Computing thermodynamic and transport properties using machine learning potentials

Bingqing Cheng

University of Cambridge
ICTP Virtual School, Jan 2021

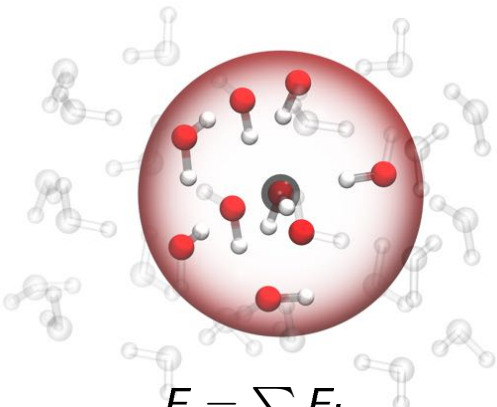
Collaborators: Jorg Behler, Jan Gerit Brandenburg, Michele Ceriotti, Christoph Dellago, Edgar Engel, Daan Frenkel, Bartomeu Monserrat, Aleks Reinhardt

Outline

- A ML potential for water.
- Hydrogen under high pressure.
- The extent of locality in MLP.
  - Computing heat conductivity.
  - Extracting ice-like local environments from liquid water.

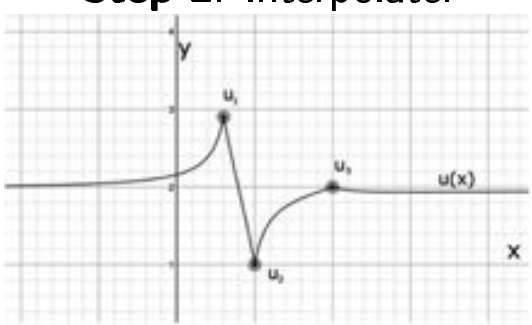
Construct ML potentials

Step 1: Collect environments.



$$E = \sum E_i$$

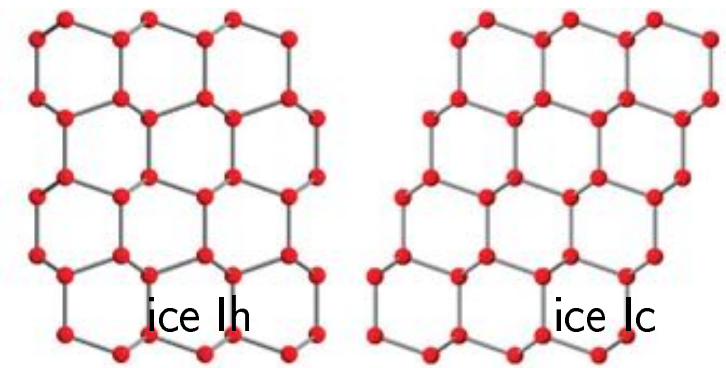
Step 2: Interpolate.



The mystery about water

- Densest at 4 degree Celsius.
- Ice floats on water.
- Unusually high melting point.
- Differences between light (H_2O) and heavy water (D_2O).
- At least 18 ice phases.

- Why hexagonal ice (Ih) is more stable than cubic ice (Ic)?



Training set for bulk water

revPBE0-D3 reference:

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

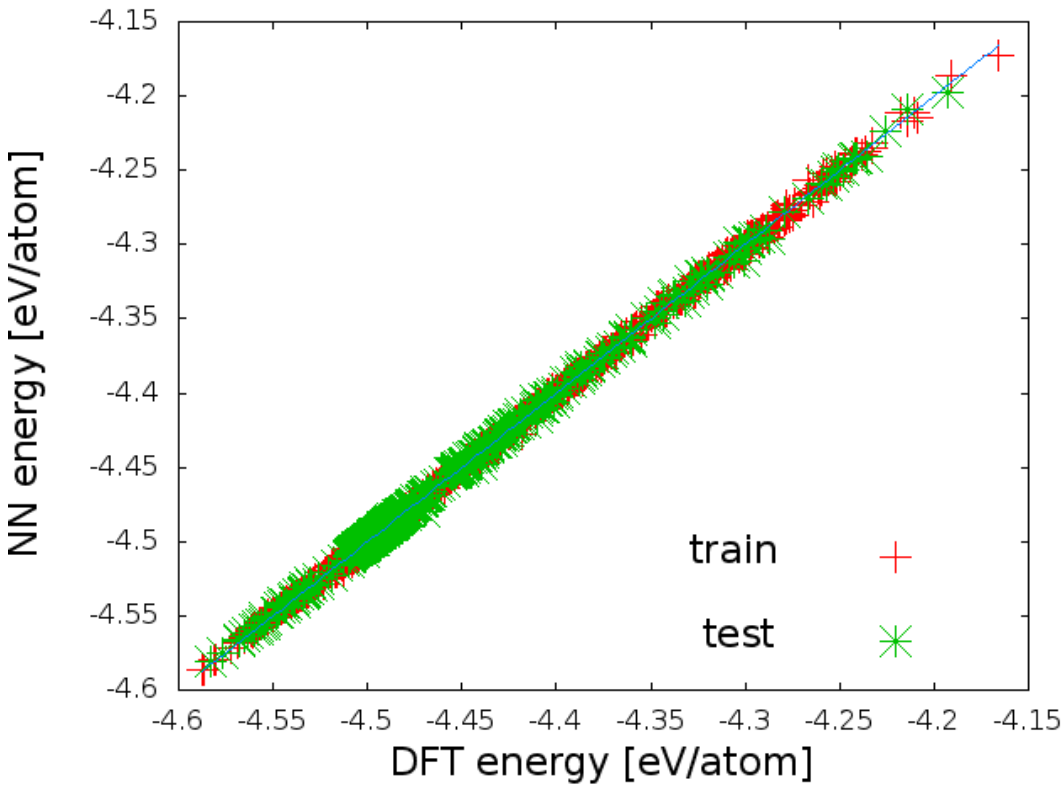
Trained using BP neural network.
[Behler Parrinello PRL 2008; Morawietz, et al. PNAS 2016]

Training set:

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

Neural network potential for water

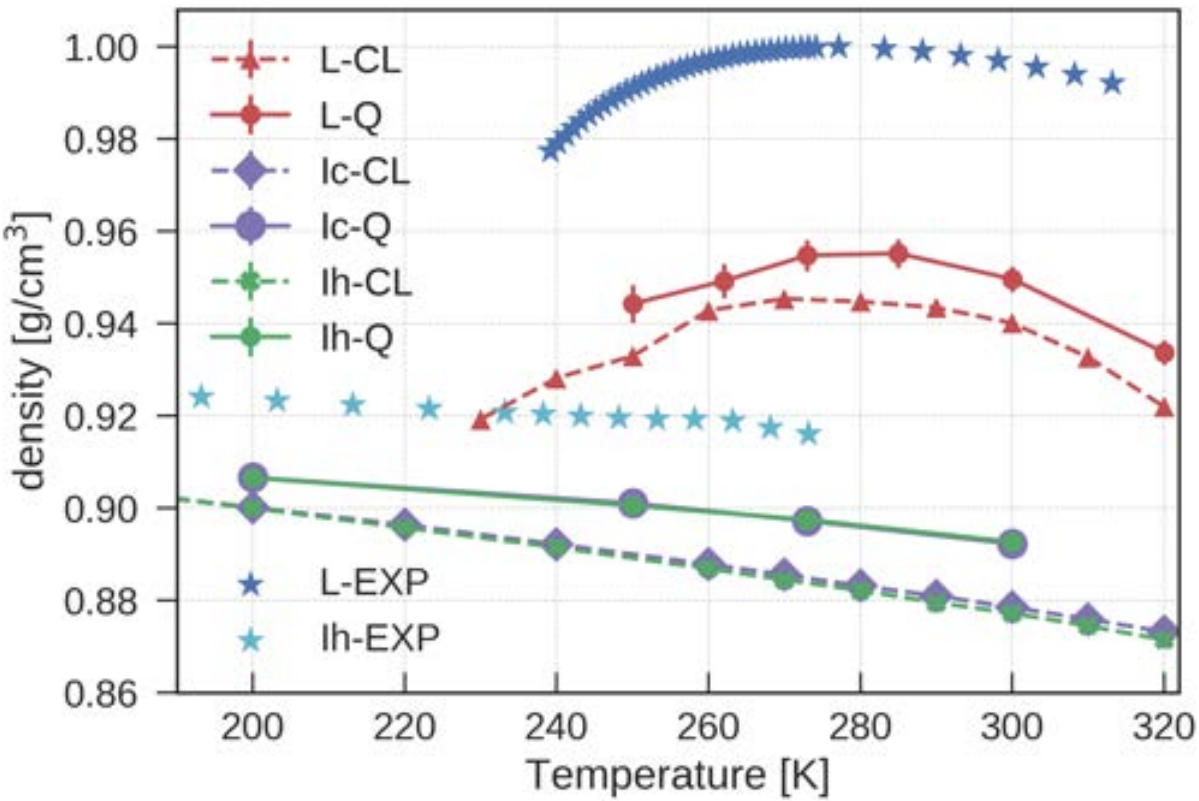
[Cheng, Engel, Behler, Dellago & Ceriotti PNAS 2019]



revPBE0+D3 DFT functional

Neural network potential for water

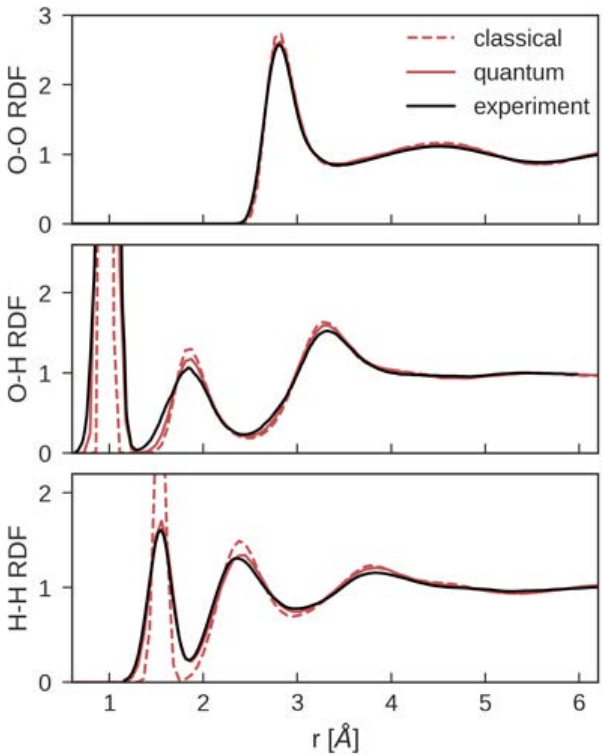
[Cheng, Engel, Behler, Dellago & Ceriotti PNAS 2019]



revPBE0+D3 DFT functional

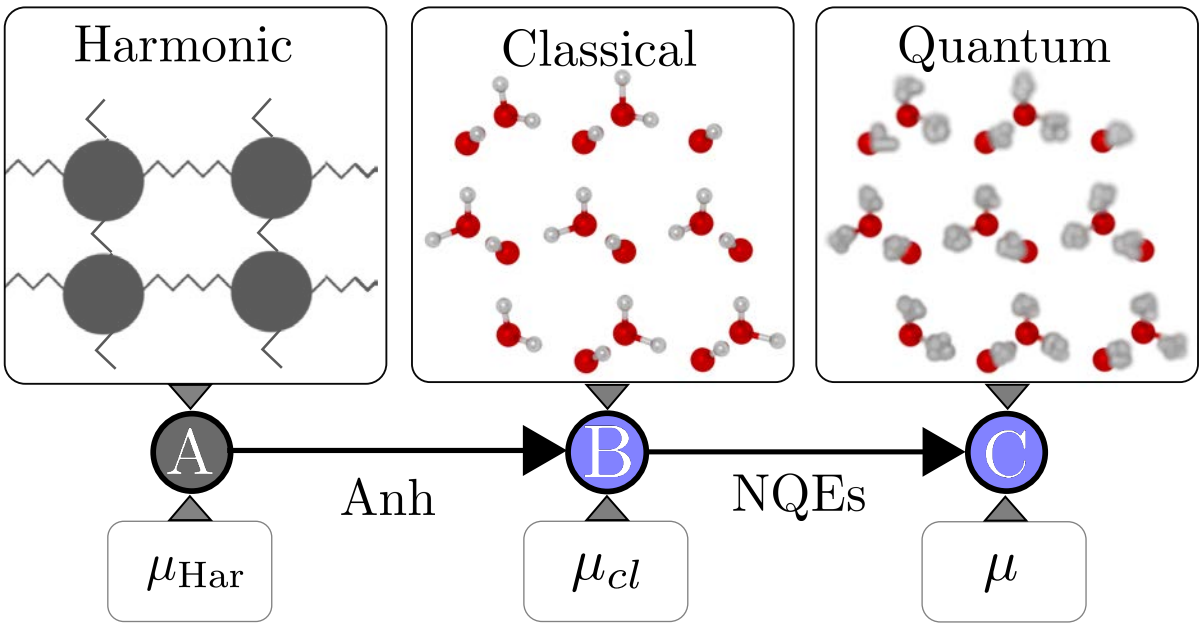
Neural network potential for water

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



revPBE0+D3 DFT functional

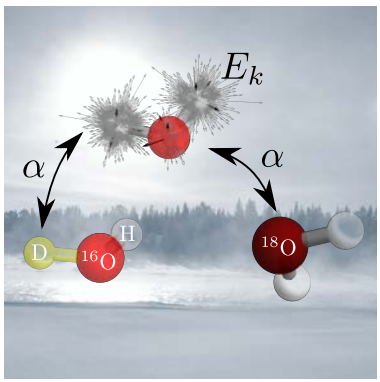
Thermodynamic integration



Transform continuously between systems A and B via a parameter λ,

$$F_A - F_B = \int_{\lambda_A}^{\lambda_B} \frac{dF(\lambda)}{d\lambda} d\lambda$$

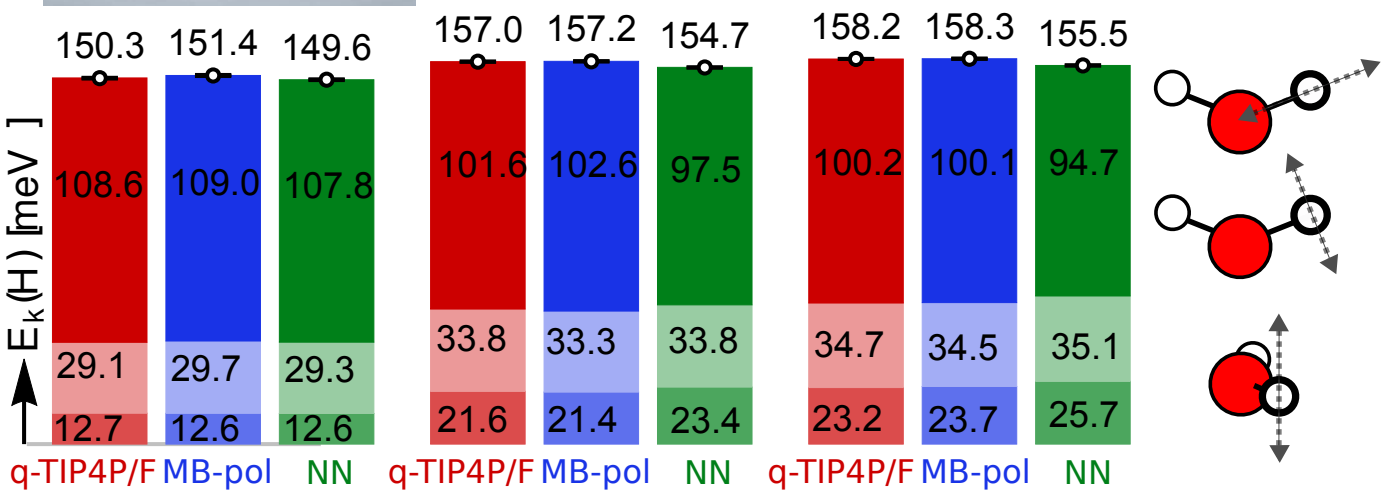
From classical to quantum nuclei



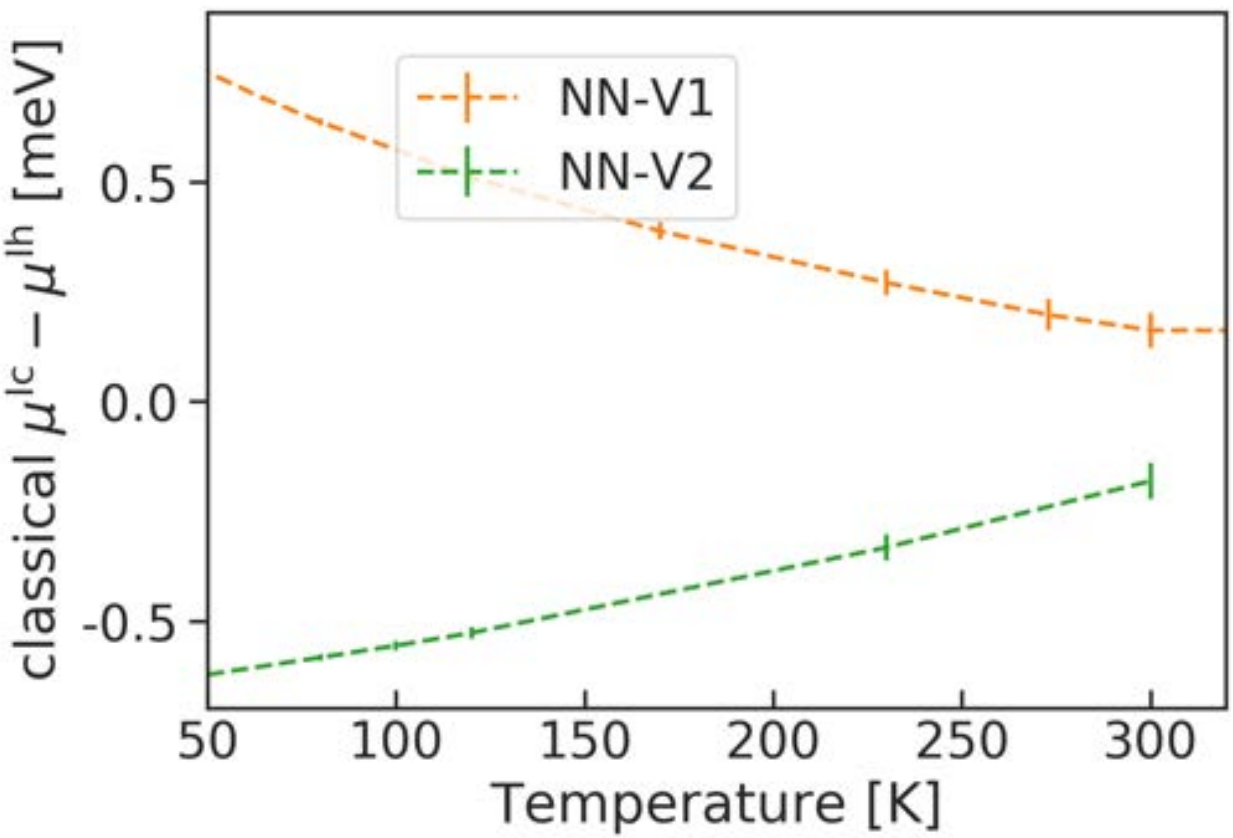
[Cheng & Ceriotti JCP 2014;
Cheng, Behler & Ceriotti JPCL 2016]

$$\frac{dG}{du} = -\frac{\langle E_k(u)\rangle}{u}$$

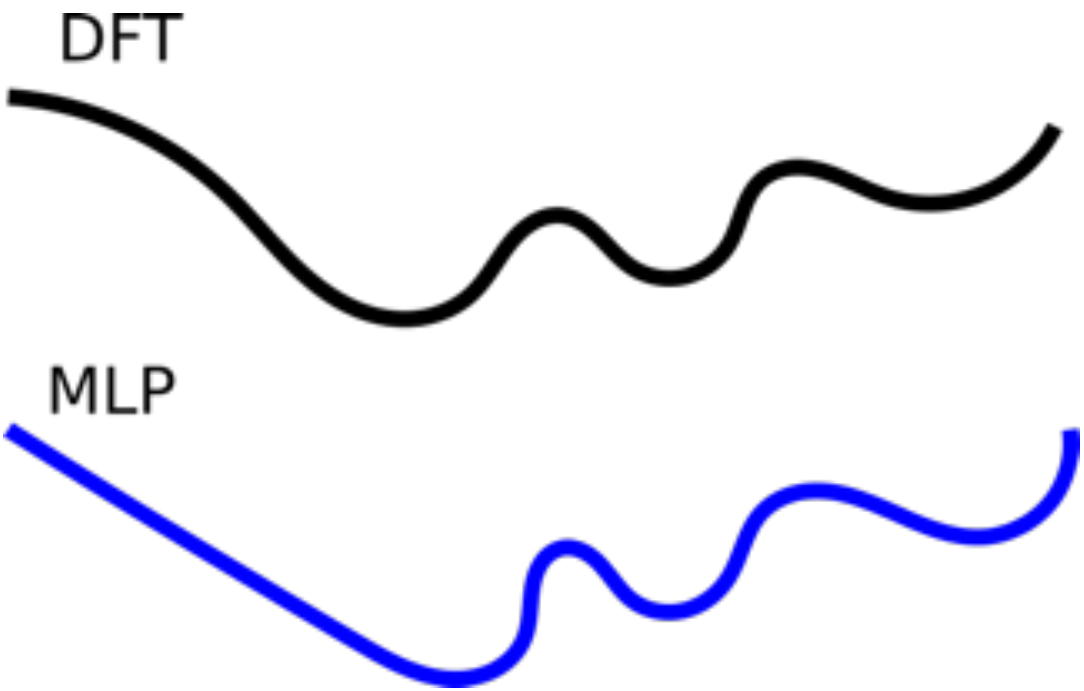
$$G_{qm} - G_{cl} = \int_m^\infty du \frac{\langle E_k(u)\rangle}{u}$$



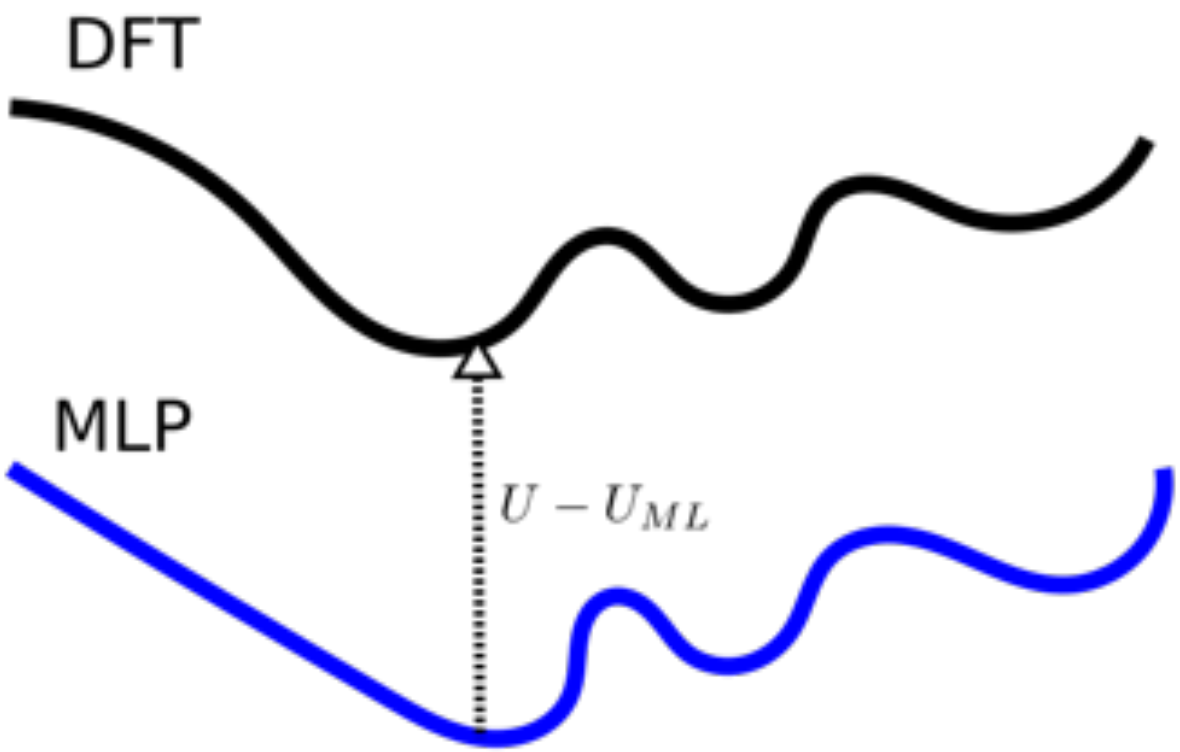
Comparison between two NN potentials



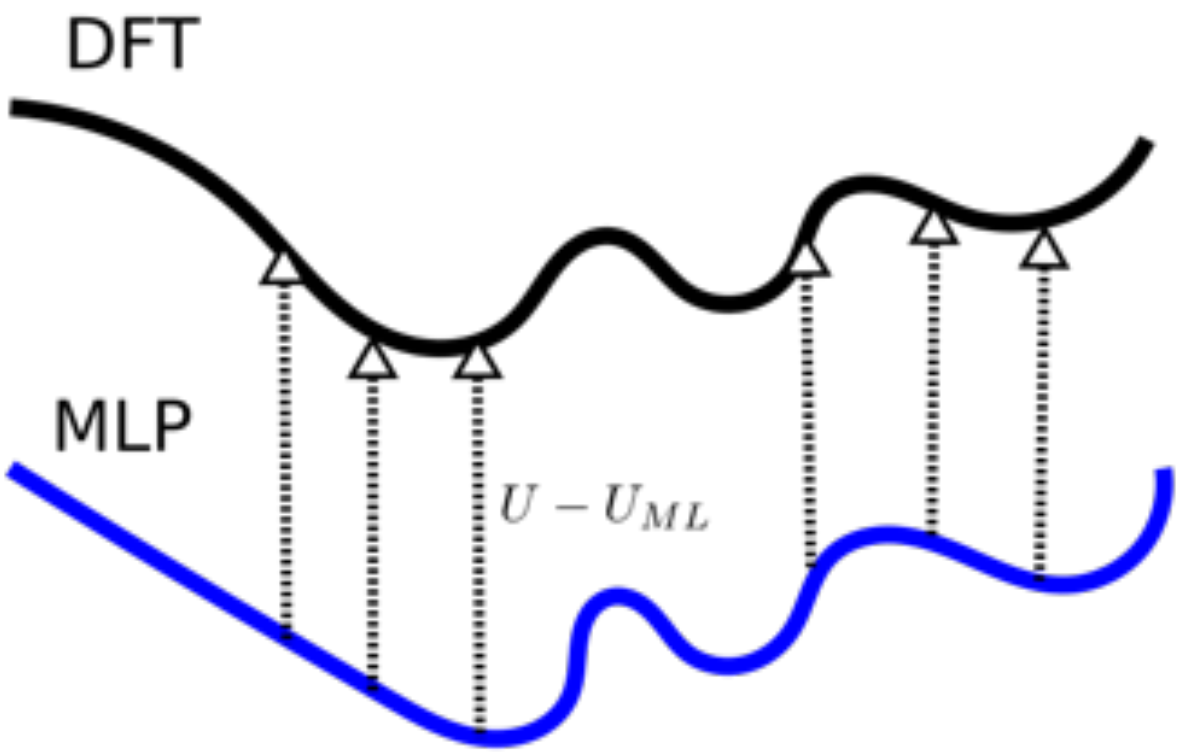
Promote NNP to DFT



Promote NNP to DFT



Promote NNP to DFT



From neural network potential to DFT

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

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

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

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:

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

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

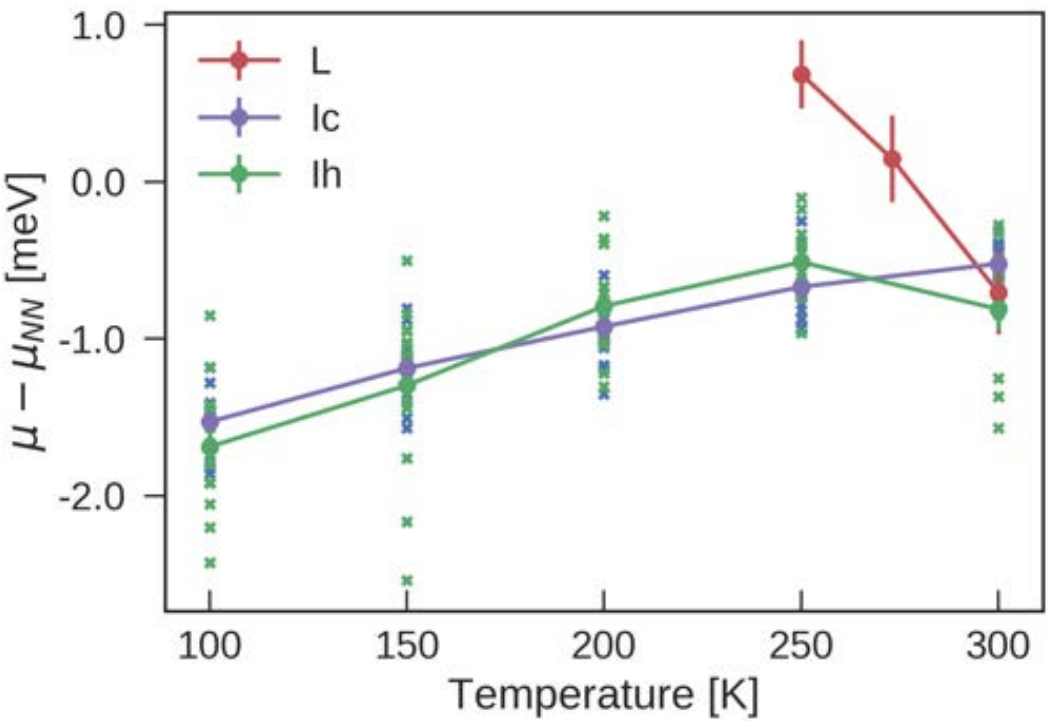
$$G - G_{ML} = -k_B T \ln \left\langle \exp\left[\frac{U(\mathbf{q}) - U_{ML}(\mathbf{q})}{k_B T}\right] \right\rangle_{ML}$$

Free energy perturbation method

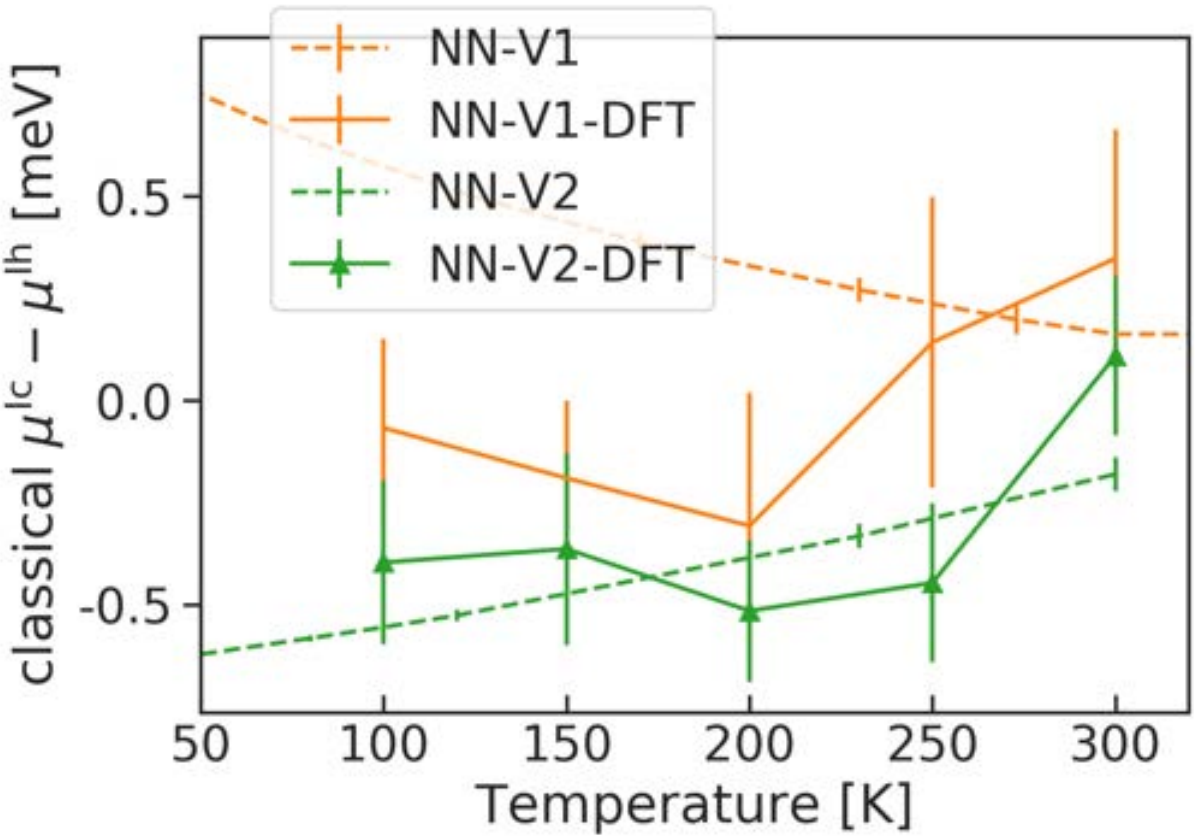
From neural network potential to DFT

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

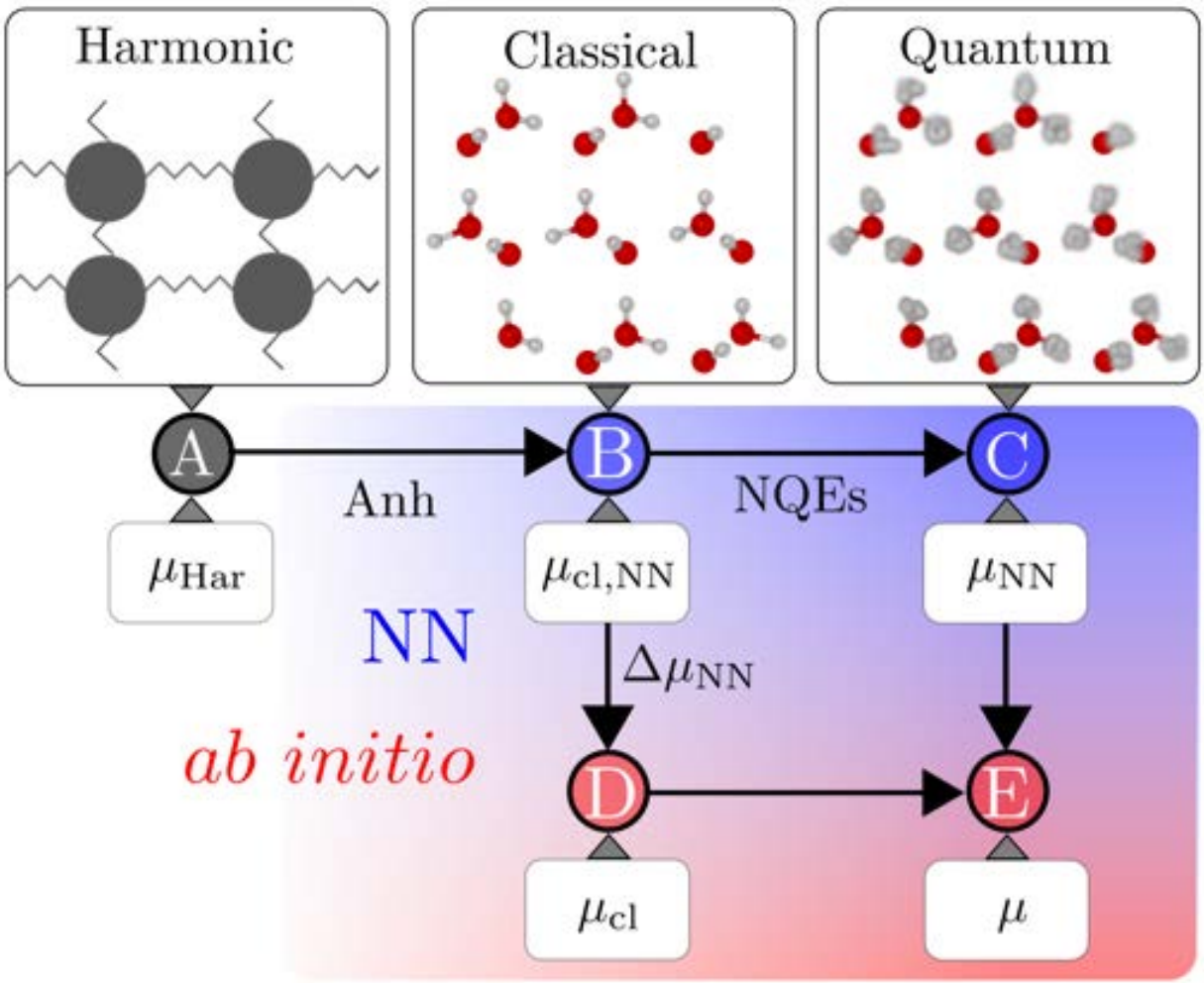
$$\Delta\mu = (G - G_{ML})/N$$



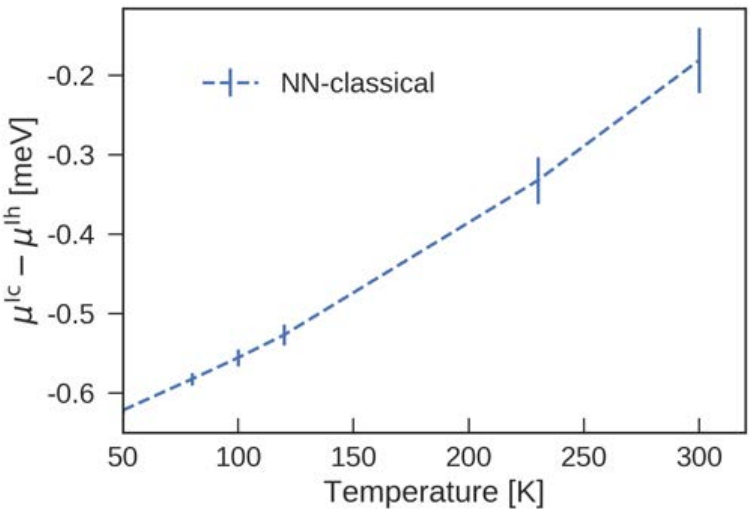
Comparison between two NN potentials



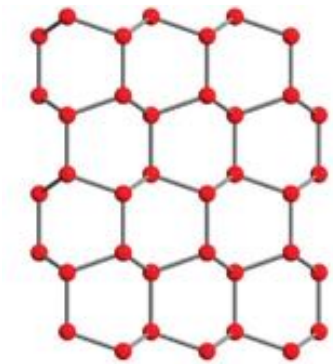
The workflow of ab initio thermodynamics



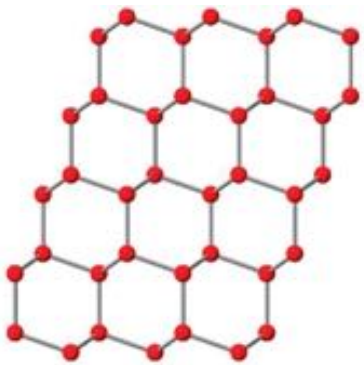
Chemical potentials of hexagonal and cubic ice



ice Ih

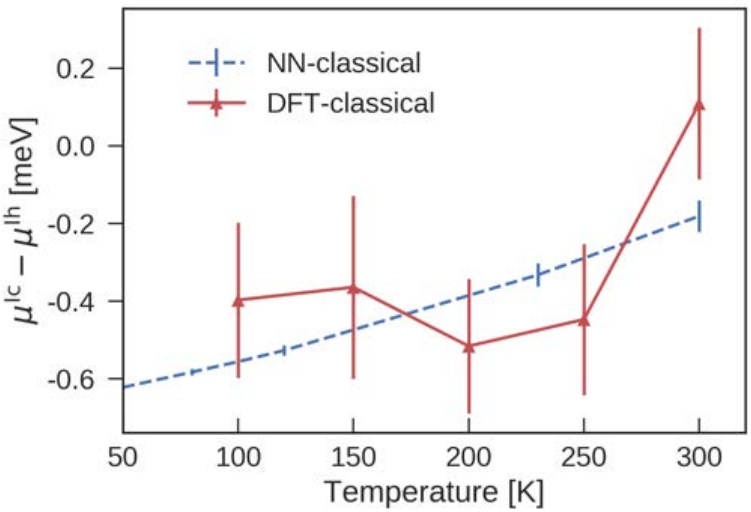


ice Ic

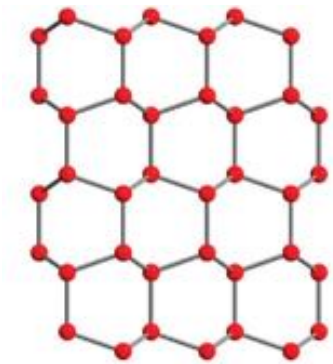


[http://www.phase-trans.msm.cam.ac.uk/dendrites.html]

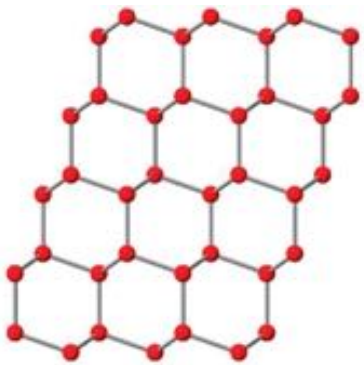
Chemical potentials of hexagonal and cubic ice



ice Ih

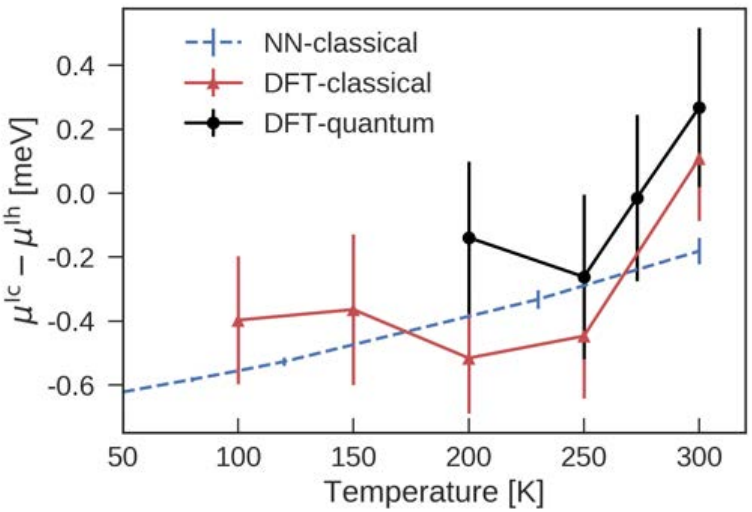


ice Ic

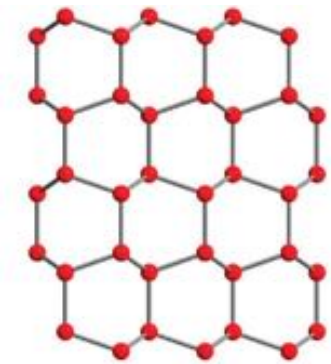


[http://www.phase-trans.msm.cam.ac.uk/dendrites.html]

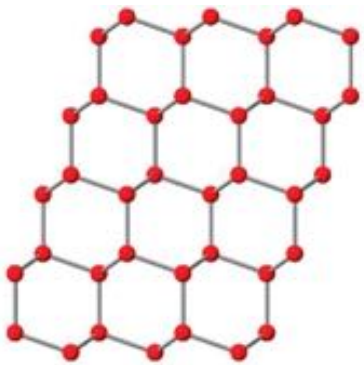
Chemical potentials of hexagonal and cubic ice



ice Ih



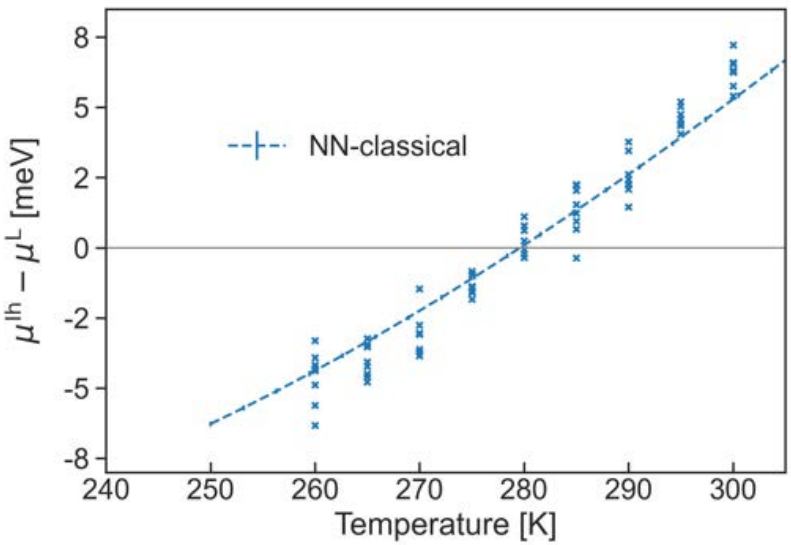
ice Ic



- NQEs significantly stabilize hexagonal ice.
- Hexagonal ice is more stable.

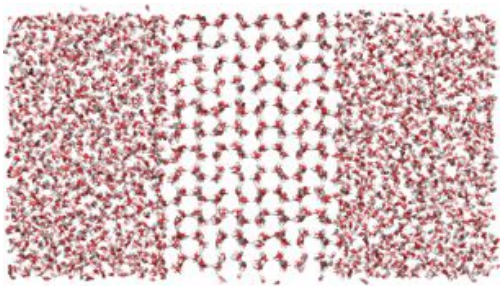
[http://www.phase-trans.msm.cam.ac.uk/dendrites.html]

Chemical potentials of ice and liquid water



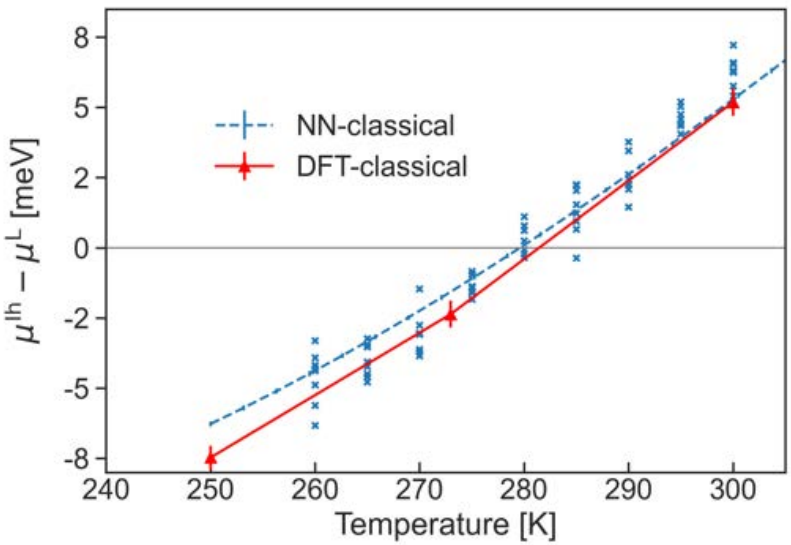
Umbrella simulation using NN potential:

$$\mathcal{H}_{biased}(\mathbf{q}) = \mathcal{H}_{ML}(\mathbf{q}) + \frac{\kappa}{2}\left(\Phi - \bar{\Phi}\right)^2$$



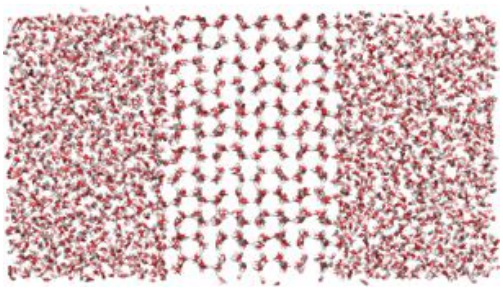
- The difference between the melting points of H_2O and D_2O agrees with experiment (3.82 K).
- D_2O and classical water have almost the same chemical potential, why?
- Error in melting point of H_2O is about 1%.

Chemical potentials of ice and liquid water



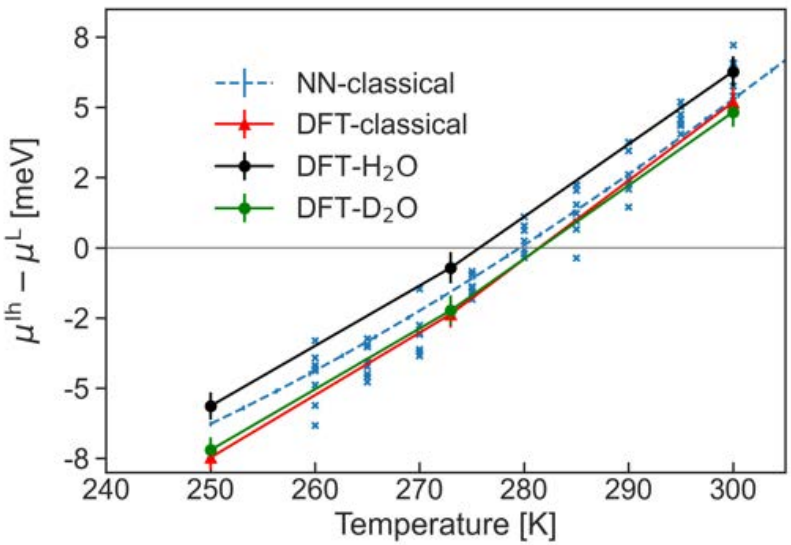
Umbrella simulation using NN potential:

$$\mathcal{H}_{biased}(\mathbf{q}) = \mathcal{H}_{ML}(\mathbf{q}) + \frac{\kappa}{2}\left(\Phi - \bar{\Phi}\right)^2$$



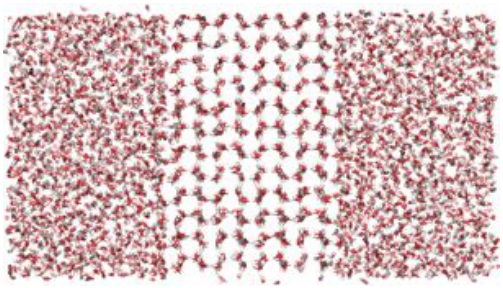
- The difference between the melting points of H_2O and D_2O agrees with experiment (3.82 K).
- D_2O and classical water have almost the same chemical potential, why?
- Error in melting point of H_2O is about 1%.

Chemical potentials of ice and liquid water



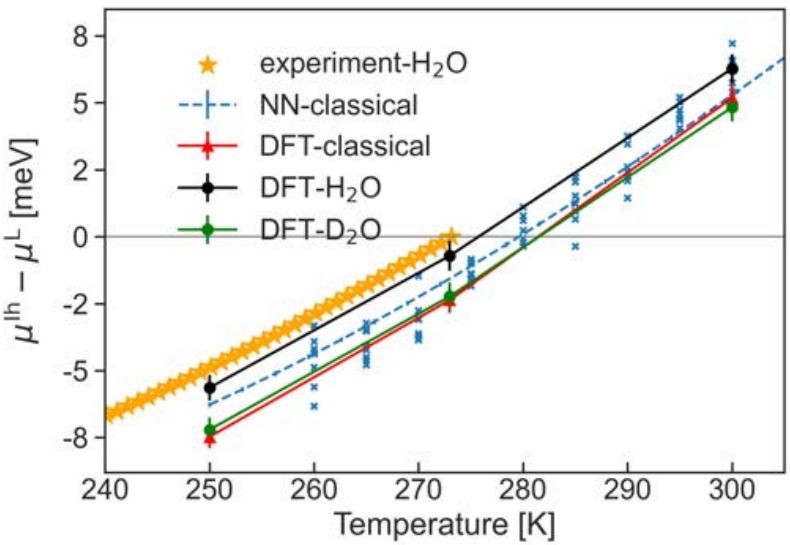
Umbrella simulation using NN potential:

$$\mathcal{H}_{biased}(\mathbf{q}) = \mathcal{H}_{ML}(\mathbf{q}) + \frac{\kappa}{2}\left(\Phi - \bar{\Phi}\right)^2$$



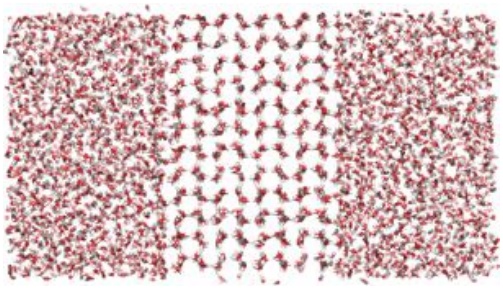
- The difference between the melting points of H_2O and D_2O agrees with experiment (3.82 K).
- D_2O and classical water have almost the same chemical potential, why?
- Error in melting point of H_2O is about 1%.

Chemical potentials of ice and liquid water



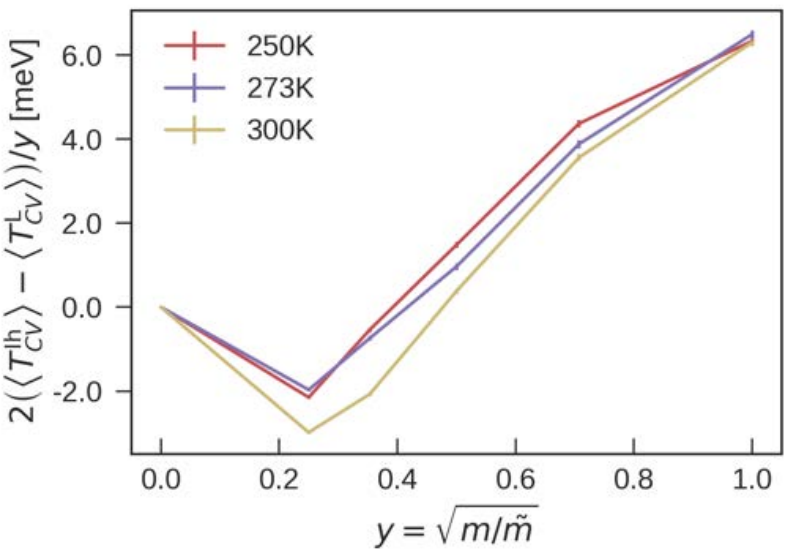
Umbrella simulation using NN potential:

$$\mathcal{H}_{biased}(\mathbf{q}) = \mathcal{H}_{ML}(\mathbf{q}) + \frac{\kappa}{2}\left(\Phi - \bar{\Phi}\right)^2$$



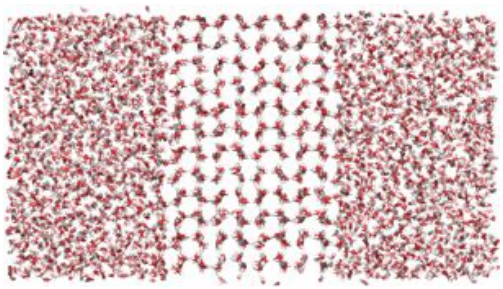
- The difference between the melting points of H_2O and D_2O agrees with experiment (3.82 K).
- D_2O and classical water have almost the same chemical potential, why?
- Error in melting point of H_2O is about 1%.

Chemical potentials of ice and liquid water



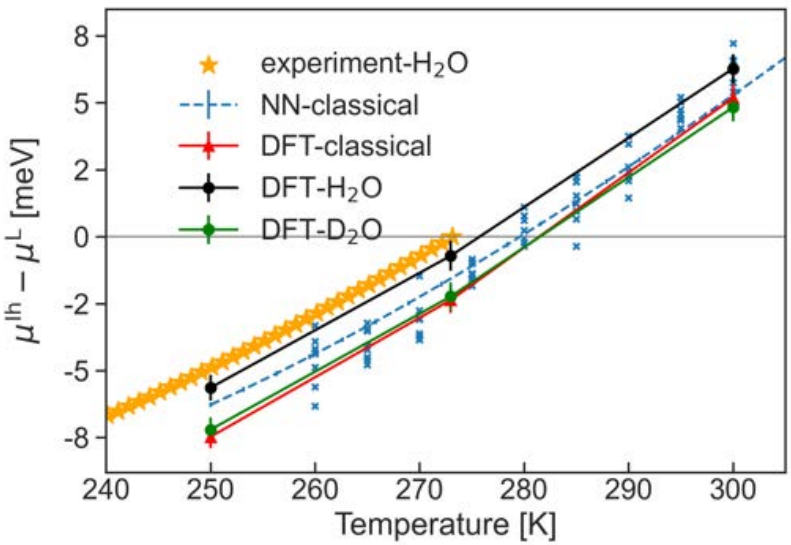
Umbrella simulation using NN potential:

$$\mathcal{H}_{biased}(\mathbf{q}) = \mathcal{H}_{ML}(\mathbf{q}) + \frac{\kappa}{2}\left(\Phi - \bar{\Phi}\right)^2$$



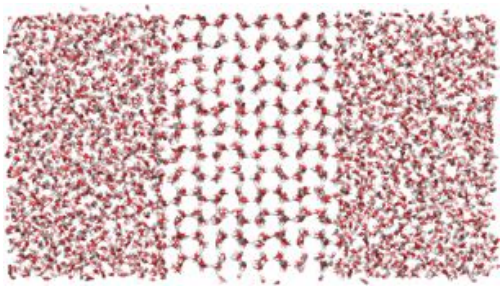
- The difference between the melting points of H_2O and D_2O agrees with experiment (3.82 K).
- D_2O and classical water have almost the same chemical potential, why?
- Error in melting point of H_2O is about 1%.

Chemical potentials of ice and liquid water



Umbrella simulation using NN potential:

$$\mathcal{H}_{biased}(\mathbf{q}) = \mathcal{H}_{ML}(\mathbf{q}) + \frac{\kappa}{2}\left(\Phi - \bar{\Phi}\right)^2$$



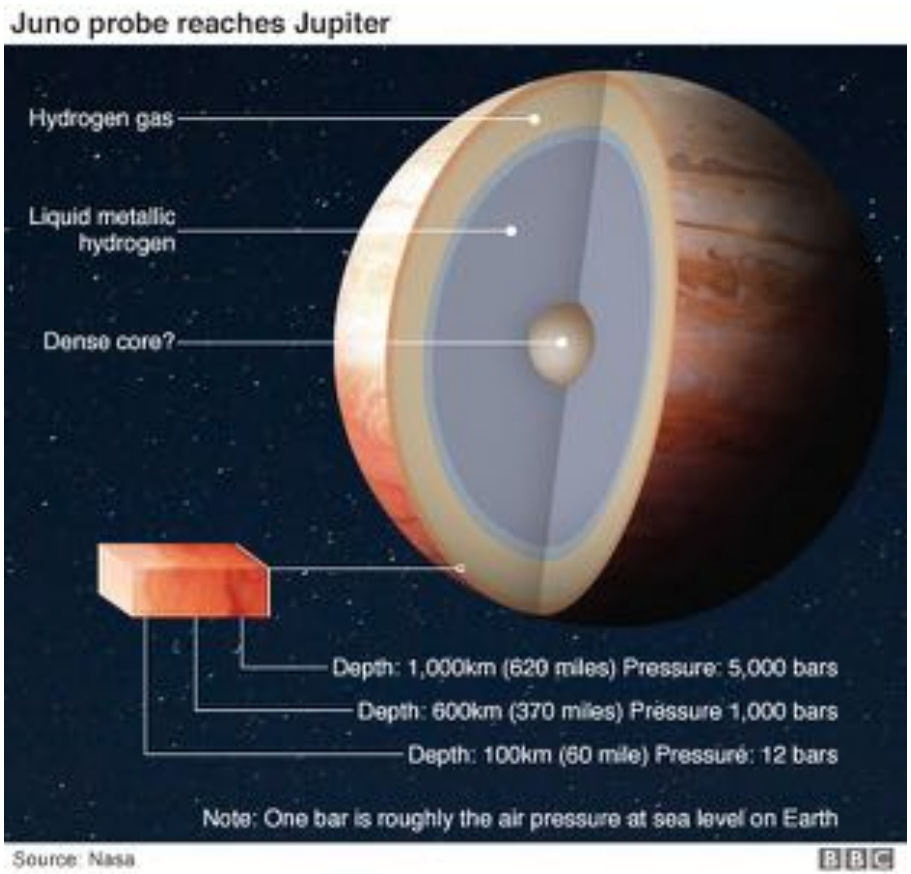
- The difference between the melting points of H_2O and D_2O agrees with experiment (3.82 K).
- D_2O and classical water have almost the same chemical potential, why?
- Error in melting point of H_2O is about 1%.

Outline

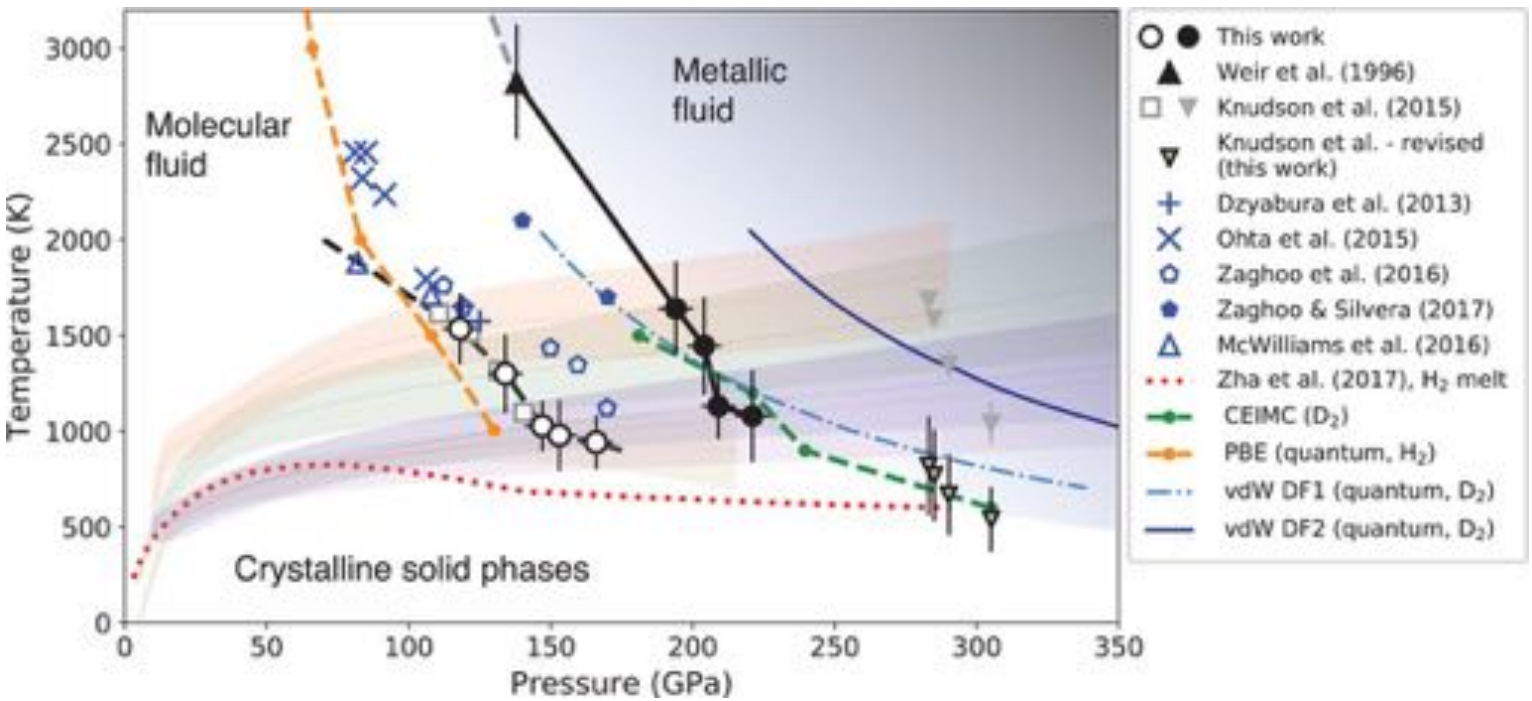
- A ML potential for water.
- Hydrogen under high pressure.
- The extent of locality in MLP.
  - Computing heat conductivity.
  - Extracting ice-like local environments from liquid water.

High pressure hydrogen

- Center of giant planets.
- Exotic properties, e.g. room-temperature superconductor.

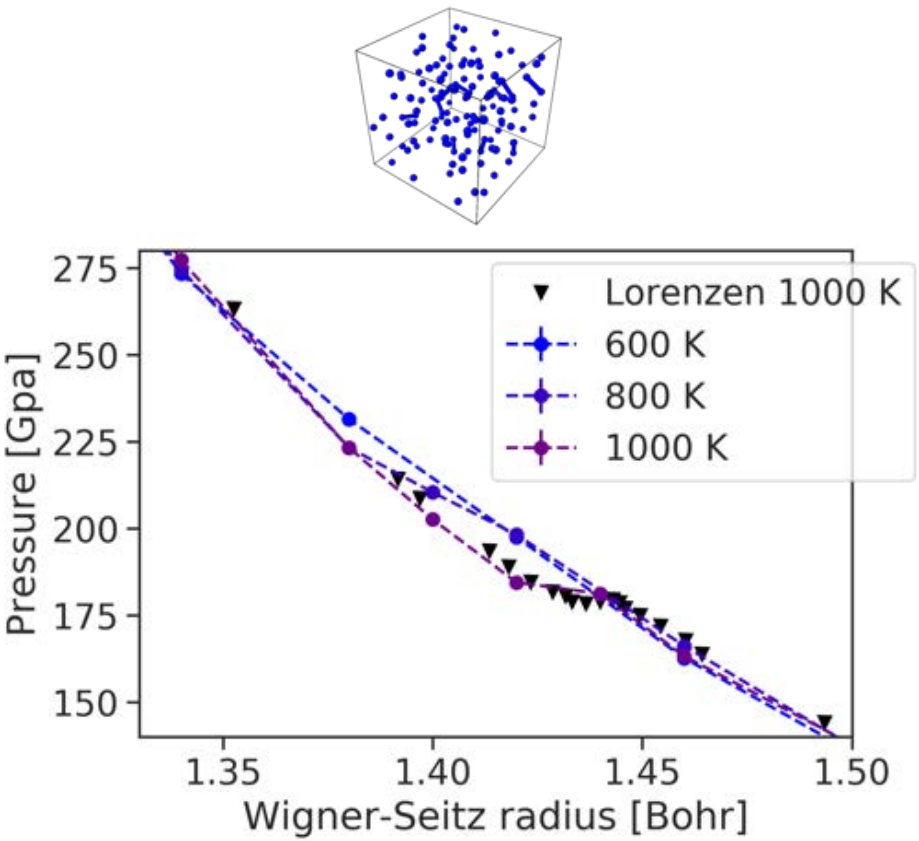


Experimental measurements of liquid-liquid transition (LLT)



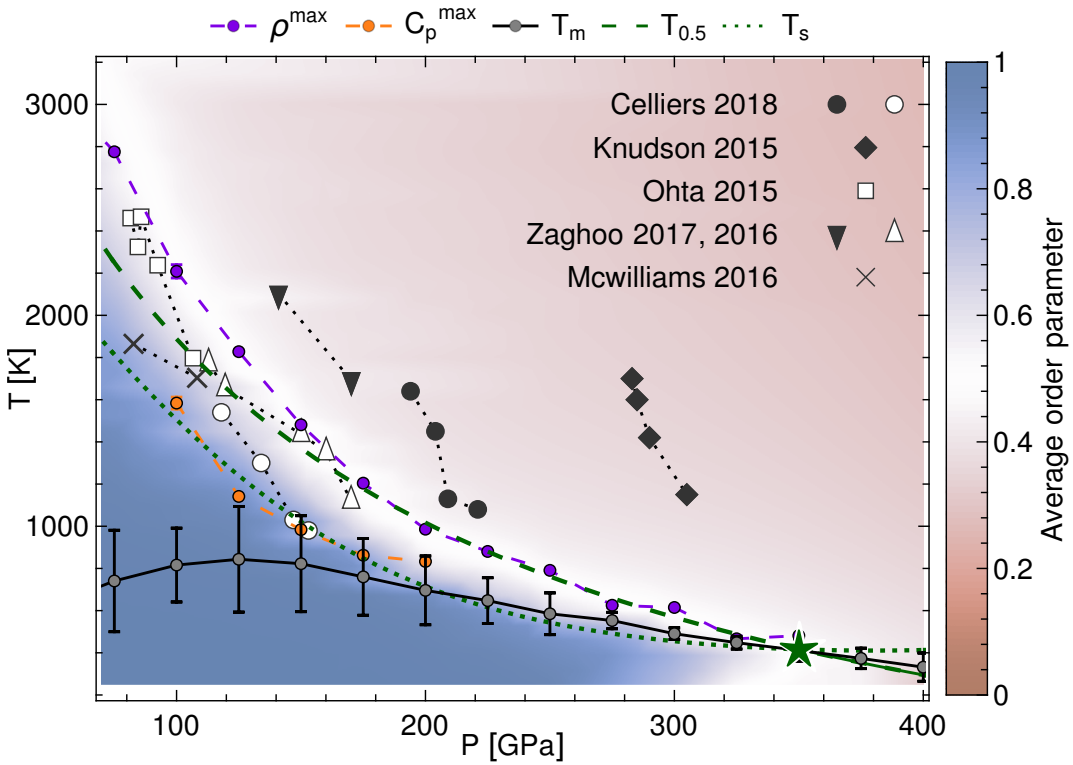
[Celliers et al., Science 361, 677–682 (2018)]

Probing LLT using DFT-MD calculations



Phase diagram of high P hydrogen from MLP

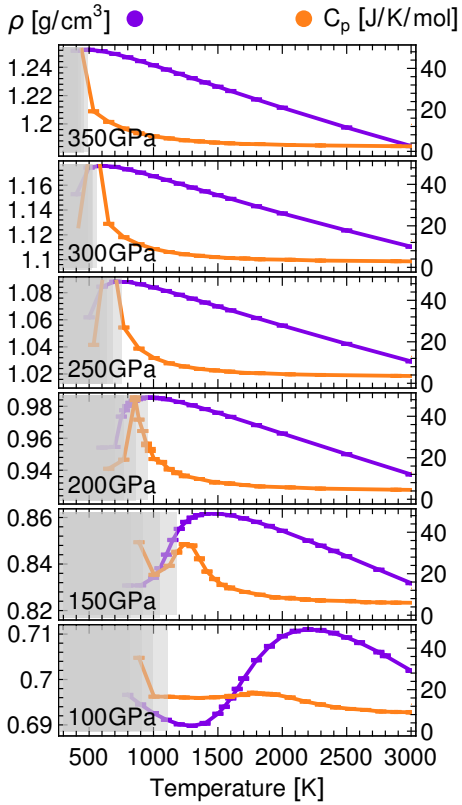
[Cheng, Mazzola, Pickard & Ceriotti Nature 2020]



ML potentials trained using PBE DFT
(we also trained MLPs based on BLYP DFT and quantum Monte Carlo)

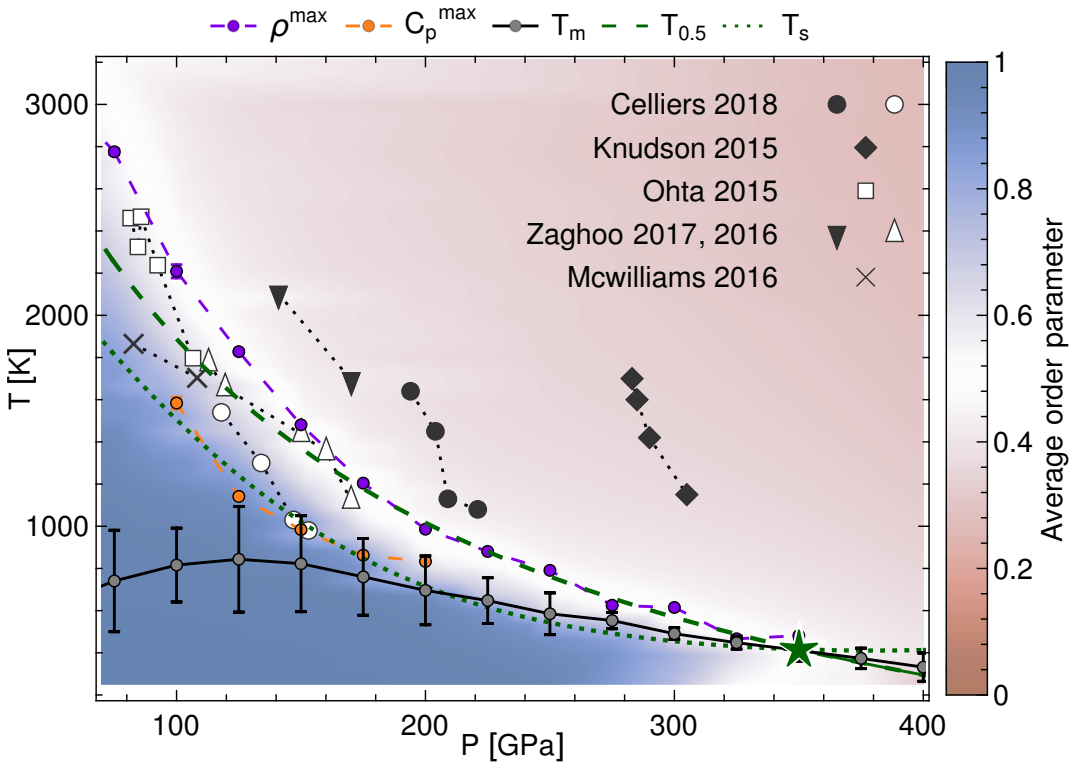
Phase diagram of high P hydrogen from MLP

[Cheng, Mazzola, Pickard & Ceriotti Nature 2020]



Phase diagram of high P hydrogen from MLP

[Cheng, Mazzola, Pickard & Ceriotti Nature 2020]



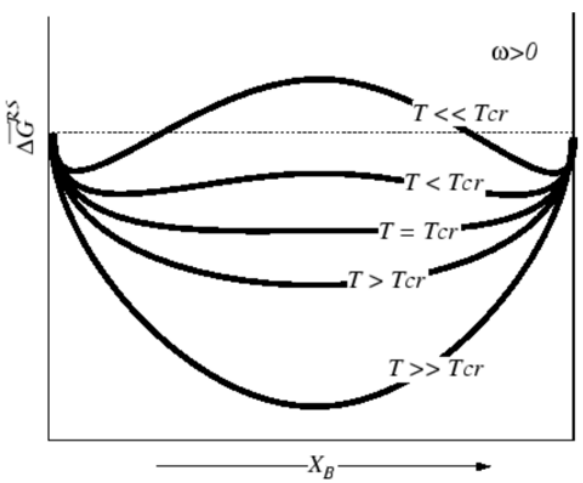
ML potentials trained using PBE DFT
(we also trained MLPs based on BLYP DFT and quantum Monte Carlo)

Phase diagram of high P hydrogen from MLP

[Cheng, Mazzola, Pickard & Ceriotti Nature 2020]

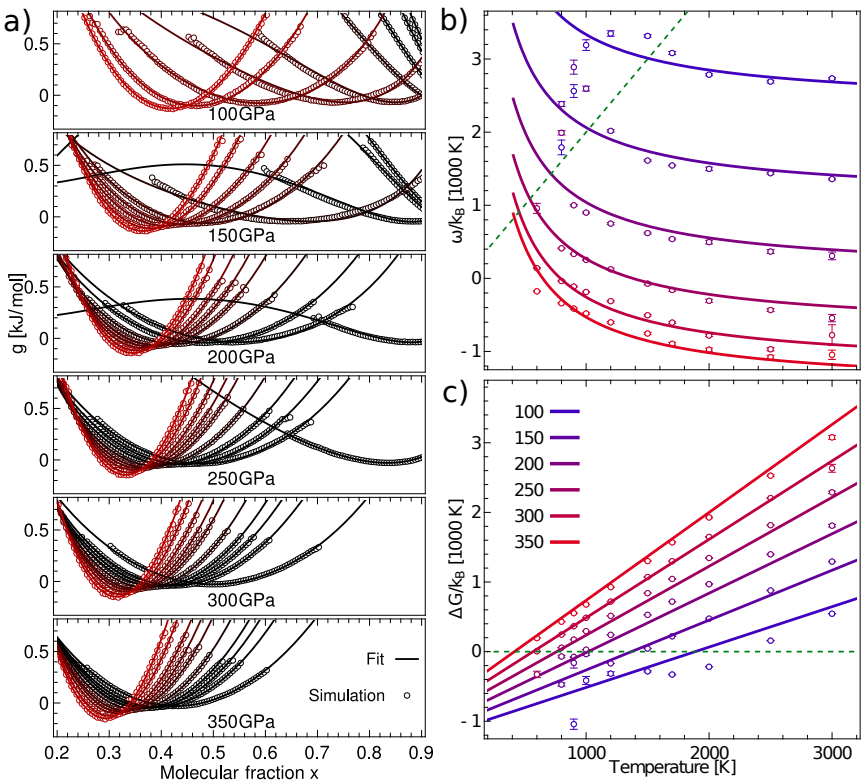
Regular solution model:

$$g(x) = x\Delta g + k_B T x \ln(x) + k_B T (1-x)\ln(1-x) + \omega x(1-x)$$



Phase diagram of high P hydrogen from MLP

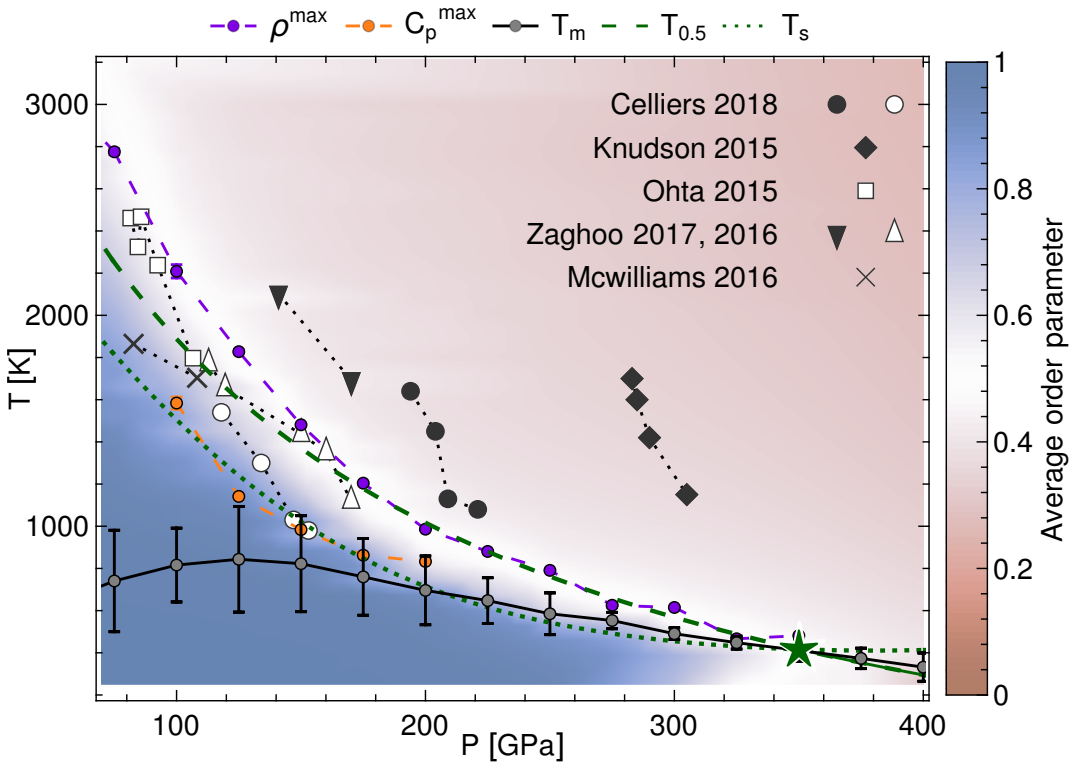
[Cheng, Mazzola, Pickard & Ceriotti Nature 2020]



(metadynamics simulations for computing free energy profiles $g(x)$)

Phase diagram of high P hydrogen from MLP

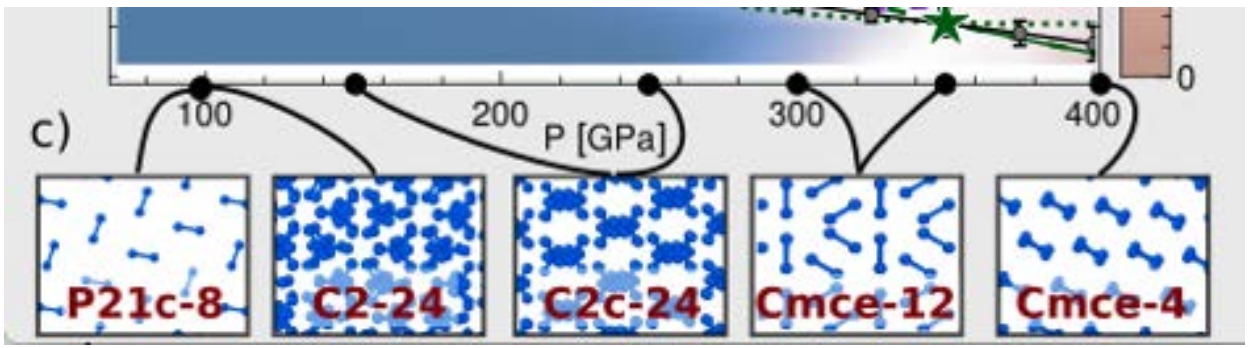
[Cheng, Mazzola, Pickard & Ceriotti Nature 2020]



ML potentials trained using PBE DFT
(we also trained MLPs based on BLYP DFT and quantum Monte Carlo)

Phase diagram of high P hydrogen from MLP

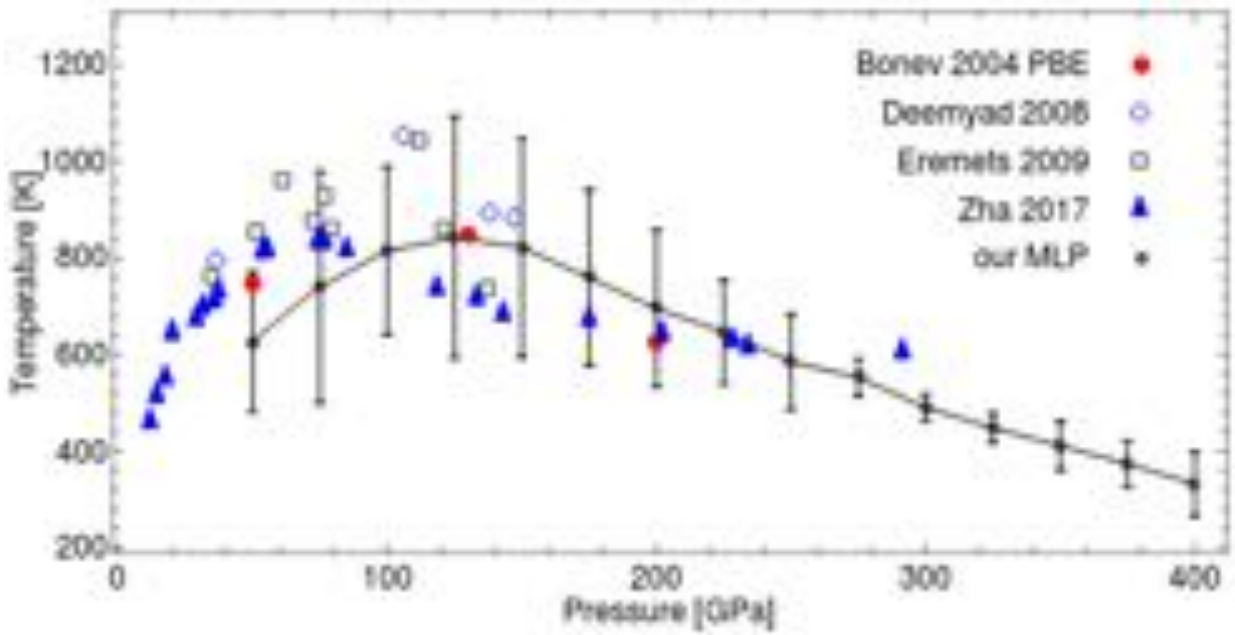
[Cheng, Mazzola, Pickard & Ceriotti Nature 2020]



The MLP correctly captures the ground state crystal structures at different pressures.

Phase diagram of high P hydrogen from MLP

[Cheng, Mazzola, Pickard & Ceriotti Nature 2020]



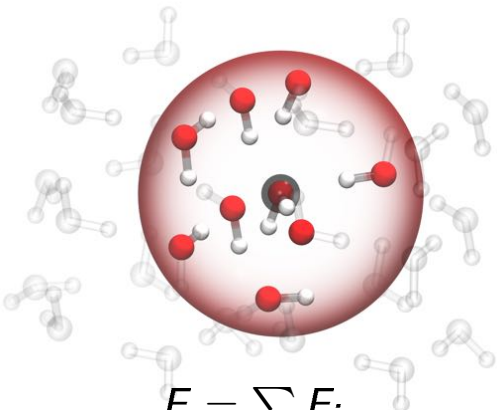
Comparison between the melting lines predicted by the MLP, DFT calculations, and experiments.

Outline

- A ML potential for water.
- Hydrogen under high pressure.
- The extent of locality in MLP.
  - Computing heat conductivity.
  - Extracting ice-like local environments from liquid water.

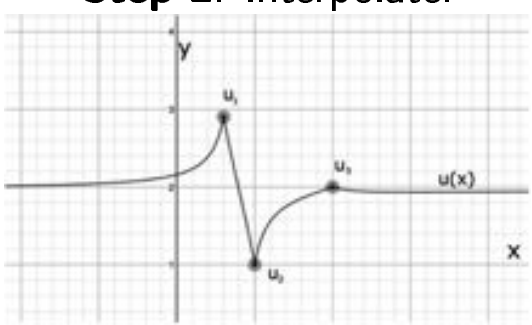
Construct ML potentials

Step 1: Collect environments.

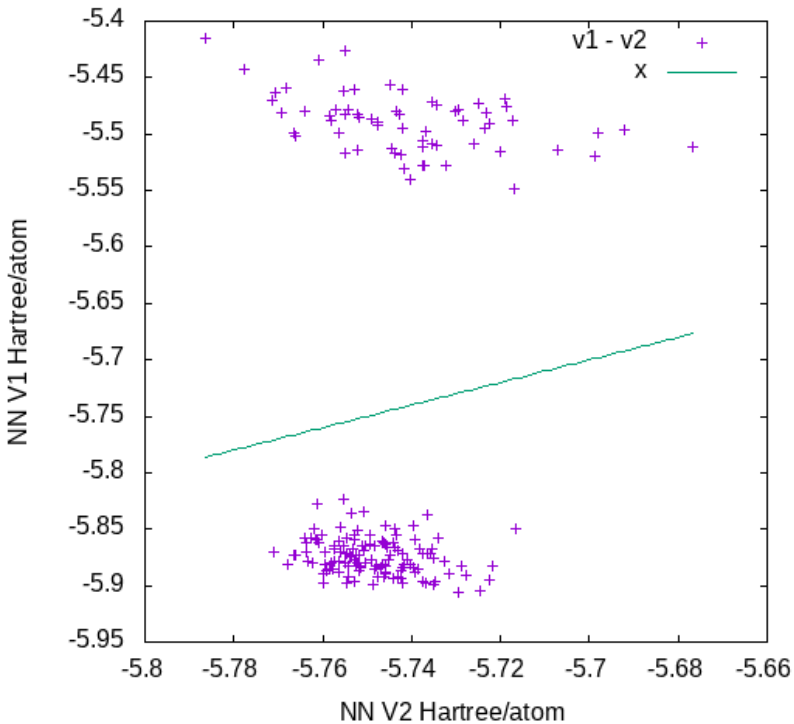


$$E = \sum E_i$$

Step 2: Interpolate.

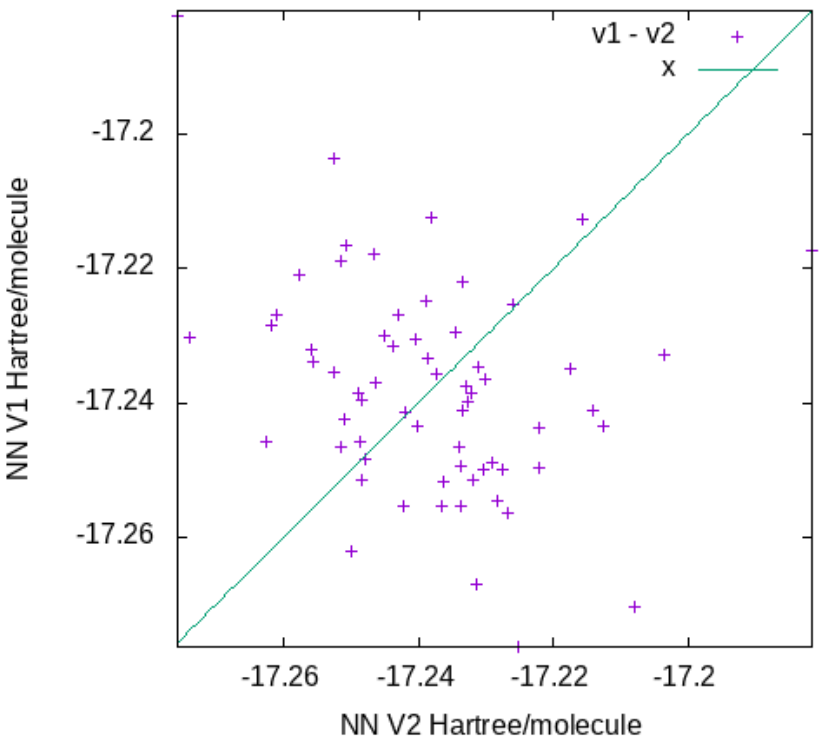


Use the atomic energies of MLPs



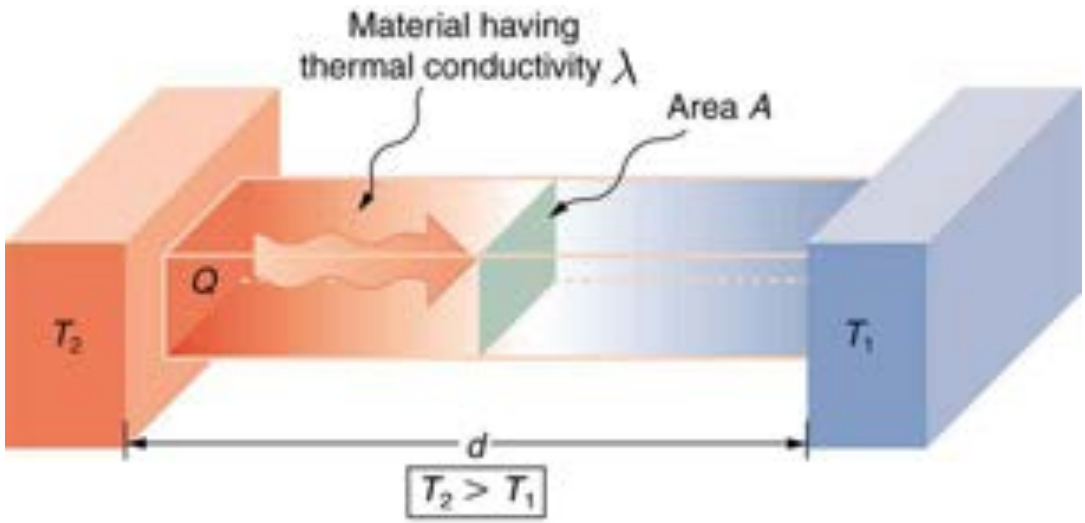
Compare the atomic energies (O, H) from two NNPs

Use the atomic energies of MLPs



Compare molecular energies from two NNPs

Why heat conductivity is important?



[Figure modified from concept-of-physics.com]

- Energy saving, heat shielding
- Energy conversion and harvesting
- Fluid mechanics, earth and planetary science

Green-Kubo for heat conductivity

The Green-Kubo (GK) relationship states that

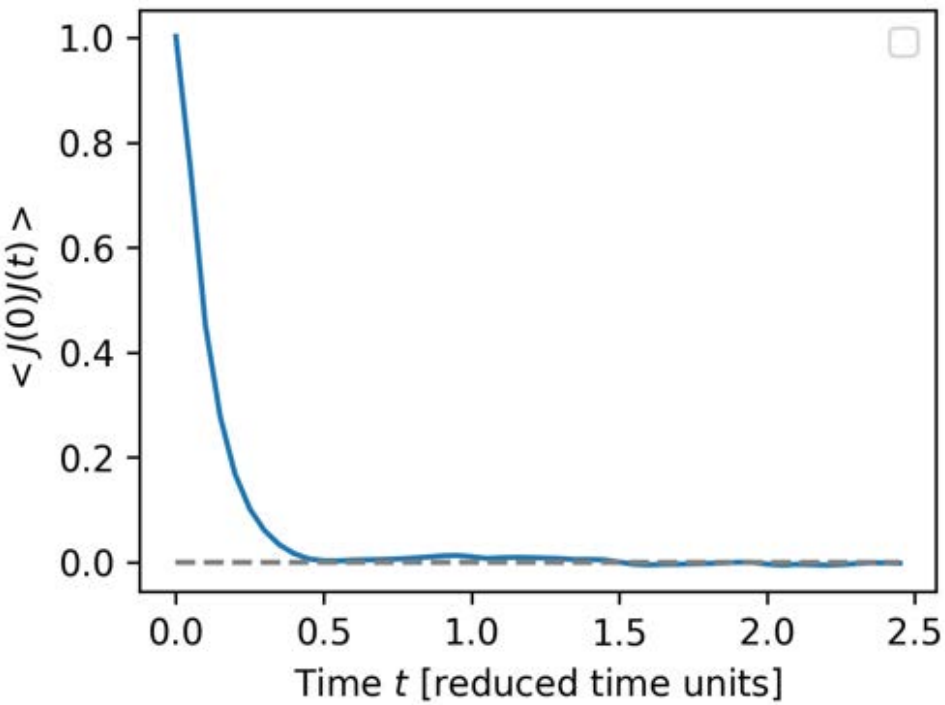
$$\lambda = \frac{1}{Vk_B T^2} \int_0^\infty dt \langle \mathbf{J}(0)\mathbf{J}(t) \rangle$$

where the heat flux is

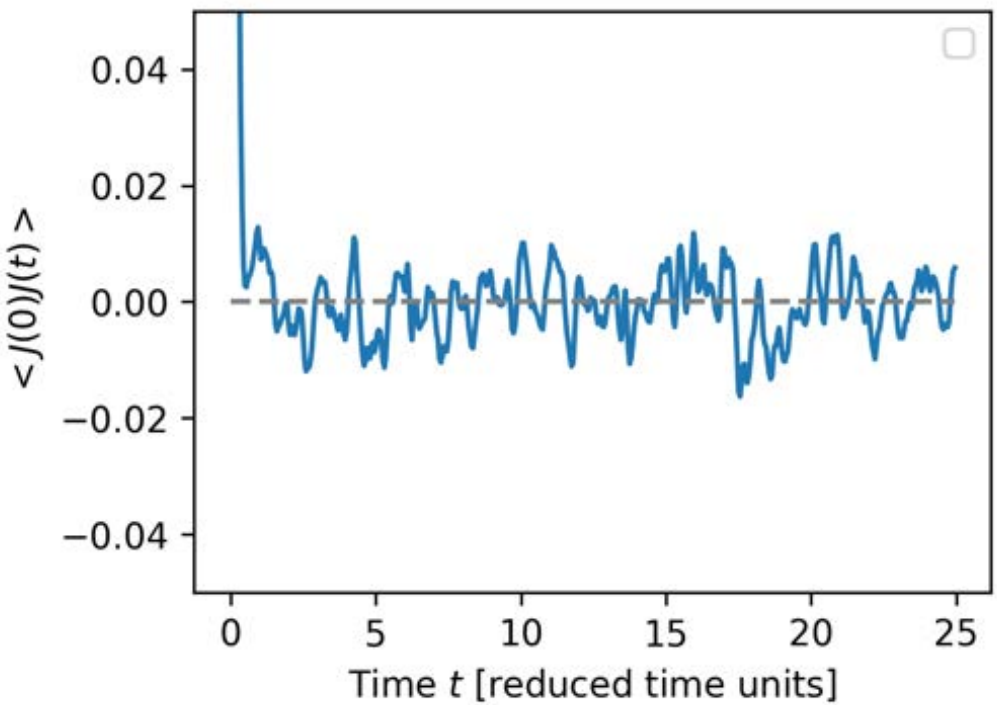
$$\mathbf{J}(t) = \sum_i^N e_i \mathbf{v}_i + \sum_{i<j} (\mathbf{F}_{ij} \cdot \mathbf{v}_i)\mathbf{r}_{ij}$$

The problem:
the atomic energy e_i and forces $\mathbf{F}_{ij}$ between two atoms are poorly defined.

The Green-Kubo integral diverges at the infinite time limit!



The Green-Kubo integral diverges at the infinite time limit!



Green-Kubo for heat conductivity

The Green-Kubo (GK) relationship states that

$$\lambda = \frac{1}{V k_B T^2} \int_0^\infty dt \langle \mathbf{J}(0)\mathbf{J}(t) \rangle$$

where the heat flux is

$$\mathbf{J}(t) = \sum_i^N e_i \mathbf{v}_i + \sum_{i<j} (\mathbf{F}_{ij} \cdot \mathbf{v}_i)\mathbf{r}_{ij}$$

The problem:
the atomic energy e_i and forces $\mathbf{F}_{ij}$ between two atoms are poorly defined.

Heat conductivity from density fluctuations

Fourier expansion of the particle density field $\rho(\mathbf{r}, t)$ in space

$$\widetilde{\rho}(\mathbf{k}, t) = \frac{1}{V}\int_V d\mathbf{r}\rho(\mathbf{r}, t)\exp(-i\mathbf{k}\cdot\mathbf{r}) = \frac{1}{V}\sum_{i=1}^{N}\exp(-i\mathbf{k}\cdot\mathbf{r}_i(t))$$

The approx. solution of the hydrodynamic equations to the 2nd order in $\mathbf{k}$:

$$C_{\widetilde{\rho}}(\mathbf{k}, t) = \int_0^t dt\langle\widetilde{\rho}(\mathbf{k}, 0)\widetilde{\rho}(\mathbf{k}, t)\rangle$$

$$= \rho_k^2\left[\frac{\gamma-1}{\gamma}\exp(-\frac{\lambda}{c_P}k^2 t) + \frac{1}{\gamma}\exp(-\Gamma k^2 t)\cos(c_s k t)\right]$$

Two poles: heat mode $-i\frac{\lambda}{c_P}\mathbf{k}^2$ and sound mode $c_s k - i\Gamma\mathbf{k}^2$.

λ: **heat conductivity** c_s: speed of sound Γ: sound attenuation $\gamma = c_P/c_V$

Hydrodynamic equations

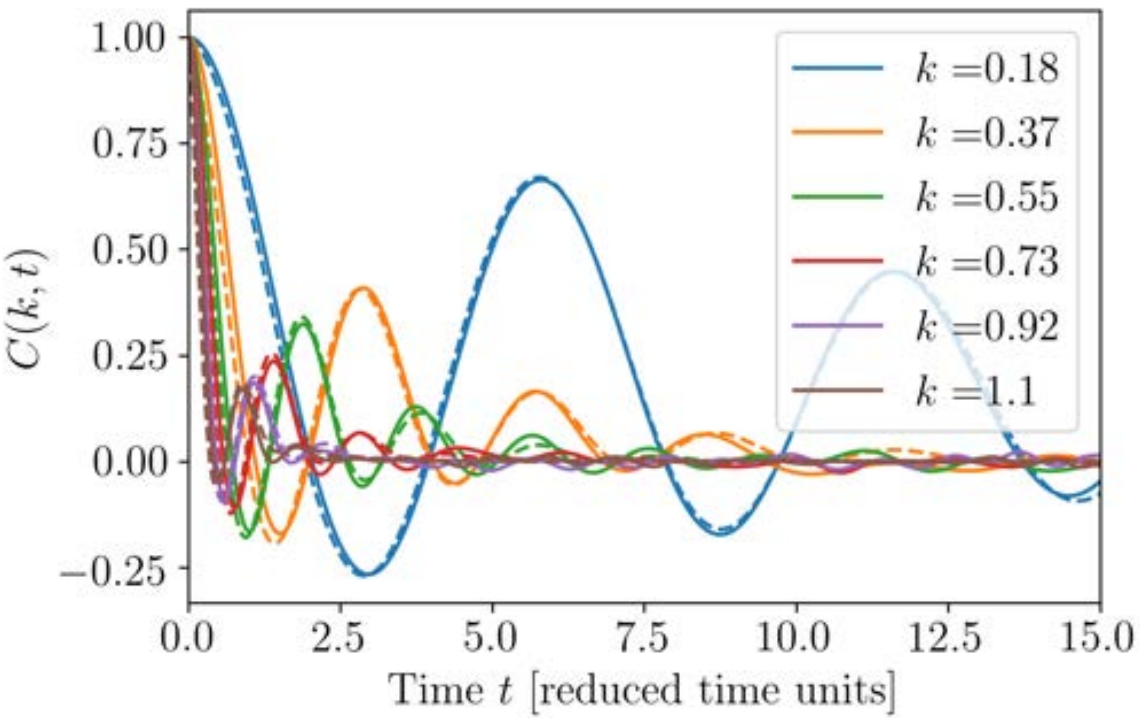
Taking **k** along the z-axis with $k = |\mathbf{k}|$, the coupled hydrodynamic equations for the three quantities $\widetilde{\rho}(\mathbf{k}, t)$, $\widetilde{\mathbf{j}}(\mathbf{k}, t)$ and $\widetilde{T}(\mathbf{k}, t)$ can be written as

$$\begin{pmatrix} -is & 0 & ik & 0 & 0 \\ 0 & -is + ak^2 & \frac{T\beta_V ik}{\rho^2 c_v} & 0 & 0 \\ \frac{ik}{\rho m \chi_T} & \frac{\beta_V ik}{m} & -is + bk^2 & 0 & 0 \\ 0 & 0 & 0 & -is + \nu k^2 & 0 \\ 0 & 0 & 0 & 0 & -is + \nu k^2 \end{pmatrix} \begin{pmatrix} \mathcal{L}\{\widetilde{\rho}\}(s) \\ \mathcal{L}\{\widetilde{T}\}(s) \\ \mathcal{L}\{\widetilde{j}_z\}(s) \\ \mathcal{L}\{\widetilde{j}_x\}(s) \\ \mathcal{L}\{\widetilde{j}_y\}(s) \end{pmatrix} = \begin{pmatrix} \widetilde{\rho} \\ \widetilde{T} \\ \widetilde{j}_z \\ \widetilde{j}_x \\ \widetilde{j}_y \end{pmatrix}$$

where $\mathcal{L}\{\ldots\}(s)$ denotes a Laplace transformation

Benchmark on Lennard-Jones

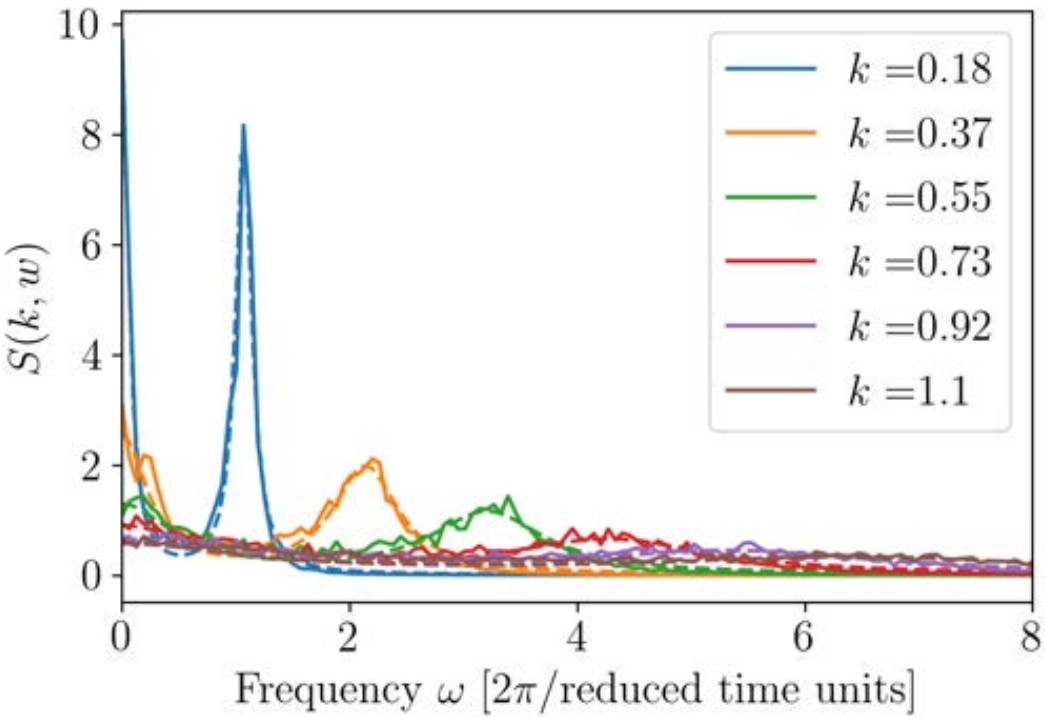
[Cheng & Frenkel PRL 2020]



Dashed lines: results from MD simulations
Solid lines: fits to the hydrodynamic equation.

Benchmark on Lennard-Jones

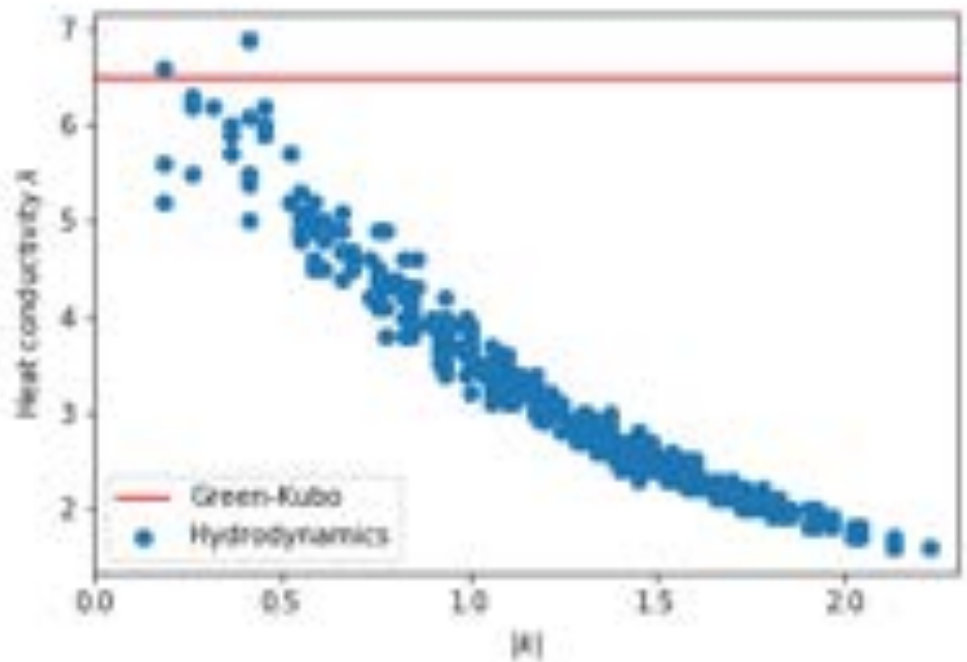
[Cheng & Frenkel PRL 2020]



Dashed lines: results from MD simulations
Solid lines: fits to the hydrodynamic equation.

Benchmark on Lennard-Jones

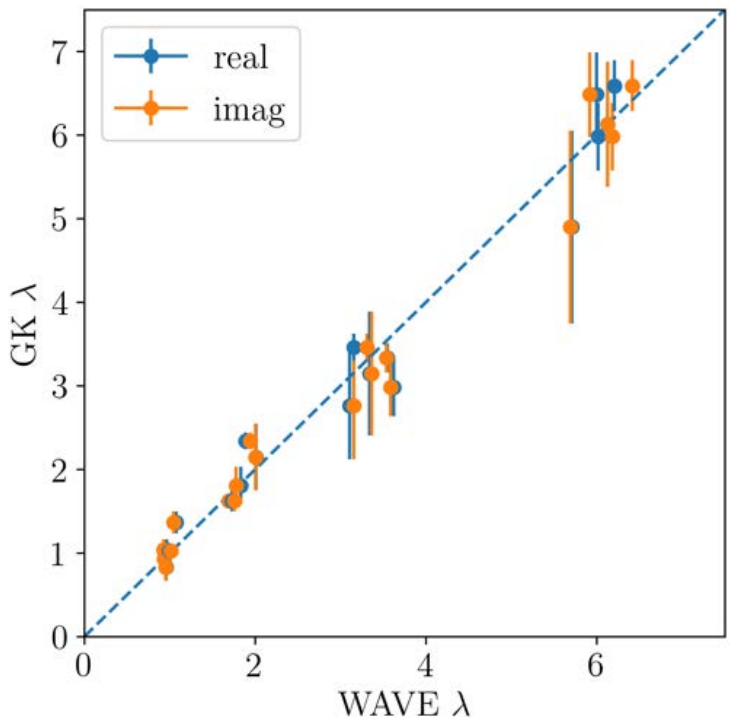
[Cheng & Frenkel PRL 2020]



Blue dots: heat conductivity computed using WAVE at different **k**.
Solid line: Green-Kubo

Benchmark on Lennard-Jones

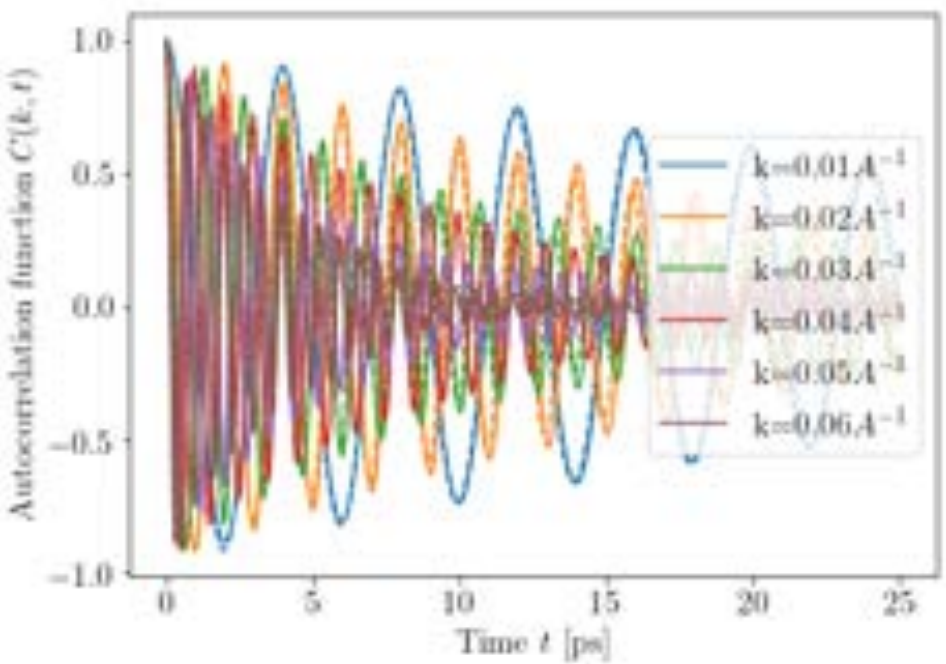
[Cheng & Frenkel PRL 2020]



WAVE Vs. Green-Kubo
at different thermodynamic conditions

Heat conductivity of high-pressure hydrogen fluid

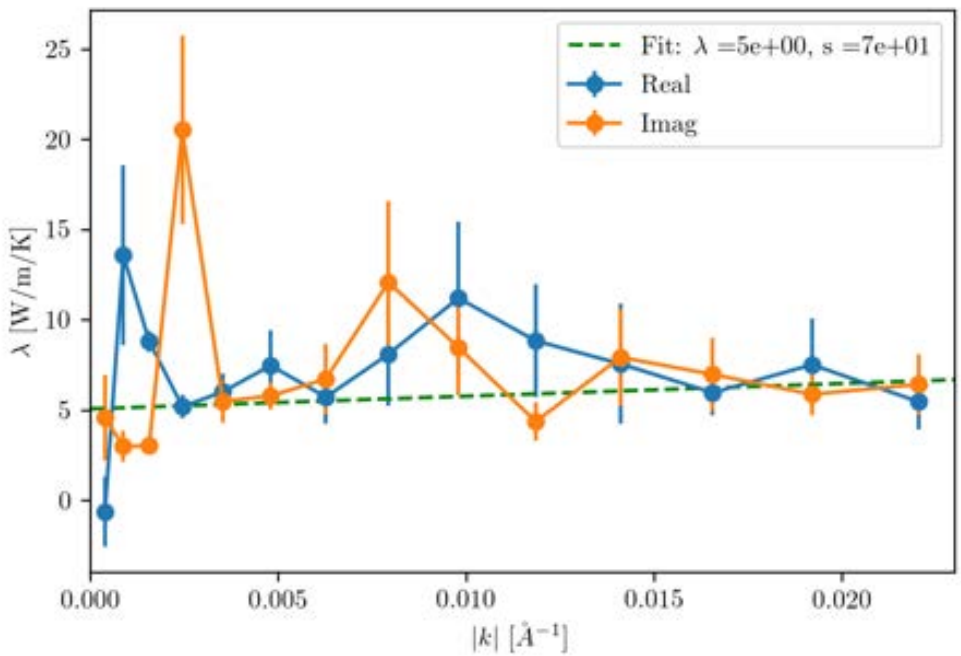
[Cheng & Frenkel PRL 2020]



From MD trajectories generated using a MLP at 2000 K, 33 GPa

Heat conductivity of high-pressure hydrogen fluid

[Cheng & Frenkel PRL 2020]



From MD trajectories generated using a MLP at 2000 K, 33 GPa

Outline

- A ML potential for water.
- Hydrogen under high pressure.
- The extent of locality in MLP.
  - Computing heat conductivity.
  - Extracting ice-like local environments from liquid water.

Training set for bulk water

revPBE0-D3 reference:

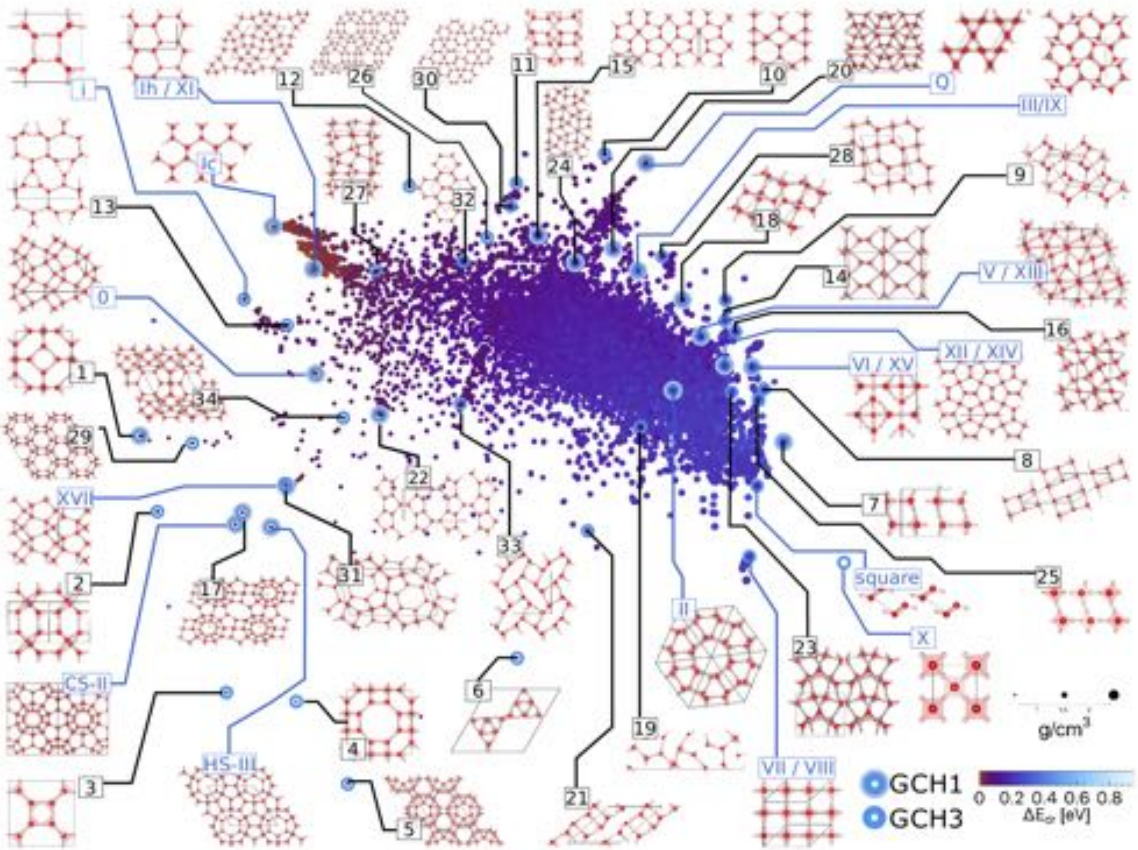
- AIMD and PIMD simulations [Marsalek & Markland JPCL 2017]
- Benchmarks with CCSD(T) and DMC [Brandenburg 2019]

Trained using BP neural network.
[Behler Parrinello PRL 2008; Morawietz, et al. PNAS 2016]

Training set:

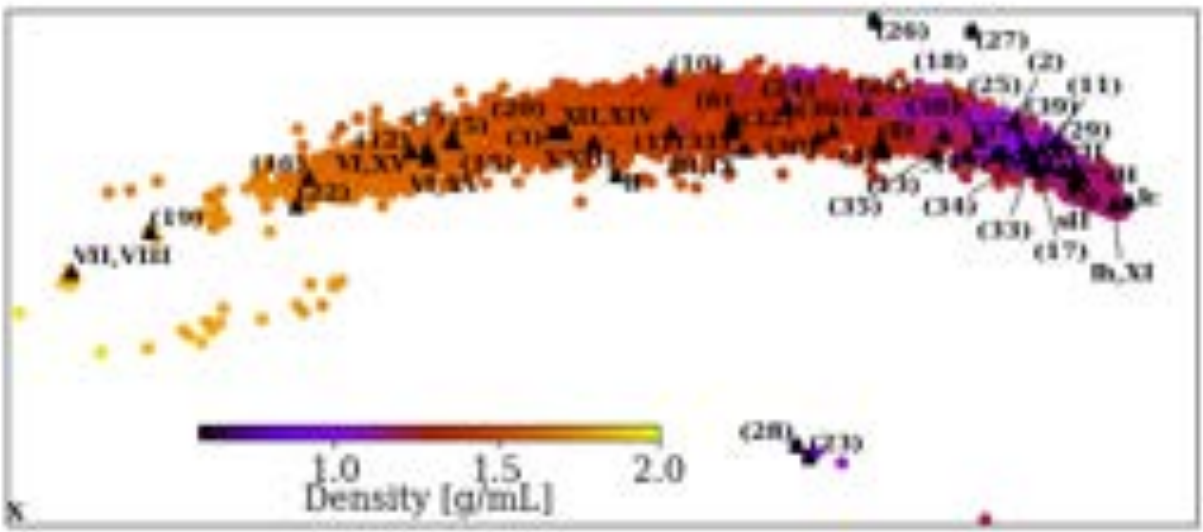
- 1593 bulk **liquid** 64 water molecules (energy + forces)
- 1000 structures from quenches at a wide range of densities.
- 593 structures from PIMD.

54 (experimental+hypothetical) ice phases



[Engel, et al. Nature Comm. 2018]

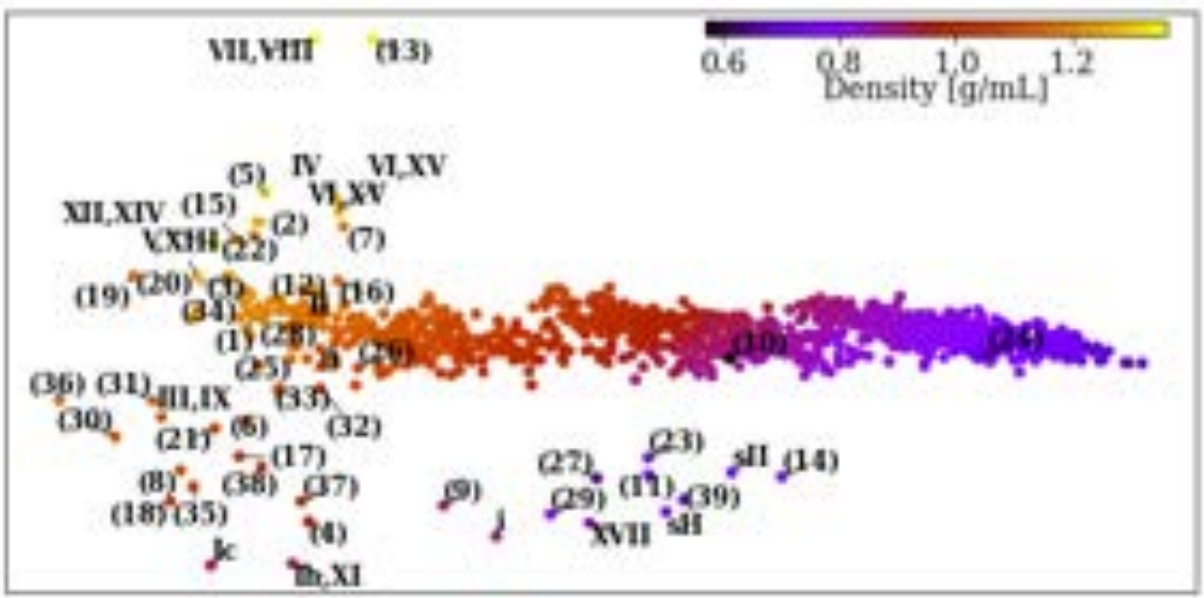
54 (experimental+hypothetical) ice phases



PCA map for the ice phases

PCA map of Ice and water

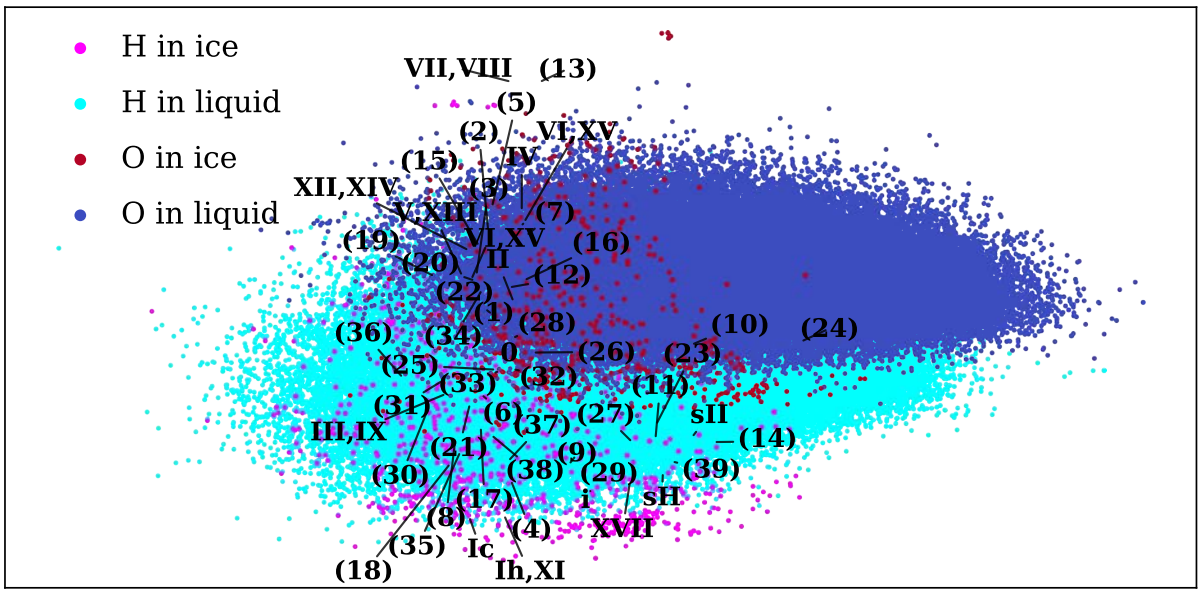
[Monserrat, Brandenburg, Engel & Cheng Nat Commun 2020]



1,000 liquid water and 54 (experimental+hypothetical) ice phases

PCA map of Ice and water

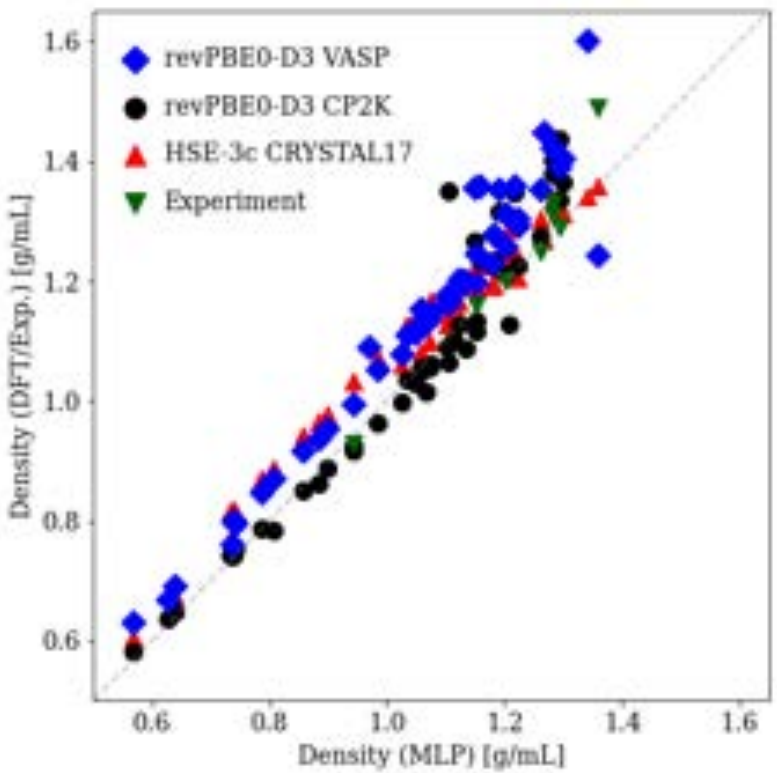
[Monserrat, Brandenburg, Engel & Cheng Nat Commun 2020]



Cutoff radius for atomic environments: 6 Angstrom

The MLP trained only on liquid water can predict properties of ice phases

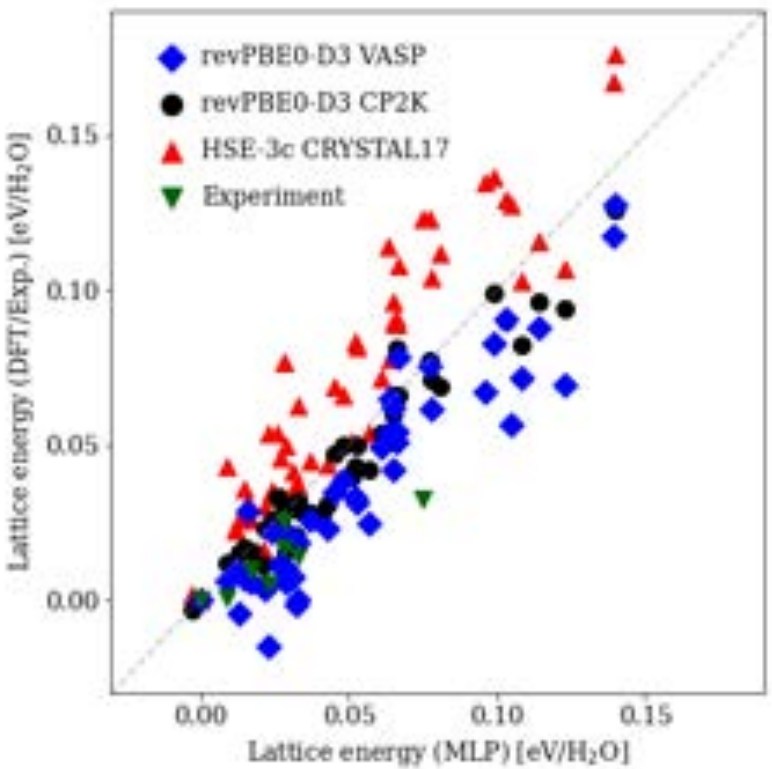
[Monserrat, Brandenburg, Engel & Cheng Nat Commun 2020]



revPBE0-D3 (CP2K), revPBE0-D3 (VASP), HSE-3c (CRYSTAL07)

The MLP trained only on liquid water can predict properties of ice phases

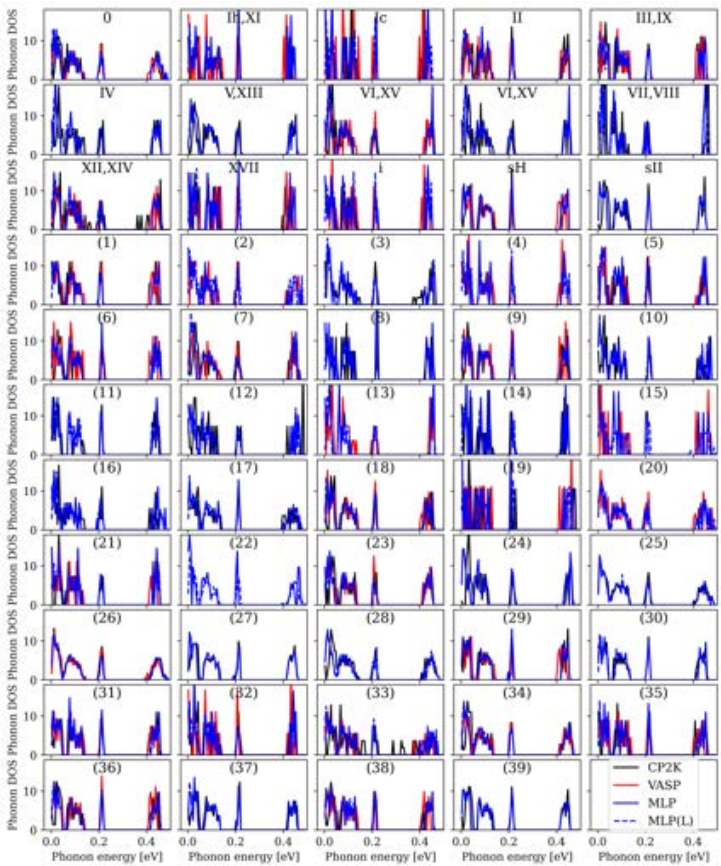
[Monserrat, Brandenburg, Engel & Cheng Nat Commun 2020]



revPBE0-D3 (CP2K), revPBE0-D3 (VASP), HSE-3c (CRYSTAL07)

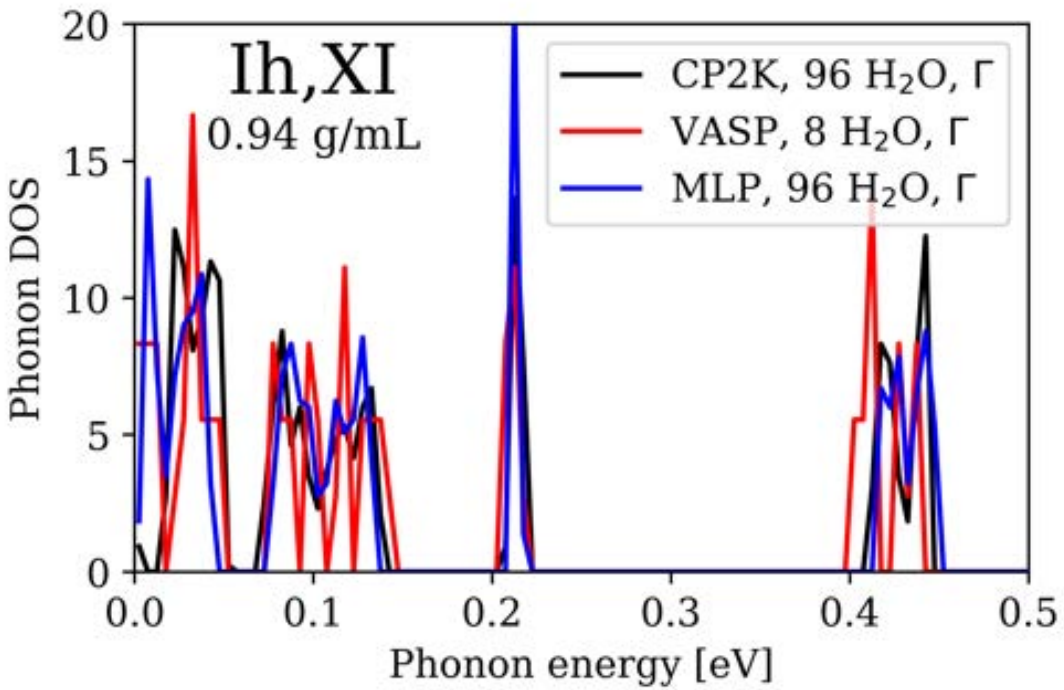
The MLP predicts the phonon density of states

[Monserrat, Brandenburg, Engel & Cheng Nat Commun 2020]



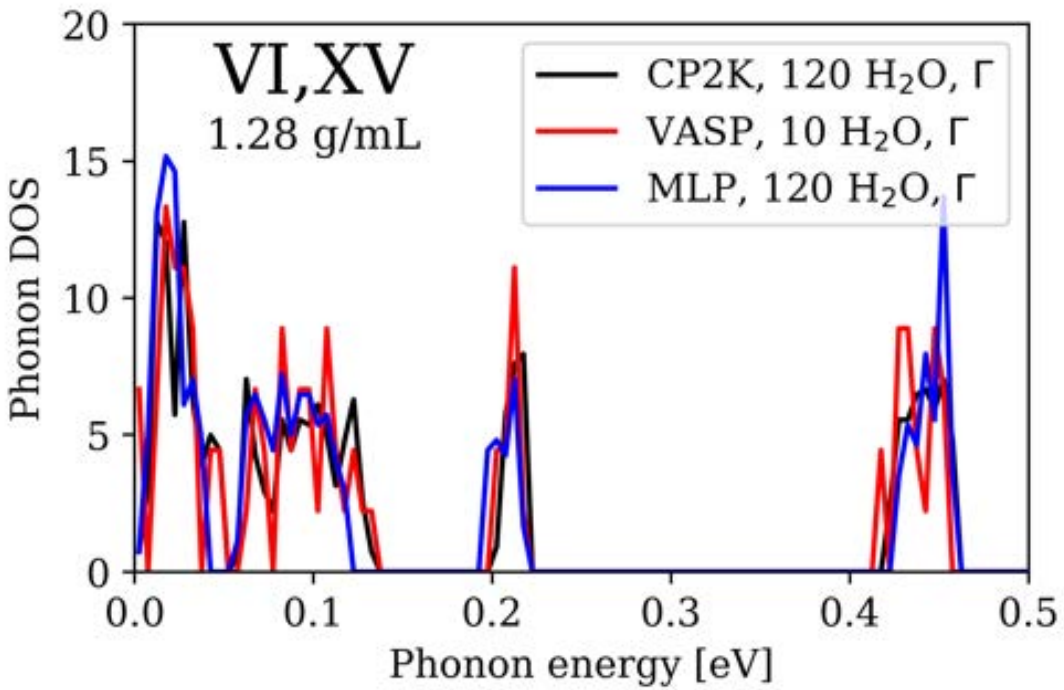
The MLP predicts the phonon density of states

[Monserrat, Brandenburg, Engel & Cheng Nat Commun 2020]



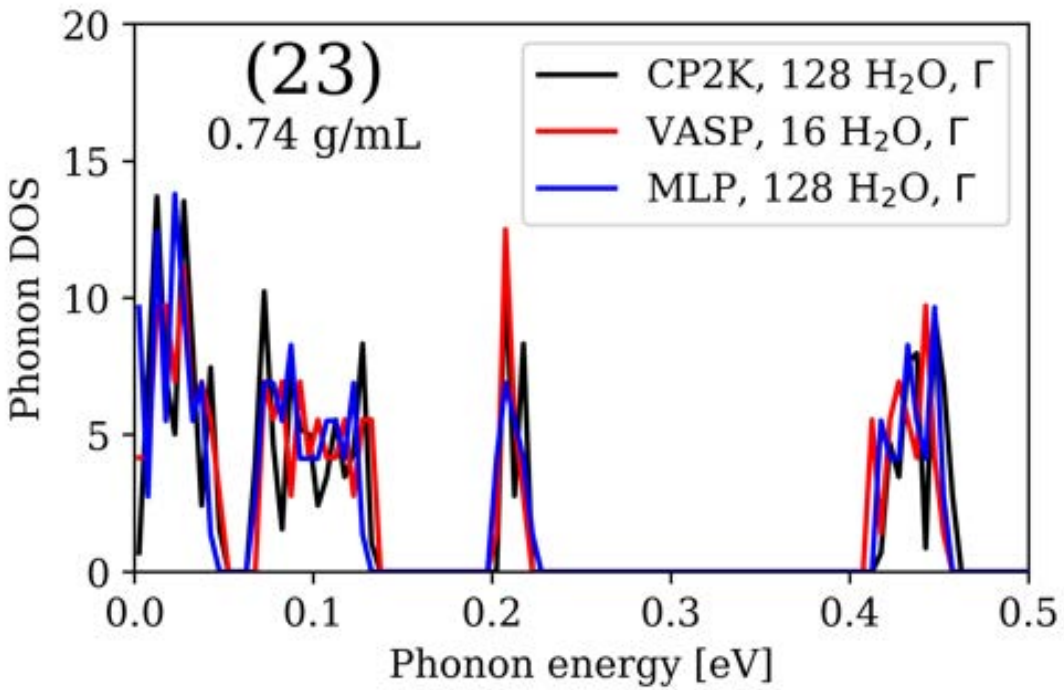
The MLP predicts the phonon density of states

[Monserrat, Brandenburg, Engel & Cheng Nat Commun 2020]



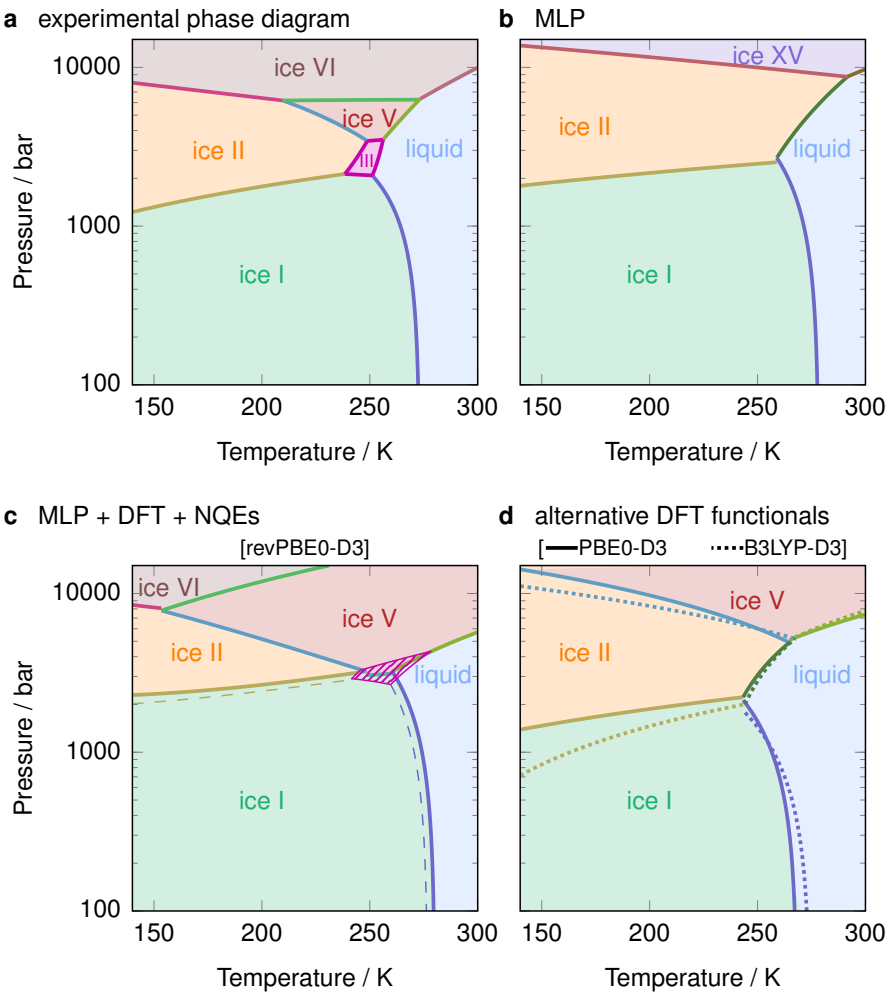
The MLP predicts the phonon density of states

[Monserrat, Brandenburg, Engel & Cheng Nat Commun 2020]



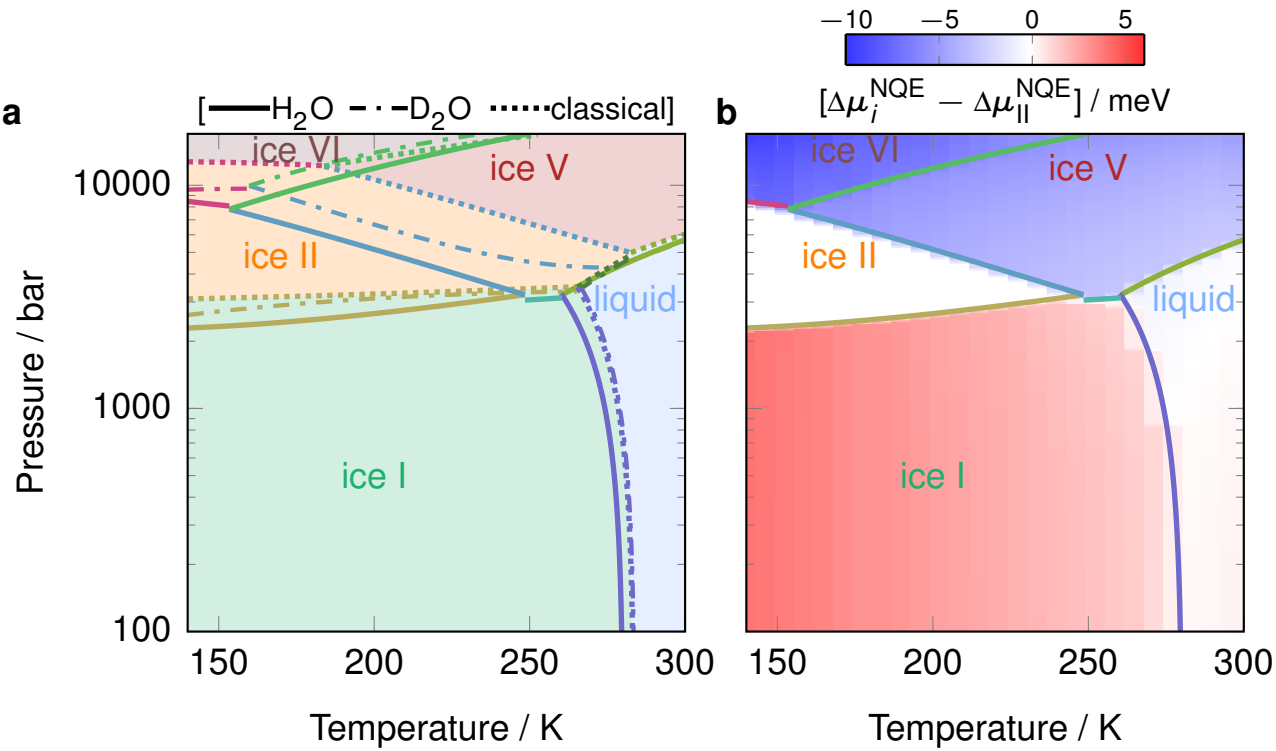
Hybrid DFT phase diagrams of water

[Reinhardt & Cheng Nat Commun (in press)]



Hybrid DFT phase diagrams of water

[Reinhardt & Cheng Nat Commun (in press)]



revPBE0-D3, averaged over different proton disordered states, including NQEs

Take home messages

- Machine learning potentials (MLPs) may transform computational chemistry. (We can compute ab initio phase diagrams now!)
- We don't fully understand MLPs, and there is a lot of room for improvement.
- Perhaps we need to develop simulation methods that are tailored for MLPs.

https://github.com/BingqingCheng/ASAP

Contributors: Ryan-Rhys Griffiths, Tamas Stenczel, Bonan Zhu, Felix Faber, Noam Bernstein

- Low-dimensional embedding, regression
- Sparsification
- Clustering, kernel density estimation