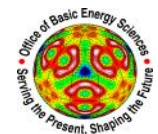


TOWARDS DIFFUSION MONTE CARLO ACCURACY ACROSS CHEMICAL SPACE WITH SCALABLE Δ -QML

ANOUAR BENALI, MARK DEWING, KEVIN GASPERICH, YE LUO, HYEONDEOK SHIN

Computational Science Division (CPS), Argonne National Laboratory, Lemont IL, 60439, USA



CENTER FOR PREDICTIVE SIMULATION OF FUNCTIONAL MATERIALS (CPSFM)
US-DEPARTMENT OF ENERGY - BES COMPUTATIONAL MATERIALS SCIENCES PROGRAM



Anouar Benali



Ye Luo



Hyeondeok Shin



Mark Dewing



Kevin Gasperich

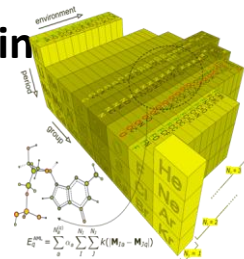
Choice of Theoretical accuracy for dataset training (beyond DFT methods)

The predictive accuracy of quantum machine learning (QML) models trained on quantum chemistry data and used for the navigation of chemical compound space (CCS) is inherently limited by the predictive accuracy of the approximations used within the underlying quantum theory.

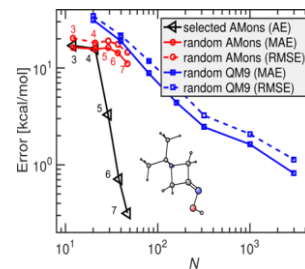
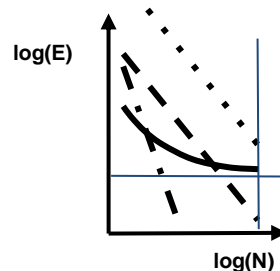
Quantum Monte Carlo methods

- Scales as N^3 with the system size and linearly with the number of compute nodes.
- Highly accurate for solids and molecules (within 1kcal/mol from experimental measurements).
- Fully exascale with highly efficient GPU optimizations

Choice of training dataset: Atom in Molecule: “AM-on”



Elementary → **building blocks** → **structure** → **effect**
Letters → words → sentence → meaning
Atoms → **AMONS** → molecule → property



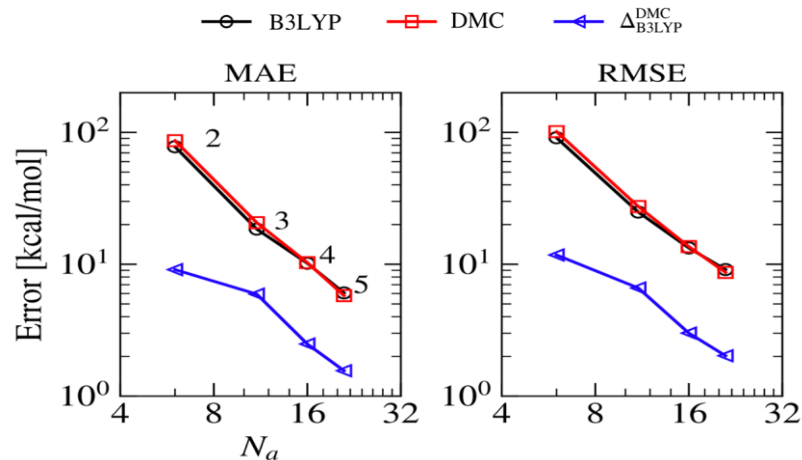
Collaborators:

- Anatole von Lilienfeld, Vector Institute
- Rajeev Assary, MSD, Argonne
- Ben Blaisik, Ian Foster, DSL, Argonne
- Paul RC. Kent, ORNL

Amon based Δ -QML on QMC

Converged DMC estimates of atomization energies for 1'175 amons with $N \leq 5$ from amon dictionary ...

and for 50 random QM9 molecules with $N = 9$



N_a : Number of amons used in the training

