Atomistic simulations with machine learning based interatomic potentials

Characterizing materials at the atomic scale is essential across various fields, from identifying clean energy sources and understanding chemical reactions in catalysis to simulating metallic alloys and studying nanomaterials. This level of analysis provides crucial insights into material performance, particularly under operando conditions, driving advancements in efficiency and broadening the scope of potential applications.
Obtaining an experimental characterization of these systems with atomic resolution is extremely challenging, while *ab initio* atomistic simulations are, in principle, an ideal tool to investigate dynamical phenomena on the atomic scale. Despite their potential, the prohibitive cost of *ab initio* molecular dynamics has limited their application to study realistic systems. Recently, machine learning (ML)-based interatomic potentials have emerged as a valuable solution to reconcile the accuracy of *ab initio* methods with the efficiency of classical force field. These ML potentials are trained to reproduce the energy and forces from a large set of quantum mechanical calculations, and they can be optimized on small system sizes and then used to simulate much larger systems for long (e.g. nanoseconds) timescales. Despite their promise, the construction of these potentials for complex, multicomponent reactive systems represents a complex endeavor, requiring a comprehensive training set that incorporates all relevant configurations. To this end, enhanced sampling methods such as metadynamics, allow to speed up the process of collecting configurations and generating uncorrelated structures which cover the complete energy landscape.

In this PhD course, we aim to present a comprehensive overview of the current state-of-the-art methodologies in the field. We will review various approaches for training machine learning (ML) potentials and collecting the required training datasets, while highlighting the advantages and limitations of each approach, depending on the specific application. Additionally, the course will feature illustrative applications aimed at deepening participants' understanding of structural dynamics, chemical reaction kinetics, and transport phenomena in energy storage systems and in heterogeneous catalysis.

The outline of the classes (2 hours each) will be:

1) *ab initio* molecular dynamics and enhanced sampling: general introduction
2) Machine learning interatomic potentials I: theoretical foundations and implementation
3) Machine learning interatomic potentials II: building the training set
4) Hand-on section 1: construction of the dataset with DFT and/or foundation models
5) Hand-on section 2: case studies with active learning

The hands-on sections will focus on the full training procedure of a ML-potential on a few case studies, using the Quantum Espresso, DeepMD-Kit, MACE and LAMMPS softwares.