Master thesis project: Solving Quantum Mechanical Problems with Machine Learning

Overview

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 parameterize 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.

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 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, new developments in the fields of quantum simulations and machine learning have emerged, promising to

enable simulations far beyond those which are classically tractable.

The approaches to machine learning are many, but are often split into two main categories. In *supervised learning* we know the answer to a problem, and let the computer deduce the logic behind it. On the other hand, *unsupervised learning* is a method for finding patterns and relationship in data sets without any prior knowledge of the system. Some authours also operate with a third category, namely *reinforcement learning*. This is a paradigm of learning inspired by behavioural psychology, where learning is achieved by trial-and-error, solely from rewards and punishment. In this thesis, the aim is to explore new developments in the field of machine learning, with an emphasis on unsupervided learning. Much of the work here and its implementations is motivated by the recent article of Carleo and Troyer in Science, 2017. In particular we have extended their work, which focused on spin-like quantum mechanical systems, to systems of interacting bosons and fermions confined to move in trapping potentials, with the harmonic oscillator as one of the foremost examples. In this work, we will start with quantum Monte Carlo methods, with an emphasis on Variational Monte Carlo methods. This approach to studies of complicated interacting many-particle systems has been widely used in almost all fields of physics where first principle calculations are employed. It provides an important starting point for almost exact solutions to Schrödinger's equation for many interacting particles using so-called Green's function Monte Carlo methods.

A variational Monte Carlo (VMC) calculation is based on an ansatz for say the ground state wave function. For fermionic systems this ansatz is often composed of a single-particle part (via a so-called Slater determinant which accounts for the anti-symmetry) and a correlated part, normally called the Jastrow factor. For bosonic systems there may also be a product function of single-particle functions and a Jastrow factor that aims at incorporating correlations beyond a mean field. These trial wave functions are thereafter used in an optimization procedure where various variational parameters are optmized in order to find a minimum for expectation values like the energy and the variance.

Constructing both the single-particle part and the Jastrow part can often be complicated and tedious. In systems like interacting nucleons, the correlation part of the wave function contains often complicated two- and three-body operators that require dedicated code developments. Similalry, for systems of atoms and molecules (and nucleons as well), the single-particle part is often constructed using mean-field methods like Hartree-Fock theory.

The aim here is to see whether methods inspired from Machine Learning can do an equally good job as the standard approach to VMC calculations, this time however with trial wave functions determined by neural networks. These trial wave functions are based on what in the literature is called Boltzmann machines. These functions contain several parameters which are used to find an energy minimum and thereby the optimal solution for the energy.

A typical machine learning algorithm consists of three basic ingredients, a dataset $\mathbf{x}$ (could be some observable quantity of the system we are studying), a model which is a function of a set of parameters α that relates to the dataset, say a likelihood function $p(\mathbf{x}|\alpha)$ or just a simple model $f(\alpha)$, and finally a so-

called *cost* function $\mathcal{C}(\mathbf{x}, f(\alpha))$ which allows us to decide how well our model represents the dataset.

We seek to minimize the function $\mathcal{C}(\mathbf{x}, f(\alpha))$ by finding the parameter values which minimize $\mathcal{C}$. Thus, VMC calculations serve both as input to exact solutions via Green's functions methods and employ similar optimization approaches as employed in machine learning.

Progress plan and milestones

The aims and progress plan of this thesis are as follows

- ☑ Set up a VMC code with a RBM-wavefunction for two electrons trapped in a harmonic oscillator
- ☑ Expand to systems of more electrons by implementing a Slater Determinant which takes the anti-symmetry
- ☐ Optimize the basis using Hartree-Fock
- ☐ Compare the optimized-RBM wavefunction to an optimized wavefunction obtained from standard VMC for harmonic oscillators. Onebody- and twobody densities are central quantities. For this we plan to make use of existing VMC code, for instance the implementation of Morten Ledum or Jørgen Høgberget.
- ☐ Move to atomic and molecular systems where Slater Type Orbitals (STOs) are used as the basis set expanded in Gaussian Type Orbitals (GTOs). This part will be based on Henrik Eiding's thesis work, *Ab Initio Studies of Molecules*. Repeat critical evaluation of Boltzmann machine.
- ☒ Investigate the electron gas and the helium gas
- ☒ Try to implement a variance optimization as done in VMC studies by Filippi and Umrigar (2005), see discussion in slides.
- ☒ Compare the results obtained with Boltzmann machines with so-called Shadow wave function approaches, see `https://arxiv.org/abs/1404.6944` and references therein.

The ✓means that the milestone is already reached. ✗means that we probably do not have time for it, but if we do it will be considered as well.

The thesis is expected to be handed in August/September 2019.

Details about work so far

A VMC code with RBM-wavefunction was developed already in the course FYS4411, and after the summer I spent some time optimizing the code and make it more general. Thereafter, a Slater determinant was implemented taking the

first exponential part of the wavefunction in addition to Hermite polynomials. The Hermite polynomials were added based on the exact SPFs in a harmonic oscillator, but some even more general polynomials would might be interesting to try as well?

The initial RBM-wavefunction reads

$$\Psi_T(x_1,\dots,x_M;\boldsymbol{a},\boldsymbol{b},\boldsymbol{W}) = \exp\left(-\sum_{i=1}^{M}\frac{(x_i-a_i)^2}{2\sigma^2}\right)\prod_{j=1}^{N}\left(1+\exp\left(b_j+\sum_{i=1}^{M}\frac{W_{ij}X_i}{\sigma^2}\right)\right) \tag{1}$$

Inspired by the Hermite functions, we use

$$\psi_n(x;a) = H_n(x)\exp\left(-\frac{(x-a)^2}{2\sigma^2}\right) \tag{2}$$

as our basis set and treat the product as a Jastrow factor. Since all terms in the Slater determinant contain the factor

$$\exp\left(-\sum_{i=1}^{M}\frac{(x_i-a_i)^2}{2\sigma^2}\right), \tag{3}$$

this can be factorized out resulting a Slater determinant consisting of Hermite polynomials only. For an electron system, the Slater determinant is known to be separable in a spin-up part and a spin-down part, saving us from a lot of computation.

For inspiration I've read some existing thesises, especially Alocias' and Vilde Flugsrud's thesises for more information about the RBM-wavefunction, Jørgen Høgberget's thesis for details about VMC and Slater determinant implementation and lately Henrik Eiding's thesis about molecular systems. I was unsure which systems to investigate, and read about the electron gas and the helium gas before I decided to go for atomic and molecular systems.

Apart from this, I have spent a lot of time working with the courses FYS4480-*Quantum mechanics for many-particle systems* and FYS4155-*Applied data analysis and machine learning* this semester, and I'm sure they both will be useful when writing my thesis.

Details about further work

Now as I've decided the further path, my motivation is on top. The first thing I will do, is to optimize the basis set using Hartree-Fock. After that, some more calculations on the harmonic oscillator is needed before we turn to molecular systems. Comparing results from the RBM-wavefunction to results obtained using standard VMC will be one of the main focuses.

The plan is to study how we can use RBM-wavefunctions in computations of the ionization energy of atoms and molecules. The way we will do it, is to create a wavefunction similar to the Gaussian Type Orbitals (GTOs) using a RBM, and then expand a Slater Type Orbital (STO) basis in the GTO basis. The

idea is that it might give a more flexible wavefunction where the same intuition as before is not needed. The results will be compared to the results obtained by Henrik Eiding.

Some results