Data-Driven Models for Predictive Molecular Simulations

PI: Paesani; Co-PIs: Götz and Zonca

Department of Chemistry and Biochemistry, and San Diego Supercomputer Center
University of California San Diego

Abstract

The main objective of our research project involves the development and implementation of unique software elements that will enable, for the first time, computer simulations on both CPU and GPU architectures using our many-body molecular dynamics (MB-MD) methodology. These software elements will be made publicly available to the scientific community through an integrated platform that combines automated computational tools based on supervised learning, volunteer distributed computing through a specialized BOINC server-client infrastructure, and robust plugins for the OpenMM toolkit for molecular simulations. MB-MD is built upon a rigorous many-body expansion of the interaction energies, which is used to develop fully transferable many-body representations of potential energy, dipole moment, and polarizability surfaces derived entirely from correlated electronic structure data without resorting on empirical or ad hoc parameters. MB-MD effectively represents a new paradigm for molecular simulations, and has already been shown to exhibit chemical and spectroscopic accuracy for aqueous systems from the gas to the condensed phase.

The new software elements comprise three components integrated in a unique platform:

1. A suite of computational tools for the automated generation of many-body potential energy functions from electronic structure data.
2. A BOINC client-server architecture for the calculation of the required electronic structure data through volunteer computing.
3. Independent CPU and GPU plugins for the OpenMM toolkit that enable MB-MD simulations of generic molecular systems across different phases. Besides supporting simulations at the classical level, our MB-MD plugins will also provide the necessary infrastructure for quantum simulations within the path-integral molecular dynamics and centroid molecular dynamics formalisms, enabling direct comparisons with molecular-level experimental measurements, including linear and nonlinear vibrational spectra.