Quantum-Classical Consistency of Machine Learning Force Field with OM/MM Simulations: tests with liquid of water, application to ionic liquids

Evaluating quantum–classical consistency in ionic liquids (ILs) using multiscale simulations, starts with validation on bulk water. We will perform full ab-initio simulations with CP2K, including many-body dispersion corrections, to generate reference dataset for machine learning. This dataset will be used to train Machine Learning Coarse-Graining (MLCG) models, construct force fields, and reproduce key properties.

L. Delle Site, S. Ahmadkhani*, Freie Universität Berlin, Department of Mathematics and Computer Science, Mathematics Institute, Arnimallee 9, 14195 Berlin*

In Short

- **Reference Data Generation:** Perform full *ab initio* simulations with CP2K [1], including many-body dispersion corrections (e.g., DFTD4)[2]. Generate high-quality reference datasets for training.
- **Machine Learning Model Development:** Train *Machine Learning Coarse-Graining (MLCG)* [3] models on the *ab initio* data to predict energy, polarizability, and multipole moments. Construct data-driven force fields for simulations.
- **QM/MM Multiscale Simulations:** Use MLCG force fields in QM/MM simulations with chemical potential constraints. Enable molecule exchange via the Grand Canonical Adaptive Resolution Scheme (AdResS) [4].
- **Physical Fidelity and Transferability:** Efficiently infer polarization response and long-range electrostatics. Study water and ionic liquids to assess model reliability, transferability, and physical accuracy.

The goal of this project is to develop and validate a robust multiscale simulation framework that integrates machine-learned force fields within a quantum-classical adaptive resolution scheme. Specifically, we aim to assess the quantum–classical consistency of these force fields in reproducing key physical observables in liquids, focusing initially on water as a test system for machine learning methods and subsequently extending the approach to ionic liquids.

Central to this effort is the Grand Canonical Adaptive Resolution Scheme (AdResS), which provides a rigorous and thermodynamically consistent framework to couple a quantum mechanically described region with a classical molecular reservoir. By enabling particle exchange under controlled chemical potential constraints, AdResS effectively models open systems where molecules can freely diffuse in and out of the quantum region, thus mimicking realistic environments. This allows for accurate simulations of complex phenomena such as solvation, polarization, and charge transfer in heterogeneous media while maintaining computational efficiency.

While ab initio molecular dynamics (AIMD) based on density functional theory (DFT) provides reliable quantum mechanical accuracy, it remains prohibitively expensive for large-scale or long-time simulations [5]. Machine learning (ML) methods, particularly graph neural network architectures such as SchNet combined with the Machine Learning Coarse-Graining (MLCG) framework, offer a promising alternative [3]. These ML models are trained on high-fidelity AIMD data to predict quantum observables—including energies, forces, polarizabilities, and multipole moments—enabling rapid and scalable inference of interatomic potentials and electronic responses at a fraction of the computational cost.

The primary objectives of this project, illustrated in Figure 1, as developing a Grand Canonical QM/MM simulation framework that incorporates state of the art dispersion corrections using DFTD4 for many-body effects, performing thermodynamically consistent simulations of solvated molecules within an open quantum-classical adaptive resolution setting, and training and test MLCG machine learning models to accurately reproduce key quantum mechanical properties such as energy, polarizability, and multipole moments.

The quantum (QM) region, treated with full ab initio accuracy, is embedded within a larger classical (MM) environment. AdResS facilitates molecular exchange and enforces thermodynamic consistency between these regions. Simultaneously, a machine-learned force field derived via MLCG captures quantum mechanical interactions within the classical reservoir, enabling a seamless integration of QM, ML, and MM components. This approach allows simulations to achieve both accuracy and scalability, extending accessible time and length scales

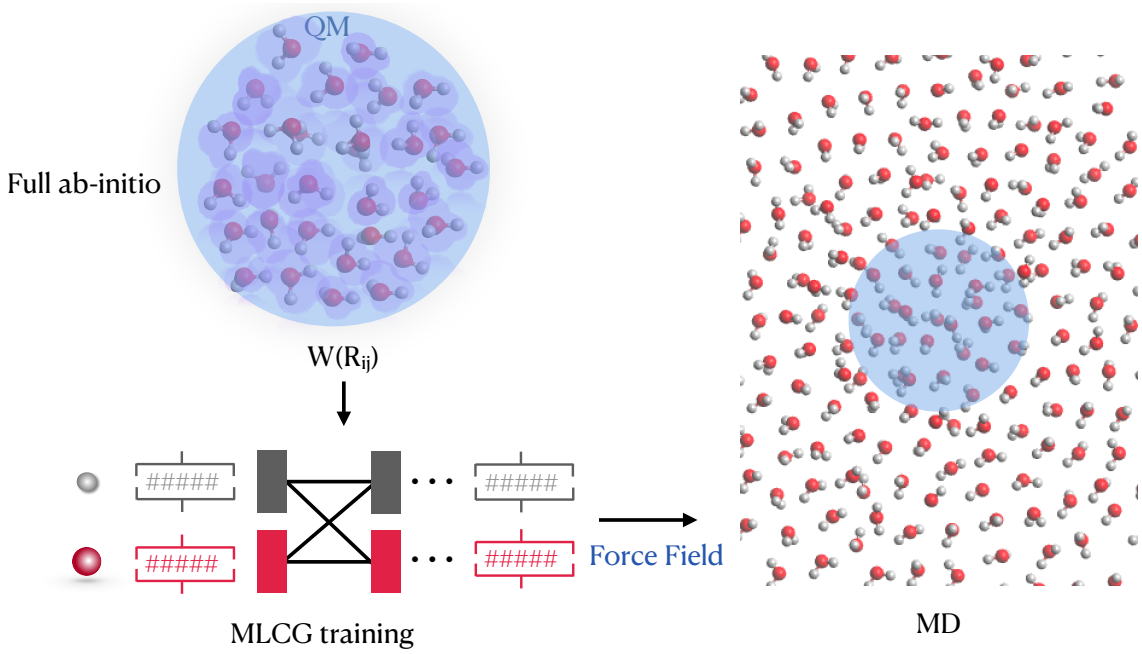

Figure 1: *Schematic overview of the multiscale simulation framework. A machine-learned force field (MLCG) trained on full ab initio data is used to model the classical (MM) region within a QM/MM simulation.*

far beyond conventional AIMD.

This project stands at the intersection of advanced quantum chemistry, cutting-edge machine learning, and adaptive resolution simulation techniques. Its innovative combination of high-accuracy many-body dispersion modeling, open-boundary QM/MM dynamics, and machine-learned force fields aims to establish a validated computational protocol for physically consistent multiscale simulations. We will thoroughly evaluate the performance and domain of applicability of MLCG models, initially focusing on liquid water and subsequently applying these methods to ionic liquids. This includes investigating when machine learning predictions remain physically meaningful and identifying conditions under which they may fail to generalize.

The outcomes will significantly deepen our understanding of solvation, polarization, dielectric response, charge transfer, and microscopic structure in complex molecular liquids. Beyond fundamental insight, these developments have broad implications for materials design, catalysis, and electrochemical applications, where accurate and efficient simulation of open molecular systems is essential.

WWW

https://www.mi.fu-berlin.de/en/math/groups/mms/members/Delle_Site.html

More Information

[1] Thomas D. Kühne, Marcella Iannuzzi, Mauro Del Ben, Vladimir V. Rybkin, Patrick Seewald, Frederick Stein, Teodoro Laino, Rustam Z. Khaliullin, Ole Schütt, Florian Schiffmann, et al., The Journal of Chemical Physics, **152**(19), 194103 (2020).

[2] Eike Caldeweyher, Jan-Michael Mewes, Sebastian Ehlert, and Stefan Grimme, Physical Chemistry Chemical Physics, **22**(16), 8499–8512 (2020).

[3] Maciej Majewski, Adrià Pérez, Philipp Thölke, Stefan Doerr, Nicholas E. Charron, Toni Giorgino, Brooke E. Husic, Cecilia Clementi, Frank Noé, and Gianni De Fabritiis, Nature Communications, **14**(1), 5739 (2023).

[4] Luigi Delle Site, Computer Physics Communications, **222**, 94–101 (2018).

[5] Sara Panahian Jand, Thomas D. Kühne, and Luigi Delle Site, Advanced Theory and Simulations, **7**(11), 2400833 (2024).

Funding

DFG Project Number DE 1140/15-1

DFG Subject Area

Theoretical/ Mathematical/ Computational physics project.