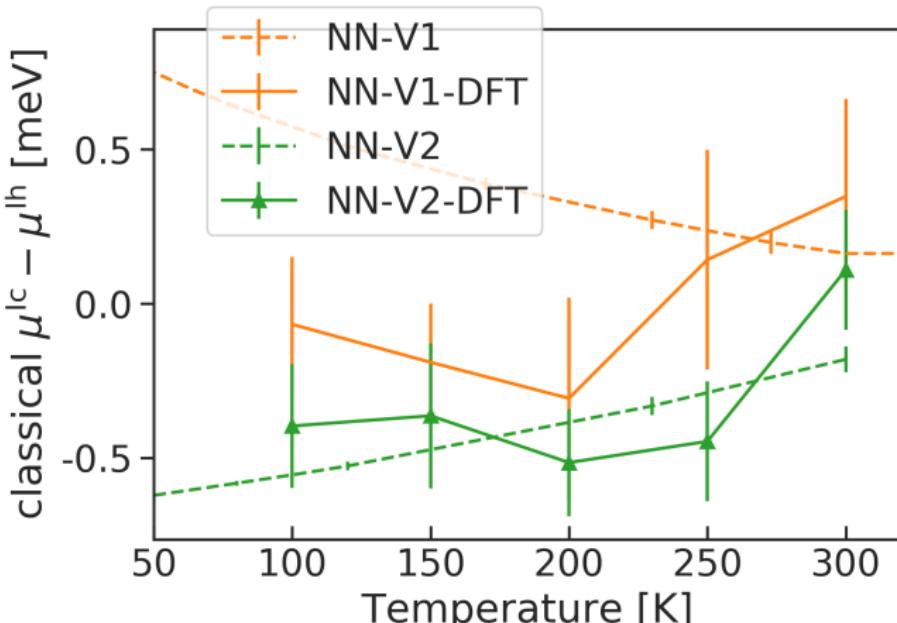
Free energy perturbation method

From Neural network potential to DFT

[Cheng, Engel, Behler, Dellago & Ceriotti PNAS 2019]

$$G - G_{ML} = -k_B T \ln \left\langle \exp \left[\frac{U(\mathbf{q}) - U_{ML}(\mathbf{q})}{k_B T} \right] \right\rangle_{ML}$$

$$\Delta\mu = (G - G_{ML})/N$$

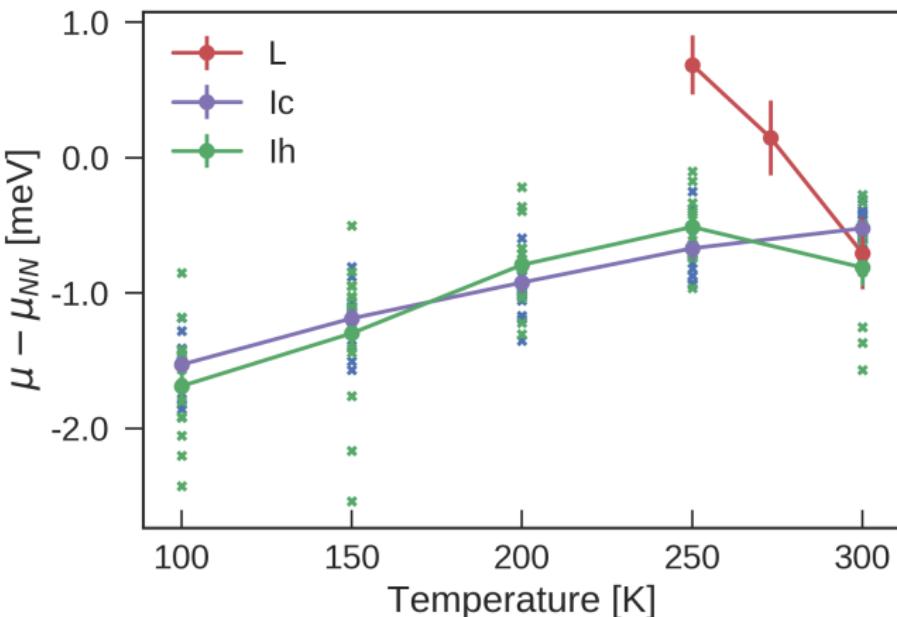


From Neural network potential to DFT

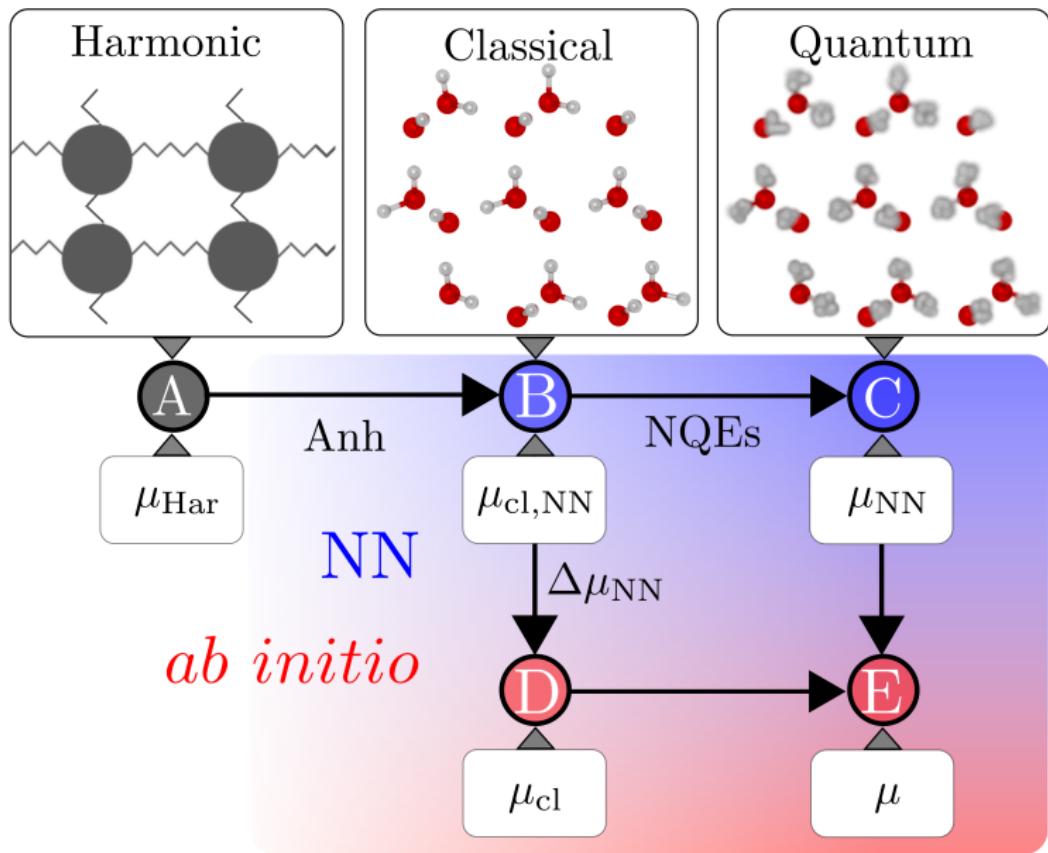
[Cheng, Engel, Behler, Dellago & Ceriotti PNAS 2019]

$$G - G_{ML} = -k_B T \ln \left\langle \exp \left[\frac{U(\mathbf{q}) - U_{ML}(\mathbf{q})}{k_B T} \right] \right\rangle_{ML}$$

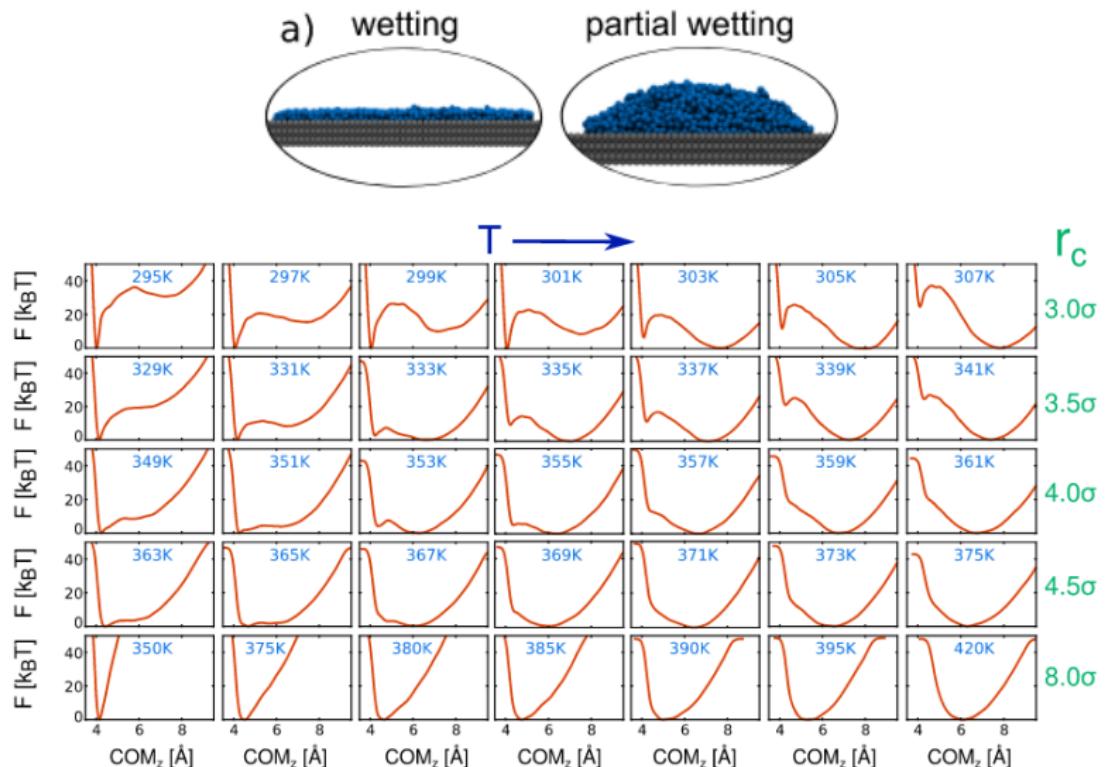
$$\Delta\mu = (G - G_{ML})/N$$



The workflow of ab initio thermodynamics



Long-range interactions are important for describing interfaces



What we will talk about:

- Machine learning for visualization, classification and analysis of structures.
 - Similarity measurement
 - Applications
- Machine learning potentials
 - How does it work?
 - An example on water.
 - Limitations.

PhD supervisor

- Michele Ceriotti, EPFL

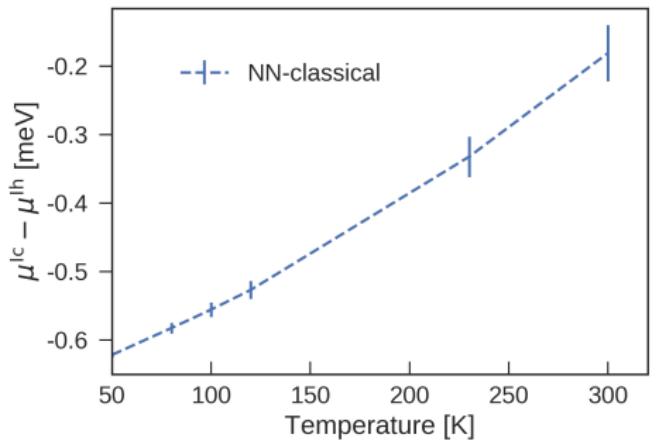
Collaborators

- Jörg Behler, Universität Göttingen
- Jan Gerit Brandenburg, Heidelberg University
- Christoph Dellago, University of Vienna
- Edgar Engel, EPFL
- Guglielmo Mazzola, IBM Zurich
- Tony Paxton, King's College London
- Chris J Pickard, University of Cambridge
- Aleks Reinhardt, University of Cambridge

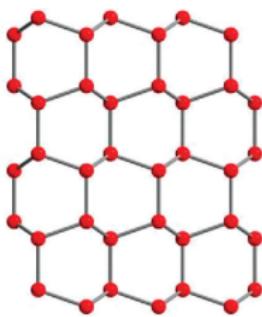


Thank you for
your attention!

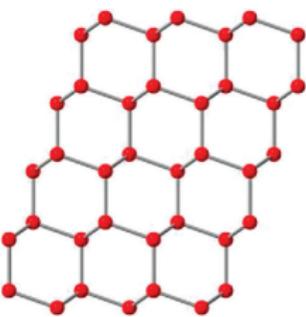
Hexagonal and cubic ice



ice Ih

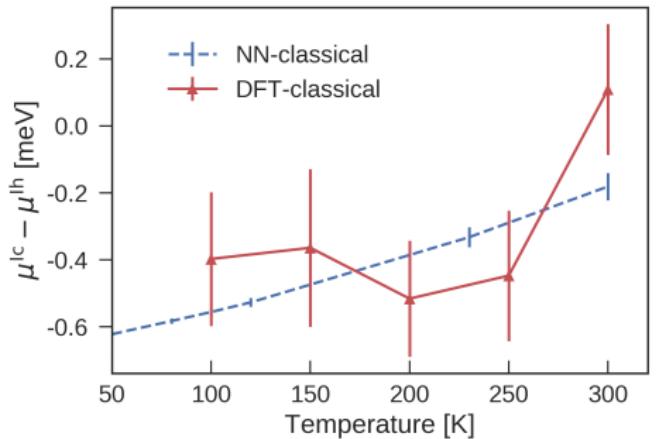


ice Ic

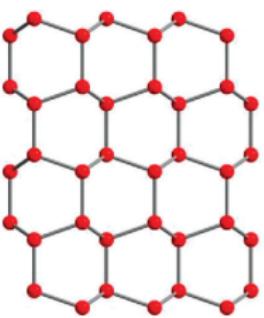


[<http://www.phase-trans.msm.cam.ac.uk/dendrites.html>]

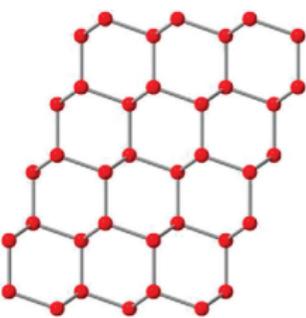
Hexagonal and cubic ice



ice Ih

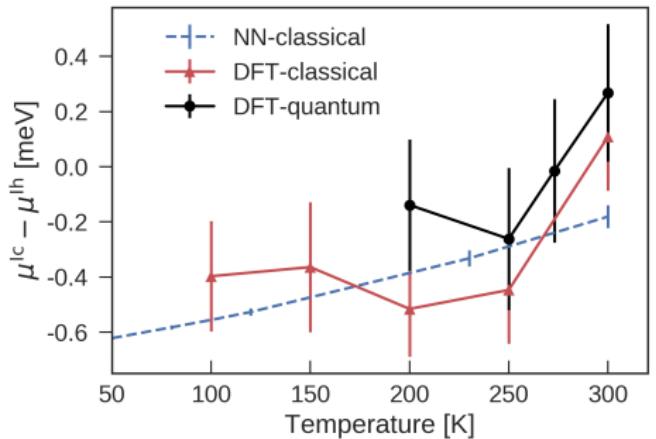


ice Ic

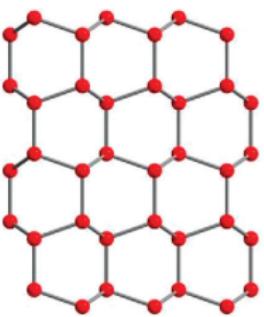


[<http://www.phase-trans.msm.cam.ac.uk/dendrites.html>]

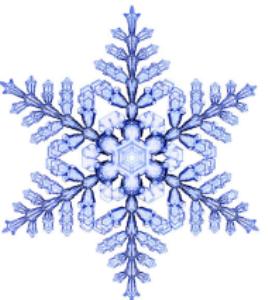
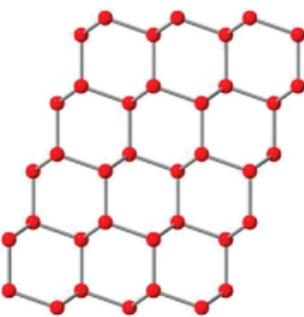
Hexagonal and cubic ice



ice Ih



ice Ic



- NQE significantly stabilize hexagonal ice.
- hexagonal ice is more stable after all!

[<http://www.phase-trans.msm.cam.ac.uk/dendrites.html>]