

Advances in machine learned potentials for molecular dynamics simulation

Kipton Barros
Los Alamos National Lab.

Physics Next, Machine Learning,
Oct 9, 2018, Riverhead, NY

Collaborators

(LANL)

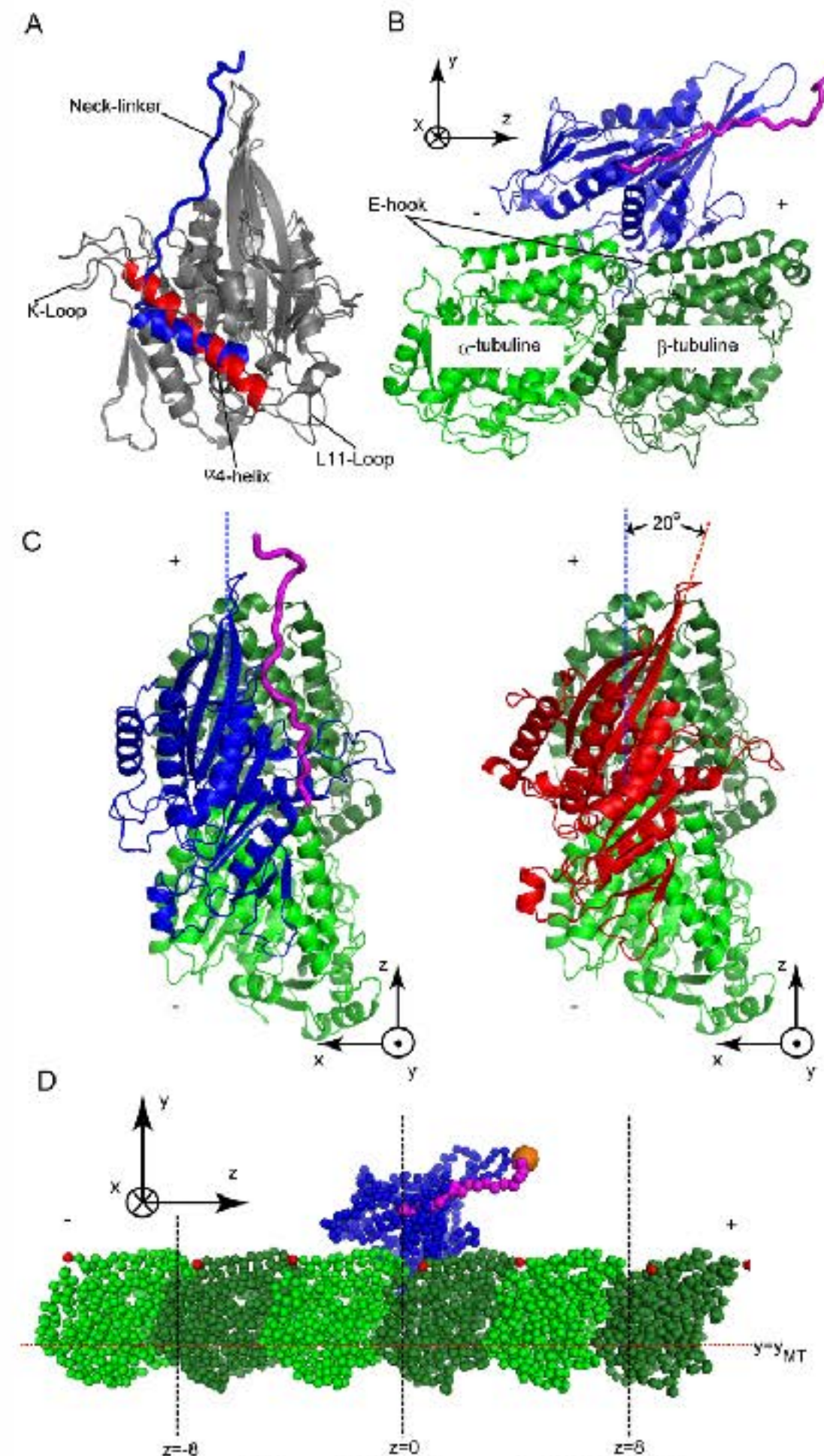
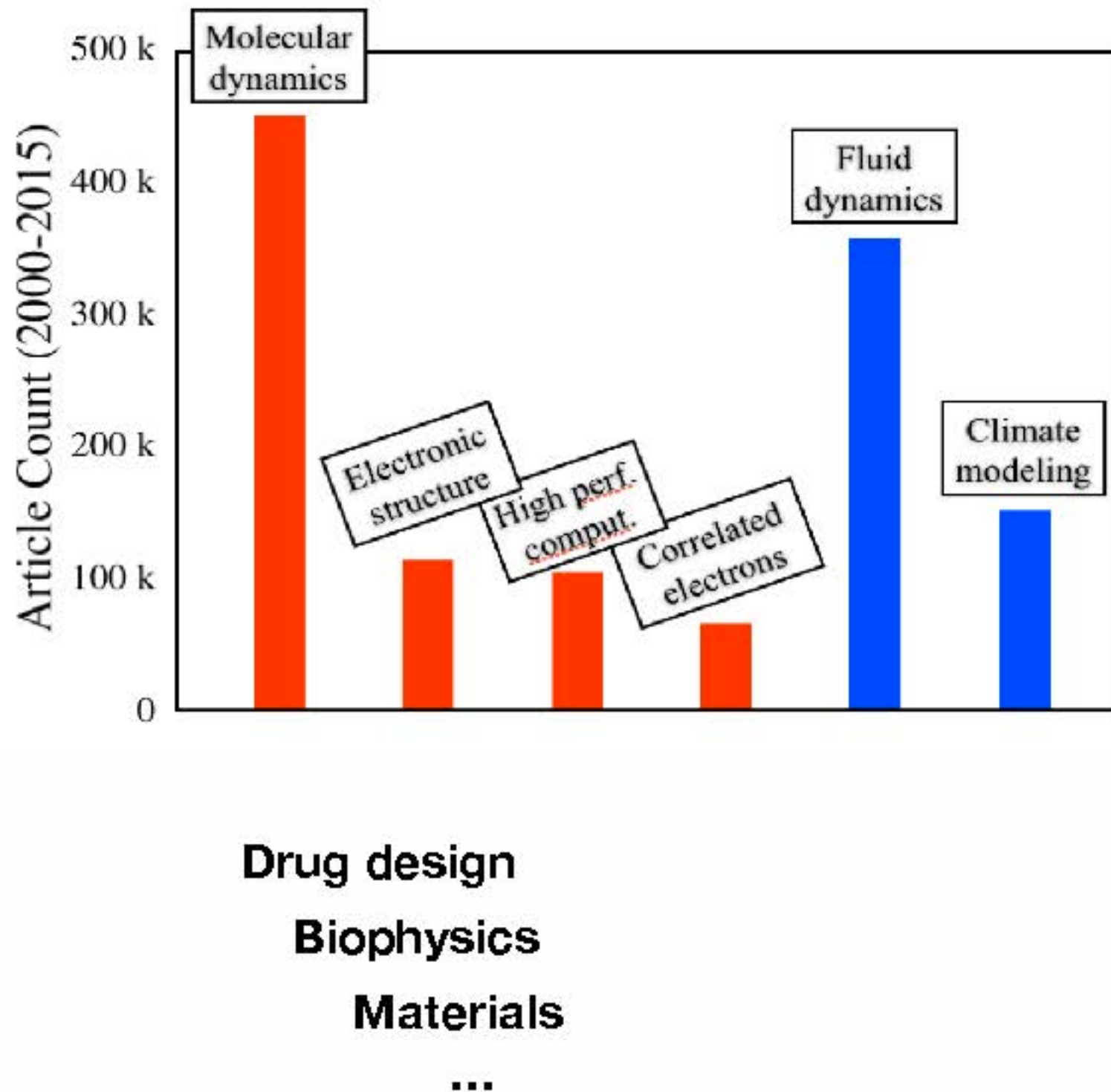
Nicholas Lubbers
Justin Smith
Andrew Sifain
Ben Nebgen
Sergei Tretiak

(UNC) Oles Isayev

(UF) Adrian Roitberg



Molecular dynamics articles



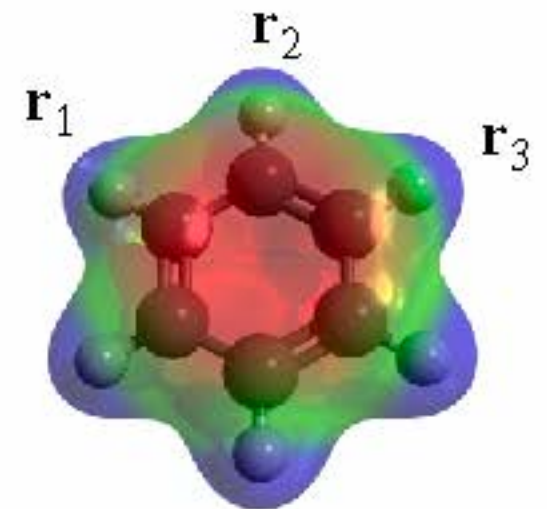
Born Oppenheimer MD

→ separation of scales

Calculate energy/forces for fixed nuclei positions

$$\mathbf{r}_1 \dots \mathbf{r}_N \mapsto E[\mathbf{r}] \quad \text{Schrödinger Eq.} \approx$$

{Hartree Fock, DFT,
Coupled Cluster, ...}



Integrate nuclei dynamics

$$\frac{d^2 \mathbf{r}_i}{dt^2} = - \nabla_i E[\mathbf{r}] / m_i$$

Not treated:

Quantum mechanics of nuclei (e.g. hydrogen)

Non-equilibrium dynamics of electrons

- von Neumann eq. for density matrix
- non-adiabatic excited state MD

Classical potential / force field

Pros: Fast

Cons:
Not very transferable,
Non-reactive
Laborious parameterization

Schrödinger Equation

Pros: Accurate,
transferable

Cons: Computationally
demanding

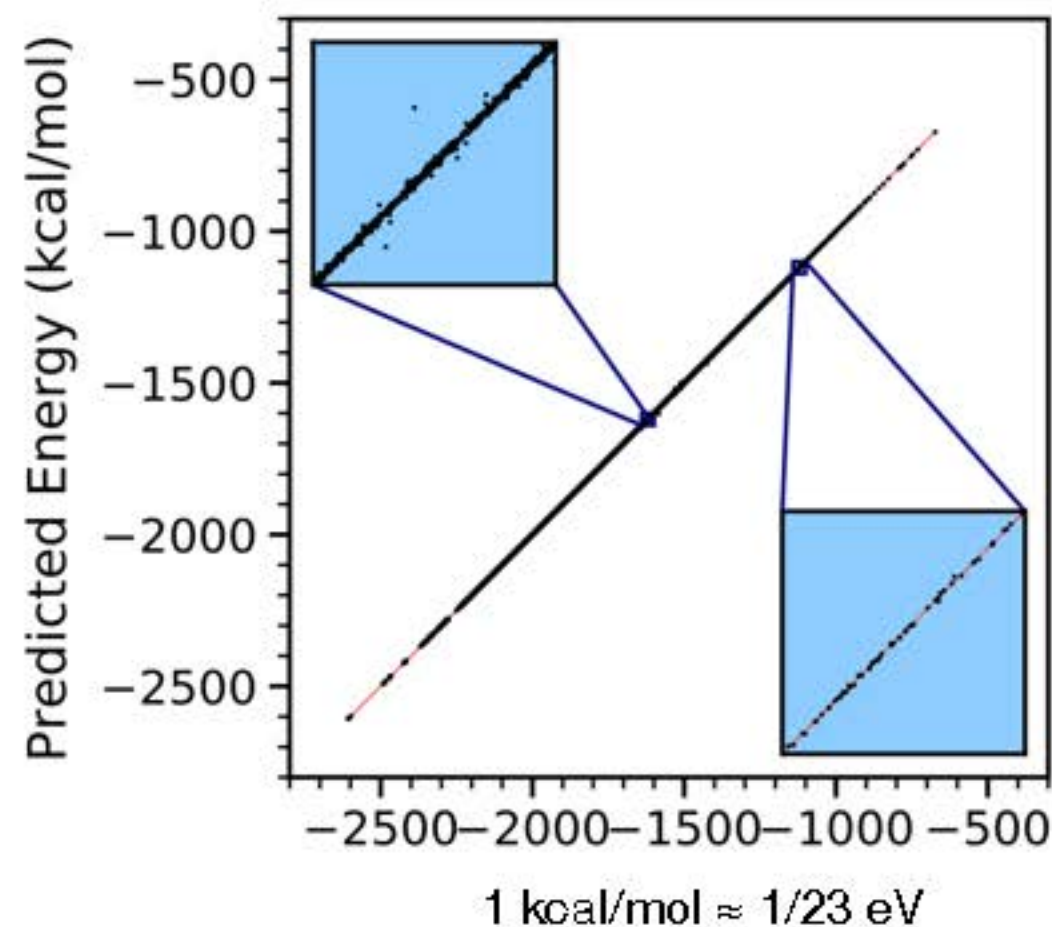
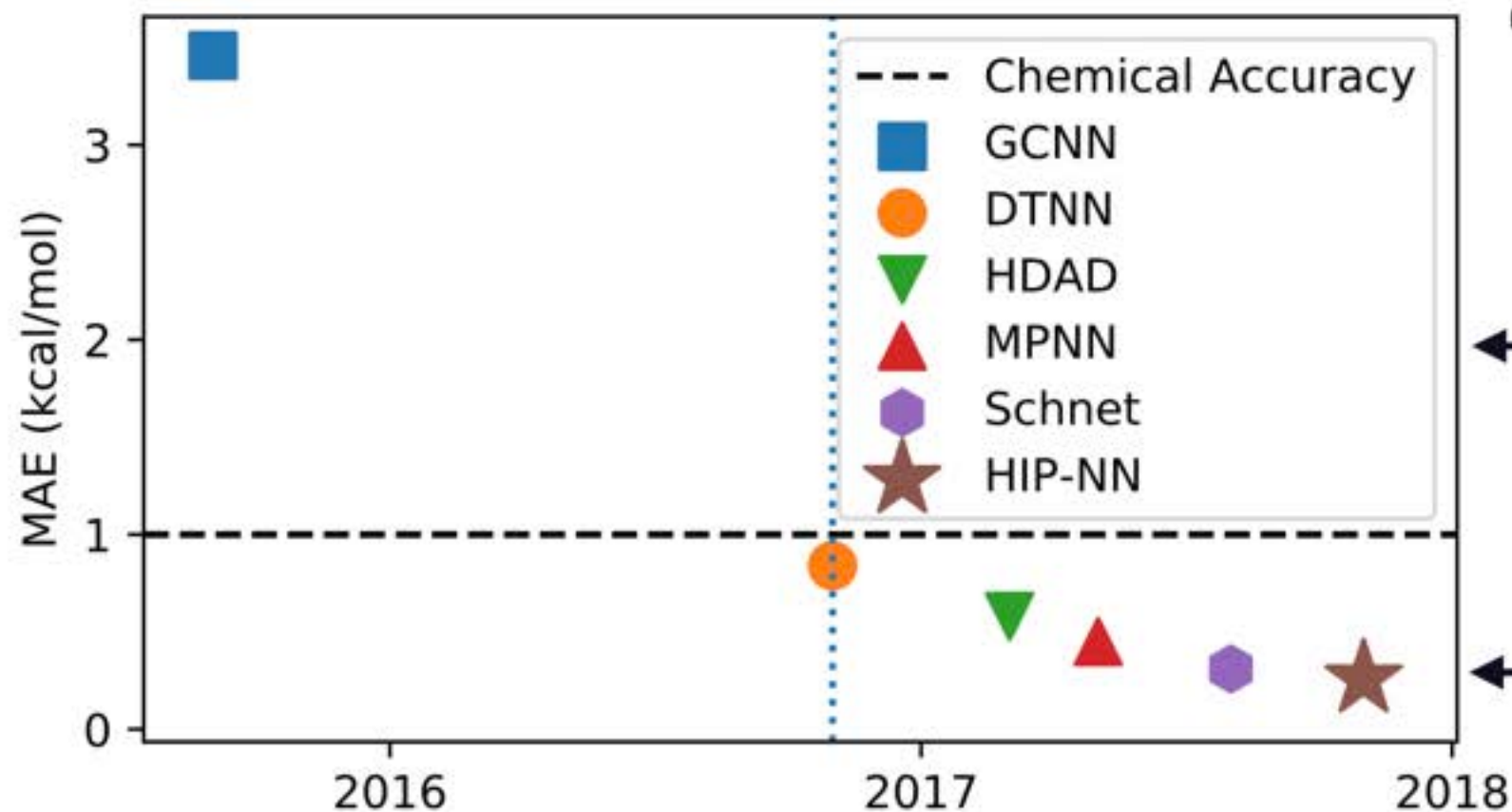
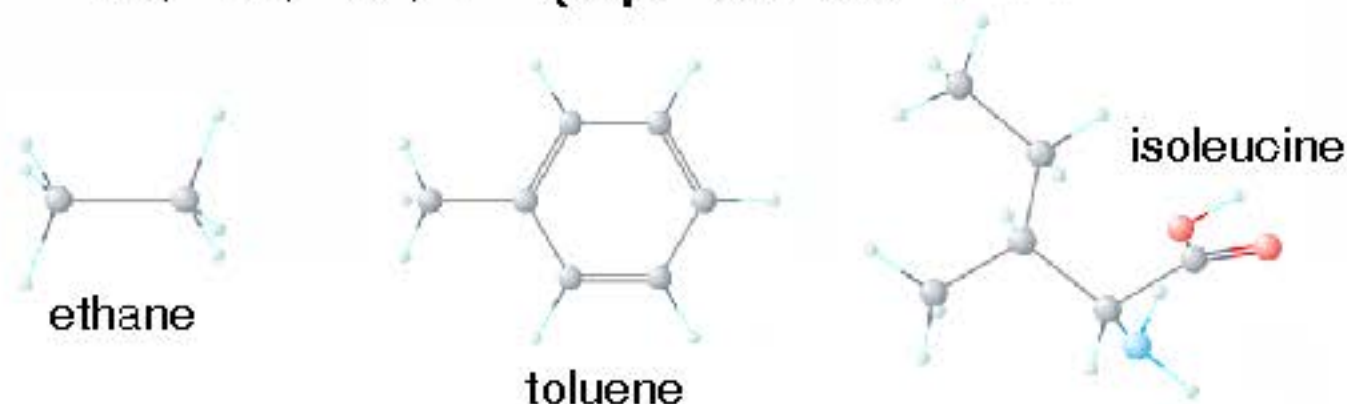
Machine learning to
emulate Schrödinger eq.

The diagram illustrates the relationship between three computational approaches. At the top left, 'Classical potential / force field' is listed with a pro of being 'Fast' and cons of being 'Not very transferable, Non-reactive' and having 'Laborious parameterization'. At the top right, the 'Schrödinger Equation' is listed with a pro of being 'Accurate, transferable' and a con of being 'Computationally demanding'. At the bottom center, 'Machine learning to emulate Schrödinger eq.' is shown in red. Two red arrows point from this central text to the 'Pros' sections of both the Classical and Schrödinger methods, indicating that machine learning aims to combine the speed of the classical method with the accuracy and transferability of the Schrödinger equation.

QM9 - 130k organic molecules

Relaxed geometries

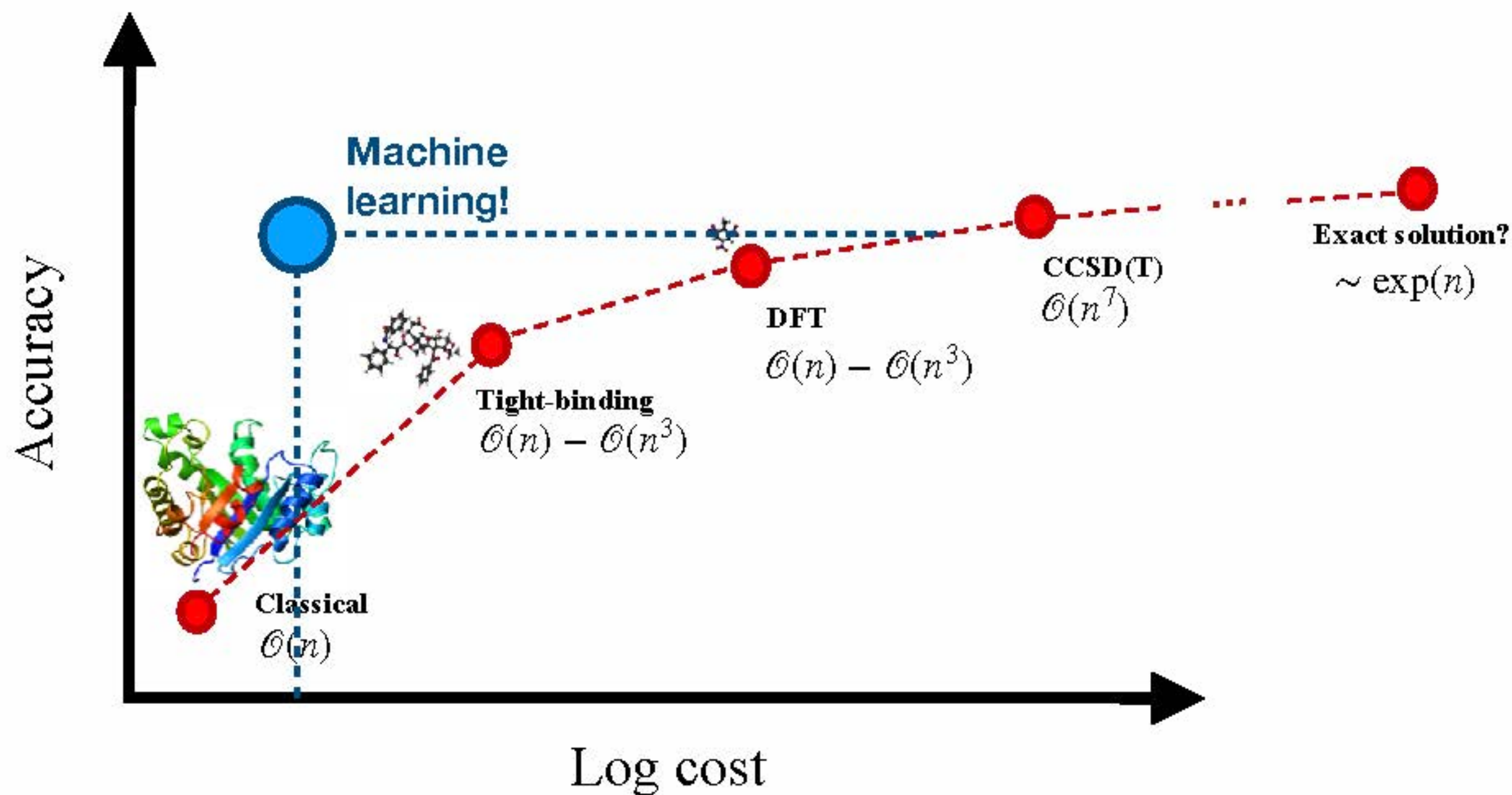
C, N, O, F (up to 9) + H



Typical
DFT Accuracy

Typical
CCSD Accuracy

Levels of quantum chemistry



Physics informed machine learning

Symmetries:

Translation

Time reversal

Rotation

Permutation

Physical principles:

Locality + Coulomb interactions

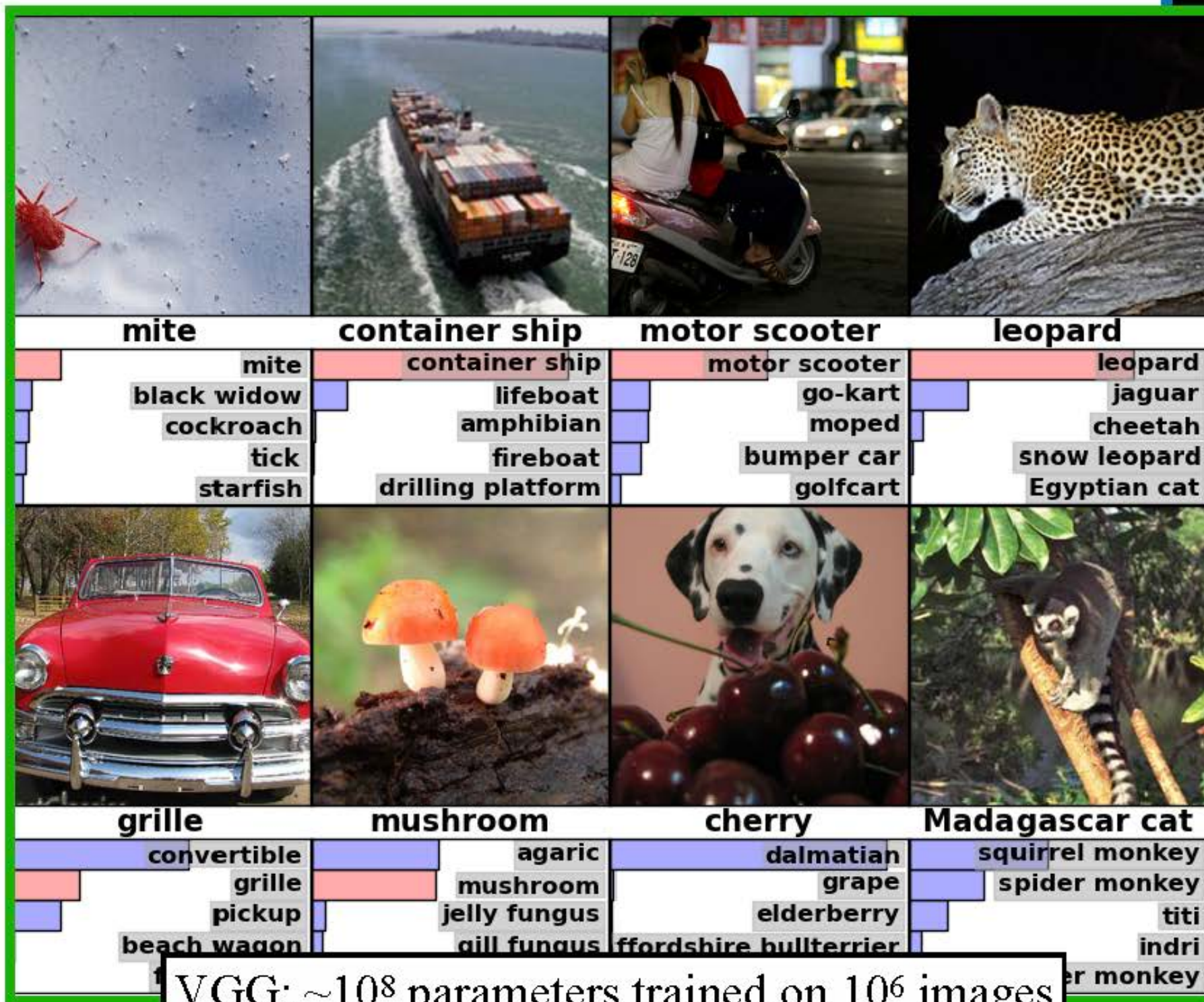
Hierarchicality (e.g. many-body expansion)

Do not constrain to partially correct physics?

Transplant techniques

Deep learning philosophy:

- *don't* hand design features
- use *large* training dataset



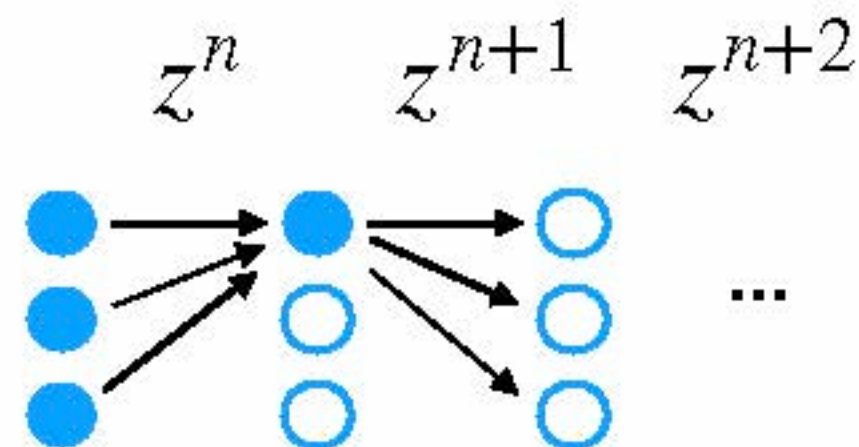
VGG: $\sim 10^8$ parameters trained on 10^6 images



AlphaGo / AlphaZero

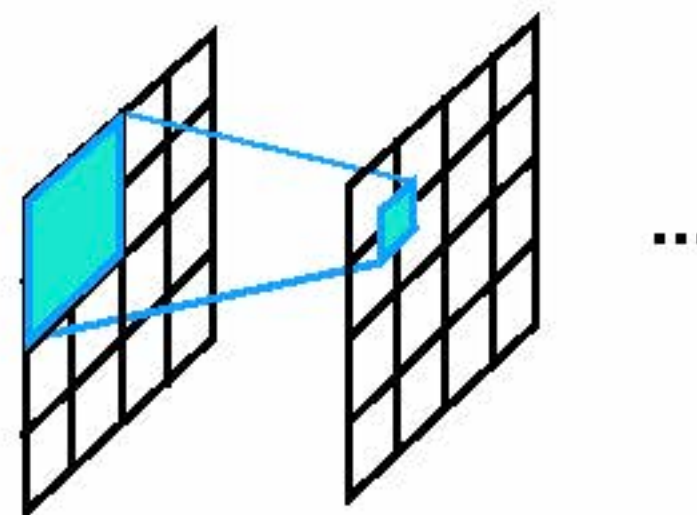
Neural Net

$$z_a^{n+1} = f_{\text{nonlin}}(\sum_b W_{ab} z_b^n + b_a)$$



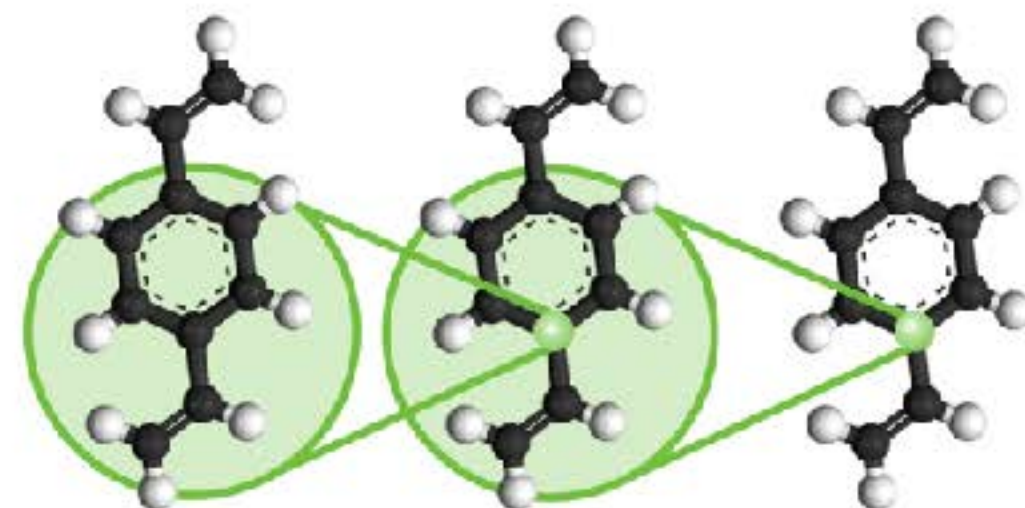
Conv Net

$$z^{n+1}(x) = f_{\text{nonlin}}[(W * z^n)(x) + b(x)]$$



Molecular generalization

$$z^{n+1}(x) = f_{\text{nonlin}}[\dots z^n \dots]$$

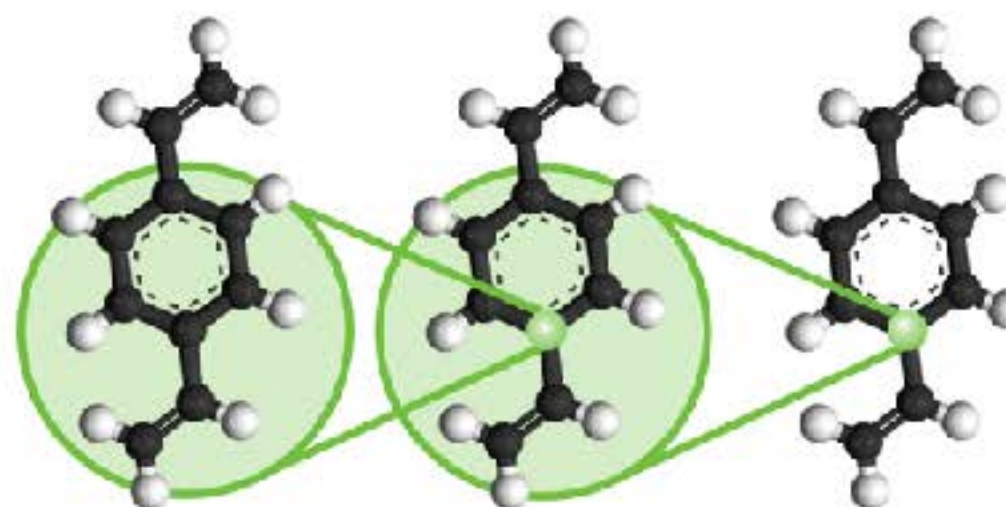


See also: DTNN, MPNN, SchNet, ...

<https://ai.googleblog.com/2017/04/predicting-properties-of-molecules-with.html>

HIP-NN

Layer $\ell = 1$ $\ell = 2$ $\ell = 3$ $\rightarrow$



Local messages

Pairwise distances

(Translation/rotation invariance)

$$\tilde{z}_{i,a}^{\ell+1} =_{\text{inter.}} f \left(\sum_{j,b} \underbrace{v_{ab}^{\ell}(r_{ij})}_{\text{Pairwise distances}} z_{j,b}^{\ell} + \sum_b \underbrace{W_{ab}^{\ell}}_{\text{Learned parameters}} z_{i,b}^{\ell} + \underbrace{B_a^{\ell}}_{\text{Learned parameters}} \right)$$

Descriptors of atom i

Learned parameters

Symmetries of ML potential

Translation,
Rotation:

Atomic positions represented
entirely with ***pairwise distances***

Time reversal:

Forces are generated as **exact
derivative** $f = -\nabla \hat{E}$.

Physical principles

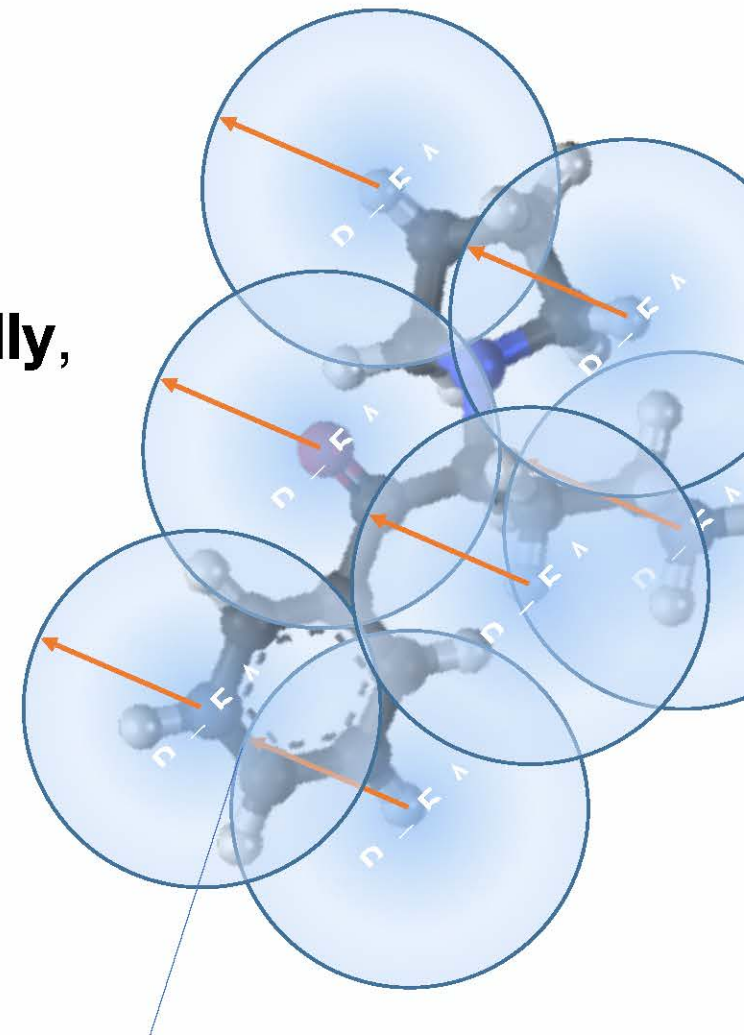
Spatial locality:

Energy decomposed **locally**,

$$\hat{E} = \sum_i \hat{E}_i$$

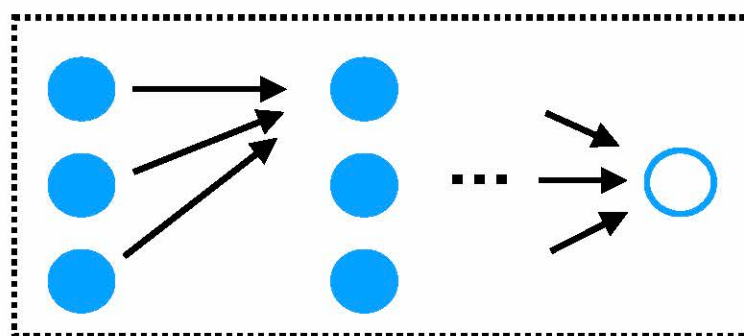
Hierarchicality:

Energy decomposed
hierarchically $\hat{E}_i = \sum_{\ell} \hat{E}_i^{\ell}$.

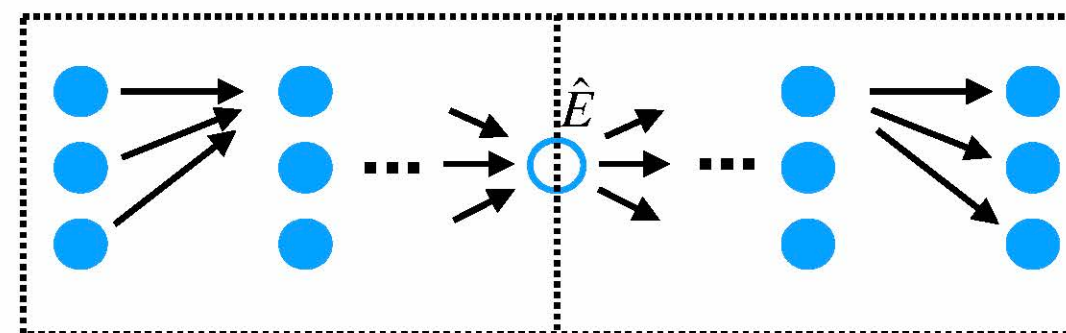


Training to energies *and* forces

Energy $\hat{E}[\mathbf{r}, w]$

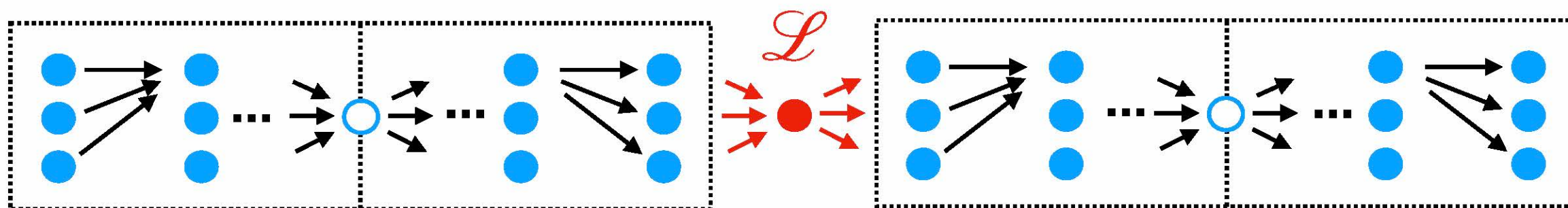


Backprop 1, forces $\nabla_{\mathbf{r}} \hat{E}$



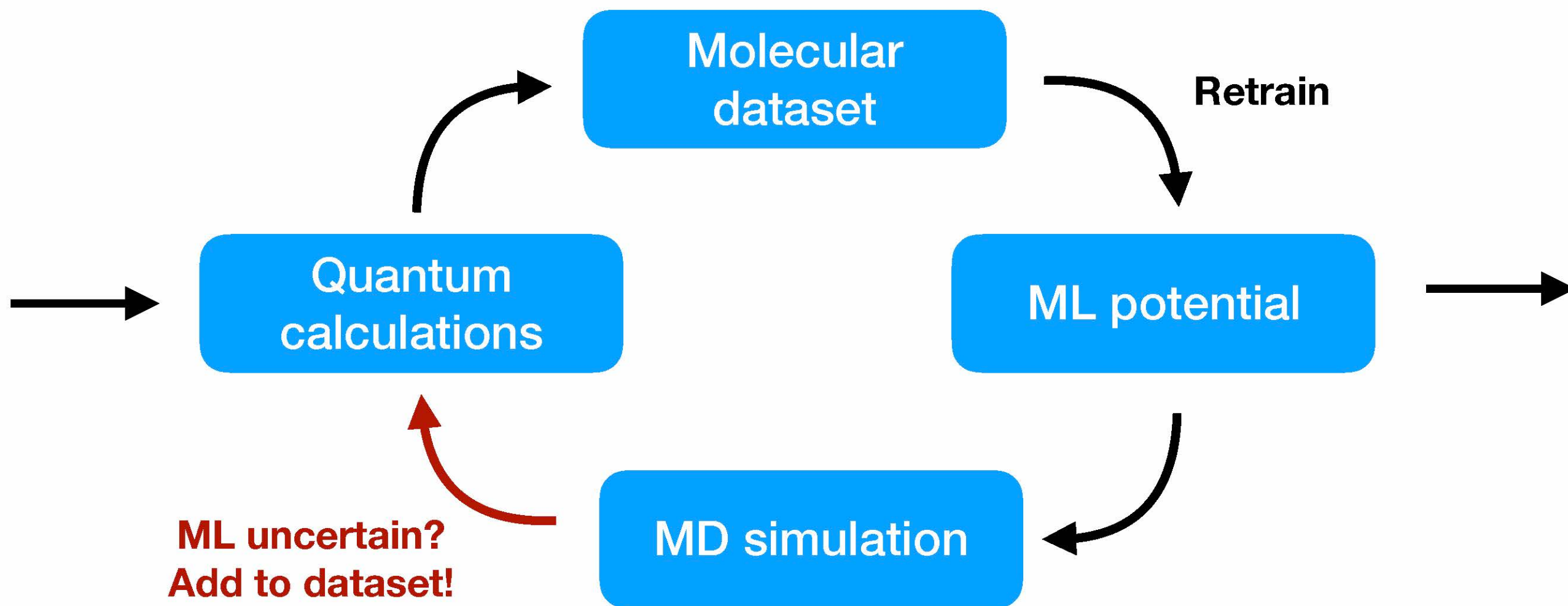
$$\text{Loss} \quad \mathcal{L} = \sum_{\text{Data}} \left[c_1 (E - \hat{E})^2 + c_2 (\nabla_{\mathbf{r}} E - \nabla_{\mathbf{r}} \hat{E})^2 \right]$$

Backprop 2, weight updates: $\nabla_w \mathcal{L}$



Active learning

Computer should help us design good datasets!

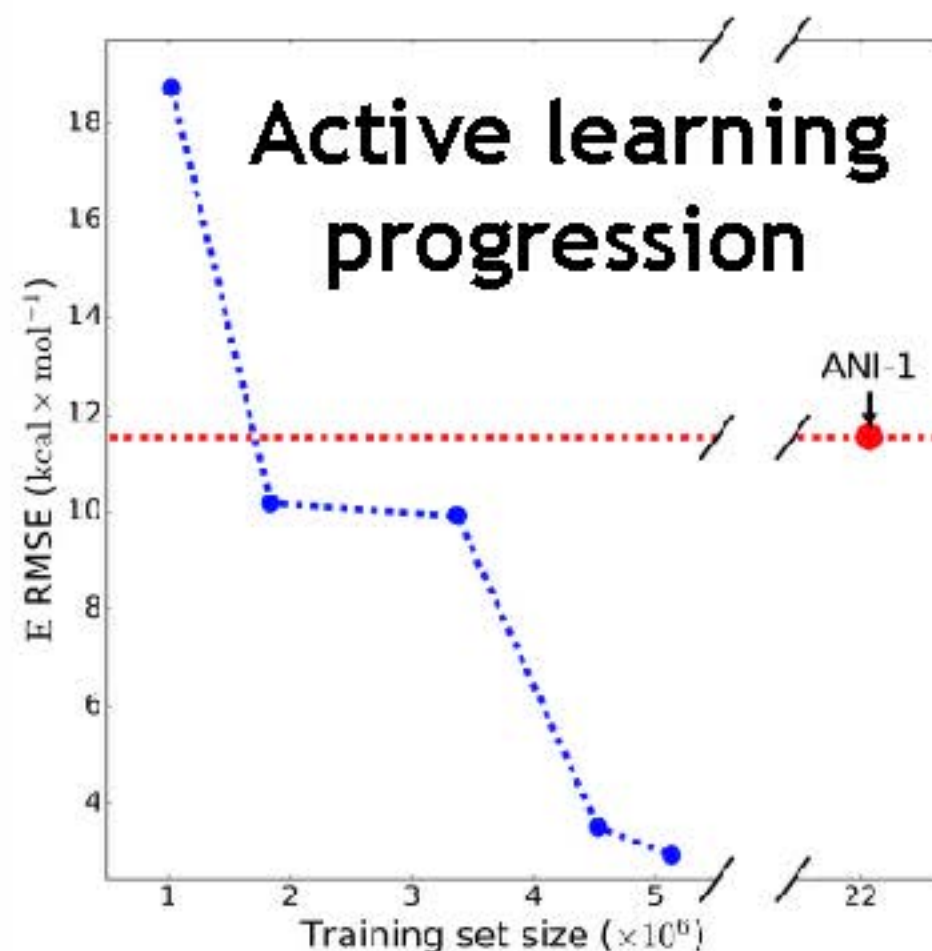


Smith et al., “Less is more: sampling chemical space with active learning”
<https://arxiv.org/abs/1801.09319>

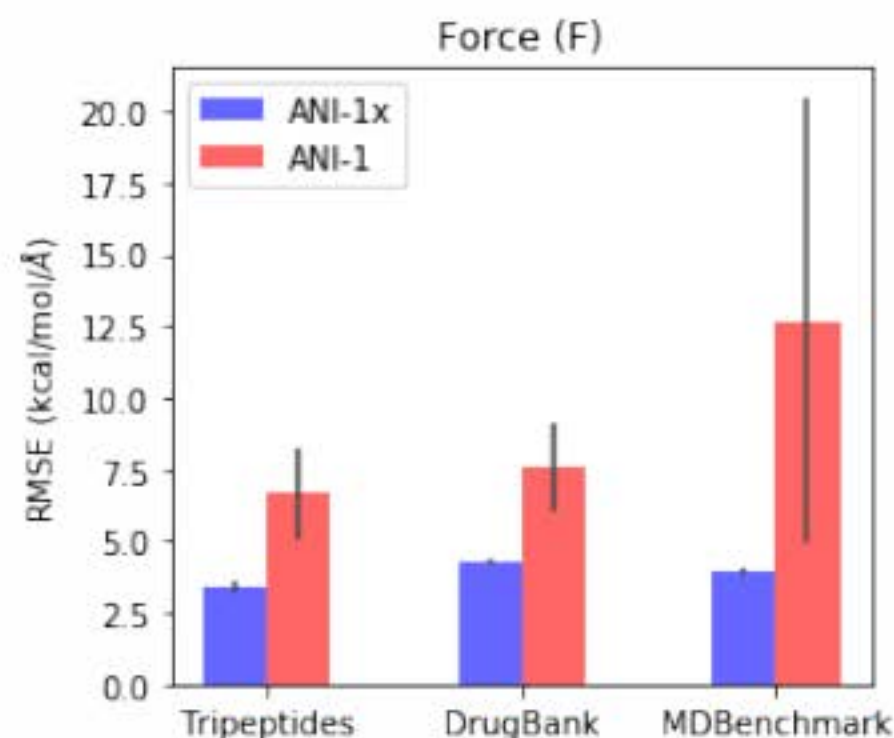
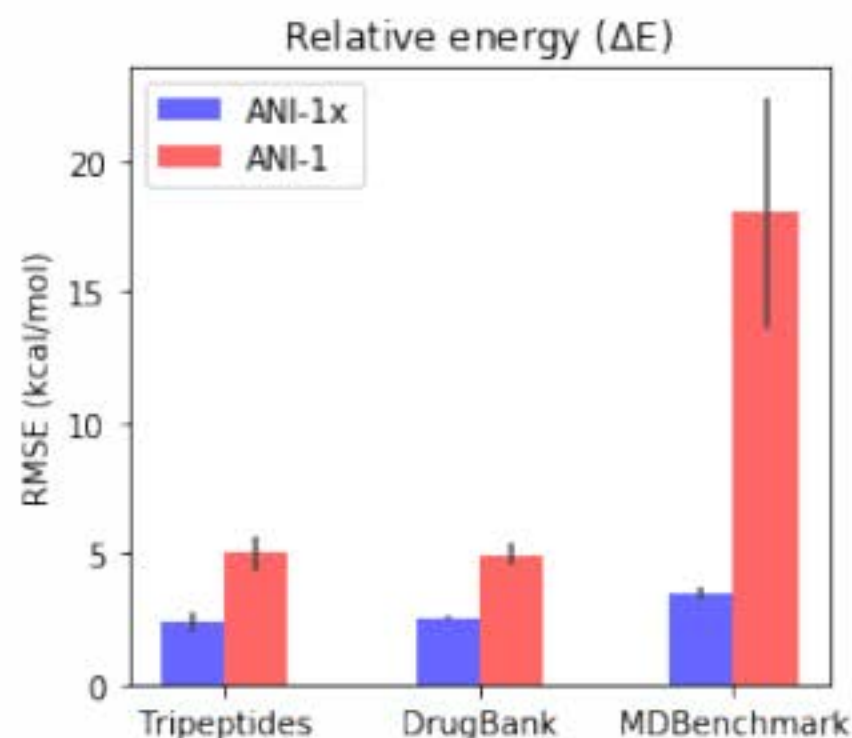
Active learning vs random sampling

Dataset sizes

ANI-1	ANI-1x
22M	5M



Errors

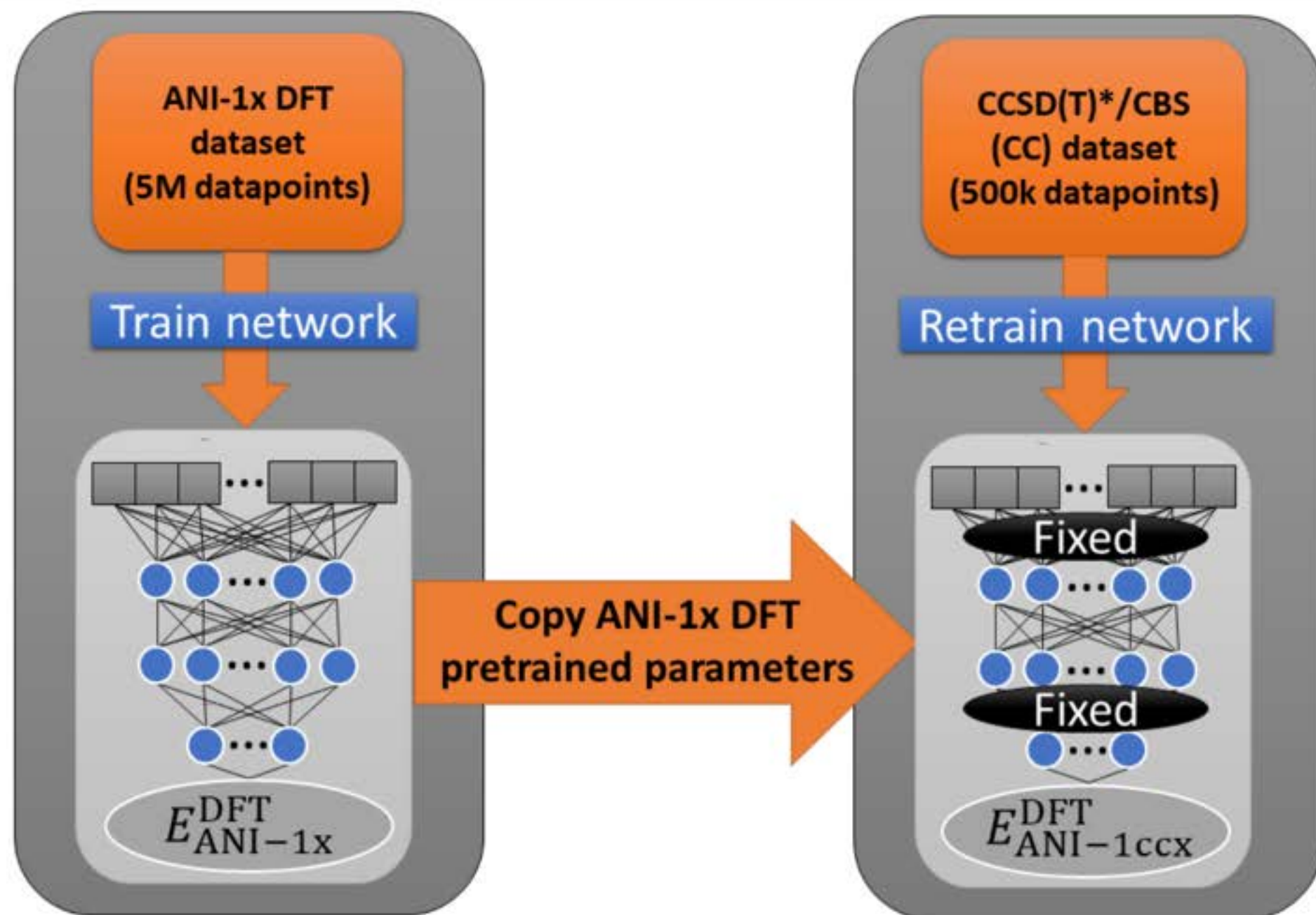


JS Smith, et al.; *The Journal of Chemical Physics*, (2018), 148 (24), 241733

Transfer learning

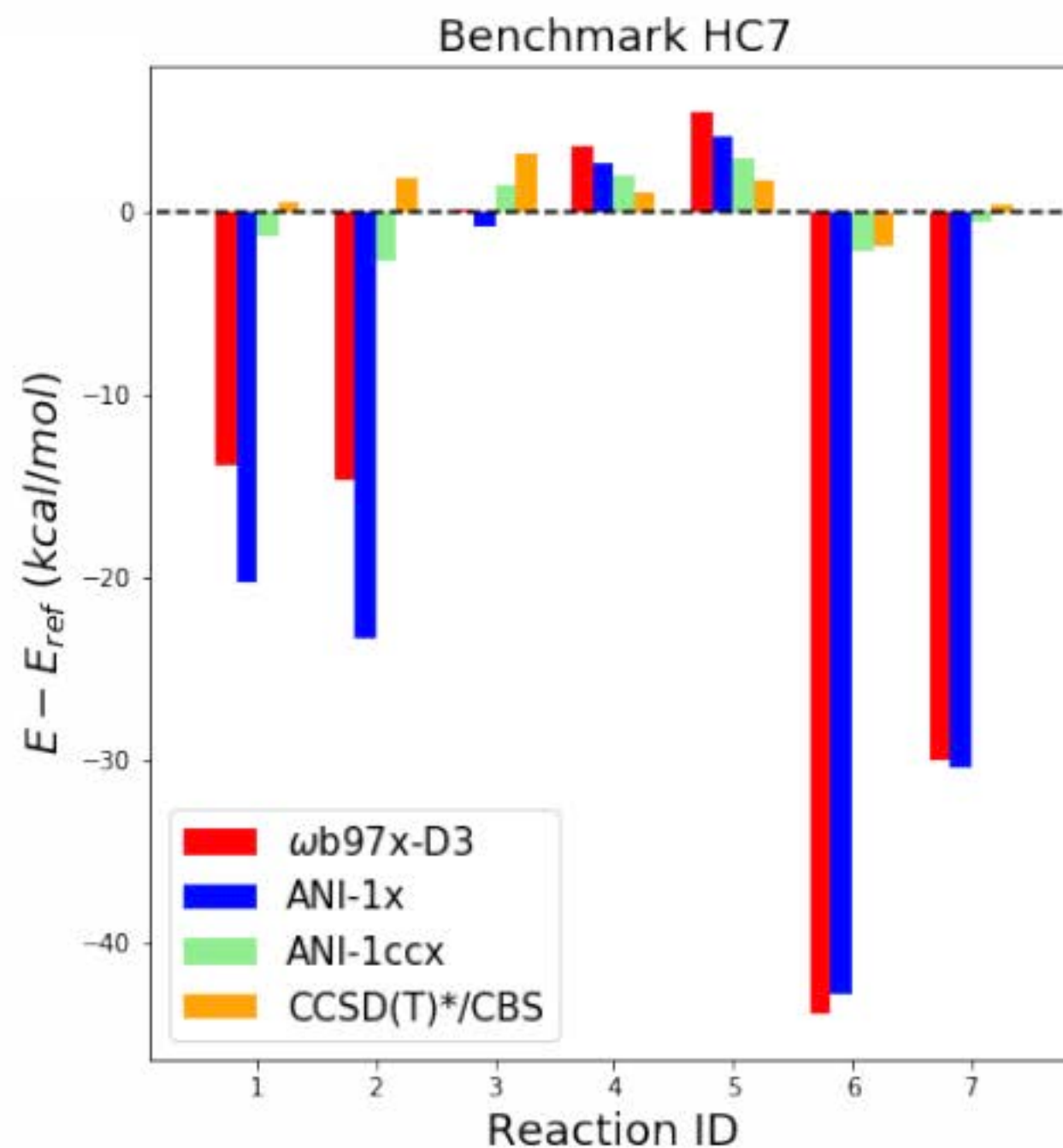
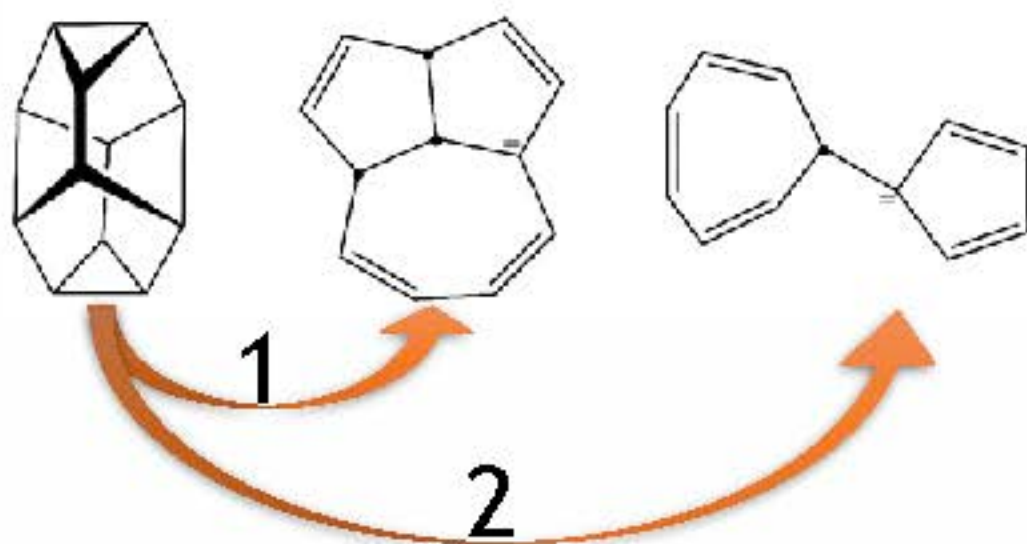
Combine lots of DFT data with some high accuracy CCSD data

- Subsample 10% of ANI-1x training data (0.5M of 5M)
- Recompute CCSD(T)/CBS level
- 340k parameters fixed, re-train 60k
- 10^7 faster than DFT



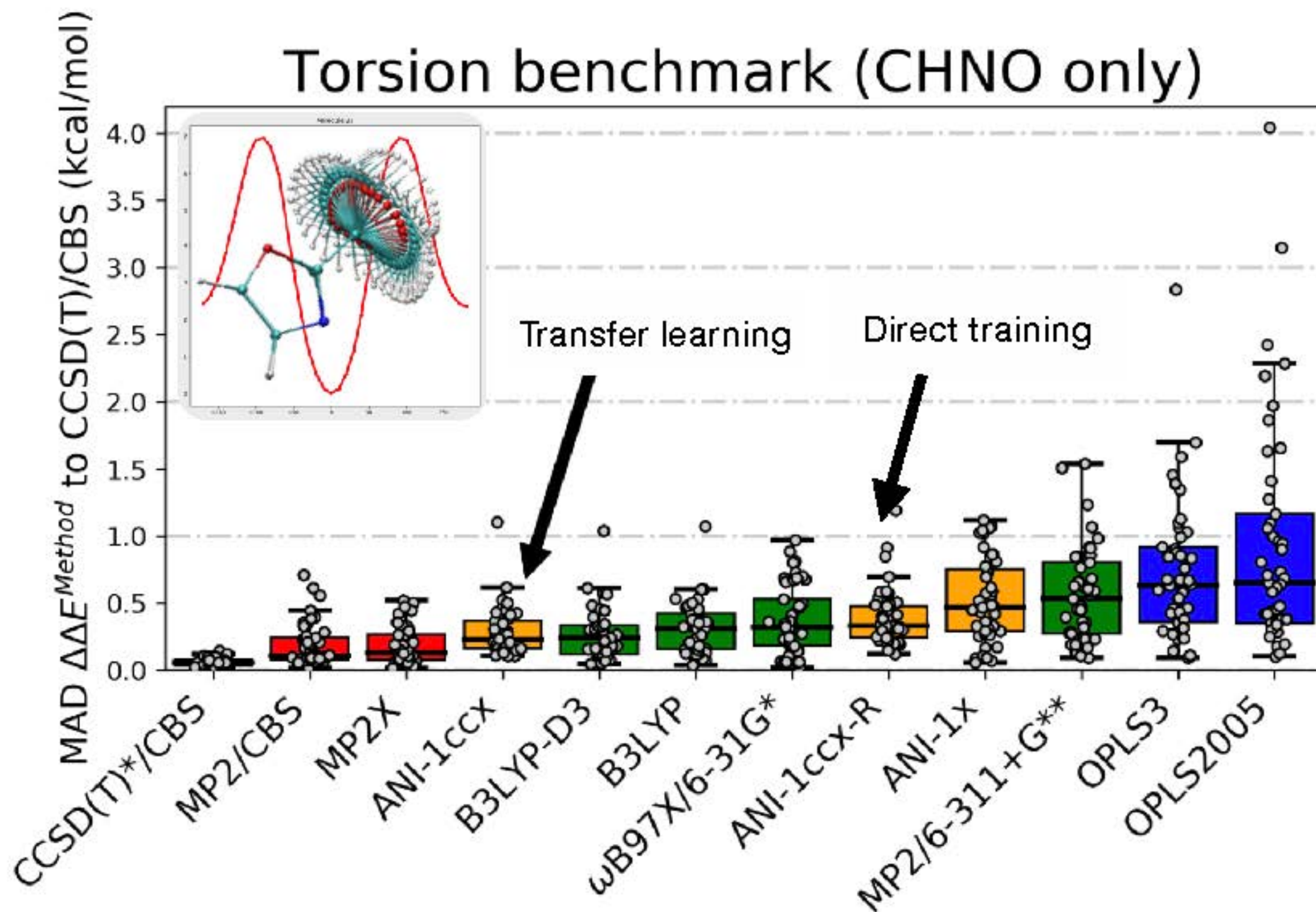
Hydrocarbon reaction energy benchmark

Examples



Reference data: Peverati, R.; Zhao, Y.; Truhlar, D. G., *J. Phys. Chem. Lett.* **2011**, 2 (16), 1991-1997.

Torsion benchmark (CHNO only)



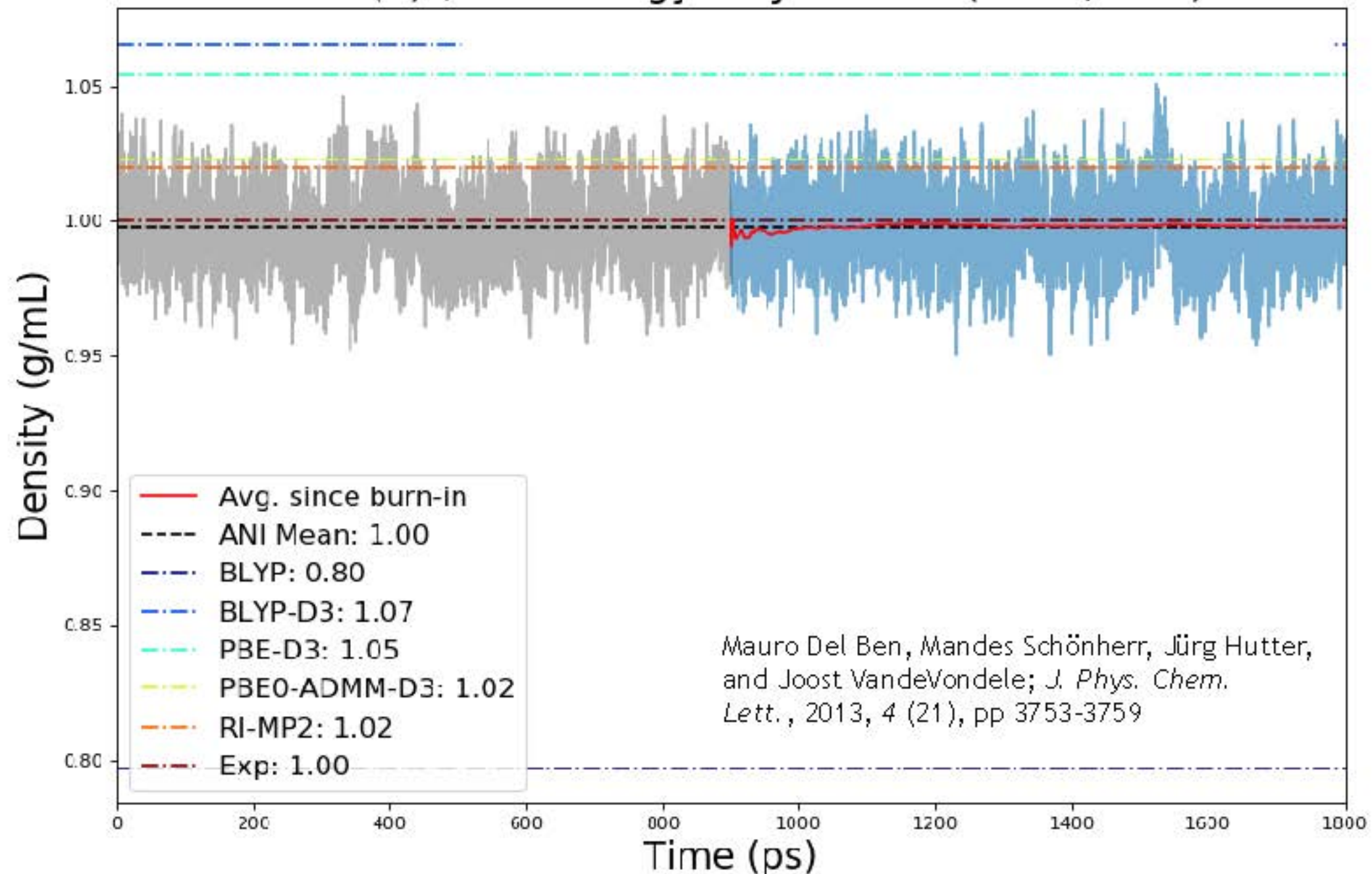
Sellers, B. D.; James, N. C.; Gobbi, A. A Comparison of Quantum and Molecular Mechanical Methods to Estimate Strain Energy in Druglike Fragments. *J. Chem. Inf. Model.* **2017**, 57 (6), 1265–1275.

Density of bulk water

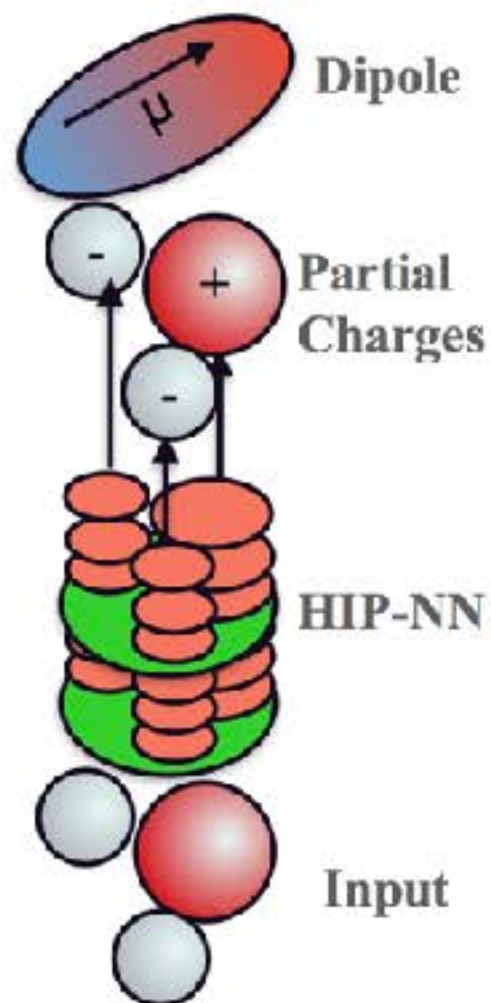
NPT conditions: 295K; 1Bar

267 waters

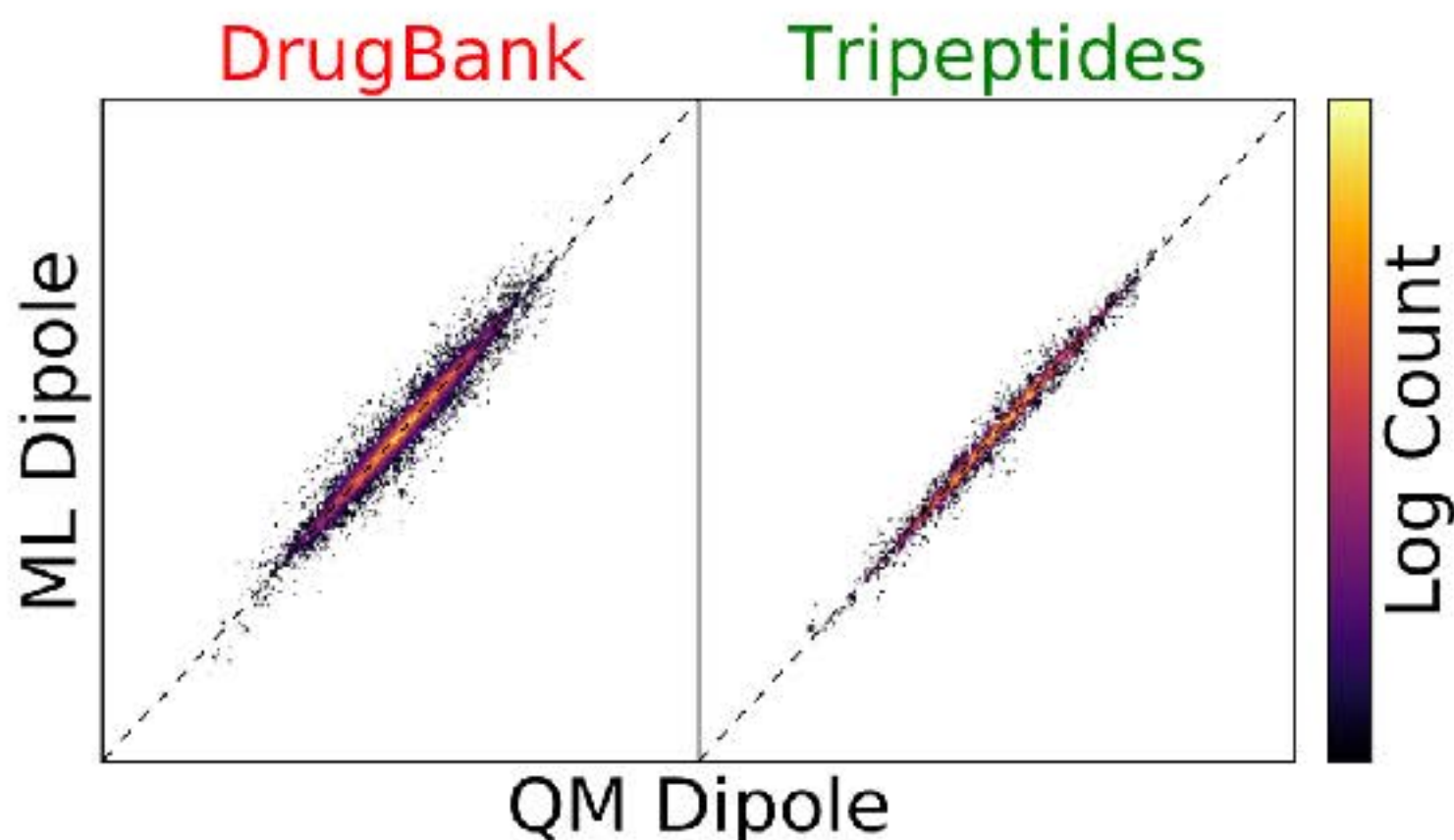
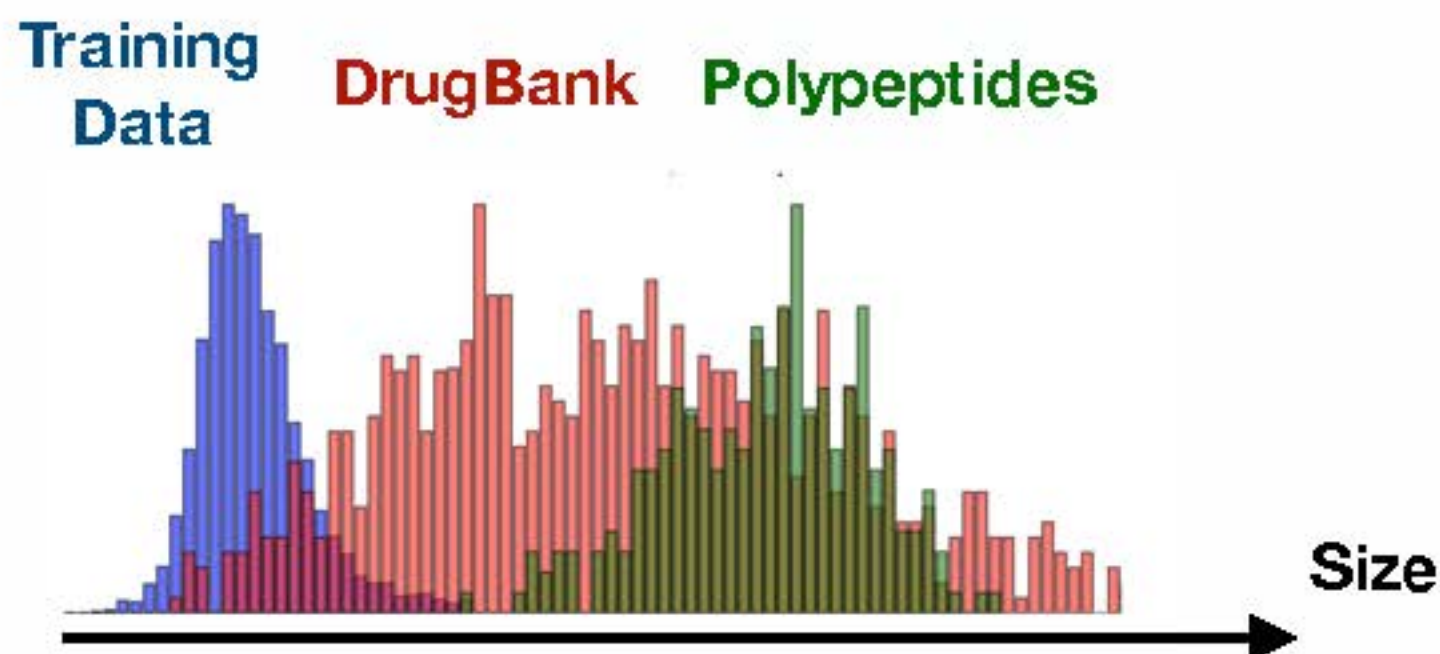
CCSD(T)*/CBS energy only trained (295K;1bar)



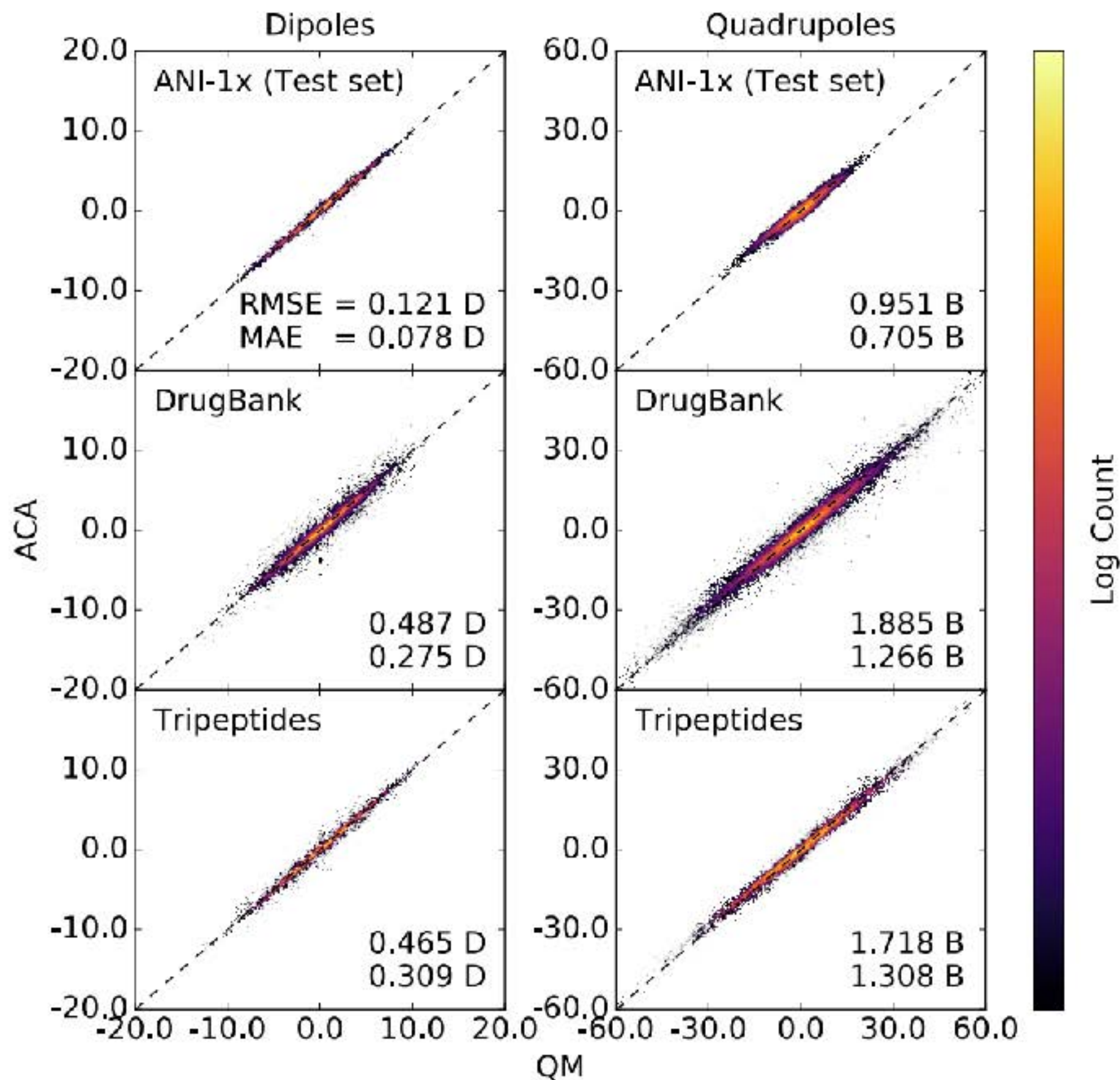
Inferring local charges to produce dipoles



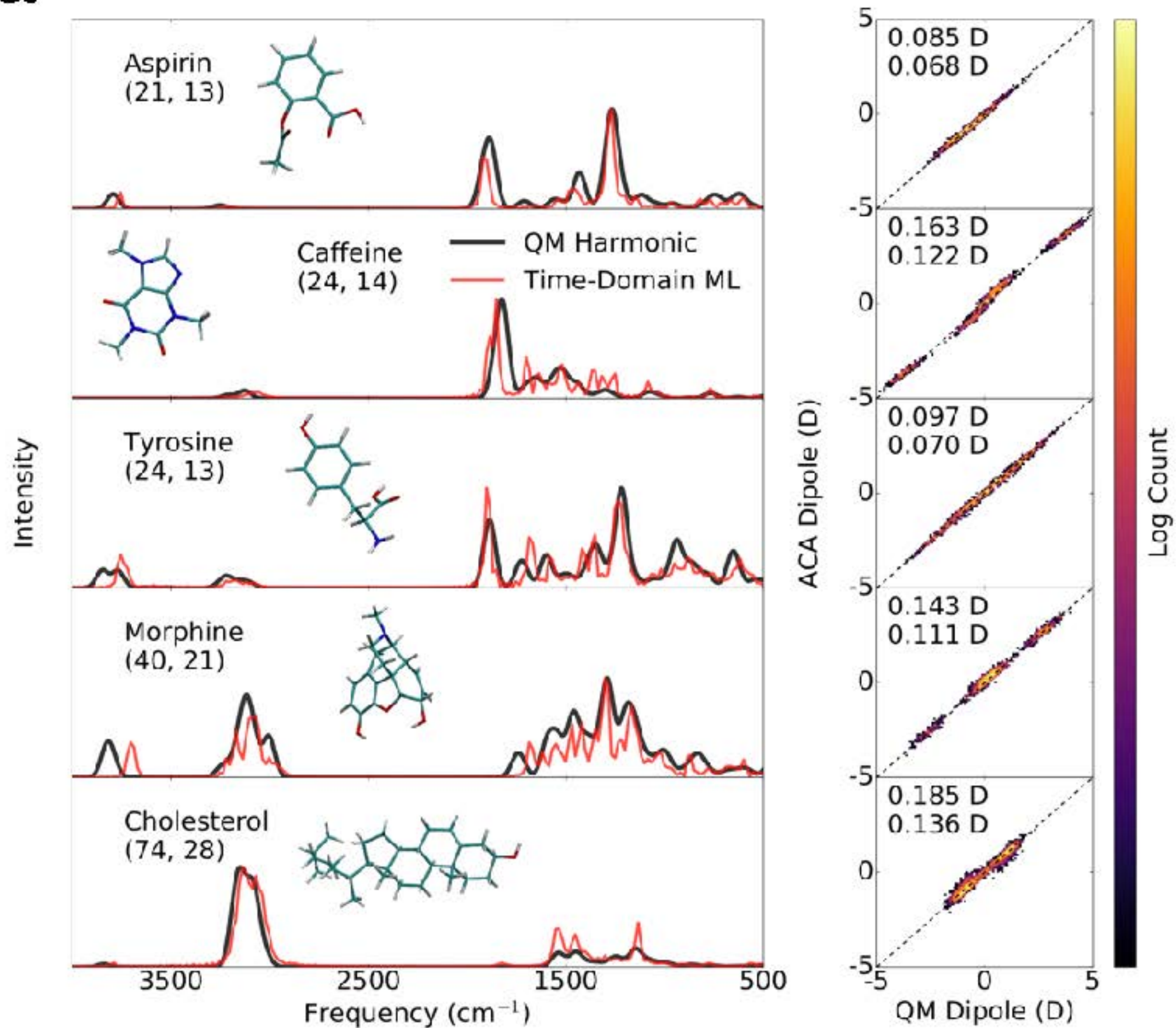
Sifain et al., "Discovering a Transferable Charge Assignment Model Using Machine Learning" [ChemRxiv:6638981]



Inferred charges generalize to *quadrupoles*



IR Spectra



Conclusions

Machine learning to *emulate quantum chemistry* works great!

Future directions?

Accounting for electron *dynamics*?

Joint-training to more information from wave-function?

Use ML to predict effective quantum Hamiltonians?

Better transfer learning? Can we incorporate data from experiment (e.g. phase diagrams)?

Kipton Barros

kbarros@lanl.gov

cnls.lanl.gov/~kbarros



Justin Smith



Nicholas Lubbers

Lubbers et al., “*Hierarchical modeling of molecular energies using a deep neural network*” [arXiv:1710.00017]

Smith et al., “*Less is more: sampling chemical space with active learning*” [arXiv:1801.09319]

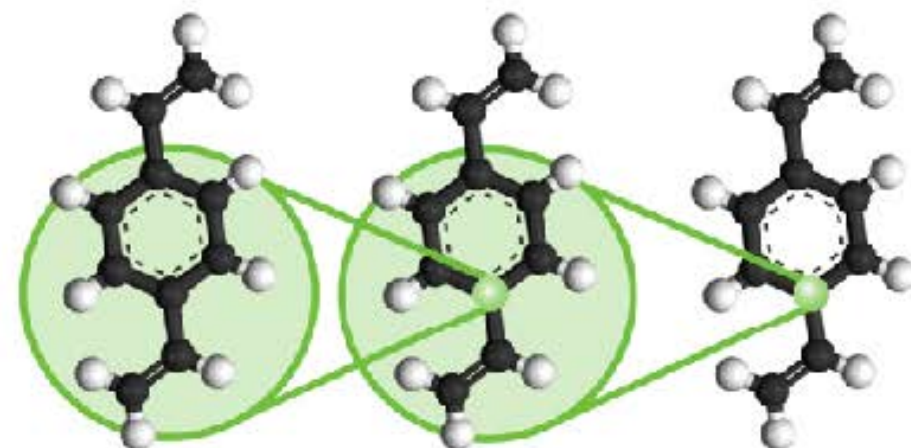
Smith et al., “*Outsmarting Quantum Chemistry Through Transfer Learning*” [ChemRxiv:6744440]

Sifain et al., “*Discovering a Transferable Charge Assignment Model Using Machine Learning*” [ChemRxiv:6638981]

Entire HIP-NN

$$E \approx \hat{E} = \sum_{i=1}^{N_{\text{atom}}} \hat{E}_i. \quad \hat{E}_i = \sum_{\ell} w_a^{\ell} z_{i,a}^{\ell}$$

**Locality/
Hierarchicality**



$$f(x) = \log(1 + e^x).$$

Activation fn.

$$\tilde{z}_{i,a}^{\ell+1} = f \left(\sum_{j,b} v_{ab}^{\ell}(r_{ij}) z_{j,b}^{\ell} + \sum_b W_{ab}^{\ell} z_{i,b}^{\ell} + B_a^{\ell} \right) \quad \text{Interactions}$$

$$z_{i,a}^{\ell+1} = \sum_b \left(\tilde{W}_{ab}^{\ell} \tilde{z}_b^{\ell+1} + \tilde{M}_{ab}^{\ell} z_{i,b}^{\ell} \right) + \tilde{B}_a^{\ell},$$

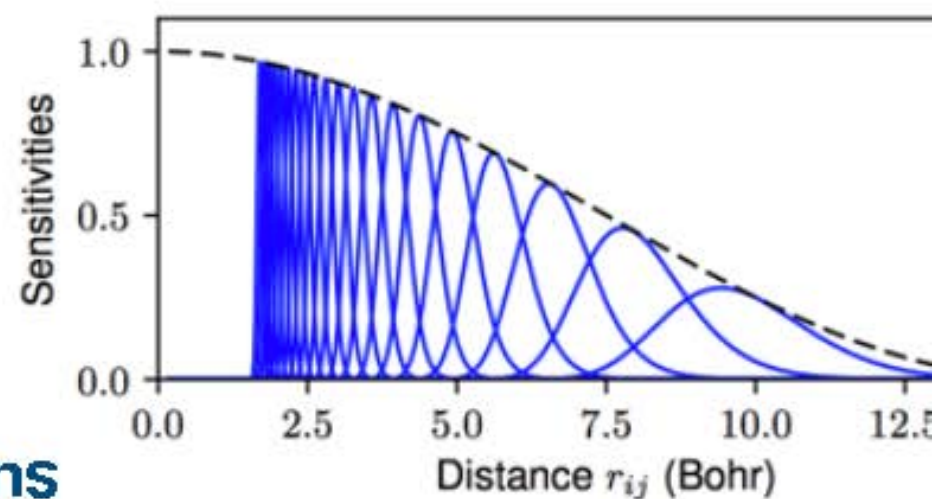
Res-net

$$v_{ab}^{\ell}(r_{ij}) = \sum_{\nu} V_{\nu,ab}^{\ell} s_{\nu}^{\ell}(r_{ij}),$$

Basis expansion

$$s_{\nu}^{\ell}(r) = \exp \left[-\frac{\left(r^{-1} - \mu_{\nu,\ell}^{-1} \right)^2}{2\sigma_{\nu,\ell}^{-2}} \right] \varphi_{\text{cut}}(r).$$

Sensitivity functions



$$\mathcal{L} = \sum_{\text{Data}} \left[c_1 (E - \hat{E})^2 + c_2 (\nabla E - \nabla \hat{E})^2 \right] \quad \text{Loss fn.}$$