

From Quantum computing to Machine Learning

Morten Hjorth-Jensen and Anders Malthe-Soerensen^{1,2}

¹Department of Physics, University of Oslo

²Center for Computing in Science Education, University of Oslo

Nov 29, 2017

Strategic importance

This letter states our intent to submit a proposal for **two FISP PhD fellowships** at the Department of Physics of the University of Oslo.

The aim is to lay the foundation for the development of new activities in Computational Science and Physics at the University of Oslo, with a strong focus on computational technologies for the future, from quantum computing to machine learning.

The proposed activities are expected to lead to an application for a future center of excellence (The Norwegian SFF program) in Computational Physics and Science application, where focus on new computational approaches for studying complex systems will play a central role.

Quantum Computing and Machine Learning are two of the most promising approaches for studying complex physical systems where several length and energy scales are involved. Traditional many-particle methods, either quantum mechanical or classical ones, face huge dimensionality problems when applied to studies of systems with many interacting particles. To be able to define properly effective potentials for realistic molecular dynamics simulations of billions or more particles, requires both precise quantum mechanical studies as well as algorithms that allow for parametrizations and simplifications of quantum mechanical results. Quantum Computing offers now an interesting avenue, together with traditional algorithms, for studying complex quantum mechanical systems. Machine Learning on the other hand allows us to parametrize these results in terms of classical interactions. These interactions are in turn suitable for large scale molecular dynamics simulations of complicated systems spanning from subatomic physics to materials science and life science.

In addition, **Machine Learning** plays nowadays a central role in the analysis of large data sets in order to extract information about complicated correlations. This information is often difficult to obtain with traditional methods. For

example, there are about one trillion web pages; more than one hour of video is uploaded to YouTube every second, amounting to 10 years of content every day; the genomes of 1000s of people, each of which has a length of 3.8×10^9 base pairs, have been sequenced by various labs and so on. This deluge of data calls for automated methods of data analysis, which is exactly what machine learning provides. Developing activities in these frontier computational technologies is thus of strategic importance for our capability to address future science problems.

The problems we target satisfy the dual criteria of being integral to the fundamental understanding of complex physical systems and also compelling a major conceptual advance in method or theory with broad applications to other science fields.

Synergies and impact. Based on the above arguments, the long-term perspective for this application is to lay the foundation for a new center of excellence in research, where Computational Science and Physics play a central role. Since the problems to be addressed are complex multiscale scientific problems, the center will, by the very nature of the problems, be multidisciplinary.

Furthermore, this activity will have strong educational outcomes and is expected to be tightly linked with the newly established center of excellence in education, the **Center for Computing in Science Education (CCSE)**. This center aims at developing a novel research program on assessment methods in science educations, with large implications for the way research in science education is conducted. Machine Learning will play a central role there as well when analyzing large data sets, providing us thereby with a truly quantitative approach to science education. We expect thus that the two PhD fellows involved with this project will benefit from cross-fertilizations and synergies with the activity at the CCSE and viceversa.

The present proposal will also lay the foundation, in close collaboration with CCSE and the Computational Physics activity at the Department of Physics, for the establishment of an educational Marie-Curie training program on Computational Science and Physics, with a strong focus on the above research topics. Furthermore, together with CCSE, one of the aims of this graduate training program is to educate the next generation of university teachers and science educators (educating the next generation of school teachers), providing them with the computational competences and skills adequate for the scientific and educational challenges of the 21st century.

Finally, we expect that this research will attract new talents as well as create several research synergies at the Department of Physics and the College of Natural Science (MN-fakultetet) at the University of Oslo. We foresee that this project will enable several research collaborations in fields like Materials Science, Condensed Matter Physics, Subatomic Physics and Life Science. Through the newly established Master of Science program **Computational Science**, we expect to recruit the best talents.

Background and previous experience

This research is strongly rooted in the activities of the award winning Computational Physics group at the Department of Physics. Since 2003, close to 70 students have finalized their Master of Science theses, and approximately half of these students have continued with PhD studies. These highly qualified candidates testify to a strong and robust educational environment. We have presently several excellent PhD candidates who are finalizing or will finalize their Master of Science theses. Many of these candidates have applied and developed Machine Learning methods to solve complex and interacting many-particle systems. These activities have a strong overlap with the research programs of the two principal investigators (Hjorth-Jensen and Malthe-Sørensen). We foresee also strong links with the CCSE.

The two PIs have a long-standing experience in developing and applying algorithms for quantum mechanical and molecular dynamics simulations of many interacting particles, with applications spanning from dense subatomic matter to the physics of complex materials.

The proposal will also benefit from extensive collaborations with leading Universities and Laboratories in Northern America (Oak Ridge National Laboratory, Michigan State University of University of Washington).

Research Program

The research program will focus on two main topics:

- **Quantum Computing** for studies of quantum mechanical systems with many particles (PhD 1). The focus is on
 - Subatomic matter
 - Atomic and molecular physics, with applications to Materials Science and Life Science
- **Machine Learning** focusing on the training of potentials for Molecular dynamics studies of large numbers of interacting particles (PhD 2), with applications to systems in
 - Condensed Matter Physics,
 - Materials Science and
 - Life Science

Another possible project which links the two main topics is to use Machine Learning algorithms to study quantum mechanical systems. Here one can think of using restricted Boltzmann machines and supervised learning to improve and optimize correlations in classical quantum mechanical many-body methods. A further approach is to use supervised learning to optimize density functionals. Density Functional theory is widely used to describe complicated systems in Condensed Matter Physics, Materials Science and Life Science.

The two projects offer several synergies. The most likely one will use results from quantum mechanical simulations (both for nuclear systems and many-electron systems) as training data for a Machine Learning approach to molecular dynamics simulations. The latter require the development of reliable and complicated two- and three-body potentials. These potentials are in turn used in classical simulations of large systems, typically with billions of interacting constituents, providing thereby a proper first principle approach to multiscale science problem.

Finally, as mentioned above, a Machine Learning approach to Science Education, as represented by the CCSE, offers a link for further synergies and testbeds of algorithms from Machine Learning.

Why Multiscale Science? 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. Such a program involves insights and collaborations across disciplines in order to foster progress.

The computations required for accurate modeling and simulation of large-scale systems with microscopic 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 reactions processes. However, in for example Materials Science, 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 of many-particle systems. 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. The ultimate goal is a reversible bottom-up, top-down approach, based on first principles QM, to characterize properties of complex many-particle systems and processes at a hierarchy of length and timescales. This will improve our ability to design, analyze, and interpret experimental results, perform model-based prediction of phenomena, and to control precisely the multi-scale nature of many-particle systems for multiple applications.

Our approach to many of the above problems involves the development of new algorithms and approaches based on quantum computing and machine learning.

From Quantum Computing to Machine Learning. Enabling simulations of large-scale many-body systems is a long-standing problem in scientific computing. Quantum many-body interactions define the structure of the universe, from nucleons and nuclei, to atoms, molecules, and even stars. Since the discovery of quantum mechanics, a lot of progress has been made in understanding the dynamics of certain many-body systems. While some of our insight comes from a small set of analytically solvable models, numerical simulations have become a mainstay in our understanding of many-body dynamics. The progress in numerical simulations has accelerated in the last few decades with the advent of modern high performance computing (HPC) and clever developments in classical simulation algorithms such as, quantum Monte Carlo, large-scale diagonalization approaches, Coupled-Cluster theory and other renormalization schemes. Despite the monumental advances, classical simulation techniques are reaching fundamental limits in terms of the size of the quantum systems that can be processed. Fortunately, the disruptive new field of quantum simulations has emerged, promising to enable simulations far beyond those which are classically tractable. In particular, scientific applications concerned with simulations of interacting fermions on a lattice are poised to reap the benefits of quantum simulations. Mathematical models of interacting fermions naturally extend to describe vastly different physics such as that of correlated electronic and the correlated nuclear systems.

Recent progress in quantum computing as well as digital and analog Quantum Algorithms (QAs) promise to enable the exciting possibility of performing simulations that are beyond the reach of all existing and future classical supercomputers. Despite the progress, there is still a gap between the resources required by state-of-the-art QA and the resources offered by available and near-future quantum hardware. It may take decades of quantum hardware development and engineering before the current QAs will outperform classical exascale class simulations. Therefore, to impact scientific computing on a more relevant time scale, improving the scalability and efficiency of quantum simulation algorithms is of the highest importance. This proposal is aimed at:

- bridging the gap between QAs and hardware by developing a novel algorithmic approach to quantum simulations and

- developing a software toolchain that enables utilization of the proposed QA as a part of the existing HPC simulation codes.

Our QA combines individual algorithmic strengths of analog and digital QA and is designed to operate on a combination of quantum digital and analog hardware. This approach reduces hardware requirements, making quantum simulations of interacting fermions possible with the existing and near-term quantum devices.

The main focus here will be on extending classical many-body methods like large-scale diagonalization approaches and Coupled-Cluster theory using Quantum Computing algorithms, with applications to correlated electron materials and nuclear properties. Correlated electron materials exhibit a myriad of interesting phenomena that arise from the strong electron-electron Coulomb interaction. In particular, electron interactions are responsible for magnetism, superconductivity and quantum critical behavior. To understand these phenomena, the electrons cannot be treated as non-interacting particles in a mean field. One has to explicitly consider the many-body dynamics arising from the interactions between the electrons.

Nuclear properties that can be directly observed include ground-state binding energies, nuclear spins, parity, excited state energies, spins and parity, and transitions branching ratios between excited states, and various decay modes. Experimental information comes from increasingly sophisticated experiments which continue to push into regions previously unexplored. Because of their connection to nucleosynthesis, the current experimental emphasis focuses on weakly bound, neutron rich nuclei. Planned and new experimental nuclear physics facilities will provide tremendous new capabilities for exploring these neutron rich nuclei. The experimental progress in probing visible matter at its extremes will pose several challenges to theory, amongst these the capability to study systematically systems with exponentially many degrees freedom. This applies to electronic systems as well. Exploring new algorithms and possibilities is thus essential if we wish to make progress in our understanding of quantum mechanical systems.

Algorithms from Quantum Computing represents thus one possible path to overcome the dimensional curse of traditional manybody methods. Another alternative is to use Machine Learning algorithms to study quantum mechanical systems. Here one can think of using restricted Boltzmann machines and supervised learning to improve and optimize correlations in classical quantum mechanical many-body methods. A further approach is to use supervised learning to optimize density functionals. We have already started such studies, with rather successful results.

The results from the quantum mechanical simulations (both for nuclear systems and many-electron systems) will in turn be used as training data for a Machine Learning approach to molecular dynamics simulations. The latter require the development of complicate two- and three-body potentials. These potentials are in turn used in classical simulations of large systems, typically with billions of interacting constituents. If successful, such an approach will provide a true first principle path to multiscale science problem. We have already

developed some local expertise in optimizing potential surfaces for molecular dynamics simulations using artificial neural networks. We expect several exciting new results in the future and a proper financing of this proposal can lead to several novel developments in multiscale physics.