

Abstract

In recent years a lot of progress has been made in the study of water and its different phases. Empirical methods are progressing to the point of being better than first principle methods. These methods are however limited in their scope, and completely incapable of being generalized or of treating reactive environments. At the ab-initio level, to reach chemical accuracy, one often has to rely on hybrid functionals, thereby limiting the size of systems for which simulations are feasible.

We propose a new method that uses **artificial neural networks** (ANN) to improve density functional theory (DFT) calculations of water. Given our model's flexibility, we are able to correct the self-interaction error made by GGA functionals and accurately reproduce total energies for a variety of systems. As our ANN only relies on the **local charge density** around a single molecule the model scales to arbitrary system sizes, in spite of being trained only on small systems of up to three molecules. Furthermore, our model can in principle correctly describe dissociation processes and heterogeneous systems, thereby extending its application beyond the scope of empirical force fields.

Introduction

- Kohn-Sham density functional theory fails to effectively describe water systems in many situations.

• MB-pol:

- State of the art force-field for water
- Fitted to high level CCSD(T) calculations

- Makes use of the many-body expansion of the interaction energy

$$V_N(x_1, \dots, x_N) = \sum_a V^{1B}(x_a) + \sum_{a>b} V^{2B}(x_a, x_b) + \sum_{a>b>c} V^{3B} + \dots + V^{NB}(x_1, \dots, x_N)$$

- Very accurate but not flexible at all because empirical

- Our approach:** Use Machine Learning to correct DFT calculations and reach MB-pol accuracy with the flexibility of an ab-initio calculation

Informed Machine Learning
for
Maximal Extrapolation

Use knowledge about the **physical mechanisms** involved to pre-process and select relevant data. Model has to generalize to unseen data and scale to **arbitrary system sizes**.

Method

Features & Targets

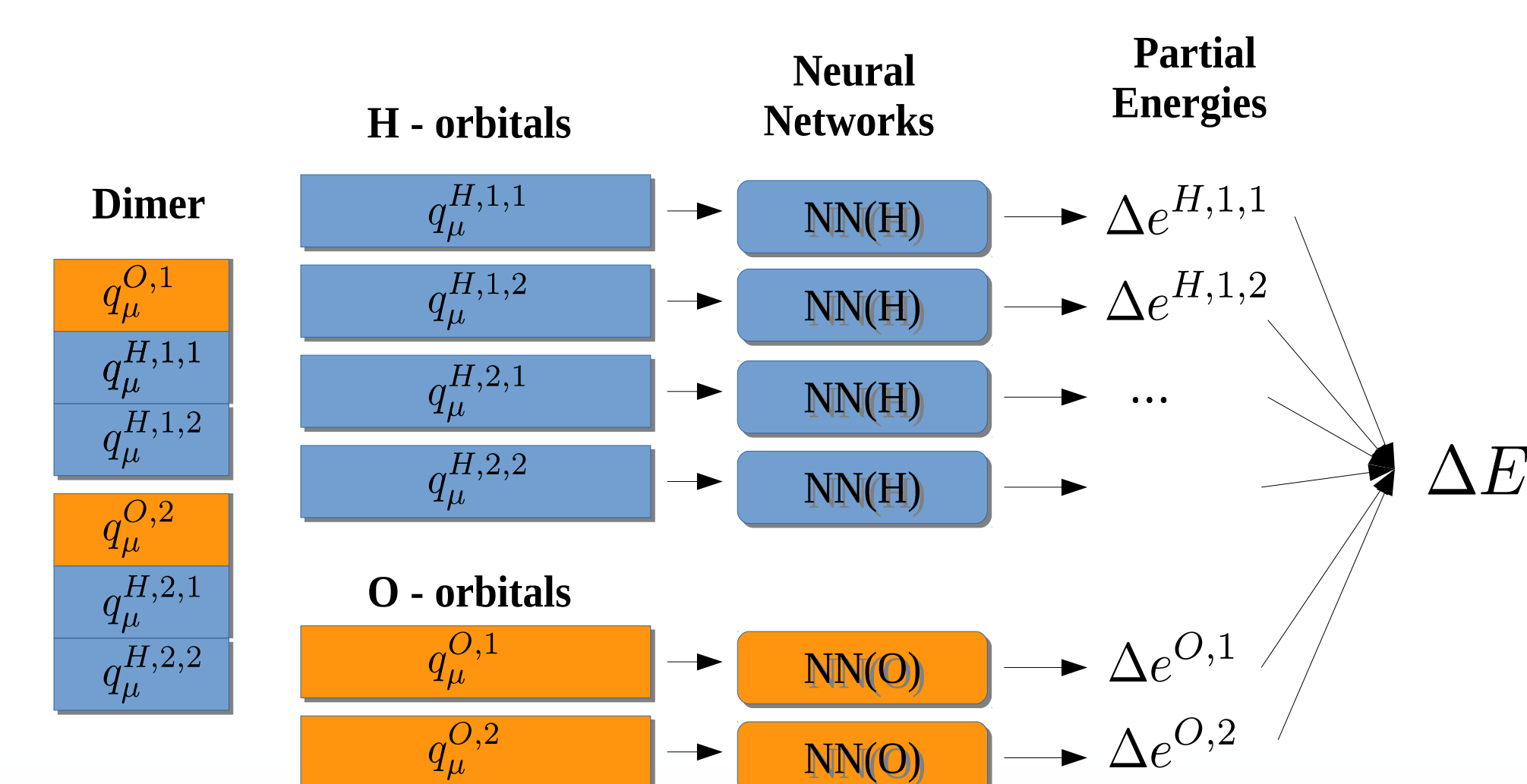
Given the density matrix $\rho_{\mu\nu}$ obtained from a converged LCAO-DFT calculation the **features** that are fed to the NN are given by

$$q_\mu = \sum_\nu \rho_{\mu\nu} S_{\mu\nu}, \quad \text{with overlap matrix } S_{\mu\nu}$$

The differences between our reference energies obtained with MB-pol and the DFT energies provide the **targets** that our model is fitted to

$$\Delta E = E_{\text{MB-pol}} - E_{\text{DFT}} \approx \Delta E_{xc}$$

Neural Network Architecture

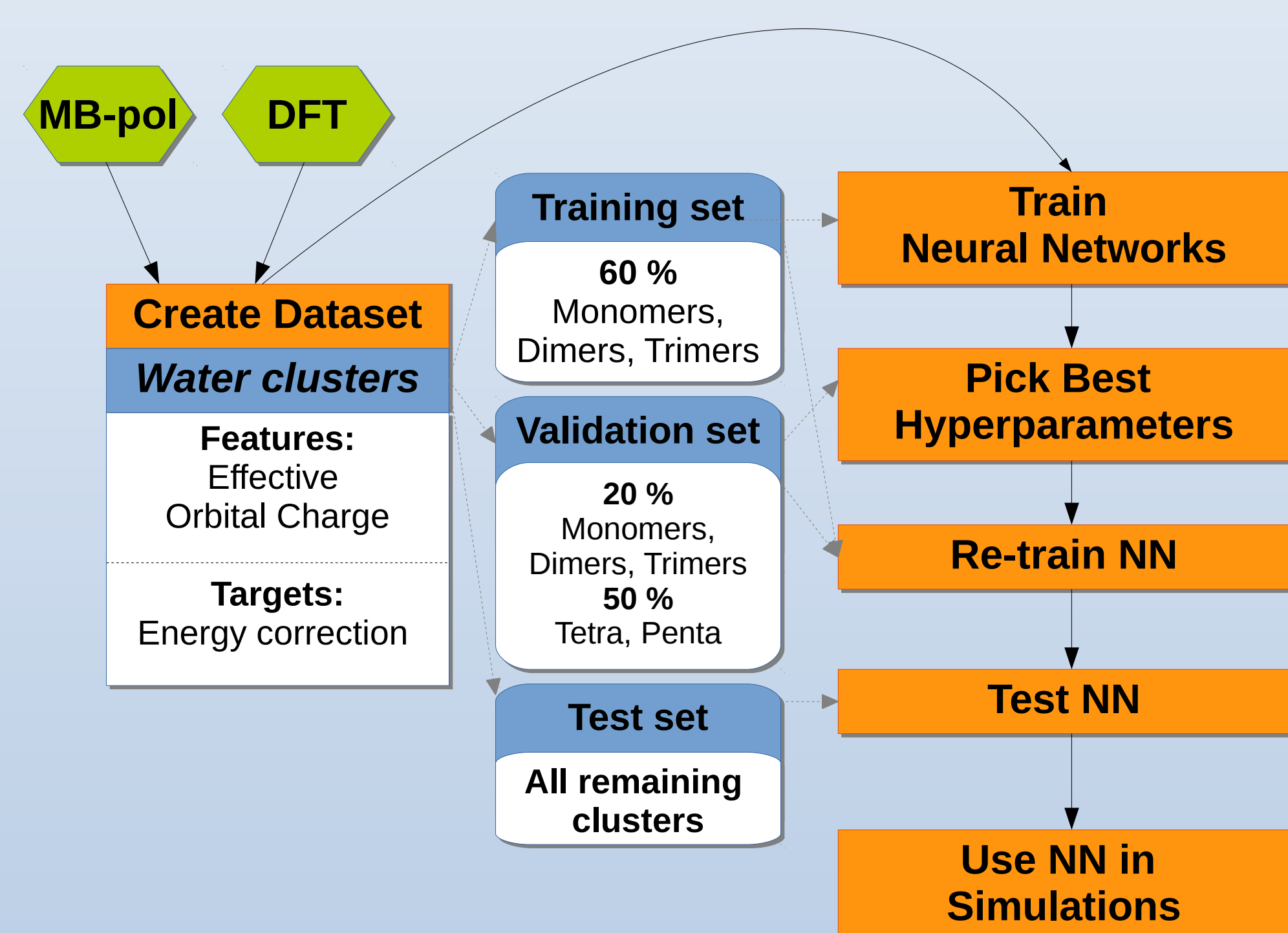


The effective orbital charges q_μ are split up into their atomic contributions and fed into separate NNs. Networks describing the same atomic species are identical. This ensures that the model respects permutation symmetry of Hydrogen atoms and that it scales to arbitrary system sizes.

Training set

- 300 Monomers - from parameter grid
- 2000 Dimers } from molecular dynamics simulations
- 1500 Trimers }

Workflow



Results

Accuracy on test set

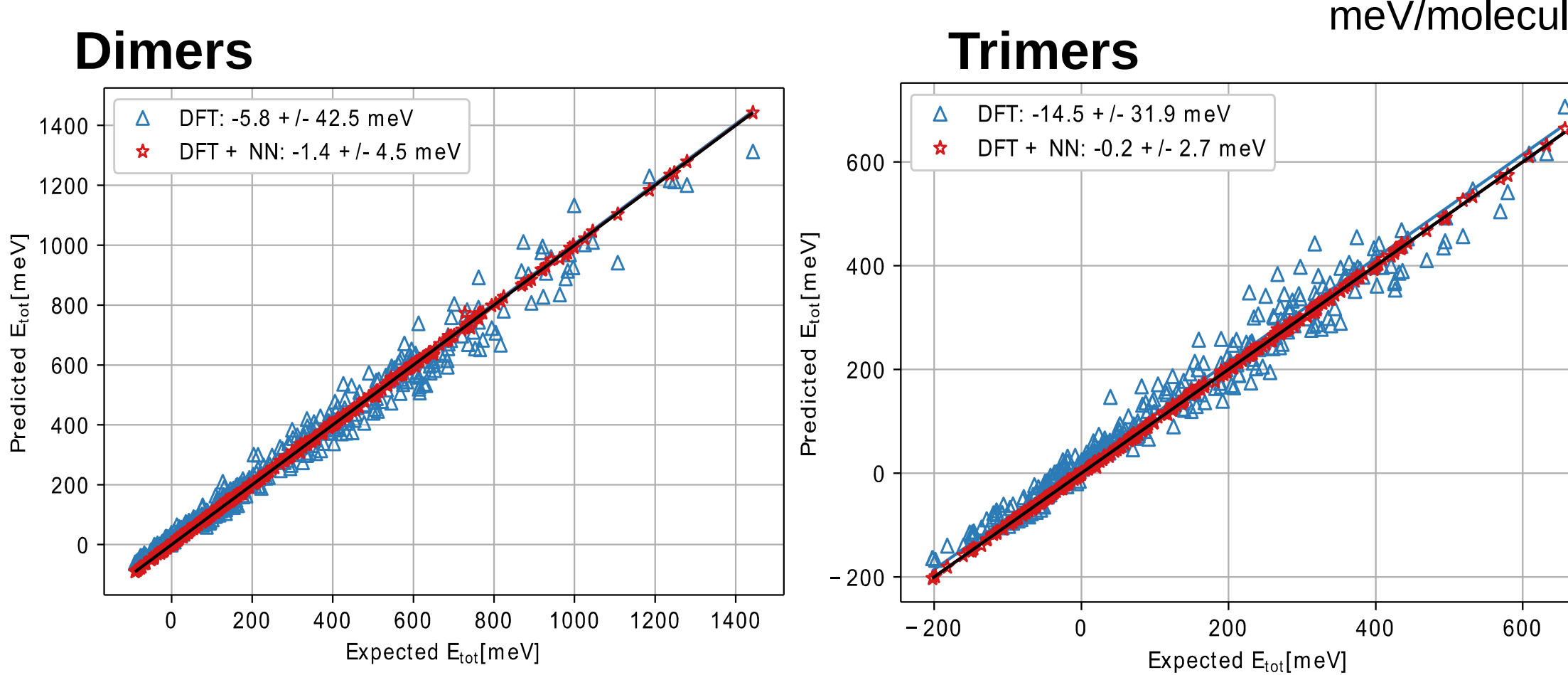
- Improvement of both total energies and relative energies of isomers.

- Being only trained on system sizes of up to three water molecules, the model extrapolates well to larger systems

$$\delta E = E_{\text{ref}} - E_{\text{pred}}$$

No. of Mol.	$\langle \delta E \rangle$		$\sqrt{\langle (\delta E)^2 \rangle} - \langle \delta E \rangle$	
	DFT	DFT + NN	DFT	DFT + NN
1	-4.2	-1.2	64.3	2.2
2	-5.8	-1.4	42.5	4.5
3	-14.5	-0.2	31.9	2.7
4	31.4	-2.4	12.6	4.8
5	31.5	0.1	12.9	3.4
8	30.6	2.2	9.8	4.4
16	33.6	8.2	8.3	2.1
24	35.7	12.4	7.7	2.1
32	39.2	13.7	5.4	1.7

Energies in meV/molecule

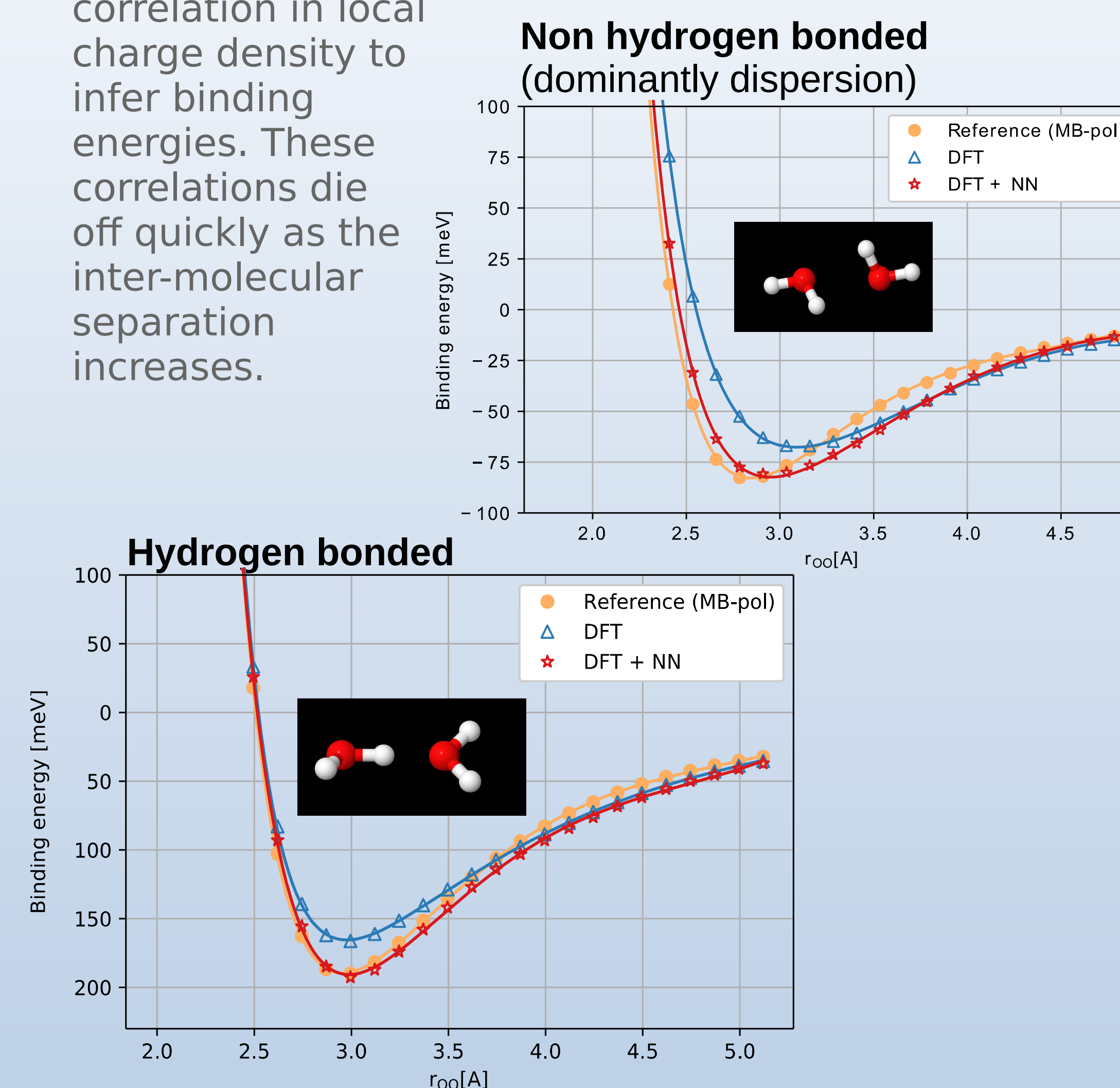


Binding energies

- In spite of being trained on local information only (the effective orbital charge) the network works well even beyond the one-body energy.

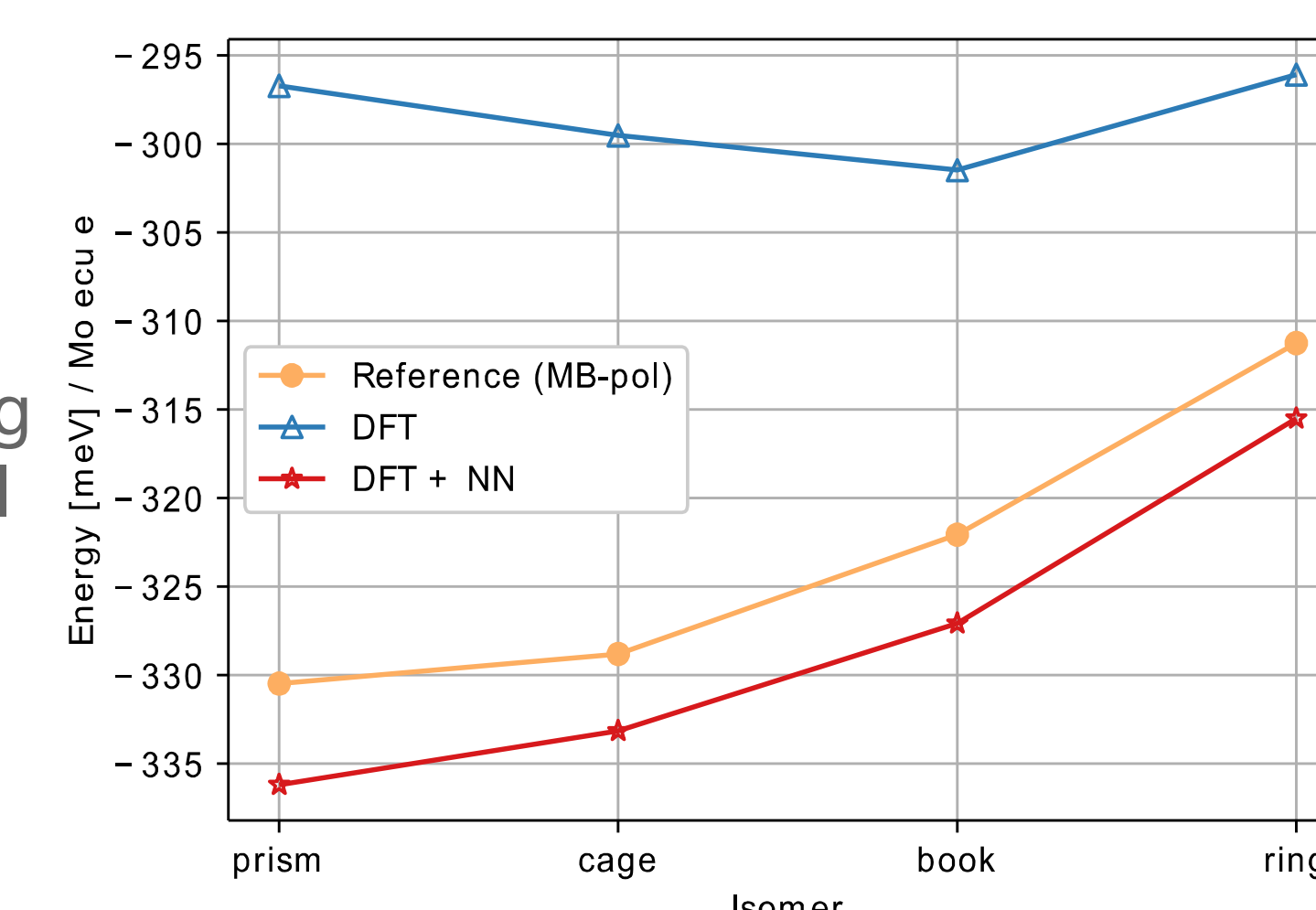
- Good agreement for short distances, under-corrects for large molecular separation

Explanation: Network uses correlation in local charge density to infer binding energies. These correlations die off quickly as the inter-molecular separation increases.



Hexamers

- Smallest water clusters that form a 3-dimensional structure
- Good for testing how well model performs for larger systems
- Important contributions from three-body energies



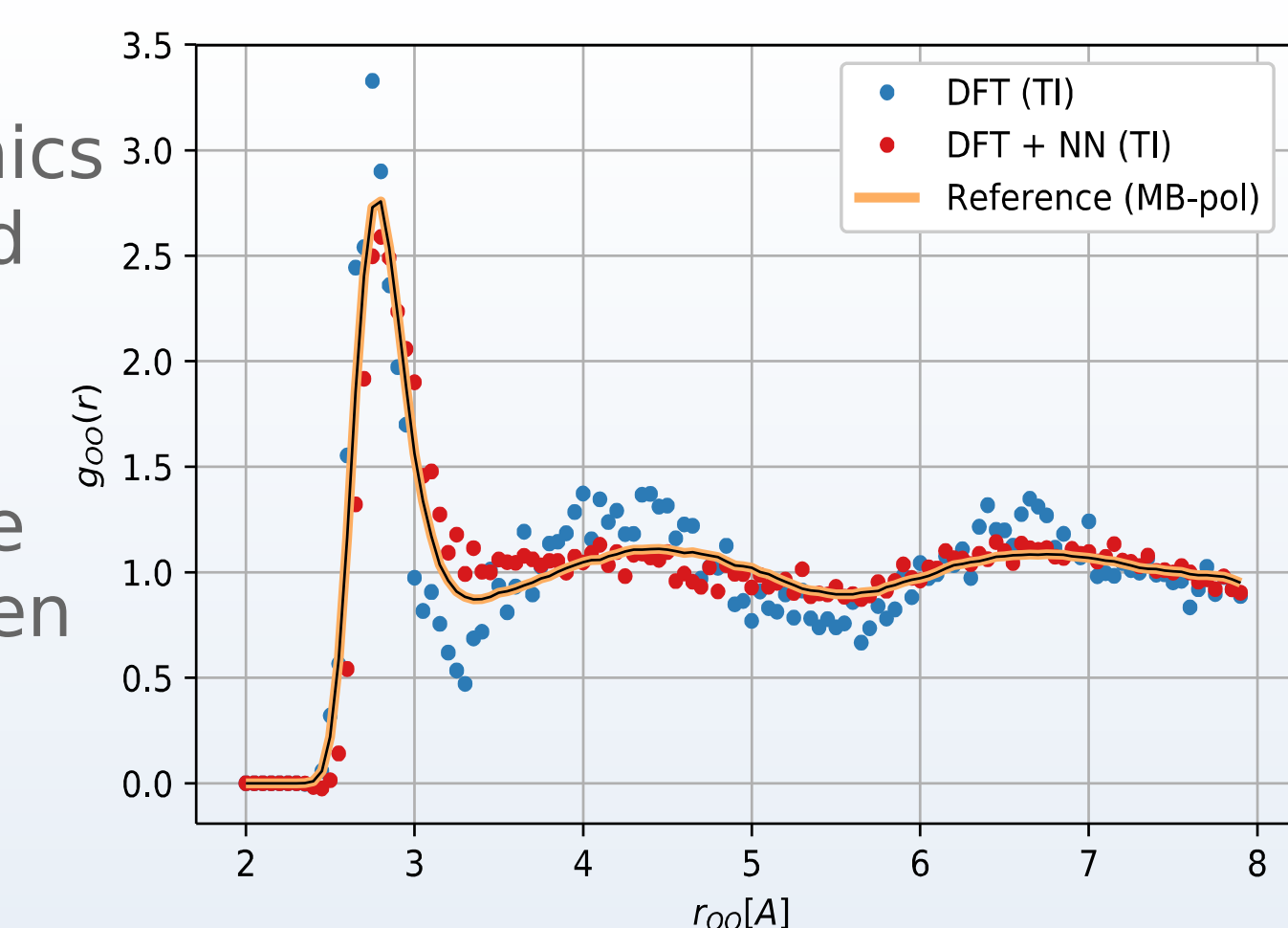
Pair correlation functions

- The response of a canonical ensemble average to a change in potential can be approximated by a Taylor expansion. This method is called **Thermodynamic Integration** (TI)

$$V_\lambda = (1 - \lambda)V_0 + \lambda V_1, \quad \text{with } V_0 = \text{DFT pot.}, \quad V_1 = \text{corrected pot.}$$

$$\Delta \langle g_{OO}(r) \rangle = -\beta [\langle g_{OO}(r) \Delta V \rangle_{\lambda=0} - \langle g_{OO}(r) \rangle \langle \Delta V \rangle_{\lambda=0}].$$

- Re-sample molecular dynamics trajectory created with MB-pol
- Neural network seems to improve the oxygen-oxygen pair-correlation function.



Conclusion

- We successfully trained a neural network that can correct the error GGA makes for water systems
- Once trained, the network can be used at negligible computational cost
- It respects physical symmetries and can potentially be generalized to treat reactive environments.
- In spite of being empirical it **exceeds force-fields in flexibility** as it only relies on the local charge density

Limitations

- Model can only be used for basis set and xc-functional that it was trained on
- So far only applicable to LCAO codes
- Charge densities cannot be corrected