

# Multiscale Physics and Machine Learning

Master of Science thesis project

Jan 22, 2018

## Multiscale physics, from quantum mechanical simulations of atoms and molecules to molecular dynamics

What is presented here is a large project which involves several Master theses projects. The various projects are listed below, opening up for several possibilities.

The aim of a program on multiscale physics is to develop a first principle approach to systems of relevance for a variety of fields, from materials science to nano-technology and biological systems and even atomic nuclei and stars. Common to all these systems is that they entail a truly multiscale physics program that involves a proper understanding of the links between the various scales, starting from quantum-mechanical first principle studies of atoms, molecules and eventually other spatially confined systems to Density functional theories and finally microscopically derived potentials to be used in molecular dynamics calculations.

The computations required for accurate modeling and simulation of large-scale systems with atomistic resolution involve a hierarchy of levels of theory: quantum mechanics (QM) to determine the electronic states; force fields to average the electronics states and to obtain atom based forces (FF), molecular dynamics (MD) based on such an FF; mesoscale or coarse grain descriptions that average or homogenize atomic motions; and finally continuum level descriptions. By basing computations on first principles QM it is possible to overcome the lack of experimental data to carry out accurate predictions with atomistic resolution, which would otherwise be impossible. Furthermore, QM provides the fundamental information required to describe quantum effects, electronically excited states, as well as reaction paths and barrier heights involved in chemical reactions processes. However, the practical scale for accurate QM today is <1,000 atoms per molecule or periodic cell (a length scale of a few nanometers) whereas the length scale for modeling supramolecular systems in biology may be in the tens of nano- meters, while elucidating the interfacial effects between grains in composite materials may require hundreds of nanometers, and modeling turbulent fluid flows or shock- induced instabilities in multilayered materials

may require micrometers. Thus, simulations of engineered materials and systems may require millions to billions of atoms, rendering QM methods impractical. Nonetheless, QM methods are essential for accurately describing atomic-level composition, structure and energy states of materials, considering the influence of electronic degrees of freedom. By incorporating time-dependent information, the dynamics of a system under varying conditions may be explored from QM-derived forces, albeit within a limited timescale ( $t < 1$  ps). The prominent challenge for theory and computation involves efficiently bridging, from QM first-principles, into larger length scales with predominantly heterogeneous spatial and density distributions, and longer timescales of simulation – enough to connect into engineering-level design variables – while retaining the appropriate accuracy and certainty. Equally challenging remains the inverse top-down engineering design problem, by which macroscopic material/process properties would be tunable from optimizing its atomic-level composition and structure. The aim of this large project is to develop breakthrough methods to staple and extend hierarchically over existing to develop the necessary tools to enable continuous lateral (multi-paradigm) and hierarchical (multiscale) couplings, between the different theories and models as a function of their length- and timescale range – a strategy often referred to as First-Principles-Based Multiscale-Multiparadigm Simulation.

To achieve these goals, several theses projects are shortly defined below. These projects open also up for several fruitful collaborations between the involved MSc students. The group in Computational Physics has long-standing experience in defining projects where several students may find a common ground, either from a formalism point of view or (and possibly and as well) phenomenological point of view.

There are several possible projects. The systems we have in mind are mainly molecules like  $\text{SiO}_2$ ,  $\text{H}_2\text{O}$ ,  $\text{CaCO}_3$  and other more complicated molecules. To model the interaction between such molecules and eventually derive microscopic interactions can be done via several MSc projects. The following projects can be defined and extended upon:

1. Start with existing codes and thesis work by previous Master of Science students at the computational physics group. Here we think in particular of using existing Hartree-Fock codes for atoms and molecules and use these to develop potential surfaces and thereby extract parametrized potentials for molecular dynamics calculations. One possible path is to start with a gas of hydrogen and compare the parametrized potentials from Hartree-Fock

theory with existing models.

1. Quantum mechanical correlations are however treated only approximatively in Hartree-Fock theory. Monte Carlo methods offer a path to compute exactly ground state properties of correlated quantum mechanical systems. Here one can use existing (developed at the Computational Physics group) or develop a variational and diffusion Monte Carlo code that computes

properties such as binding energies, root mean square radii, charge distributions and local potentials. This topic can easily form the basis for one or two MSc thesis projects. These first principle approaches can be used to develop potential surfaces, which in turn can be used to parametrize potentials for molecular dynamics studies.

2. If more complex atoms and molecules are involved, one may consider freezing the inner degrees of freedom using for example many-body perturbation theory to derive an effective interaction for electrons outside a chosen core. Alternatively one can use existing coupled cluster codes (also part of the computational quantum mechanics codes developed at the computational physics group) to derive potential surfaces and thereby parametrize potentials to use in molecular dynamics calculations.
3. The quantum mechanical calculations can in turn be used to define better density functionals for the above atoms. Density functional theory can in turn be used to study for example lattices of atoms and molecules with say periodic boundary conditions.
4. These effective potentials can in turn be compared with existing models such as the so-called ReaxFF potential. The implementation and studies of this potential for studies of say  $\text{SiO}_2$  compounds could form the basis for one or two MSc thesis projects.
5. Test several Machine Learning algorithms such as Multiperception models and the newly develop [neuroevolution](#) approach.
6. For those interested in a more high-performance computing project, a lively research area is to develop and test the applicability of neural network algorithms on high-performance computing facilities using both GPUs and CPUs.

The milestones are as follows

1. Spring 2018: Finalize relevant courses and start looking at Machine Learning approaches
2. Fall 2018: Work on thesis project
3. Spring 2019: Finalize thesis project.

The thesis is expected to be handed in May/June 2019.

**Recent thesis and relevant literature.** The recent theses of Morten Ledum, John-Anders Stende and Haakon Treider are excellent starting points for further explorations of Machine Learning approaches to Molecular dynamics studies. The recent text on [neuroevolution algorithms](#) may also be a good read.