QMC-HAMM P1: Quantum Monte Carlo database for machine learning accurate interatomic potentials of dense hydrogen

Scott Jensen, Yubo Yang, Kacper Kowalik, Hongwei Niu, Markus Holzmann, Carlo Pierleoni, Matthew Turk, and David M. Ceperley

QMC-HAMM | summit | BLUE WATERS SUSTAINED PETASCALE COMPUTING | U.S. DEPARTMENT OF ENERGY | Office of Science

ILLINOIS

Understanding the phase diagram of dense hydrogen

The phase diagram of dense hydrogen is not completely known but its properties are necessary in astrophysics to understand the structure of the giant planets. Calculating the complex phase diagram with accurate Ab initio simulations like quantum Monte Carlo (QMC) is difficult [1] because of its computational demands.

Machine learning solution

Machine-learned interatomic potentials trained on Ab initio data approach the accuracy of the Ab initio methods.

Machine learning workflow

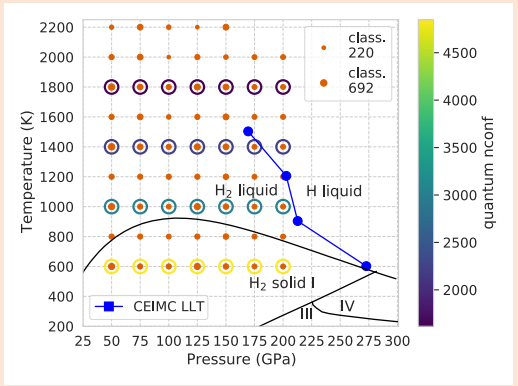
1) Generate atomic snapshots with molecular dynamics (MD) and Born-Oppenheimer path integral Monte Carlo (BOPIMC)[1].
2) Electronic properties from density functional theory (DFT) and QMC.
3) Train with DeepMD on Ab initio forces [2].
4) Large-scale MD to explore phase diagram.

Data product: QMC hydrogen database

We present the first large-scale publicly accessible database of QMC force calculations for dense hydrogen enabling more accurate machine-learned potentials.

- Database of 100,000 configurations for 96 protons.
- Generated with MD and BOPIMC
- Forces from DFT with PBE and vdW functionals
- Subset of 18,000 configurations selected for accurate QMC calculations.

Simulations grid for database



Typically, ML potentials in the literature have relied on density functional theory to generate the training data.

yt Hub

The yt Hub is a deployment of an open source data repository based on the Girder project.

- Storage
- Active and passive analysis of datasets
- Custom and queryable metadata
- Access control lists for fine-grained sharing and distribution of datasets

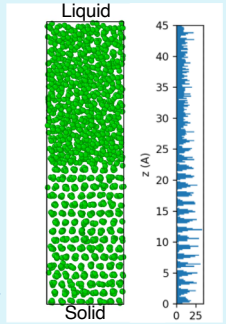
We have extended it to include QMC-HAMM-specific visualization as well as to provide on-demand Jupyterlab analysis.

Target application: Melting of molecular hydrogen

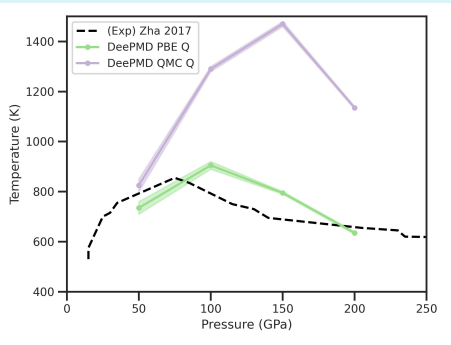
First study with database: Determine the melting line.

Previous works model hydrogen using PBE with less accuracy than QMC.

Two-phase NPT ensemble runs with 3072 protons to determine the phase.



QMC database allows for accurate calculation of phase diagram with ML potentials.



References

[1] J. M. McMahon, M. A. Morales, C. Pierleoni, and D. M. Ceperley, Rev. Mod. Phys. 84, 1607 (2012)
[2] H. Wang, L. Zhang, J. Han, and W. E., Commun. Comput. Phys. 23, 629 (2018).
[3] C. Zha, H. Liu, J.S. Tse, R.J. Hemley, Phys. Rev. Lett. 119, 075302 (2017).

Acknowledgements

This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, Computational Materials Sciences program under Award Number DE-SC0020177.