Establishing the equation of state of matter in silico: the first machine learning potential of hydrogen with QMC accuracy

Andrea Tirelli, Giacomo Tenti, Kousuke Nakano, Sandro Sorella

International School for Advanced Studies (SISSA)
Via Bonomea 265, 34136 Trieste, Italy

atirelli@sissa.it

In this work, we develop a new technique combining, for the first time, the accuracy of quantum Monte Carlo in describing the electron correlation with the efficiency of machine learning algorithms, in particular Gaussian Kernel Regression. Thanks to such improvements, we are able to obtain a machine learning potential (MLP) that is able to reproduce very accurately computationally prohibitive Quantum Monte Carlo (QMC) state of the art results. On the one hand, our work crucially relies on HPC simulations through the QMC method for the generation of high quality training datasets, which are exploited to generate highly accurate and efficient MLP. On the other hand, we believe that our work will significantly reduce by orders of magnitude the amount of HPC resources needed to obtain MD simulations with this level of accuracy.

Our technique is perfectly suited for being employed in a context where high accuracy is needed, but training data only come in small sizes, which is exactly the context of Quantum Monte Carlo. We benchmarked our approach by studying the liquid-liquid transition in Hydrogen at high pressures, whose phase diagram has been the subject of a very lively debate in the scientific community. In this context, we have obtained a MLP that is capable to detect the correct pressure plateau for certain isotherms obtained by long enough modelcular dynamics simulations, in contrast with the less accurate MLPs provided in previous works.

We believe that the present work represents a significant advance for the development of MLPs and, in particular, constitutes the first successful attempt to combine such Machine Learning techniques with advanced many-body techniques, such as Quantum Monte Carlo; this will enable the study of systems that have so far been impossible to simulate accurately, due to their intractable computational cost. As such, this work should have a considerable impact not only within the Condensed Matter and Materials Science communities, as well as Astrophysics and planetary science, but also for a possible direct technological impact for designing and predicting new materials by means of highly reliable and accurate computer simulations, combining QMC and state-of-the-art ML techniques.