Studies of Quantum Dots using Machine Learning

Even Marius Nordhagen

University of Oslo

evenmn@fys.uio.no

December 13, 2019

Outline

- Motivation
- Quantum Theory
- Machine Learning Theory
- Methods
- Software
- Results
- Conclusion

Motivation

Studies of Quantum Dots using Machine Learning

Studies of Quantum Dots using Machine Learning

Machine Learning

- “*Machine learning is the science of getting computers to act without being explicitly programmed*[1]”.

Machine Learning

- > *Machine learning is the science of getting computers to act without being explicitly programmed*[1].

- Image recognition

Machine Learning

- > *Machine learning is the science of getting computers to act without being explicitly programmed*[1].

- Image recognition

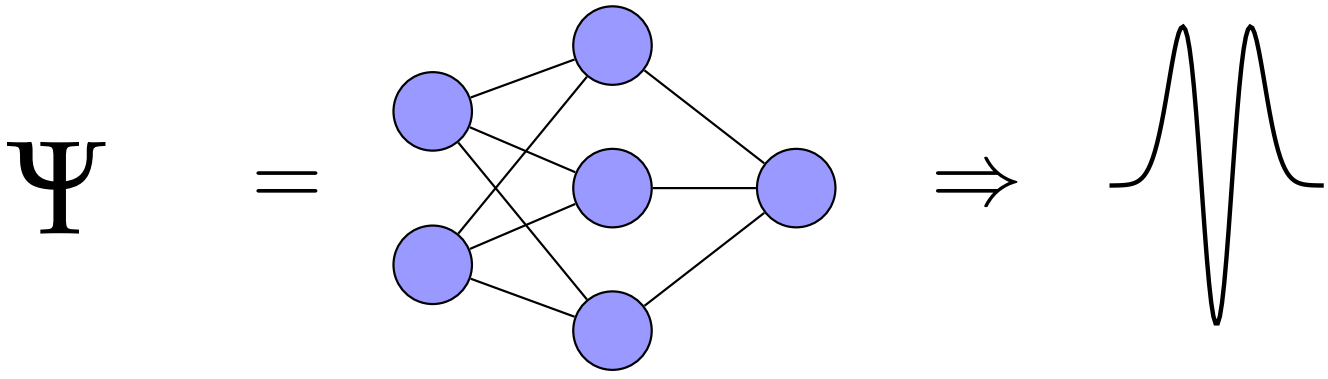
- Voice commands

Machine Learning + Quantum Mechanics

- Neural networks are eminent function approximators

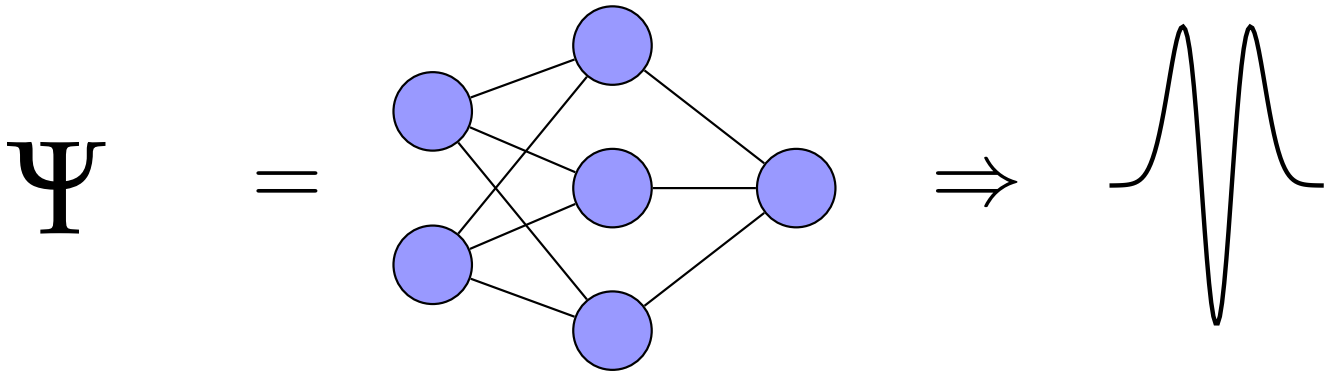
Machine Learning + Quantum Mechanics

- Neural networks are eminent function approximators



Machine Learning + Quantum Mechanics

- Neural networks are eminent function approximators



- Existing methods are reminiscent of machine learning algorithms

Studies of Quantum Dots using Machine Learning

Studies of Quantum Dots using Machine Learning

Quantum Dots

- **Technologically:**
  Quantum dots are expected to be the next big thing in display technology[2,3].

Quantum Dots

- **Technologically:**
  Quantum dots are expected to be the next big thing in display technology[2,3].

- **Experimentally:**
  Researchers have managed to study two-dimensional quantum dots in the laboratory[4].

Quantum Dots

- **Technologically:**
  Quantum dots are expected to be the next big thing in display technology[2,3].

- **Experimentally:**
  Researchers have managed to study two-dimensional quantum dots in the laboratory[4].

- **Physically:**
  An array of interesting physical phenomena can be observed in quantum dots.

Ethics in Science

- Respect for other's work

Ethics in Science

- Respect for other's work
- Reproducibility

Quantum Theory

The Schrödinger Equation

$$\hat{\mathcal{H}}\Psi = E\Psi$$

The Schrödinger Equation

$$\hat{\mathcal{H}}\Psi = E\Psi$$

$$\Downarrow$$

$$E = \frac{\int d\boldsymbol{X}\Psi^*(\boldsymbol{X})\hat{\mathcal{H}}\Psi(\boldsymbol{X})}{\int d\boldsymbol{X}\Psi^*(\boldsymbol{X})\Psi(\boldsymbol{X})}$$

The Variational Principle

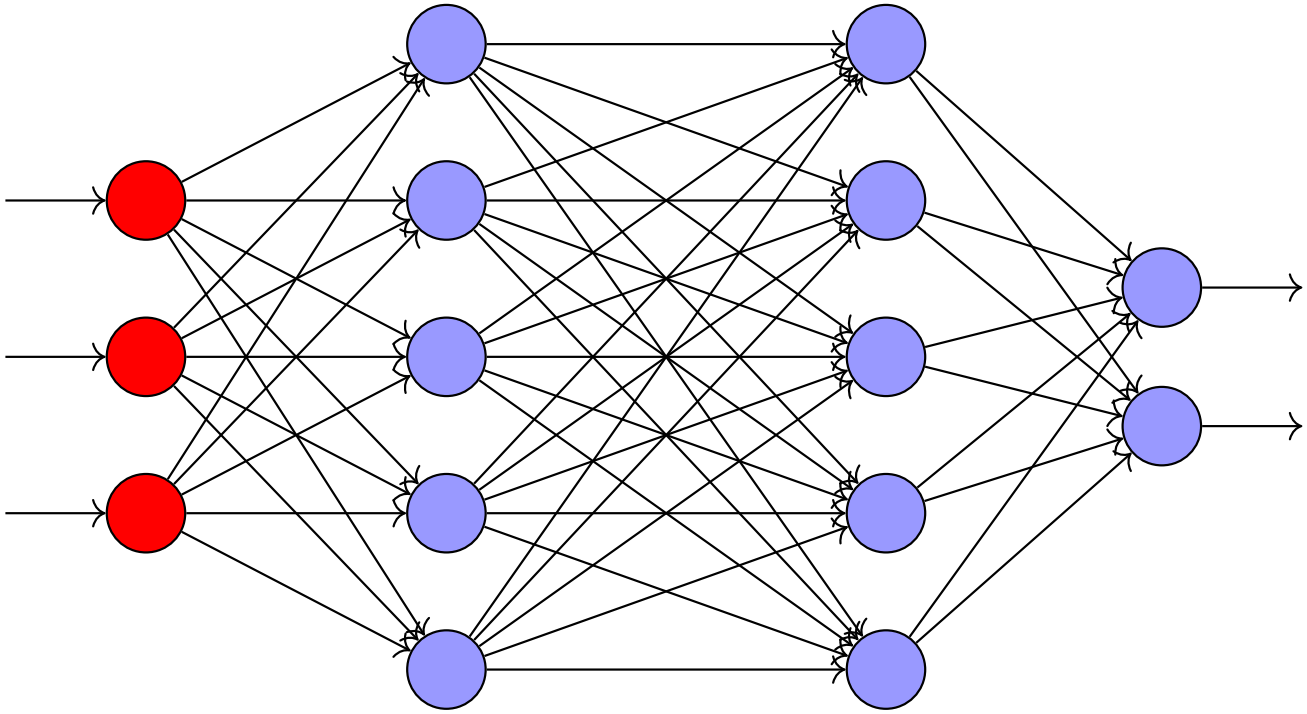
The variational principle serves as a way of finding the ground state energy. For an arbitrary trial wave function $\Psi_T(\boldsymbol{X})$, it states that the obtained energy is larger or equal to the ground state,

$$E_0 \leq E = \frac{\int d\boldsymbol{X} \Psi_T^*(\boldsymbol{X}) \hat{\mathcal{H}} \Psi_T(\boldsymbol{X})}{\int d\boldsymbol{X} \Psi_T^*(\boldsymbol{X}) \Psi_T(\boldsymbol{X})}.$$

Thus, by minimizing the obtained energy, E, we can estimate the ground state energy.

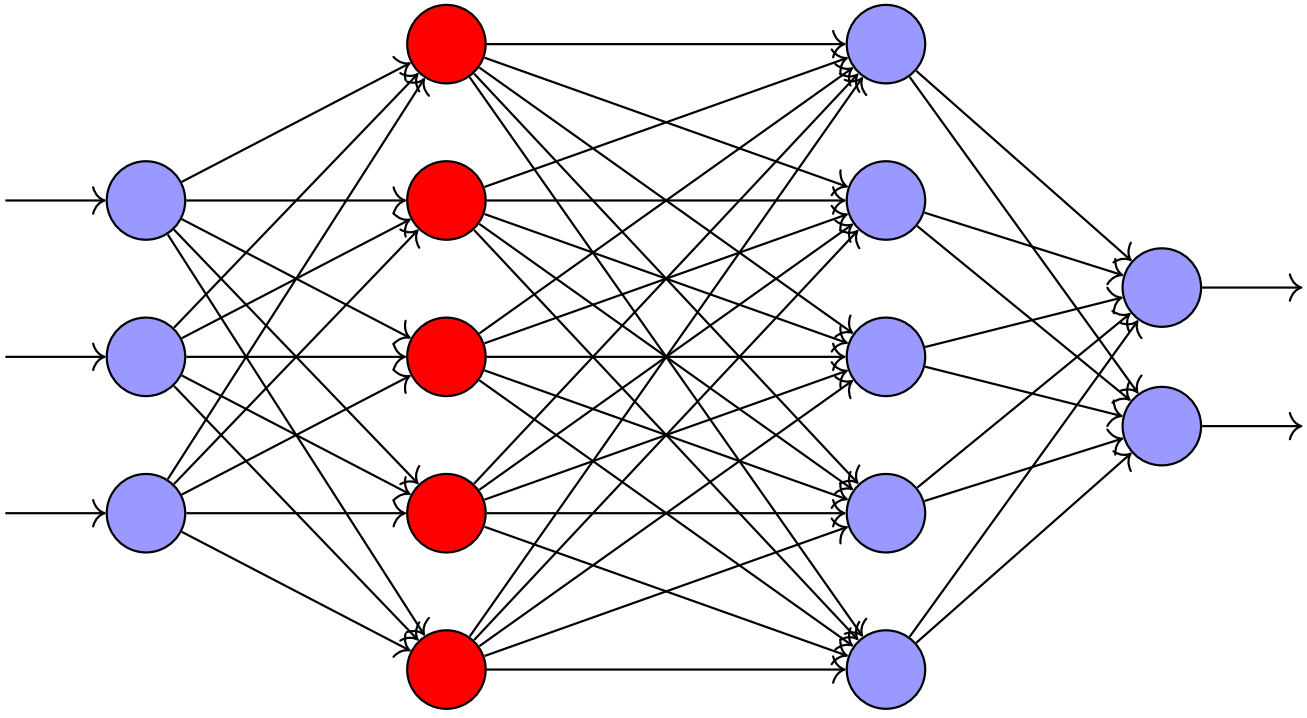
Machine Learning Theory

Feed-forward Neural Network (FNN)



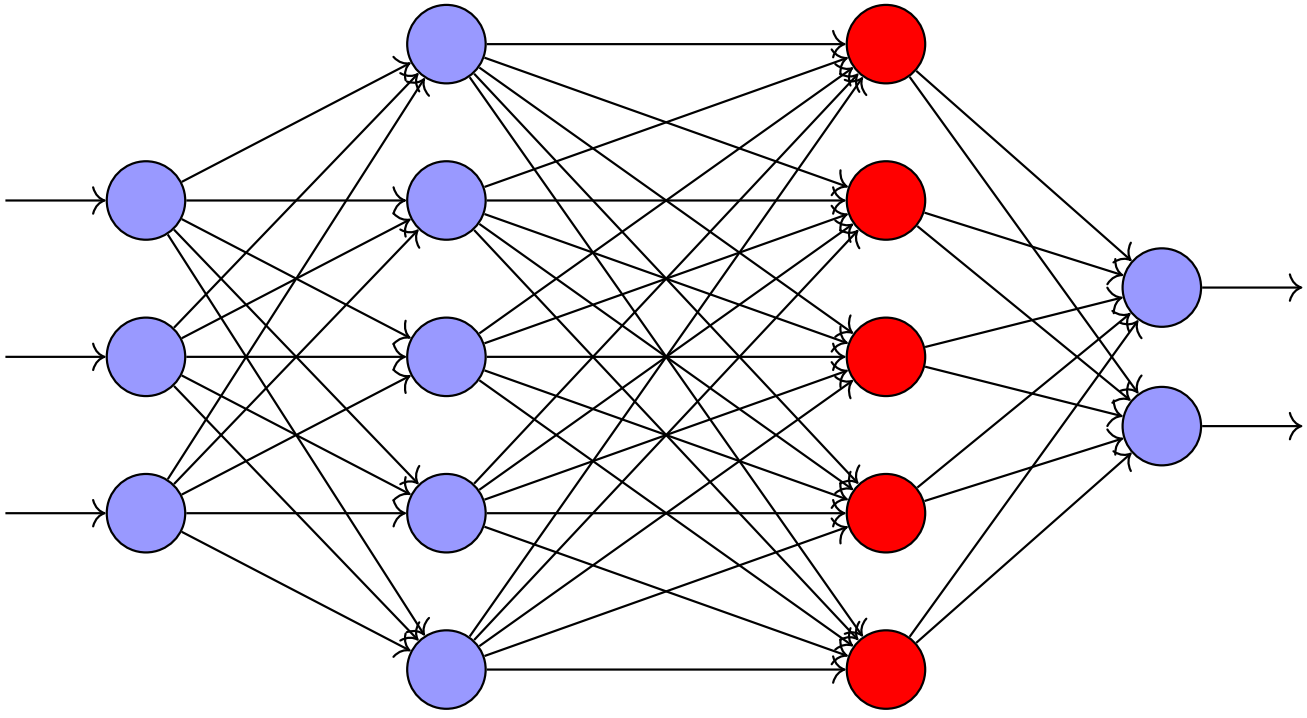
$$a_0 = x$$

Feed-forward Neural Network (FNN)



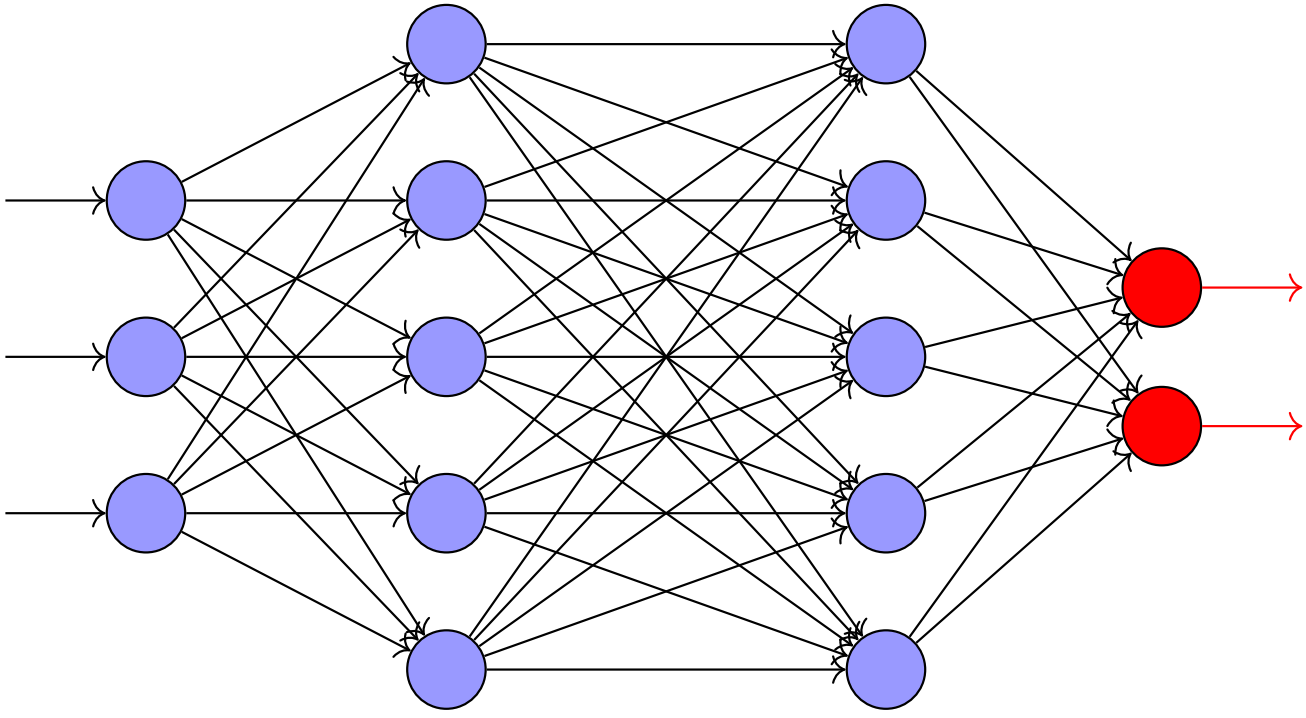
$$\boldsymbol{a}_0 = \boldsymbol{x} \qquad \boldsymbol{a}_1 = f_1(\boldsymbol{a}_0)$$

Feed-forward Neural Network (FNN)



$$\boldsymbol{a}_0 = \boldsymbol{x} \qquad \boldsymbol{a}_1 = f_1(\boldsymbol{a}_0) \qquad \boldsymbol{a}_2 = f_2(\boldsymbol{a}_1)$$

Feed-forward Neural Network (FNN)



$$\boldsymbol{a}_0 = \boldsymbol{x} \qquad \boldsymbol{a}_1 = f_1(\boldsymbol{a}_0) \qquad \boldsymbol{a}_2 = f_2(\boldsymbol{a}_1) \qquad \tilde{\boldsymbol{y}} = f_3(\boldsymbol{a}_2)$$

Cost function

- The cost function defines the error

Cost function

- The cost function defines the error
- Mean square error (MSE):

$$\mathcal{C} = \frac{1}{2}\sum_{i=1}^{n}(y - \tilde{y})^2.$$

Cost function

- The cost function defines the error
- Mean square error (MSE):

$$\mathcal{C} = \frac{1}{2}\sum_{i=1}^{n}(y - \tilde{y})^2.$$

- Attempt to minimize the cost function

Optimization Algorithms

- Minimize the cost function

Optimization Algorithms

- Minimize the cost function

- The gradient descent method:

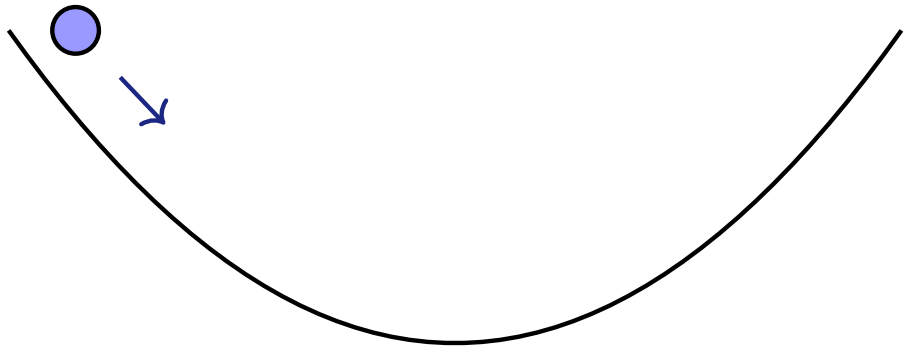
$$\theta^{+} = \theta - \frac{\partial \mathcal{C}}{\partial \theta}.$$

Optimization Algorithms

- Minimize the cost function

- The gradient descent method:

$$\theta^{+} = \theta - \frac{\partial \mathcal{C}}{\partial \theta}.$$

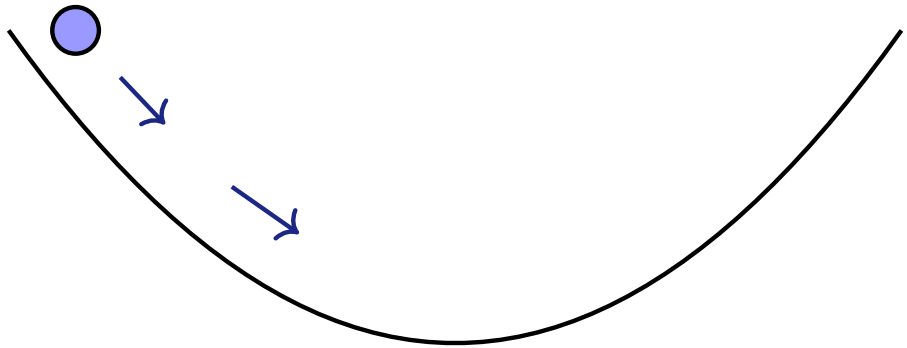


Optimization Algorithms

- Minimize the cost function

- The gradient descent method:

$$\theta^{+} = \theta - \frac{\partial \mathcal{C}}{\partial \theta}.$$

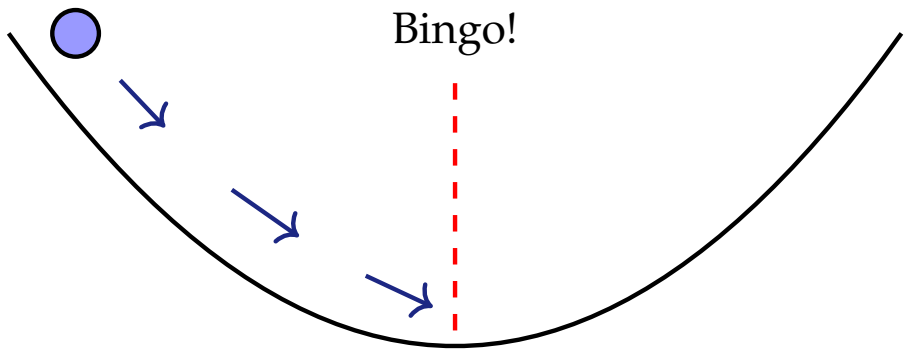


Optimization Algorithms

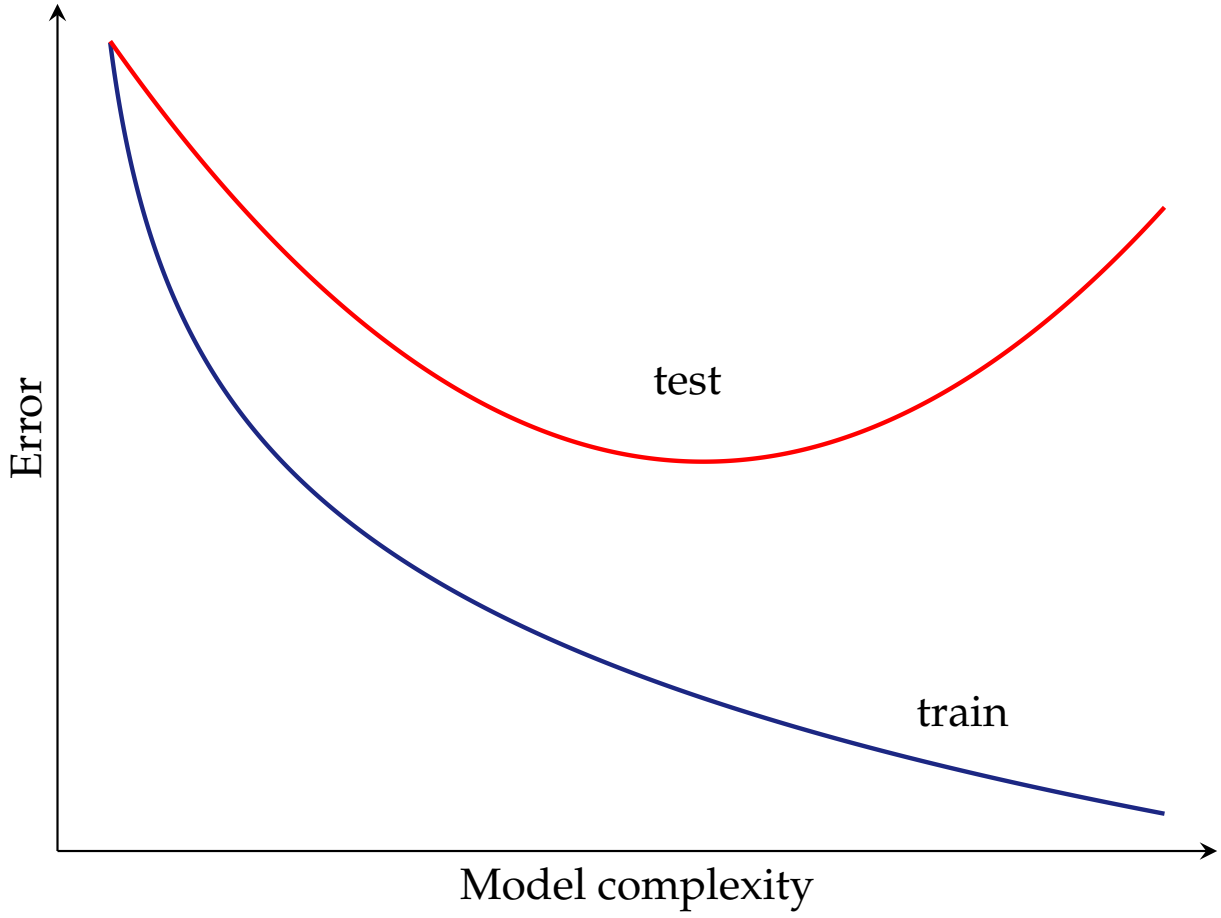
- Minimize the cost function

- The gradient descent method:

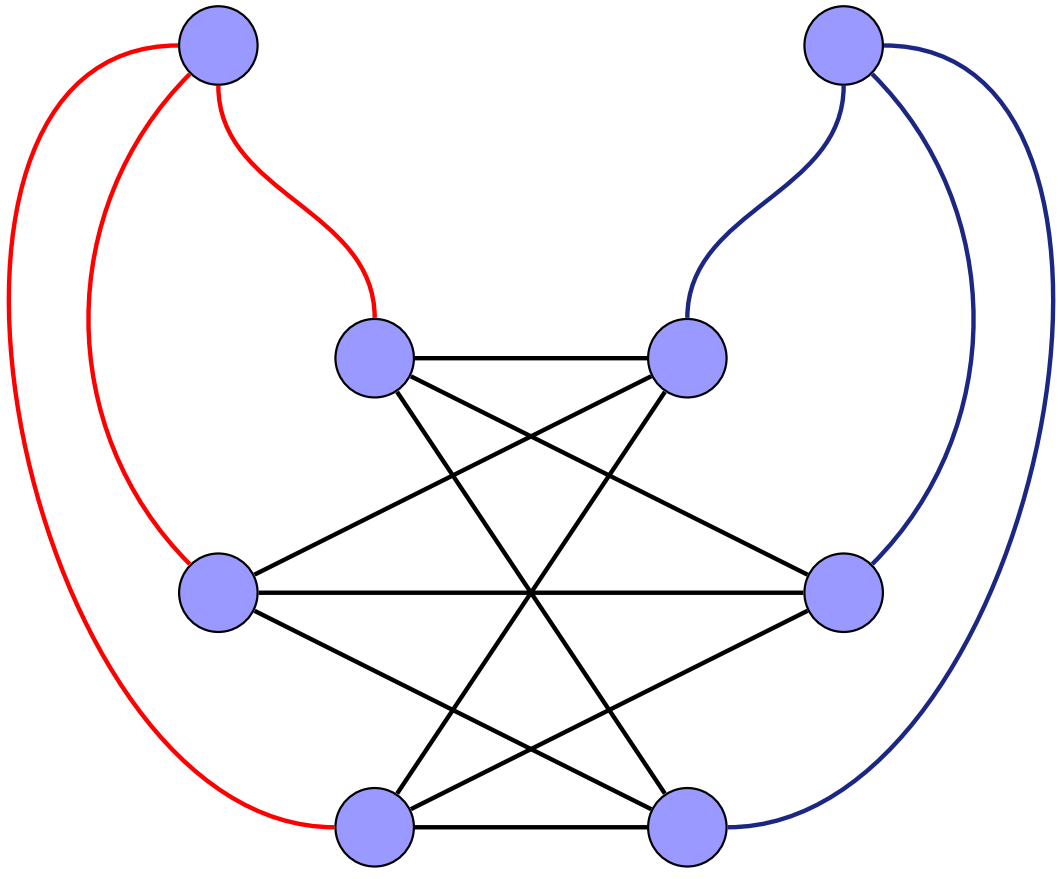
$$\theta^{+} = \theta - \frac{\partial \mathcal{C}}{\partial \theta}.$$



Find Appropriate Complexity



Restricted Boltzmann Machines



Methods

Variational Monte Carlo (VMC)

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

$$E_0 < E_{\mathrm{VMC}} = \frac{\int d\boldsymbol{R}\Psi_T(\boldsymbol{R})^*\hat{\mathcal{H}}\Psi_T(\boldsymbol{R})}{\int d\boldsymbol{R}\Psi_T(\boldsymbol{R})^*\Psi_T(\boldsymbol{R})},$$
$$= \int d\boldsymbol{R}E_L(\boldsymbol{R})P(\boldsymbol{R}),$$

with

$$E_L(\boldsymbol{R}) = \frac{1}{\Psi_T(\boldsymbol{R})}\hat{\mathcal{H}}\Psi_T(\boldsymbol{R}) \quad \wedge \quad P(\boldsymbol{R}) = \frac{\Psi_T(\boldsymbol{R})^*\Psi_T(\boldsymbol{R})}{\int d\boldsymbol{R}\Psi_T(\boldsymbol{R})^*\Psi_T(\boldsymbol{R})}$$

Monte Carlo Integration

We attempt to solve the integral by sampling from the probability density function $P(\boldsymbol{R}) \propto \Psi_T(\boldsymbol{R})^* \Psi_T(\boldsymbol{R})$:

$$E_{\mathrm{VMC}} = \int d\boldsymbol{R} E_L(\boldsymbol{R}) P(\boldsymbol{R}),$$

$$\approx \frac{1}{M} \sum_{i=1}^{M} E_L(\boldsymbol{R}_i).$$

Trial Wave Function Ansatz

The Slater-Jastrow function is the *de facto* standard trial wave function for electronic structure systems,

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

where the Slater matrix,

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

contains all the single-particle functions.

Single-particle Functions

The Hermite functions,

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

are often used as the single-particle functions for quantum dots. 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

As suggested by Carleo & Troyer[5], 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

$$\phi_n(\boldsymbol{r}) \propto H_n(\sqrt{\omega}\boldsymbol{r})P(\boldsymbol{r};\boldsymbol{\theta})$$

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

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

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

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

Our Trial Wave Function Ansätze

- $\Psi_{\mathrm{RBM}}(\boldsymbol{R}) = |\hat{D}_{\mathrm{RBM}}(\boldsymbol{R})|$
- $\Psi_{\mathrm{RBM+SJ}}(\boldsymbol{R}) = |\hat{D}_{\mathrm{RBM}}(\boldsymbol{R})| J(\boldsymbol{R};\boldsymbol{\beta})$
- $\Psi_{\mathrm{RBM+PJ}}(\boldsymbol{R}) = |\hat{D}_{\mathrm{RBM}}(\boldsymbol{R})| J(\boldsymbol{R};\beta)$
- $\Psi_{\mathrm{VMC}}(\boldsymbol{R}) = |\hat{D}_{\mathrm{Gauss}}(\boldsymbol{R})| J(\boldsymbol{R};\beta)$

Software

Aims

- Efficient
- Flexible
- Easy to use

Results

Ground State Energy

Number of electrons: $N = 2$. Frequency: ω.

ω	RBM	RBM+SJ	RBM+PJ	VMC	HF *	Exact †
1/6	0.7036(1)	0.67684(7)	0.66715(6)	0.66710(1)	0.768675	2/3
0.28	1.07050(4)	1.03470(7)	1.021668(7)	1.02192(1)	1.14171	
1	3.0803(2)	3.02108(5)	2.999587(5)	2.99936(1)	3.16190	3

*Computation of the Hartree-Fock limit by Mariadason, 2018 [6].

†Semi-analytical ground state energy calculated by Taut, 1993 [7].

Ground State Energy

Number of electrons: $N = 20$. Frequency: ω.

ω	RBM	RBM+SJ	RBM+PJ	VMC	HF [‡]	DMC [§]
0.1	30.824(2)	30.567(3)	30.1553(9)	30.0403(2)	31.1902	29.9779(1)
1.0	159.428(3)	156.816(4)	156.104(1)	155.8900(4)	158.004	155.8822(1)

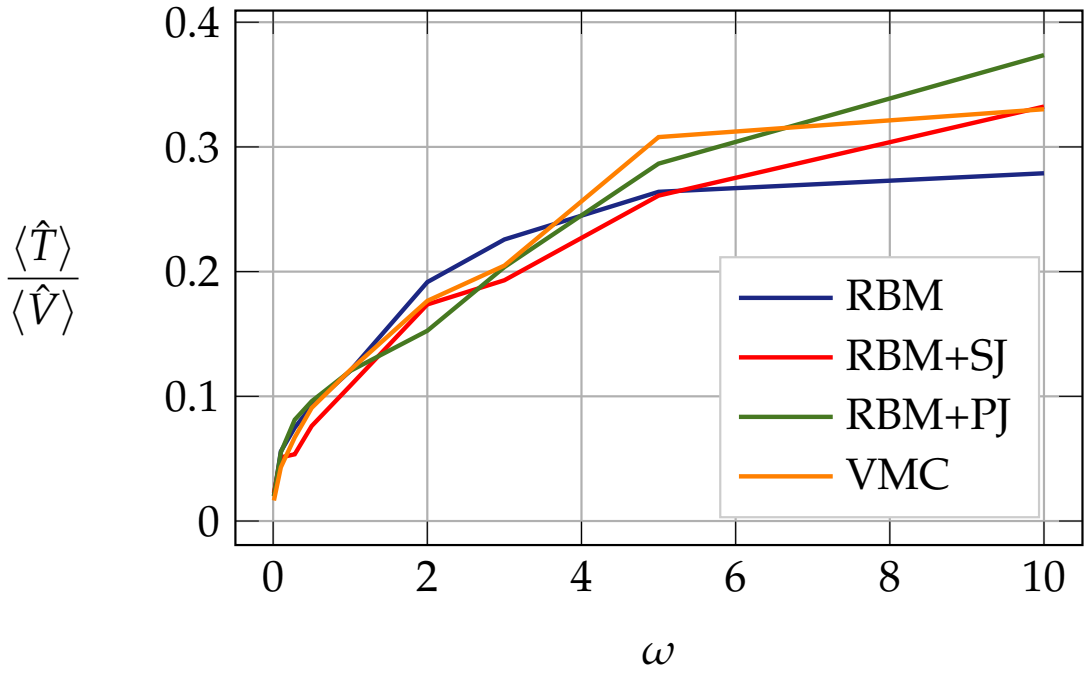
[‡]Computation of the Hartree-Fock limit by Mariadason, 2018 [6].

[§]Ground state energy estimate using the diffusion Monte Carlo method. By Høgberget, 2013 [8].

Energy distribution

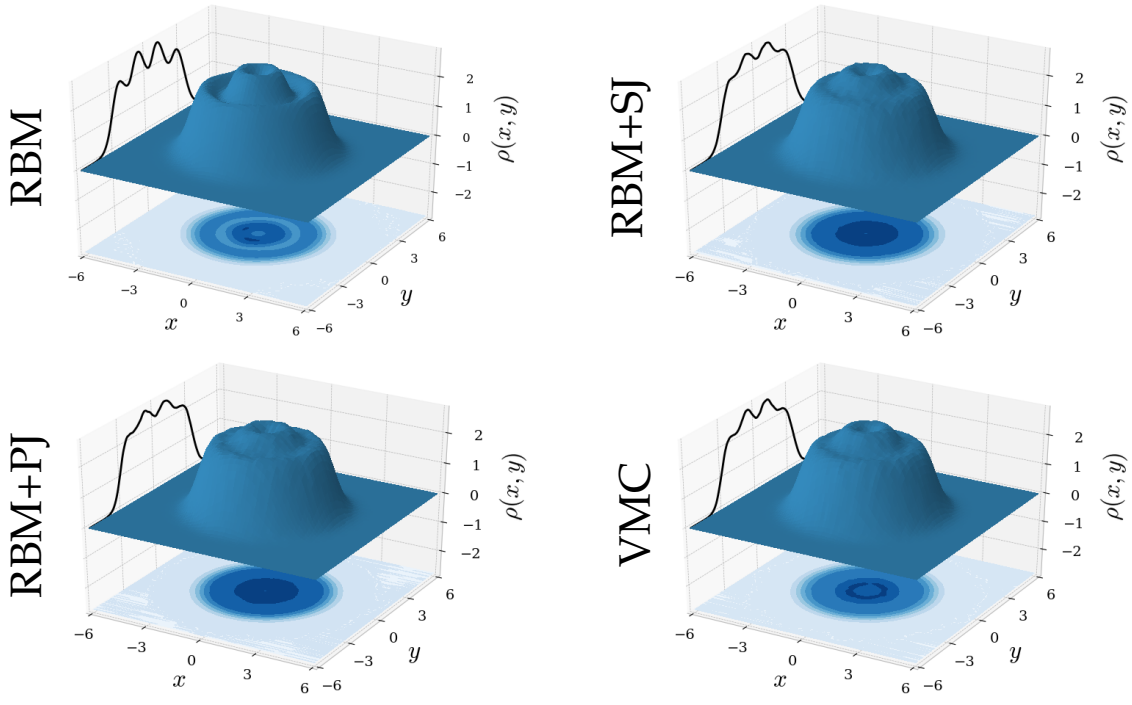
Number of electrons: $N = 20$. Frequency: ω.

Ratio between the kinetic energy, $\langle \hat{T} \rangle$, and the total potential energy, $\langle \hat{V} \rangle$.



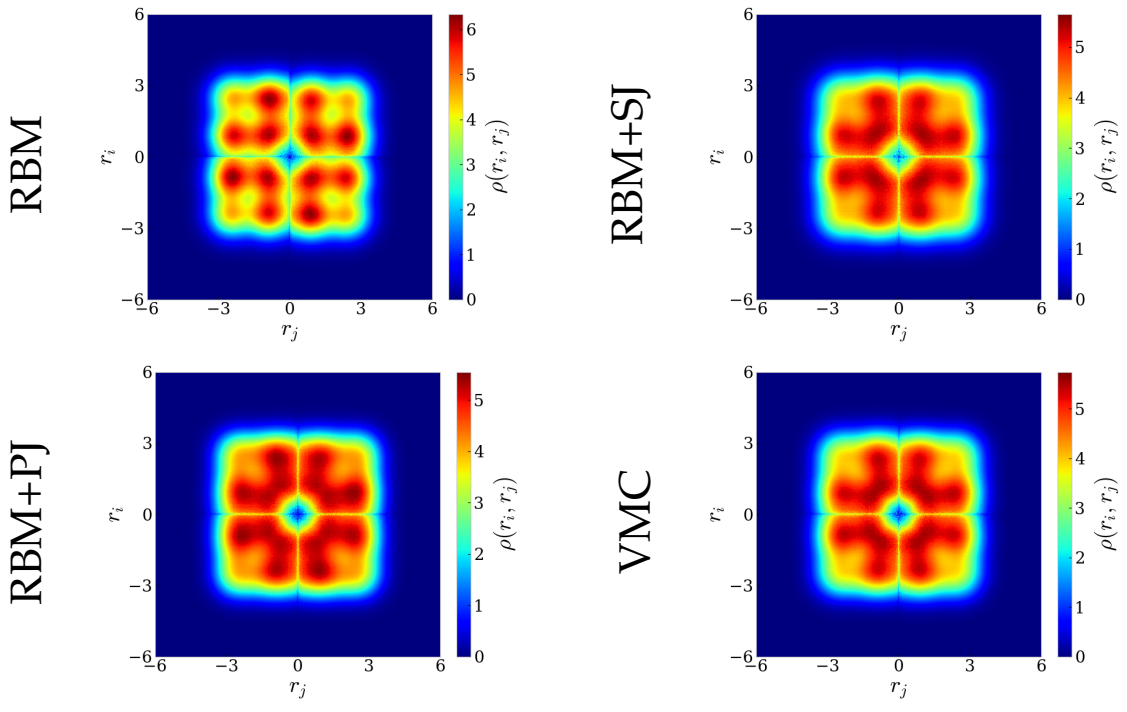
One-body density

Number of electrons: $N = 20$. Frequency: $\omega = 1.0$.



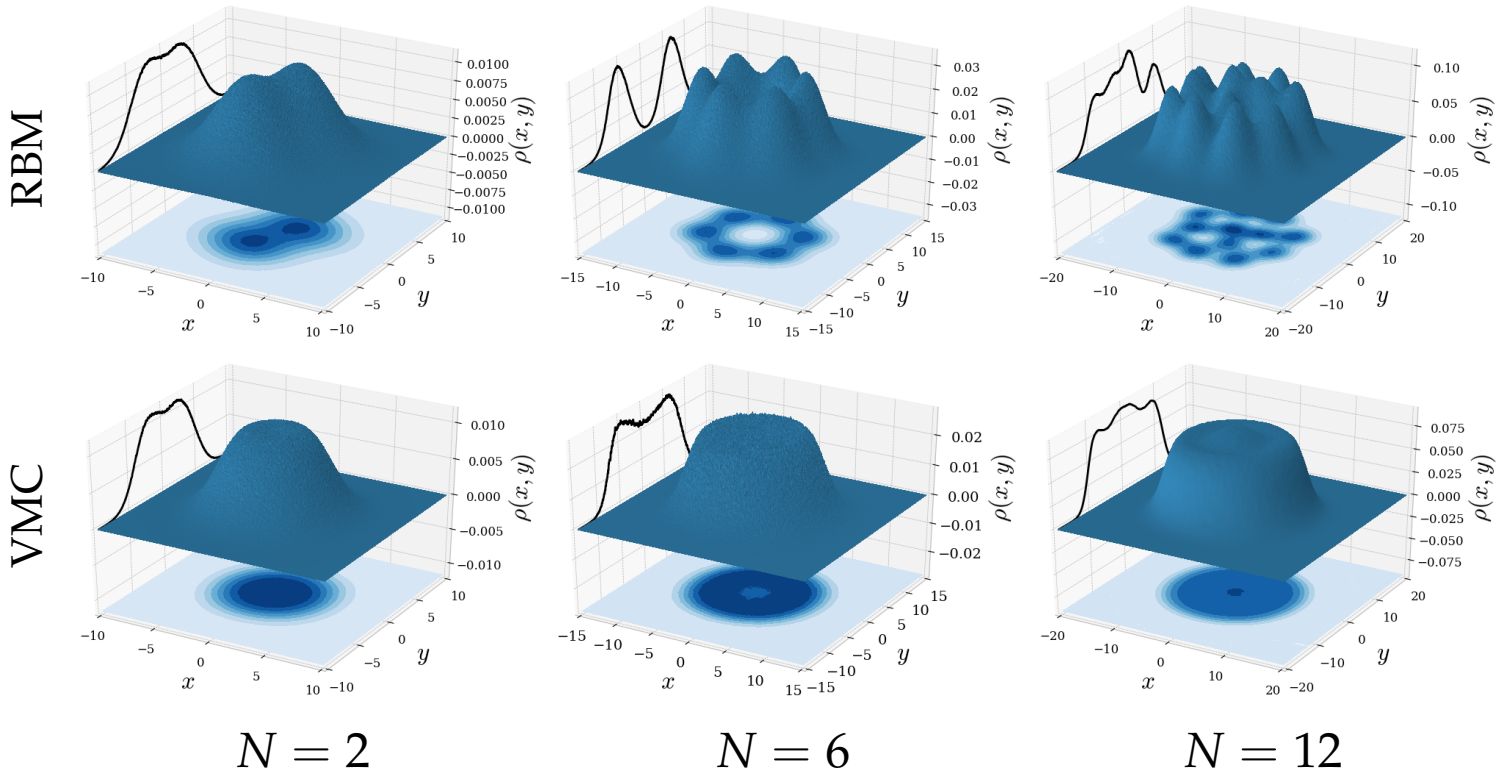
Two-body density

Number of electrons: $N = 20$. Frequency: $\omega = 1.0$.



Low-frequency dots

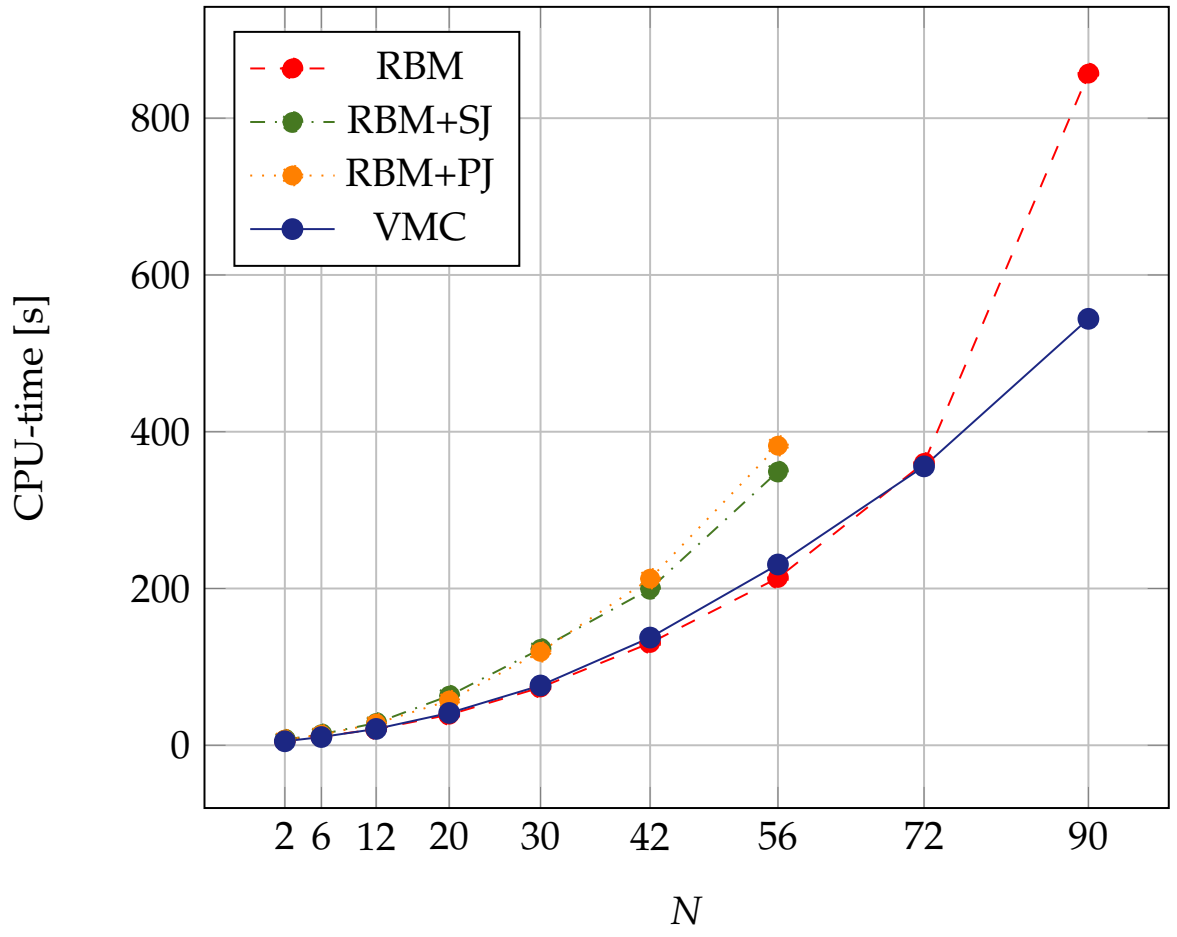
Number of electrons: N. Frequency: $\omega = 0.1$.



UiO : University of Oslo

Computational Cost

Number of electrons: N.



Conclusions

Conclusions

- The RBM ansatz is able to account for most of the correlations
- The RBM+PJ ansatz might give a better ground state estimate of small quantum dots, compared to the traditional VMC ansatz
- For larger quantum dots, RBM+PJ gives slightly larger energy than VMC
- The energy distribution is different for the different methods, indicating different electron configurations.
- The ground state energy might not be the best way to evaluate the various ansätze

Future Work

- Investigate restricted Boltzmann machines with other architectures
- Pass more information to the restricted Boltzmann machine
- Apply the method on more complex systems

Thank you!

References

1. *Machine Learning - Stanford University.* http://mlclass.stanford.edu/.
2. *Samsung QLED TV | The Frame | Premium UHD TV.* http://www.samsung.com/global/tv/blog/why-are-quantum-dot-displays-so-good/.
3. Manders, J. R. *et al.* 8.3: Distinguished Paper: Next-Generation Display Technology: Quantum-Dot LEDs. *SID Symposium Digest of Technical Papers* **46,** 73 (2015).
4. Brunner, K., Abstreiter, G., Böhm, G., Tränkle, G. & Weimann, G. Sharp-Line Photoluminescence and Two-Photon Absorption of Zero-Dimensional Biexcitons in a GaAs/AlGaAs Structure. *Physical Review Letters* **73,** 1138 (1994).
5. Carleo, G. & Troyer, M. Solving the Quantum Many-Body Problem with Artificial Neural Networks. *Science* **355,** 602 (2017).
6. Mariadason, A. A. *Quantum Many-Body Simulations of Double Dot System.* MA thesis (2018).
7. Taut, M. Two electrons in an external oscillator potential: Particular analytic solutions of a Coulomb correlation problem. *Physical Review A* **48,** 3561 (1993).
8. Høgberget, J. *Quantum Monte-Carlo Studies of Generalized Many-body Systems.* MA thesis (2013).

Test

Test

TEST

Test

Test

TEST

Test

Test

TEST