Solving Many-body Quantum Problem using Machine Learning

by

Even M. Nordhagen

THESIS

for the degree of

MASTER OF SCIENCE

Faculty of Mathematics and Natural Sciences
University of Oslo

September 19, 2018

Abstract

Abstract should be written here

Acknowledgements

Acknowledgements should be written here

-One of my mottos used to be that everything has a reason, which I used to point out whenever people were talking about things that apparently could not be described immediately. It was first when I learned about quantum mechanics that I understood I was wrong, and that is one of the reasons why quantum mechanics catched me so hardly.

Contents

List of abbreviations

Letters		Meaning
RBM	-	Restricted Boltzmann Machine
MC	-	Monte Carlo
VMC	-	Variational Monte Carlo
DMC	-	Diffusion Monte Carlo
ML	-	Machine Learning
WF	-	Wave Function
SPF	-	Single Particle Function

Table 1: List of symbols used with explaination.

Source Code

The source code is given in `https://github.com/evenmn`

1 Introduction

Write introduction here
    - Introduce the wavefunction - Mention the uncertainity principle and also quantum entanglement to catch the readers interest

2 Background theory

Here I might present basic quantum mechanics briefly.

2.1 Hamiltonian and wavefunction

Quantum quantities can be found by solving eigenvalue equations with the wavefunction as eigenfunction.

$$\hat{\mathrm{H}}\Psi_n(\boldsymbol{x}) = \epsilon_n \Psi_n(\boldsymbol{x}) \tag{1}$$

2.2 Born-Oppenheimer approximation

The Born-Oppenheimer approximation is the assumption that we can split the Hamiltonian in a one-body part and a two-body part, and calculate the energies separately.

$$\hat{\mathrm{H}} = \sum_{i=1}^{P}(-\frac{1}{2}\nabla_i^2 + \frac{1}{2}\omega^2 r_i^2) + \sum_{i<j}\frac{1}{r_{ij}} \tag{2}$$

In atomic units.

2.3 Wavefunction properties

Assume we have a permutation operator $\hat{P}$ which switches two coordinates in the wave function,

$$\hat{P}\Psi_n(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_i, \ldots, \boldsymbol{x}_j, \ldots, \boldsymbol{x}_M) = p\Psi_n(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_j, \ldots, \boldsymbol{x}_i, \ldots, \boldsymbol{x}_M), \tag{3}$$

where p is just a factor which comes from the transformation. If we again apply the $\hat{P}$ operator, we should switch the same coordinates back, and we expect to end up with the initial wave function. For that reason, $p = \pm 1$. [1]

---

[1]This was true until 1976, when J.M. Leinaas and J. Myrheim discovered the anyon, https://www.uio.no/studier/emner/matnat/fys/FYS4130/v14/documents/kompendium.pdf.

The particles that have an antisymmetric (AS) wavefunction under exchange of two coordinates are called fermions, named after Enrico Fermi, and have half integer spin. On the other hand, the particles that have a symmetric (S) wavefunction under exchange of two coordinates are called bosons, named after Satyendra Nath Bose, and have integer spin.

It turns out that because of their antisymmetric wavefunction, two identical fermions cannot be found at the same position at the same time, known as the Pauli principle. This causes some difficulties when dealing with multiple fermions, because we always need to ensure that the total wavefunction becomes zero if two identical particles happen to be at the same position. To do this, we introduce a Slater determinant as described in the next chapter. in this particular project, we are going to focus on fermions, just because they are more difficult to handle, but much of the theory and implementation applies for bosons as well.

Read https://manybodyphysics.github.io/FYS4480/doc/pub/secondquant/html/secondquant bs.html

2.4   Slater determinant

For a system of more particles we can define a total wavefunction, which is a composition of all the single particle wavefuncions (SPF) and contains all the information about the system. The way we compile the SPFs needs to be based on Pauli's exclusion, which states that two identical fermions cannot possibly be in the same state at the same time. If this happens, we set the total wavefunction to zero, which is done by defining the wavefunction as a determinant.

Consider a system of two identical fermions with SPFs ϕ_1 and ϕ_2 at positions $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$ respectively. The way we define the wavefunction of the system is then

$$\Psi_T = \begin{vmatrix} \phi_1(\boldsymbol{r}_1) & \phi_2(\boldsymbol{r}_1) \\ \phi_1(\boldsymbol{r}_2) & \phi_2(\boldsymbol{r}_2) \end{vmatrix} = \phi_1(\boldsymbol{r}_1)\phi_2(\boldsymbol{r}_2) - \phi_2(\boldsymbol{r}_1)\phi_1(\boldsymbol{r}_2), \tag{4}$$

which is set to zero if the particles are at the same position. This is called a Slater determinant, and yields the same no matter how big the system is.

Notice that we denote the wavefunction with the $'T'$, which indicates that it is a trial wavefunction. We do this because the spin part is avoided with ψ as the radial parts only, thus this wavefunction is not the true wavefunction. We will look closer at how we can factorize out the spin part later. The spin part is assumed to not affect the energies. For bosons the total wavefunction is defined similarly, but with no negative signs since the Pauli principle does not apply for fermions.

A general Slater determinant for a system of N particles takes the form

$$\Psi(\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_N) = \begin{vmatrix} \psi_1(\boldsymbol{r}_1) & \psi_2(\boldsymbol{r}_1) & \ldots & \psi_N(\boldsymbol{r}_1) \\ \psi_1(\boldsymbol{r}_2) & \psi_2(\boldsymbol{r}_2) & \ldots & \psi_N(\boldsymbol{r}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_1(\boldsymbol{r}_N) & \psi_2(\boldsymbol{r}_N) & \ldots & \psi_N(\boldsymbol{r}_N) \end{vmatrix} \tag{5}$$

where the ψ's are the true single particle wavefunctions, which are the tensor products

$$\psi = \phi \otimes \xi \tag{6}$$

with ξ as the spin part.

2.4.1 Electron system

For our purpose we will study fermions with spin $\sigma = \pm 1/2$ only, which can be seen as an electron gas. In this particular case, the SPFs can be arranged in spin-up and spin-down parts, such that the Slater determinant can be simplied to

$$\Psi(\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_N) = \begin{vmatrix} \phi_1(\boldsymbol{r}_1)\xi_\uparrow & \phi_1(\boldsymbol{r}_1)\xi_\downarrow & \ldots & \phi_{N/2}(\boldsymbol{r}_1)\xi_\downarrow \\ \phi_1(\boldsymbol{r}_2)\xi_\uparrow & \phi_1(\boldsymbol{r}_2)\xi_\downarrow & \ldots & \phi_{N/2}(\boldsymbol{r}_2)\xi_\downarrow \\ \vdots & \vdots & \ddots & \vdots \\ \phi_1(\boldsymbol{r}_N)\xi_\uparrow & \phi_1(\boldsymbol{r}_N)\xi_\downarrow & \ldots & \phi_{N/2}(\boldsymbol{r}_N)\xi_\downarrow \end{vmatrix}. \tag{7}$$

This is called the wavefunction ansatz, because assumptions are raised, like two particles with opposite spins are found to be at the same position all the time, i.e, an equal number of fermions have spin up as spin down, are applied. Further the...

SHOULD END UP WITH SPLITTED DETERMINANTS HERE

For a detailed walkthrough, see appendix I in REF(The Stochastic Gradient Approximation: an application to Li nanoclusters, Daniel Nissenbaum).

2.5 Jastrow factor

2.5.1 CUSP condition

3 What I have done

As mentioned before, our wavefunction is inspired by machine learning in the sense that ... Neural networks ... Boltzmann machine ...

JOINT PROBABILITY

For a quantum harmonic oscillator system, we know that the hermite functions are the true eigenfunctions. In two dimensions, they go as

$$\phi_{n_{x1},n_{x2}}(x_1, x_2) = H_{n_{x1}}(x_1) H_{n_{x2}}(x_2) \exp\left(-\frac{x_1^2 + x_2^2}{2}\right) \qquad (8)$$

where $H_n(x)$ is the hermite polynomial of n'th order. We are going heavily inspired by this when setting up our SPFs, but in spirit of machine learning, we need to add variational variables:

$$\phi_{n_{x1},n_{x2}}(x_1, x_2, a_1, a_2) = H_{n_{x1}}(x_1 - a_1) H_{n_{x2}}(x_2 - a_2) \exp\left(-\frac{(x_1 - a_1)^2 + (x_2 - a_2)^2}{2\sigma^2}\right)$$
$$(9)$$

Notice that we subtract a variable a_i from every position x_i, and that we also have added another variable σ. Since x_i always is related to a_i, we introduce shorthand notation $x_i - a_i \equiv xa_i$. We then set up a slater determinant, and as described in section 2.4 we can split it up in a spin-up part and a spin-down part:

$$D^{\uparrow} = \begin{vmatrix} \phi_{0,0}(xa_1, xa_2) & \phi_{0,0}(xa_3, xa_4) & \phi_{0,0}(xa_5, xa_6) & \dots \\ \phi_{1,0}(xa_1, xa_2) & \phi_{1,0}(xa_3, xa_4) & \phi_{1,0}(xa_5, xa_6) & \dots \\ \phi_{0,1}(xa_1, xa_2) & \phi_{0,1}(xa_3, xa_4) & \phi_{0,1}(xa_5, xa_6) & \dots \\ \vdots & \vdots & \vdots & \ddots \end{vmatrix} \qquad (10)$$

SIMILARLY FOR $D^{\downarrow}$. When inserting the SPFs, one can observe that the exponential part can be factorized out, and we and up with determinants that consist of Hermite polynomials only:

$$\Psi_T(\vec{x}, \vec{a}, \vec{b}, \hat{W}, \sigma) = \frac{1}{Z} \det(\hat{D}^{\uparrow})' \det(\hat{D}^{\downarrow})' e^{\sum_i^M \frac{(xa_i)^2}{2\sigma^2}} \prod_j^N (1 + e^{b_j + \sum_i M \frac{X_i W_{ij}}{\sigma^2}})$$
$$(11)$$

With

$$\det(\hat{D}^{\uparrow})' = \begin{vmatrix} H_0(xa_1)H_0(xa_2) & H_0(xa_3)H_0(xa_4) & H_0(xa_5)H_0(xa_6) & \dots \\ H_1(xa_1)H_0(xa_2) & H_1(xa_3)H_0(xa_4) & H_1(xa_5)H_0(xa_6) & \dots \\ H_0(xa_1)H_1(xa_2) & H_0(xa_3)H_1(xa_4) & H_0(xa_5)H_1(xa_6) & \dots \\ \vdots & \vdots & \vdots & \ddots \end{vmatrix}$$
$$(12)$$

3.0.1   6 fermion example

The theory above might be a bit abstract, so let's turn to an example: 6 electrons in two dimensions.

$$\det(\hat{D}^\uparrow)' = \begin{vmatrix} H_0(xa_1)H_0(xa_2) & H_0(xa_3)H_0(xa_4) & H_0(xa_5)H_0(xa_6) \\ H_1(xa_1)H_0(xa_2) & H_1(xa_3)H_0(xa_4) & H_1(xa_5)H_0(xa_6) \\ H_0(xa_1)H_1(xa_2) & H_0(xa_3)H_1(xa_4) & H_0(xa_5)H_1(xa_6) \end{vmatrix} \quad (13)$$

$$\det(\hat{D}^\downarrow)' = \begin{vmatrix} H_0(xa_7)H_0(xa_8) & H_0(xa_9)H_0(xa_{10}) & H_0(xa_{11})H_0(xa_{12}) \\ H_1(xa_7)H_0(xa_8) & H_1(xa_9)H_0(xa_{10}) & H_1(xa_{11})H_0(xa_{12}) \\ H_0(xa_7)H_1(xa_8) & H_0(xa_9)H_1(xa_{10}) & H_0(xa_{11})H_1(xa_{12}) \end{vmatrix} \quad (14)$$

Observe that we have particles in the two lowest energy levels only =¿ Hermite polynomials of 1'st and 2'nd order:

$$H_0(x) = 1$$
$$H_1(x) = 2x$$

which gives

$$\det(\hat{D}^\uparrow)' = \begin{vmatrix} 1 & 1 & 1 \\ 2xa_1 & 2xa_3 & 2xa_5 \\ 2xa_2 & 2xa_4 & 2xa_6 \end{vmatrix} \quad (15)$$

$$\det(\hat{D}^\downarrow)' = \begin{vmatrix} 1 & 1 & 1 \\ 2xa_7 & 2xa_8 & 2xa_9 \\ 2xa_{10} & 2xa_{11} & 2xa_{12} \end{vmatrix} \quad (16)$$