Studies of Quantum Dots using Machine Learning

Even Marius Nordhagen

University of Oslo

evenmn@fys.uio.no

December 3, 2019

Outline

- Motivation
- Quantum Theory
- Machine Learning
- Methods
- Results
- Conclusion
- (Code)

Motivation

Why quantum dots?

- Technology
- Theoretically
- Experimentally

Why is it challenging?

- Quantum many-body problem
- Fermi-Dirac statistics

How to overcome the challenges?

- Efforts..
- gh
- Our approach: Machine Learning

Theory

The Time-independent Schrödinger Equation

A stationary quantum mechanical system is described by

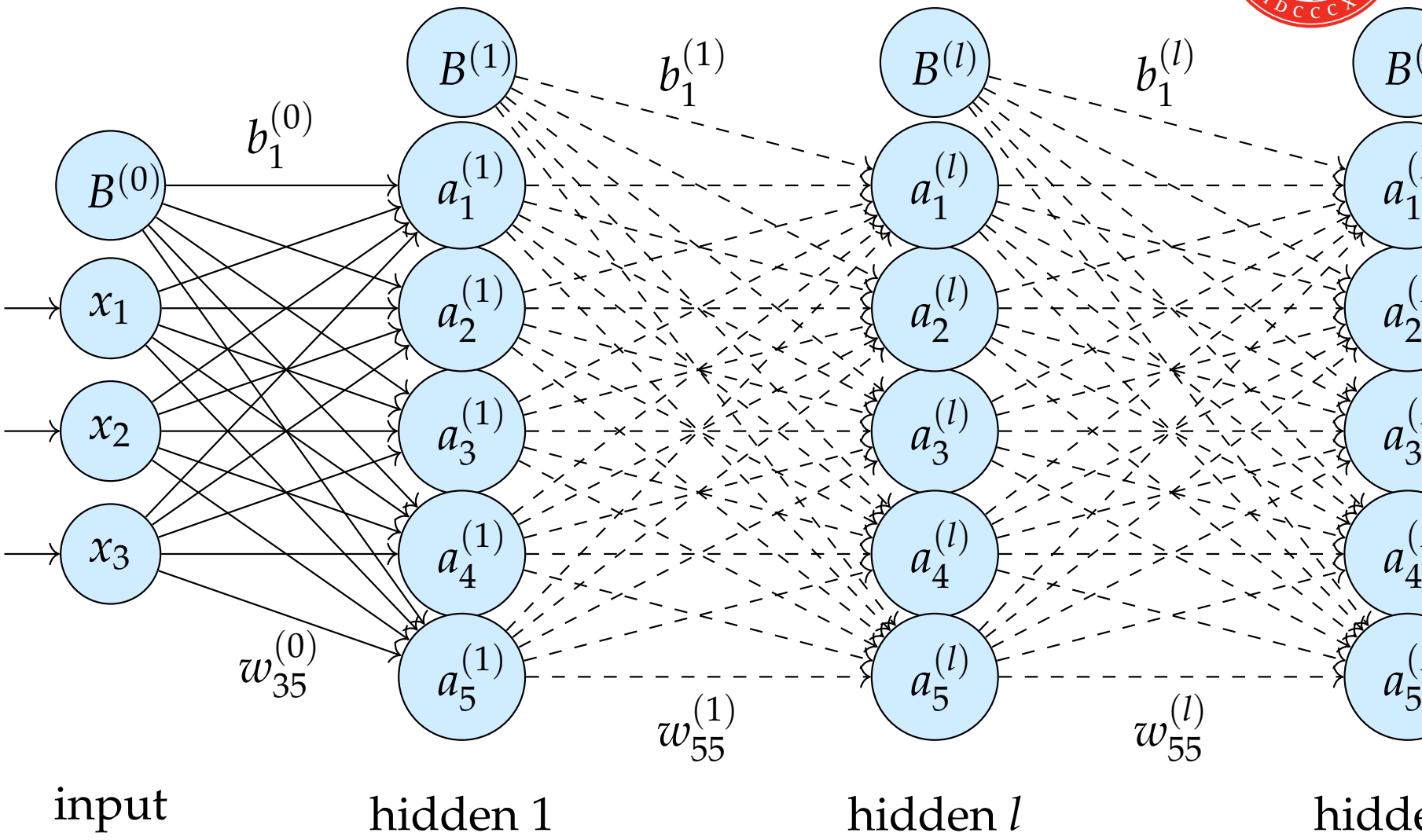
$$E_n = \frac{\int d\boldsymbol{r} \Psi_n^*(\boldsymbol{r}) \hat{\mathcal{H}} \Psi_n(\boldsymbol{r})}{\int d\boldsymbol{r} \Psi_n^*(\boldsymbol{r}) \Psi_n(\boldsymbol{r})}$$

which gives the energy of state n.

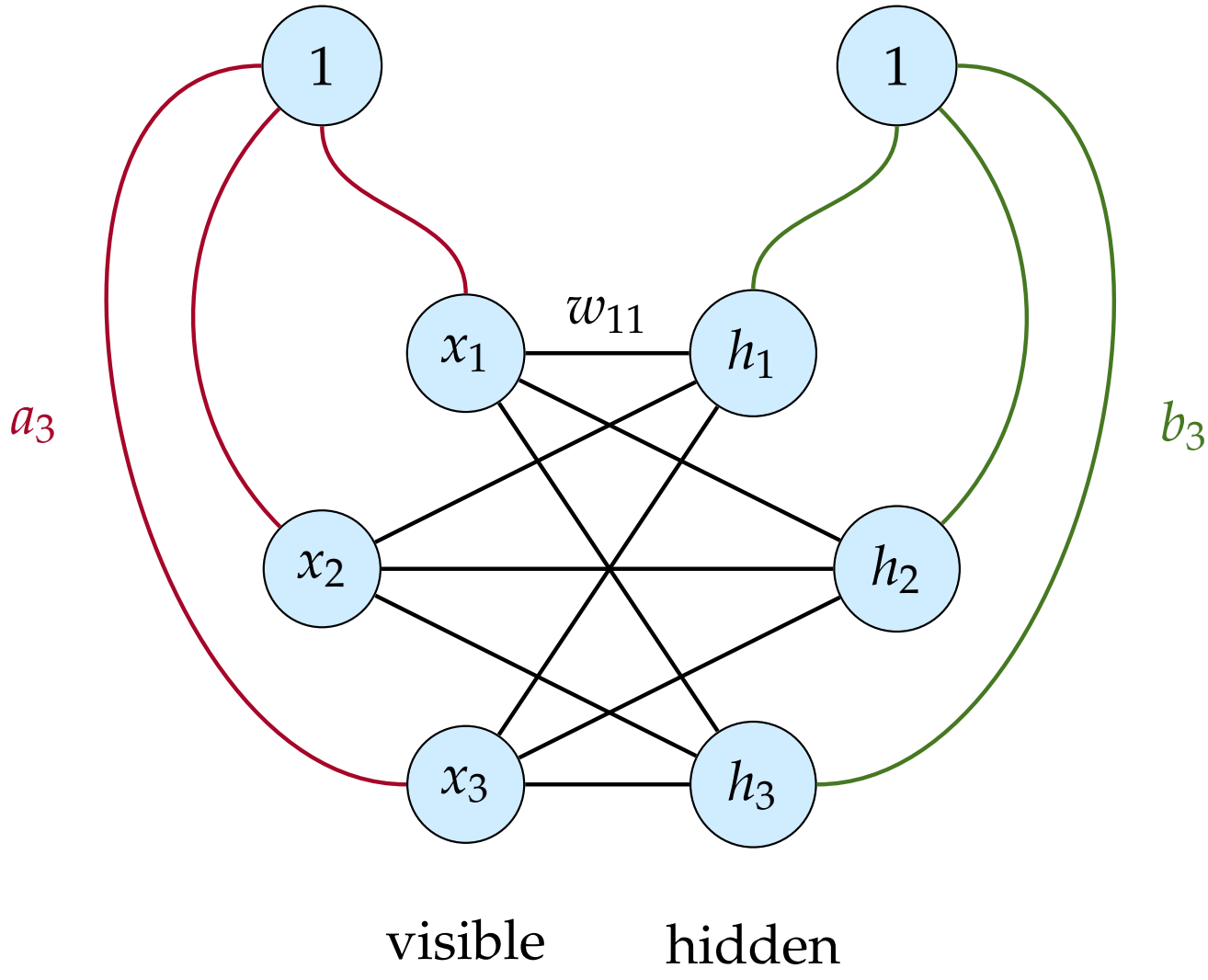
Machine Learning

> *"Machine learning is the science of getting computers to act without being explicitly programmed."*

Feed-forward Neural Network



Restricted Boltzmann machines



Methods

Variational Monte Carlo (VMC)

Exploit the variational principle in order to obtain the ground state energy

$$
\begin{aligned}
E_0 < E_{\mathrm{VMC}} &= \frac{\int d\boldsymbol{r}\Psi_T(\boldsymbol{r})^*\hat{\mathcal{H}}\Psi_T(\boldsymbol{r})}{\int d\boldsymbol{r}\Psi_T(\boldsymbol{r})^*\Psi_T(\boldsymbol{r})} \\
&= \int d\boldsymbol{r}\, \underbrace{\frac{\Psi_T(\boldsymbol{r})^*\Psi_T(\boldsymbol{r})}{\int d\boldsymbol{r}\Psi_T(\boldsymbol{r})^*\Psi_T(\boldsymbol{r})}}_{P(\boldsymbol{r})} \cdot \underbrace{\frac{1}{\Psi_T(\boldsymbol{r})}\hat{\mathcal{H}}\Psi_T(\boldsymbol{r})}_{E_L(\boldsymbol{r})}
\end{aligned}
$$

Monte Carlo Integration

We attempt to solve the integral by sampling from the probability density function $P(\boldsymbol{r})$

$$E_{\mathrm{VMC}} = \int d\boldsymbol{r} E_L(\boldsymbol{r}) P(\boldsymbol{r})$$
$$\approx \frac{1}{M} \sum_{i=1}^{M} E_L(\boldsymbol{r}_i)$$

Trial Wave Function

$$P(\boldsymbol{r}) \propto \Psi_T(\boldsymbol{r})^* \Psi_T(\boldsymbol{r})$$

Use the Slater-Jastrow function as our trial wave function

$$\Psi_T(\boldsymbol{r}) = |\hat{D}(\boldsymbol{r})| J(\boldsymbol{r})$$

where the Slater matrix, $\hat{D}(\boldsymbol{r})$, contains all the single-particle functions

$$\hat{D}(\boldsymbol{r}) = \begin{pmatrix} \phi_1(\boldsymbol{r}_1) & \phi_2(\boldsymbol{r}_1) & \dots & \phi_N(\boldsymbol{r}_1) \\ \phi_1(\boldsymbol{r}_2) & \phi_2(\boldsymbol{r}_2) & \dots & \phi_N(\boldsymbol{r}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_1(\boldsymbol{r}_N) & \phi_2(\boldsymbol{r}_N) & \dots & \phi_N(\boldsymbol{r}_N) \end{pmatrix}$$

Single-particle Functions

The Hermite functions,

$$\phi_n(\boldsymbol{r}) \propto H_n(\boldsymbol{r}) \exp\left(-\frac{1}{2}\alpha\omega|\boldsymbol{r}|^2\right),$$

are used as the single-particle functions for quantum dots in standard VMC. The Gaussian can be factorized out from the Slater determinant.

$$|\hat{D}(\boldsymbol{r};\alpha)| \propto \exp\left(-\frac{1}{2}\alpha\omega|\boldsymbol{r}|^2\right) \begin{vmatrix} H_1(\boldsymbol{r}_1) & H_2(\boldsymbol{r}_1) & \dots & H_N(\boldsymbol{r}_1) \\ H_1(\boldsymbol{r}_2) & H_2(\boldsymbol{r}_2) & \dots & H_N(\boldsymbol{r}_2) \\ \vdots & \vdots & \ddots & \vdots \\ H_1(\boldsymbol{r}_N) & H_2(\boldsymbol{r}_N) & \dots & H_N(\boldsymbol{r}_N) \end{vmatrix}$$

Restricted Boltzmann Machine

We use the marginal distribution of the visible units as the single-particle functions in the Slater determinant, and see if them can model the correlations

$$\phi_n(\boldsymbol{r}) \propto H_n(\boldsymbol{r}) P(\boldsymbol{r}; \boldsymbol{a}, \boldsymbol{b}, \boldsymbol{W})$$

where $P(\boldsymbol{r})$ is the marginal distribution of the visible units.

$$|\hat{D}(\boldsymbol{r}; \boldsymbol{a}, \boldsymbol{b}, \boldsymbol{W})| \propto P(\boldsymbol{r}; \boldsymbol{a}, \boldsymbol{b}, \boldsymbol{W}) \begin{vmatrix} H_1(\boldsymbol{r}_1) & H_2(\boldsymbol{r}_1) & \dots & H_N(\boldsymbol{r}_1) \\ H_1(\boldsymbol{r}_2) & H_2(\boldsymbol{r}_2) & \dots & H_N(\boldsymbol{r}_2) \\ \vdots & \vdots & \ddots & \vdots \\ H_1(\boldsymbol{r}_N) & H_2(\boldsymbol{r}_N) & \dots & H_N(\boldsymbol{r}_N) \end{vmatrix}$$

Jastrow Factor

The Jastrow factor is added to account for the correlations
Simple Jastrow factor

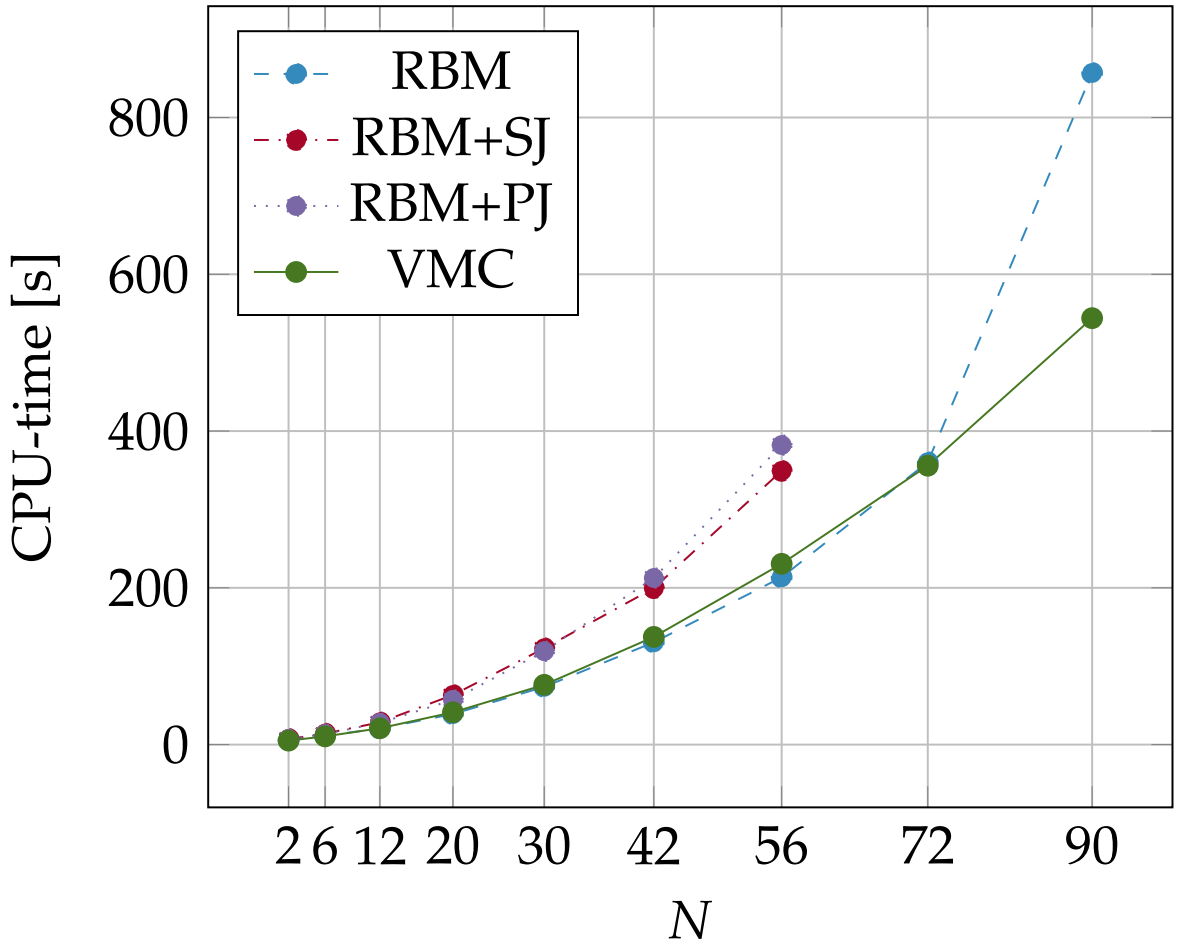
$$J(\boldsymbol{r};\boldsymbol{\beta}) = \exp\left(\sum_{i=1}^{N}\sum_{j>i}^{N}\beta_{ij}r_{ij}\right).$$

Padé-Jastrow factor

$$J(\boldsymbol{r};\beta) = \exp\left(\sum_{i=1}^{N}\sum_{j>i}^{N}\frac{a_{ij}r_{ij}}{1+\beta r_{ij}}\right).$$

Results

Computational Cost



Conclusion

Retrospect

- RBM is able to account for most of the correlations
- RBM+PJ implies to give a lower ground state energy and model the correlations better than a traditional VMC
- RBM+SJ is both more expensive and less accurate than its fellow methods, and we see no reason to choose it

Future Work

- Repeat the exercise using spherical coordinates - interactions are easier to model in spherical coordinates
- Check the ability of modeling the three-body correlations, considering nuclear systems
- Reduce the computational cost

Thank you!