XACS
Xiamen Atomistic Computing Suite
XACScloud.com

MLatom.com

Δ-learning

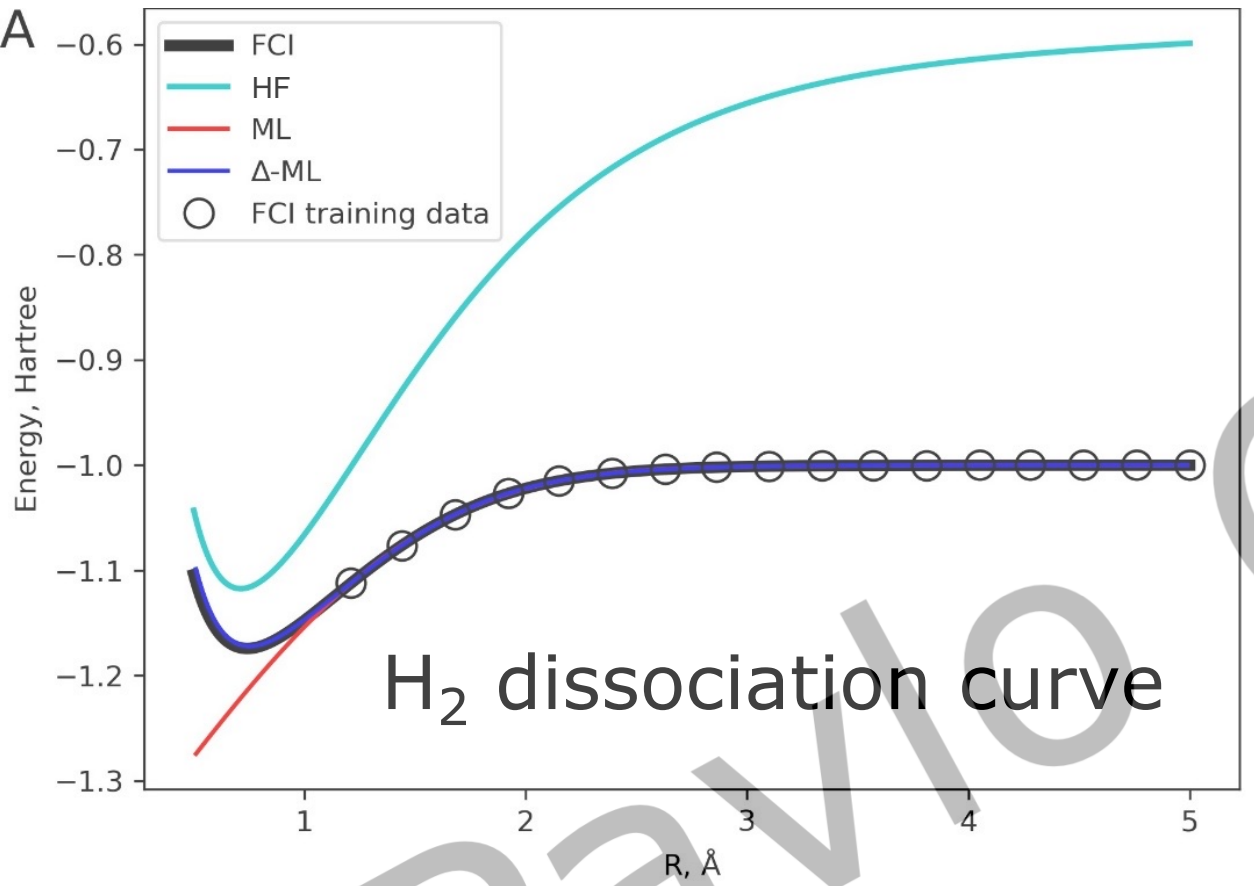
Pavlo O. Dral

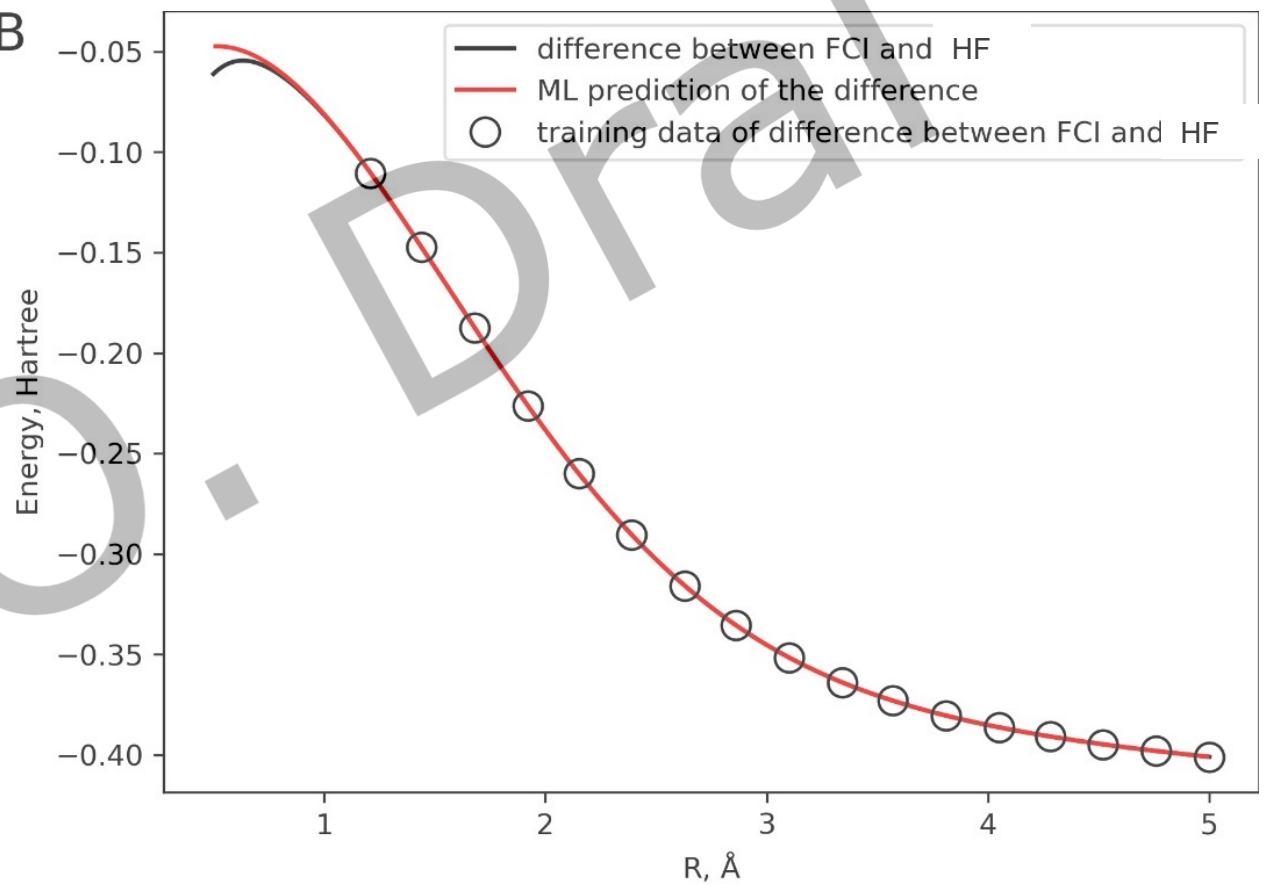
Xiamen University, P.R. China

Visiting Professor in
Nicolaus Copernicus University, Poland

5 July 2024

Multi-fidelity data: Δ-learning





Δ-learning: R. Ramakrishnan, P. O. Dral, M. Rupp, O. A. von Lilienfeld, *J. Chem. Theory Comput.* **2015**, *11*, 2087

Figure: Pavlo O. Dral, Tetiana Zubatiuk, Bao-Xin Xue, Learning from multiple quantum chemical methods: Δ-learning, transfer learning, co-kriging, and beyond. In *Quantum Chemistry in the Age of Machine Learning*, Pavlo O. Dral, Ed. Elsevier: **2023**. Paperback ISBN: 9780323900492

Example ML-delta.

Train the KREG model on the differences between the FCI and HF/STO-3G energies **interatomic distances larger or equal 0.8 Angstrom** of H_2. These differences are residual error of HF/STO-3G wrt FCI.

Estimate the residual errors with this KREG delta-learning model for all internuclear distances from 0.5 to 5 Angstrom. Add them back to HF/STO-3G energies to obtain an estimate of FCI energies with such a delta-learning model.

See the previous tasks how to train and predict.

Plot the potential energy curve:

1. How does this delta-learning model perform?
2. Compared to the KREG model you trained previously directly on the FCI data for the same training set (e.g., in the previous task)?

Required data files (full potential energy curve from 0.5 to 5 Angstrom - you need to cut them as needed):

- E_FCI_451.dat - full CI/aug-cc-pV6Z energies.
- H2_HF.en - HF/STO-3G energies.
- h2.xyz - geometries.
- R_451.dat - interatomic distances for your convenience (they can be extracted from `h2.xyz` file).

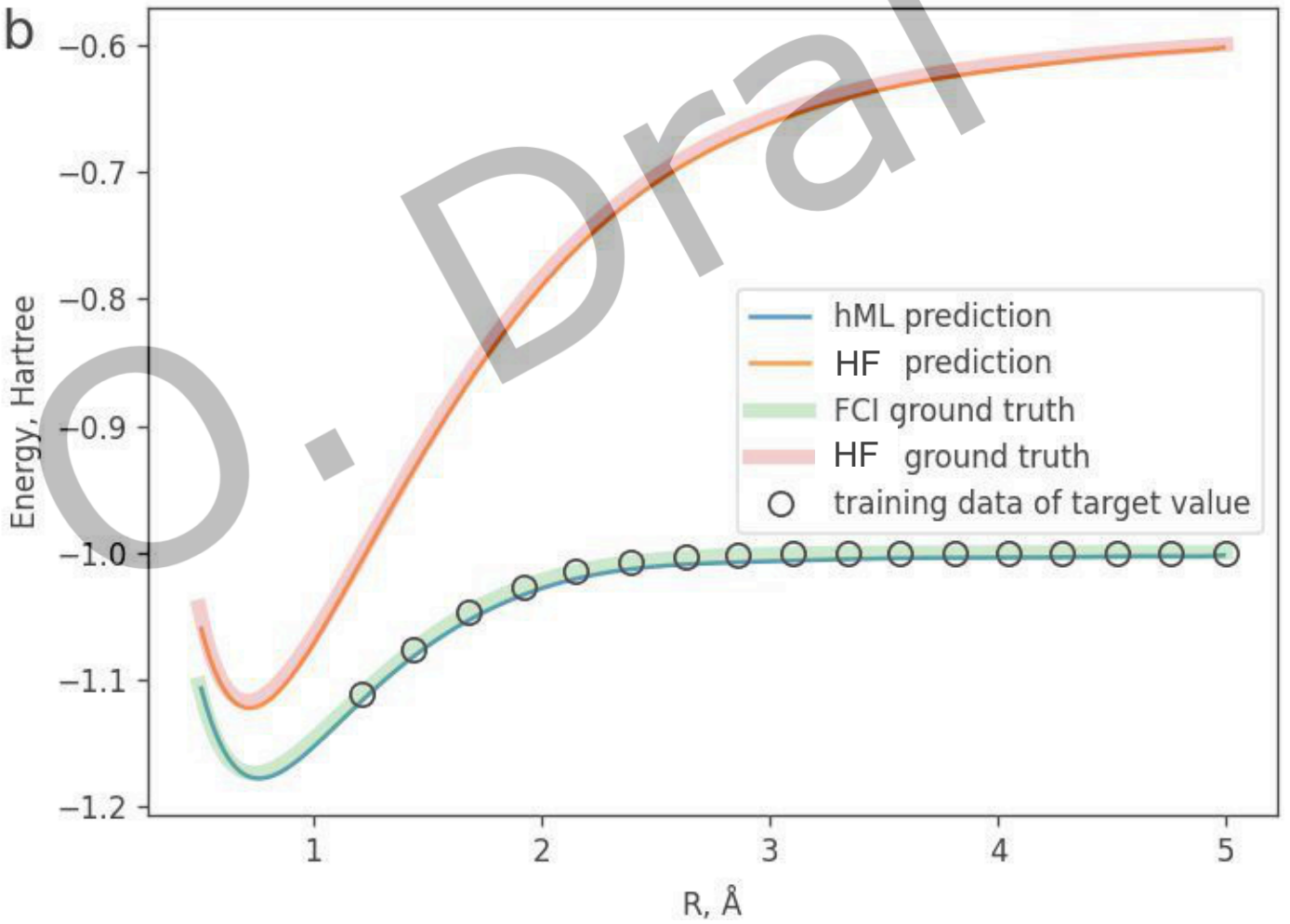
Hierarchical ML (hML): generalized Δ-learning

Automatic procedure to find optimal training points for each delta-model

$$\hat{y}_{\Sigma,i} = \sum_M \hat{y}_{M,i}\left(N_{\mathrm{tr},M}\right)$$

hML for H_2

$$\mathrm{hML} = \Delta_0^{\mathrm{HF}} + \Delta_{\mathrm{HF}}^{\mathrm{FCI}}$$



Hierarchical ML: P. O. Dral, A. Owens, A. Dral, G. Csányi, *J. Chem. Phys.* **2020**, *152*, 204110

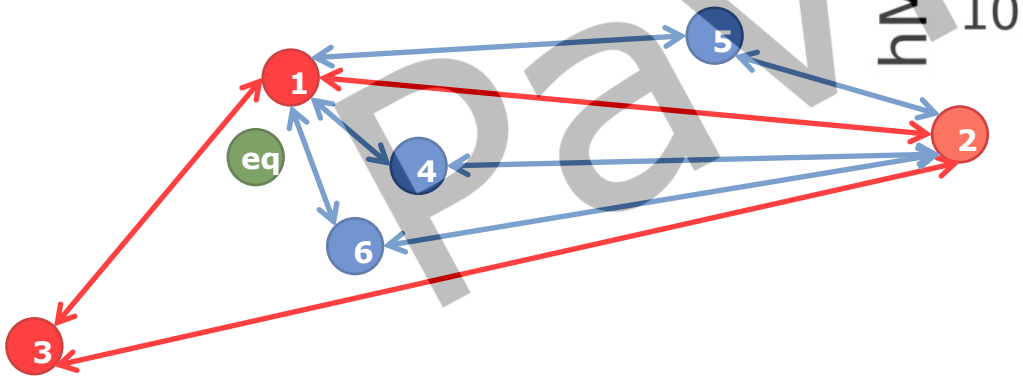
Figure: Pavlo O. Dral, Tetiana Zubatiuk, Bao-Xin Xue, Learning from multiple quantum chemical methods: Δ-learning, transfer learning, co-kriging, and beyond. In *Quantum Chemistry in the Age of Machine Learning*, Pavlo O. Dral, Ed. Elsevier: **2023**. Paperback ISBN: 9780323900492

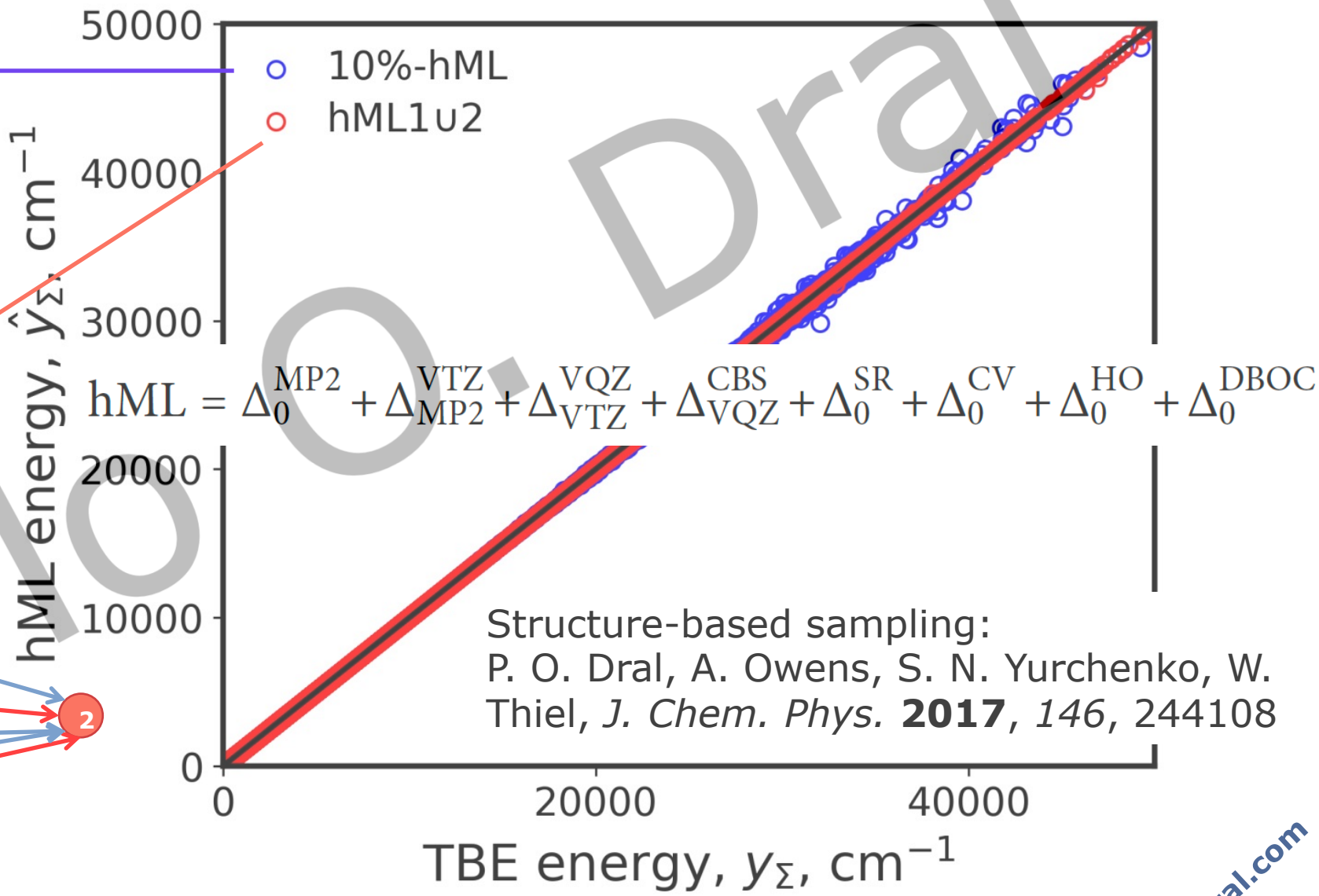
Hierarchical ML for CH_3Cl PES

Pure ML-model trained on **saved 90% of CPU-time**

Weighted RMSE:
3.49 cm⁻¹ = 0.01 kcal/mol

Δ-ML models:
saved 99% of CPU-time
Weighted RMSE:
1.12 cm⁻¹ = 0.003 kcal/mol



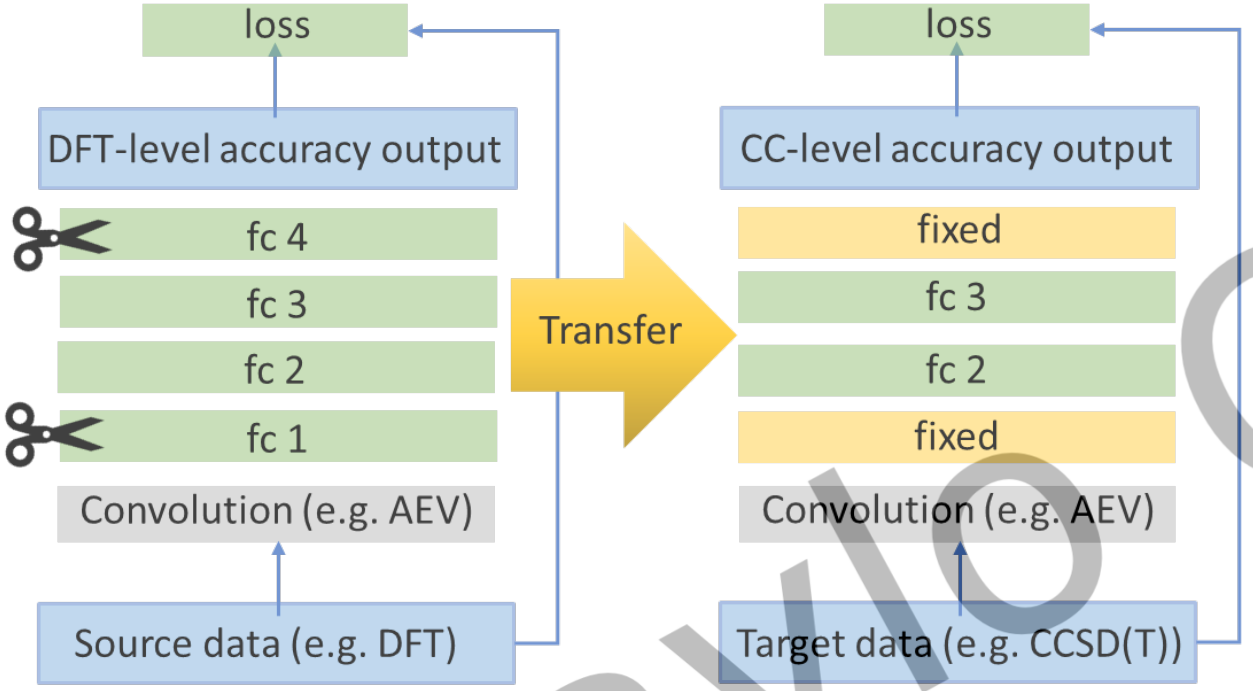


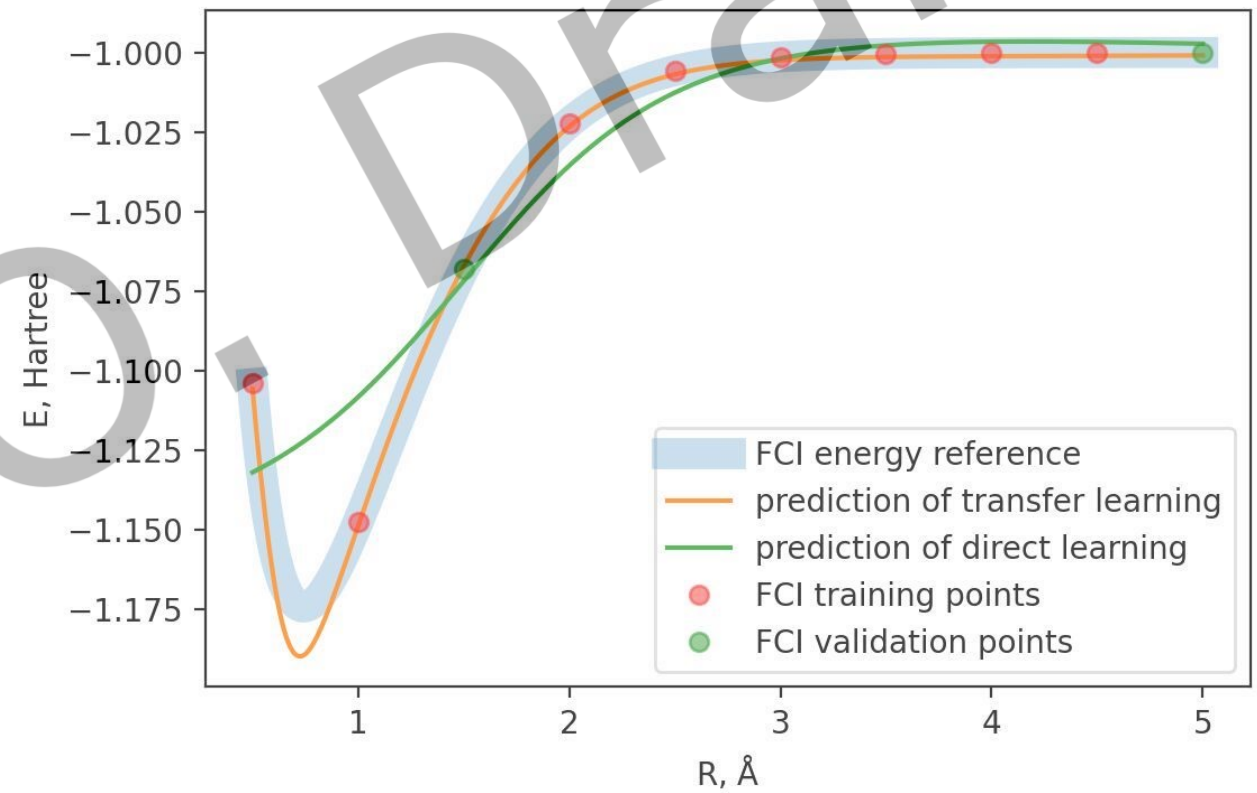
$$\mathrm{hML} = \Delta_0^{\mathrm{MP2}} + \Delta_{\mathrm{MP2}}^{\mathrm{VTZ}} + \Delta_{\mathrm{VTZ}}^{\mathrm{VQZ}} + \Delta_{\mathrm{VQZ}}^{\mathrm{CBS}} + \Delta_0^{\mathrm{SR}} + \Delta_0^{\mathrm{CV}} + \Delta_0^{\mathrm{HO}} + \Delta_0^{\mathrm{DBOC}}$$

Structure-based sampling:
P. O. Dral, A. Owens, S. N. Yurchenko, W. Thiel, *J. Chem. Phys.* **2017**, *146*, 244108

P. O. Dral, A. Owens, A. Dral, G. Csányi, *J. Chem. Phys.* **2020**, *152*, 204110

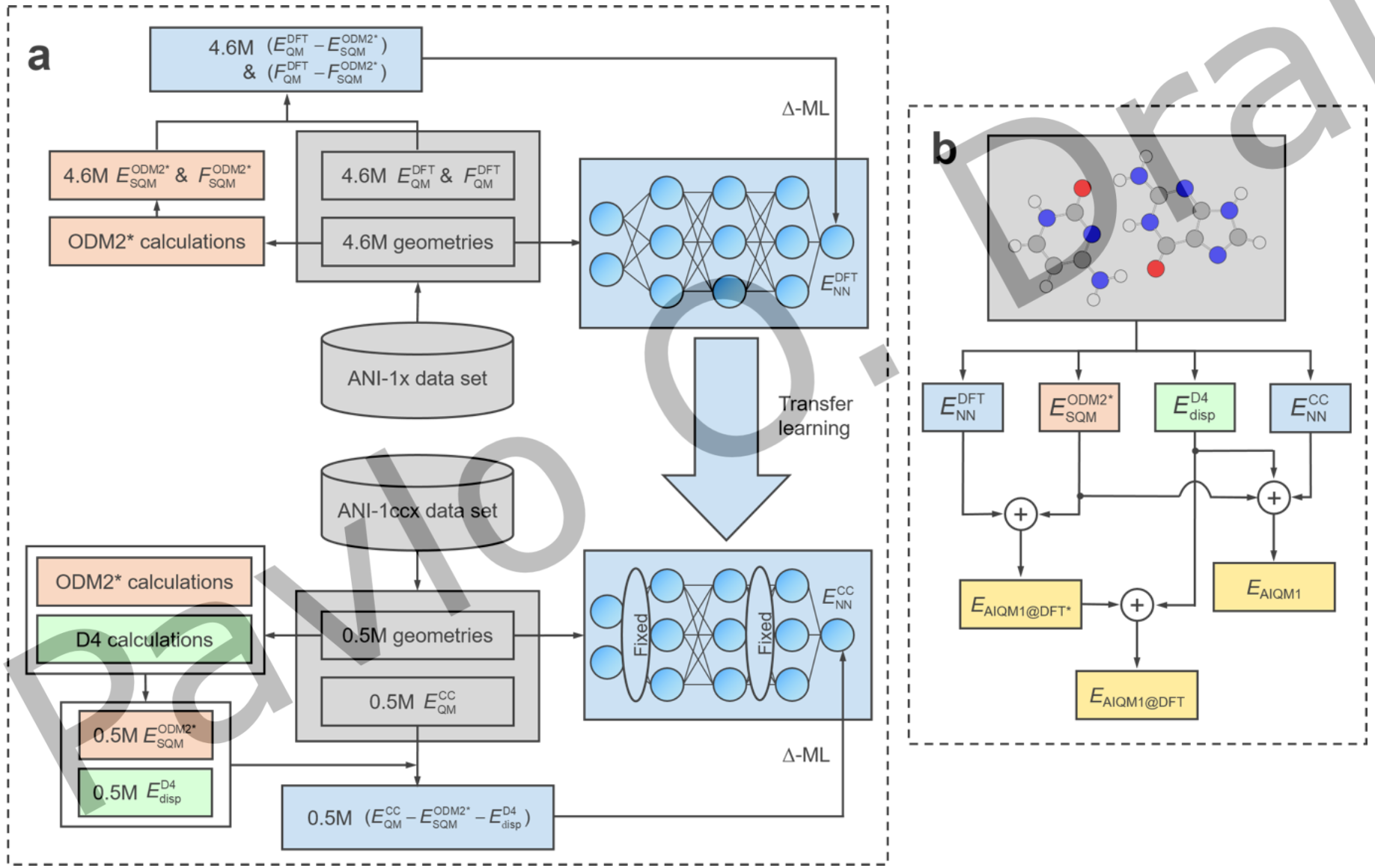
Multi-fidelity data: Transfer learning





Figures: Pavlo O. Dral, Tetiana Zubatiuk, Bao-Xin Xue, Learning from multiple quantum chemical methods: Δ-learning, transfer learning, co-kriging, and beyond. In *Quantum Chemistry in the Age of Machine Learning*, Pavlo O. Dral, Ed. Elsevier: **2023**. Paperback ISBN: 9780323900492

Design



P. Zheng, R. Zubatyuk, W. Wu, O. Isayev, P. O. Dral, *Nat. Commun.* **2021**, *12*, 7022