

BOLTZIEMANMACHINE YO!

Johan Nereng

Department of Physics, University of Oslo, Norway

Spring, 2020

Abstract

her kjem det greier

Author's comments: Lalalla

1 Introduction

method and the "automated blocking" algorithm from [?].

In order to write this project paper and the code required to produce the results, I used a variety of tools, including: C++, Python 3.7.7, NumPy [?], as well as a number of books, web-pages and articles - of which most are listed under [references](#). All the code required to reproduce the results may be found on my [github page](#) .

2 Material and methods

2.1 System: Electrons in 2D isotropic HO

The system consists of P electrons in a D dimensional isotropic harmonic oscillator (HO) potential, with the following idealized total Hamiltonian, when using natural units, ($\hbar = c = e = m_e = 1$), and energies in atomic units a.u:

$$H = \sum_{i=1}^P \left(-\frac{1}{2} \nabla_i^2 + \frac{1}{2} \omega^2 r_i^2 \right) + \sum_{i=1}^P \sum_{j=1}^i \frac{1}{r_{ij}}, \quad (1)$$

Where ω is the oscillator frequency ω , $r_{ij} = |\mathbf{r}_1 - \mathbf{r}_2|$, and $r_i = \sqrt{r_{ix}^2 + r_{iy}^2}$. Furthermore,

$$\hat{H}_0 = \sum_{i=1}^P \left(-\frac{1}{2} \nabla_i^2 + \frac{1}{2} \omega^2 r_i^2 \right),$$

is the standard HO part of the Hamiltonian, while

$$\hat{H}_1 = \sum_{i < j} \frac{1}{r_{ij}},$$

is the interactive part, where $r_{ij} = |\mathbf{r}_1 - \mathbf{r}_2|$, and $r_i = \sqrt{r_{ix}^2 + r_{iy}^2}$.

2.2 Boltzmann

2.2.1 Wave Function

Using $M = DP$ visible nodes, and N hidden nodes;

$$\begin{aligned} \Psi(\mathbf{X}) &= F_{rbm}(\mathbf{X}) \\ &= \frac{1}{Z} \sum_{\{h_j\}} e^{-E(\mathbf{X}, \mathbf{h})} \\ &= \frac{1}{Z} \sum_{\{h_j\}} e^{-\sum_i^M \frac{(X_i - a_i)^2}{2\sigma^2} + \sum_j^N b_j h_j + \sum_{i,j}^{M,N} \frac{X_i w_{ij} h_j}{\sigma^2}} \\ &= \frac{1}{Z} e^{-\sum_i^M \frac{(X_i - a_i)^2}{2\sigma^2}} \prod_j^N (1 + e^{b_j + \sum_i^M \frac{X_i w_{ij}}{\sigma^2}}) \end{aligned}$$

Så det går litt kjappere:

$$v(\mathbf{X}, j) = b_j + \sum_{i=1}^M \frac{X_i W_{ij}}{\sigma^2} \quad (2)$$

Så:

$$\Psi(\mathbf{X}) = \frac{1}{Z} e^{-\sum_i^M \frac{(X_i - a_i)^2}{2\sigma^2}} \prod_j^N (1 + e^{v(\mathbf{X}, j)}) \quad (3)$$

2.2.2 Local Energy

When using the Monte Carlo methods described later, the particles will be moved one at a time. After each such move, the local energy (4) will be evaluated, and used to calculate the ground energy of the system.

$$E_L(\mathbf{r}) = \frac{1}{\Psi_T(\mathbf{r})} H \Psi_T(\mathbf{r}). \quad (4)$$

As the algorithms will involve repeatedly calculating the local energy, every reduction of floating point operations (FLOPs) involved in doing so will lead to a significant speed up - which is why an analytic expression for the local energy is desirable. Using the trial wave function (3), and (4) (see Appendix 1: Analytic expression for the local energy):

Non-interacting, using $H = H_0$

$$\begin{aligned} E_L(\mathbf{r}) &= \frac{1}{\Psi_T(\mathbf{r})} \sum_{i=1}^M \left(-\frac{1}{2} \nabla_i^2 + \frac{1}{2} \omega^2 r_i^2 \right) \Psi_T(\mathbf{r}) + \frac{1}{\Psi_T(\mathbf{r})} \sum_{i < j} \frac{1}{r_{ij}} \Psi_T(\mathbf{r}) \\ &= \frac{1}{\Psi_T(\mathbf{r})} \sum_{i=1}^M \left(-\frac{1}{2} \nabla_i^2 \Psi_T(\mathbf{r}) \right) + \sum_{i=1}^M \frac{1}{2} \omega^2 r_i^2 + \sum_{i < j} \frac{1}{r_{ij}} \end{aligned} \quad (5)$$

Since

$$\frac{1}{\Psi_T} \nabla^2 \Psi_T = (\nabla \ln \Psi_T)^2 + \nabla^2 \ln \Psi_T \quad (6)$$

E_L equals

$$E_L(\mathbf{r}) = -\frac{1}{2} \sum_{k=1}^M (\nabla \ln \Psi_T)^2 + \nabla^2 \ln \Psi_T + \sum_{k=1}^M \frac{1}{2} \omega^2 r_k^2 + \sum_{i < j} \frac{1}{r_{ij}} \quad (7)$$

Where

$$\nabla \ln \Psi_T = -\frac{(X_k - a_k)}{\sigma^2} + \sum_j^N \frac{w_{kj}}{\sigma^2} \text{logistic}(-v(j)) \quad (8)$$

and

$$\nabla^2 \ln \Psi_T = -\frac{1}{\sigma^2} + \sum_j^N \frac{w_{kj}^2}{\sigma^4} \text{logistic}^2(-v(j)) \exp(-v(j)) \quad (9)$$

Or:

$$E_L(\mathbf{r}) = -\frac{1}{2} \sum_{k=1}^M (\nabla \ln \Psi_T)^2 + \nabla^2 \ln \Psi_T + \sum_{k=1}^M \frac{1}{2} \omega^2 r_k^2 + \sum_{i < j} \frac{1}{r_{ij}} \quad (10)$$

Interacting, using $H = H_0 + H_1$

As will be discussed more under importance sampling, the drift force of the particles are also needed;

$$F_i = \frac{2\nabla\Psi_T}{\Psi_T} = 2 \left[-\frac{(X_k - a_k)}{\sigma^2} + \sum_i^N \frac{w_{kj}}{\sigma^2} \frac{1}{1 + e^{-v(\mathbf{X},j)}} \right] \quad (11)$$

2.2.3 Force

Defined as

$$F_i = \frac{2\nabla\Psi_T}{\Psi_T} = 2\nabla_i \ln \Psi_T \quad (12)$$

or (see Appendix 1: Analytic expression for the local energy):

$$F_i = -2 \frac{(X_k - a_k)}{\sigma^2} + 2 \sum_j^N \frac{w_{kj}}{\sigma^2} \text{logistic}(-v(j)) \quad (13)$$

2.2.4 Updating parameters

$$\beta_{k+1} = \beta_k - \eta_k \mathbf{g}(\beta_k) \quad (14)$$

gradient of the local energy w.r.t the parameters β which can be $\mathbf{a}, \mathbf{b}, \mathbf{X}$; $\mathbf{g}(\beta) = \nabla_\beta \langle E_L \rangle$:

$$\nabla_\beta \langle E_L \rangle = 2 \left(\left\langle \frac{\bar{\Psi}_\beta}{\Psi[\beta]} E_L[\beta] \right\rangle - \left\langle \frac{\bar{\Psi}_\beta}{\Psi[\beta]} \right\rangle \langle E_L[\beta] \rangle \right)$$

∂ eller d ??? Where $\frac{\bar{\Psi}_\beta}{\Psi[\beta]} = \frac{1}{\Psi_T[\beta]} \frac{d\Psi[\beta]}{d\beta}$

2.2.5 neste greie

In this project we will deal only with a system of two electrons in a quantum dot with a frequency of $\hbar\omega = 1$. The reason for this is that we have exact closed form expressions for the ground state energy from Taut's work for selected values of ω , see M. Taut, Phys. Rev. A **48**, 3561 (1993). The energy is given by 3 a.u. (atomic units) when the interaction between the electrons is included. We can however easily extend our system to say interacting bosons as in project 1.

If only the harmonic oscillator part of the Hamiltonian is included, the so-called unperturbed part,

$$\hat{H}_0 = \sum_{i=1}^N \left(-\frac{1}{2} \nabla_i^2 + \frac{1}{2} \omega^2 r_i^2 \right),$$

the energy is 2 a.u. The wave function for one electron in an oscillator potential in two dimensions is

$$\phi_{n_x, n_y}(x, y) = A H_{n_x}(\sqrt{\omega}x) H_{n_y}(\sqrt{\omega}y) \exp(-\omega(x^2 + y^2)/2).$$

The functions $H_{n_x}(\sqrt{\omega}x)$ are so-called Hermite polynomials while A is a normalization constant. For the lowest-lying state we have $n_x = n_y = 0$ and an energy $\epsilon_{n_x, n_y} = \omega(n_x + n_y + 1) = \omega$. Convince yourself that the lowest-lying energy for the two-electron system is simply 2ω .

The unperturbed wave function for the ground state of the two-electron system is given by

$$\Phi(\mathbf{r}_1, \mathbf{r}_2) = C \exp(-\omega(r_1^2 + r_2^2)/2),$$

with C being a normalization constant and $r_i = \sqrt{r_{ix}^2 + r_{iy}^2}$. Note that the vector \mathbf{r}_i refers to the x and y position for a given particle. What is the total spin of this wave function? Find arguments for why the ground state should have this specific total spin.

2.3 Gibbs Sampling

$$\Psi_{T,Gibbs} = \sqrt{F_{RBM}(\mathbf{X})} \quad (15)$$

$$\begin{aligned} P(H_j = 1|\mathbf{X}) &= \text{logistic}(-(v(j))) \\ P(H_j = 0|\mathbf{X}) &= \text{logistic}(v(j)) \end{aligned} \quad (16)$$

where

2.4 Algorithms

Treene en RBM til å foreslå en god Ψ slik at variational principle gir et godt estimat på E_0 , det vil si lavest mulig - altså et tak.

Den foreslår en spesifikk boltzmann distribusjon. Hvorfor er det riktig? Foreslår masse posisjoner ??

Hva er poenget med maskinen? Skjønner ikke

Gi $d \times n$ x_i Interacting vs non-interacting?

- Initiate positions X , weights W , and biases b and a

- Use X as positions for VMC

When moving a particle, change d weights that are "coupled" (one particle)

Sample some stuff

Estimate E

- Use the stuff and E estimate to adjust W, b, a
- Repeat.

3 Results

3.1 Non-interacting Boson gas

4 Conclusions

Appendices

A Appendix 1.

$$\begin{aligned}\Psi(\mathbf{X}) &= F_{rbm}(\mathbf{X}) \\ &= \frac{1}{Z} \sum_{\{h_j\}} e^{-E(\mathbf{X}, \mathbf{h})} \\ &= \frac{1}{Z} \sum_{\{h_j\}} e^{-\sum_i^M \frac{(X_i - a_i)^2}{2\sigma^2} + \sum_j^N b_j h_j + \sum_{i,j}^{M,N} \frac{X_i w_{ij} h_j}{\sigma^2}} \\ &= \frac{1}{Z} e^{-\sum_i^M \frac{(X_i - a_i)^2}{2\sigma^2}} \prod_j^N (1 + e^{b_j + \sum_i^M \frac{X_i w_{ij}}{\sigma^2}})\end{aligned}$$

A.1 Analytic expression for the local energy

$$E_L = \frac{1}{\Psi_T} \hat{H} \Psi_T = \sum_{i=1}^N -\frac{1}{2} \frac{1}{\Psi_T} \nabla^2 \Psi_T + \frac{1}{2} \omega^2 r_i^2 \quad (17)$$

Where

$$\begin{aligned}\frac{1}{\Psi_T} \nabla^2 \Psi_T &= \frac{1}{\Psi_T} \nabla \left(\Psi_T \frac{1}{\Psi_T} \nabla \frac{1}{\Psi_T} \right) \\ &= \left(\frac{1}{\Psi_T} \nabla \Psi_T \right)^2 + \nabla \left(\frac{1}{\Psi_T} \nabla \Psi_T \right) \\ &= (\nabla \ln \Psi_T)^2 + \nabla^2 \ln \Psi_T\end{aligned} \quad (18)$$

First derivative:

$$\begin{aligned}
\frac{1}{\Psi_T} \nabla_k \Psi &= \nabla_k \ln \Psi_T \\
&= \nabla_k \left(\ln \frac{1}{Z} - \sum_i^M \frac{(X_i - a_i)^2}{2\sigma^2} + \sum_j^N \ln(1 + e^{b_j + \sum_i^M \frac{X_i w_{ij}}{\sigma^2}}) \right) \\
&= -\frac{(X_k - a_k)}{\sigma^2} + \sum_j^N w_{kj} \frac{\exp(b_j + \sum_i^M \frac{X_i w_{ij}}{\sigma^2})}{1 + e^{b_j + \sum_i^M \frac{X_i w_{ij}}{\sigma^2}}} \\
&= -\frac{(X_k - a_k)}{\sigma^2} + \sum_j^N \frac{w_{kj}}{\sigma^2} \frac{1}{1 + e^{-v(\mathbf{X}, j)}}
\end{aligned} \tag{19}$$

Second derivative:

$$\begin{aligned}
\nabla_k^2 \ln \Psi_T &= \nabla_k \left(-\frac{(X_k - a_k)}{\sigma^2} + \sum_j^N \frac{w_{kj}}{\sigma^2} \frac{1}{1 + e^{-b_j - \sum_i^M \frac{X_i w_{ij}}{\sigma^2}}} \right) \\
&= -\frac{1}{\sigma^2} + \sum_j^N \frac{w_{kj}^2}{\sigma^4} \frac{e^{-v(\mathbf{X}, j)}}{(1 + e^{-v(\mathbf{X}, j)})^2}
\end{aligned} \tag{20}$$

mak Thus to calculate (17), one may use (18), (19) and (20).

B Appendix 2.

B.1 Derivatives w.r.t RBM parameters

Where $\frac{\bar{\Psi}_\beta}{\Psi[\beta]} = \frac{1}{\Psi_T[\beta]} \frac{d\Psi[\beta]}{d\beta}$. For parameter k : $\frac{1}{\Psi_T} \frac{\partial \Psi_T}{\partial \beta_k} = \frac{\partial}{\partial \beta_k} \log \Psi_T$
for $\beta = \mathbf{a}$:

$$\frac{\partial}{\partial a_k} \log \Psi_T = \frac{X_k - a_k}{\sigma^2} \tag{21}$$

for $\beta = \mathbf{b}$:

$$\frac{\partial}{\partial b_k} \log \Psi_T = \frac{1}{1 + e^{-v(\mathbf{X}, k)}} \tag{22}$$

for $\beta = \mathbf{W}$:

$$\frac{\partial}{\partial W_{kl}} \log \Psi_T = \frac{X_k e^{v(\mathbf{X}, l)}}{(1 + e^{v(\mathbf{X}, l)})\sigma^2} = \frac{X_k}{(1 + e^{-v(\mathbf{X}, l)})\sigma^2} \tag{23}$$