

# VARIATIONAL MONTE CARLO METHODS FOR QUANTUM DOTS

---

Matteo Seclì

July 2015

UNIVERSITY OF TRENTO – DEPARTMENT OF PHYSICS

What are quantum dots?

The algorithm

The 2-electrons system

The 6-electrons system

Analytical derivatives

Conclusion

## WHAT ARE QUANTUM DOTS?

---

Quantum dots can be thought as **artificial atoms**.

Similarities with natural atoms:

- They are made up of electrons confined in an *attractive potential*
- They have a shell-structure with its relative **magic numbers**

Differences:

- *Shape* of the potential (2D isotropic harmonic potential)
- *Dimensions* (quantum dots are a few hundred angstroms big)
- *Structure* and source of the potential

The term “quantum dots” refers to finite fermion systems consisting of an artificial 3D confinement of a few electrons, that have a size of only a few hundred angstroms.

Although this definition is quite precise, I prefer a more immediate one: “quantum dots are *artificial atoms*”. Despite its simplicity, this definition contains a good amount of relevant information. Like natural atoms, in fact, quantum dots are made up of electrons confined in an attractive potential; and as one may guess, they show a similar shell-like structure with its relative *magic numbers*.

Such confinement is usually achieved by restricting the two-dimensional electron gas that forms at the interface between two different semiconductor materials (or heterostructure), either laterally or vertically.

However, there are some obvious differences [READ DIFFERENCES](#).

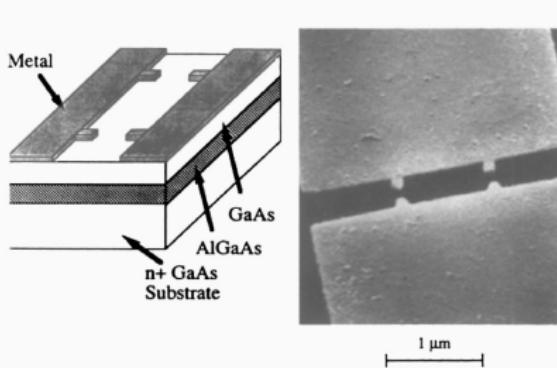
While in natural atoms the attractive potential is generated by a charged nucleus, in quantum dots there is no such a thing.

To understand how the confinement is generated in this case, we have to take a look at the internal structure of such devices.

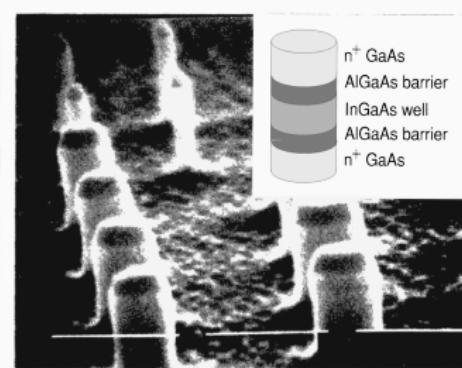
Building technique: stack up different layers of **semiconductor** materials.

Then, an electron gas forms at the interface between two layers.

The electron gas is confined by applying an electric voltage to the whole structure, either laterally or vertically.



(a) Lateral quantum dot.



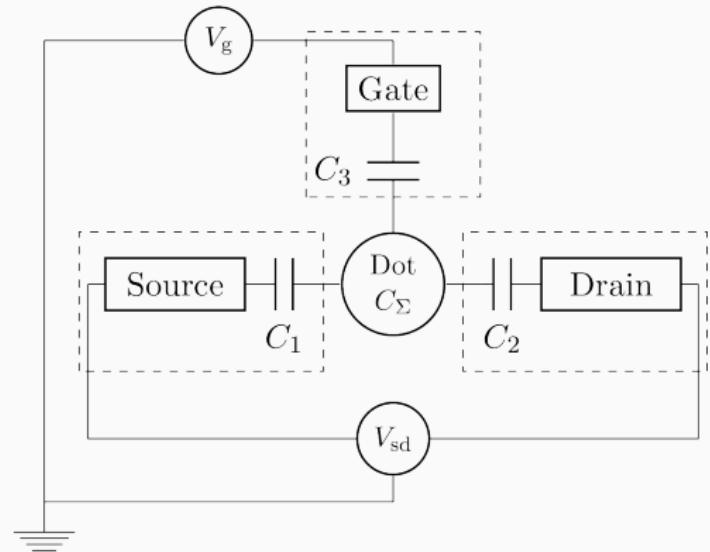
(b) Etched quantum dots.

Quantum dots are built by stacking up different layers of semiconductors. The electron gas forms at the interface of different layers, and the confinement is obtained by applying a voltage to the top metal electrodes – called *gates*. In the Figure, the white bars have a length of 0.5 μm.; so, the gates are created by lithographic patterning.

Another common method to fabricate quantum dots is to build heterostructure pillars by etching techniques. Here, the electrodes are at the top and at the bottom of the pillars.

Other fabricating processes include self-growth mechanisms, in which the growth conditions determine the form of the structure (that can be pyramidal, disk shaped or lens shaped), and cleaved-edge overgrowth, that consists in two separate Molecular Beam Epitaxy growths on a specific substrate.

A quantum dot can be schematized as a *single-electron transistor*. Its charge is a multiple of the elementary charge and transport proceeds one electron at the time.

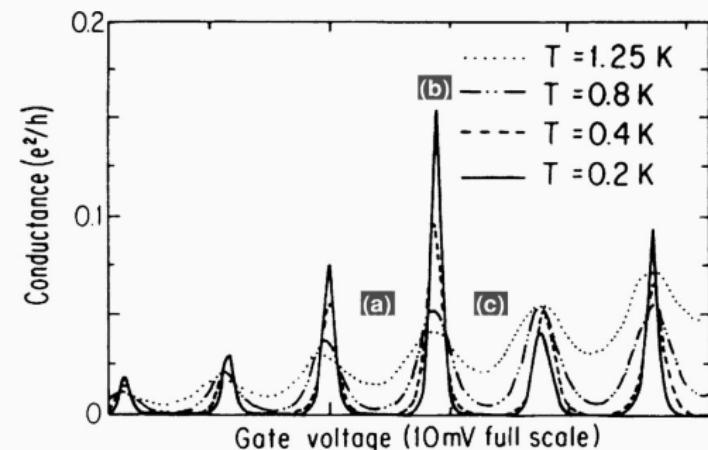


### READ THE SLIDE.

In figure, we have the scheme of a single electron transistor. The island has index 0, the source has index 1, the drain index 2 and the gate index 3. The capacitances are meant to be intrinsic capacitances of the respective electrode (source, drain or gate).

## THE COULOMB BLOCKADE EFFECT

- The electron island is weakly coupled with source and drain via tunnel barriers.
- When the thermal energy is not enough to make an electron tunnel the barriers, the conduction is blocked.
- This effect is called the **Coulomb blockade**, and shows up in *conductance* measurements.



The behavior of a quantum dot can be analyzed by studying its electron transport mechanism. In order to do that, the dot is connected to a circuit that provides a certain gate voltage, and its conductance is measured.

### READ THE SLIDE

A plot like the one in Figure (Coulomb oscillations in a lateral quantum dot) is obtained; you see that there are conductance peaks at specific gate voltages, all of them evenly spaced. This behavior can be explained by the *Coulomb blockade model*.

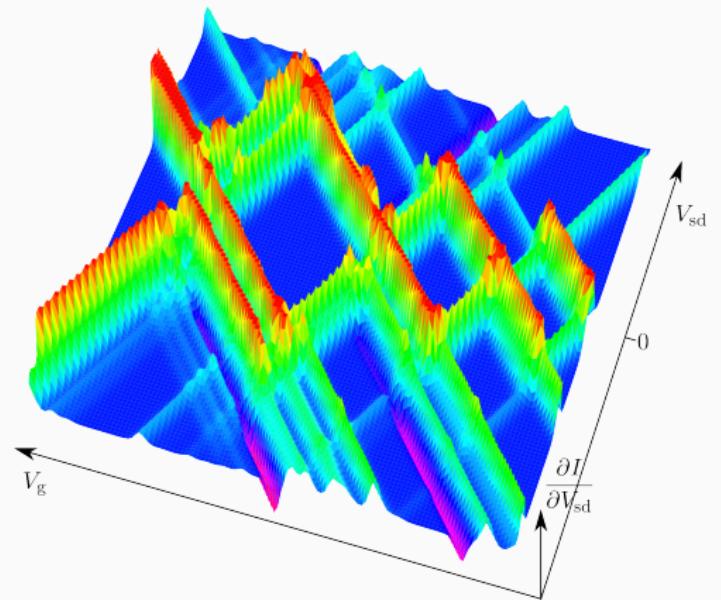
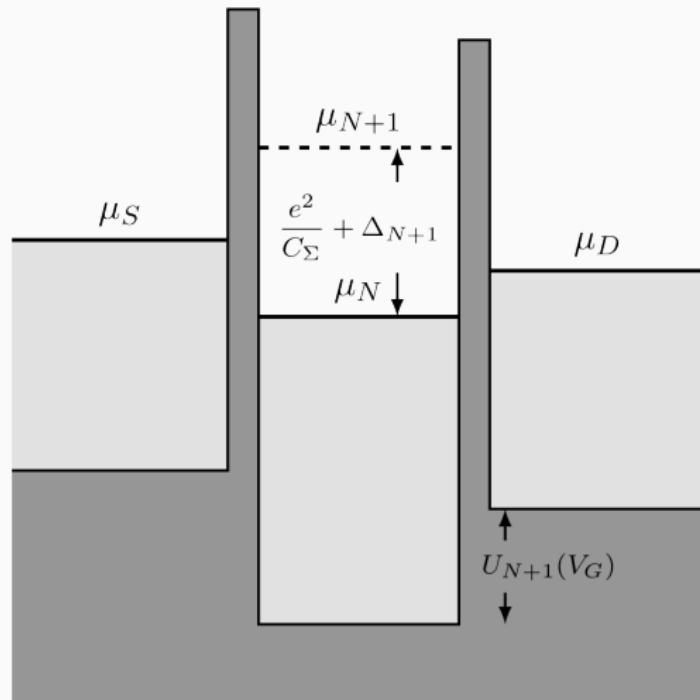


Figure 2: Differential conductance measurements in a nanowire.

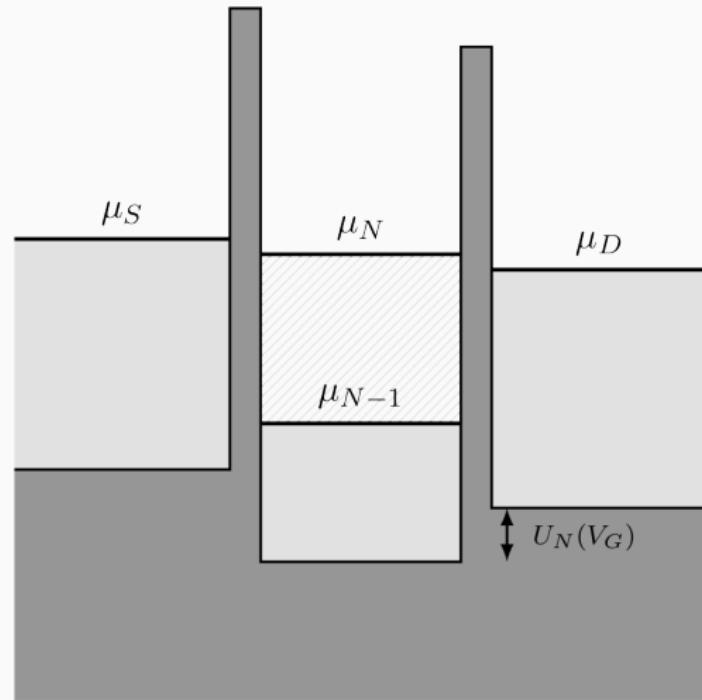
## BAND DIAGRAM



Let's see what happens in terms of the electrochemical potential.

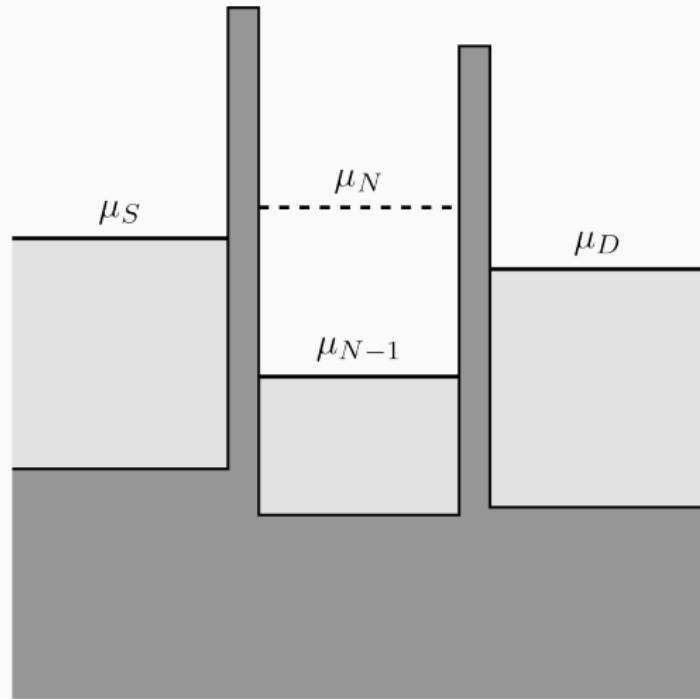
Let's consider the low-temperature, low-bias voltage case ( $eV_{\text{bias}}, k_B T \ll e^2/C_\Sigma$ ). Let's also suppose that, initially, the electrochemical potential inside the dot –  $\mu_N$  – is lower than the electrochemical potential of the drain –  $\mu_D$ : the transport is blocked due to the Coulomb blockade effect.

The number of electrons inside the dot is  $N$ .



Now, we can decrease the gate voltage: the overall effect is that we have a linear increase in  $\mu_N$  wrt the gate voltage. Eventually,  $\mu_N$  will align to  $\mu_D$  ( $\mu_N = \mu_D$ ), which means that an electron can leave the dot. At the same time, if  $\mu_S \gtrsim \mu_D$ , an electron from  $\mu_S$  can enter the dot; the overall effect is that the number of the electrons inside the dot oscillates between  $N$  and  $N - 1$ , and an electric current flows through the dot from the drain to the source. In terms of the conductance, this corresponds in a peak.

## BAND DIAGRAM



If we keep lowering the gate voltage – i.e., increasing  $\mu_N$  – until  $\mu_N > \mu_S$ , then the dot is left with only  $N - 1$  electrons. Now, the level to consider is  $\mu_{N-1}$ : you see that it's far below  $\mu_D$ , so the current is blocked again.

## THE ALGORITHM

---

The **variational principle** is a powerful tool that allows, for *any* system, to calculate an *upper bound* estimate for the *ground state* energy.

Being  $E_{\text{gs}}$  the ground-state energy,  $|\psi\rangle$  a state whatsoever and  $\hat{H}$  the Hamiltonian of the system, the principle states that

$$E_{\text{gs}} \leq \frac{\langle \psi | \hat{H} | \psi \rangle}{\langle \psi | \psi \rangle}. \quad (1)$$

We are going to use this principle by calculating the quantity on the right hand side, for a chosen *trial wave-function*  $\psi_T$ . We will come back later on how to guess a realistic  $\psi_T$ .

The Variational Principle is a method of general validity that can be used to gather information about a system with a Hamiltonian that we are unable to diagonalize. Specifically, this principle gives us an *upper bound* for the energy of the ground state. The formulation is really simple: if you pick *any* state  $|\psi\rangle$  whatsoever, then **SHOW FORMULA ON SLIDE**.

In practice, one chooses a class of states for  $|\psi\rangle$  parametrized by one or more parameters – the so-called *variational parameters*, and then calculates the quantity  $E_T$  for multiple sets of values of the variational parameters  $\alpha_1, \dots, \alpha_n$ . The lower value of  $E_T$  obtained in this way is the required upper bound. If one manages to find an even lower value for  $E_T$  with a more clever wave-function, then he has found a better upper bound for the ground state energy.

The only trouble with this method is that we never know for sure how close we are to the *actual* ground state energy; all we get for sure is just an upper bound. However, if one manages to guess a *realistic*  $|\psi\rangle$ , he often gets values for the ground state energy that miraculously match the actual ones.

Calculating  $E_T$  as it appears in the previous relation is not a piece of cake. However, we can recast that equation in a simpler form as [READ THE SLIDE](#).

In the function-representation of the states, we can recast the above relation as

$$E_{\text{gs}} \leq \int d\vec{\tau} \mathcal{P}(\vec{\tau}) E_L(\vec{\tau}) \simeq \frac{1}{n} \sum_{i=1}^n E_L(\vec{\tau}), \quad (2)$$

where

$$\mathcal{P}(\vec{\tau}) = \frac{|\psi_T|^2}{\int d\vec{\tau} |\psi_T|^2} \quad \text{and} \quad E_L(\vec{\tau}) = \frac{1}{\psi_T} \hat{H} \psi_T. \quad (3)$$

The quantity  $E_L$  is called the [local energy](#).

Let's do a brief recap.

- Our upper bound is given by  $\int d\vec{\tau} \mathcal{P}(\vec{\tau}) E_L(\vec{\tau})$
- We can calculate it as  $\frac{1}{n} \sum_{i=1}^n E_L(\vec{\tau}_i)$  for a big enough  $n$
- In other words, the idea is to sample tons of  $E_L$  values at different positions in space and then take the mean
- Since the  $E_L$ 's have a certain distribution, we should sample more points where the probability is higher and less points where the probability is lower, in order to be sure to have a consistent set of samples.
- But how can we achieve such a distribution of samples?

The solution is to use the Metropolis algorithm.

Let's suppose that we have just sampled  $E_L$  at a certain point in space, let's call it  $\vec{r}^{\text{old}}$ . Then, we make a random move to another point in space,  $\vec{r}^{\text{new}}$ .

To check whether we moved to a higher probability region, we calculate the ratio of the respective probabilities, that is

$$R \doteq \frac{\mathcal{P}(\vec{r}^{\text{new}})}{\mathcal{P}(\vec{r}^{\text{old}})} = \frac{|\psi(\vec{r}^{\text{new}})|^2}{|\psi(\vec{r}^{\text{old}})|^2}. \quad (4)$$

Then

- If  $R \geq 1$ , we accept this move because we are moving to a higher probability region and we sample  $E_L$  in the new position.
- If  $R < 1$ , we can't blindly reject this move just because we are moving to a lower probability region; after all, we also have to populate the tails of the distribution! However, we can't accept all of this kind of moves.

*Solution:* generate a random number  $r \in (0, 1)$ , and accept the move (i.e., sample  $E_L$ ) if  $r \leq R$ . Otherwise, the sample is rejected.

This procedure is known as the **brute force** Metropolis algorithm.

Let's recap it:

- Choose a starting position  $\bar{r}^{\text{old}}$  and a step-length  $l$ .
- Generate a random number  $\varepsilon$  in the interval  $(0, 1)$  and compute a new position  $\bar{r}^{\text{new}} = \bar{r}^{\text{old}} + l\varepsilon$ .
- Compute the acceptance ratio  $R = \frac{|\psi_T(\bar{r}^{\text{new}})|^2}{|\psi_T(\bar{r}^{\text{old}})|^2}$ .
- Generate a new random number  $r$  in the interval  $(0, 1)$ .
- If  $R \geq r$ , accept the step and store the position by letting  $\bar{r}^{\text{old}} = \bar{r}^{\text{new}}$ .
- If  $R < r$ , reject the step and discard the position by letting  $\bar{r}^{\text{new}} = \bar{r}^{\text{old}}$ .

So far, we haven't specified any rule for the choice of the step length. [READ LIST.](#)

**Problem:** we have *no rule* to say what a proper step length should be.

- If it's *too much*, our walker will make huge jumps all around and will sample very few points inside our distribution; the measurement will not be good.
- If it's *too less*, our walker will wander around the starting point and will not cover all the distribution space; the measurement will not be good as well.

**Solution:** as a *rule of thumb*, one can test different step lengths and then choose the one that gives an acceptance of around 50 %.

A clever approach is to implement the **importance sampling**.

*Idea:* use the knowledge about the system to improve the acceptance ratio  $R$ .

Procedure:

- Recast the Schrödinger equation a *diffusion problem*, in which the walker is pushed in those regions where the trial wave-function is larger.
- The force  $\vec{F}$  responsible for pushing the walker is called the *quantum force*, and can be expressed as

$$\vec{F} = 2 \frac{1}{\psi_T} \nabla \psi_T \quad (5)$$

An improved acceptance ratio wastes less points. A better approach is to implement the *Importance sampling*, that consists in doing more clever assumptions on the shape of the transition probability.

Since we know the system we are working on, we know that if we recast the Schrödinger equation as a diffusion problem, the transition probability is going to be a Gaussian distribution (or a modified Gaussian).

The driving force of this diffusion problem is called the *quantum force*, and can be expressed as **SHOW THE FORMULA**.

- The new position is calculated as

$$r^{\text{new}} = r^{\text{old}} + D\vec{F}(r^{\text{old}})\Delta t + \varepsilon, \quad (6)$$

where  $\Delta t$  is a chosen time-step and  $D = 0.5$  is the diffusion constant.

- The improved acceptance ratio can be expressed as

$$R = \frac{|\psi_T(r^{\text{new}})|^2 G(r^{\text{old}}, r^{\text{new}}, \Delta t)}{|\psi_T(r^{\text{old}})|^2 G(r^{\text{new}}, r^{\text{old}}, \Delta t)}. \quad (7)$$

where  $G(x, y, \Delta t)$  is the Green's function

$$G(y, x, \Delta t) = \frac{1}{(4\pi D \Delta t)^{3N/2}} \exp(-(y - x - D\Delta t F(x))^2 / 4D\Delta t). \quad (8)$$

### READ THE NEW POSITION.

Here, the random variable  $\varepsilon$  is no more uniform! For importance sampling, it's a *Gaussian* random variable.

[CONTINUE READING.](#)

## THE 2-ELECTRONS SYSTEM

---

Hamiltonian of the system:

$$\hat{H}_0 = \sum_{i=1}^2 \left( -\frac{1}{2} \nabla_i^2 + \frac{1}{2} \omega^2 r_i^2 \right), \quad (9)$$

Energy of a single electron in a 2D harmonic potential:

$$E_s = \hbar\omega(n_x + n_y + 1). \quad (10)$$

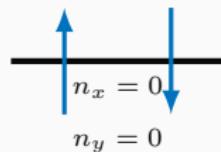
Let's apply this algorithm to two electrons in an harmonic oscillator, without the electron-electron repulsion.

The unperturbed Hamiltonian is [SHOW THE HAMILTONIAN](#).

Unlike the one-dimensional case, for two dimensions we need two quantum numbers:  $n_x$  and  $n_y$  for the  $x$  and  $y$  directions, respectively, and this causes a degeneracy in the energy levels. The principal quantum number is  $n = n_x + n_y$ .

For a single electron in a 2D harmonic potential, the energy can be expressed as [SHOW THE ENERGY](#).

Configuration:



Ground-state energy of the system, in natural units:

$$E_{\text{gs}} = 2\omega. \quad (11)$$

Exact wave-function:

$$\phi(\vec{r}_1, \vec{r}_2) = C \exp(-\omega(r_1^2 + r_2^2)/2). \quad (12)$$

This is the configuration of the system.

You see that the spatial parts of the wave-functions are equal because of the spin degeneracy, so the ground state energy of the complete system has a factor of 2 to take this thing into account.

For the same reason, the exact wave-function (or at least, its spatial part) can be expressed as the product of the single-particle wave-functions.

$C$  is a normalization constant.

Since we know the exact wave-function, the problem is trivial. We choose our trial wave-function as

$$\psi_T(\vec{r}_1, \vec{r}_2) = \exp(-\alpha\omega(r_1^2 + r_2^2)/2). \quad (13)$$

where  $\alpha$  is a **variational parameter**.

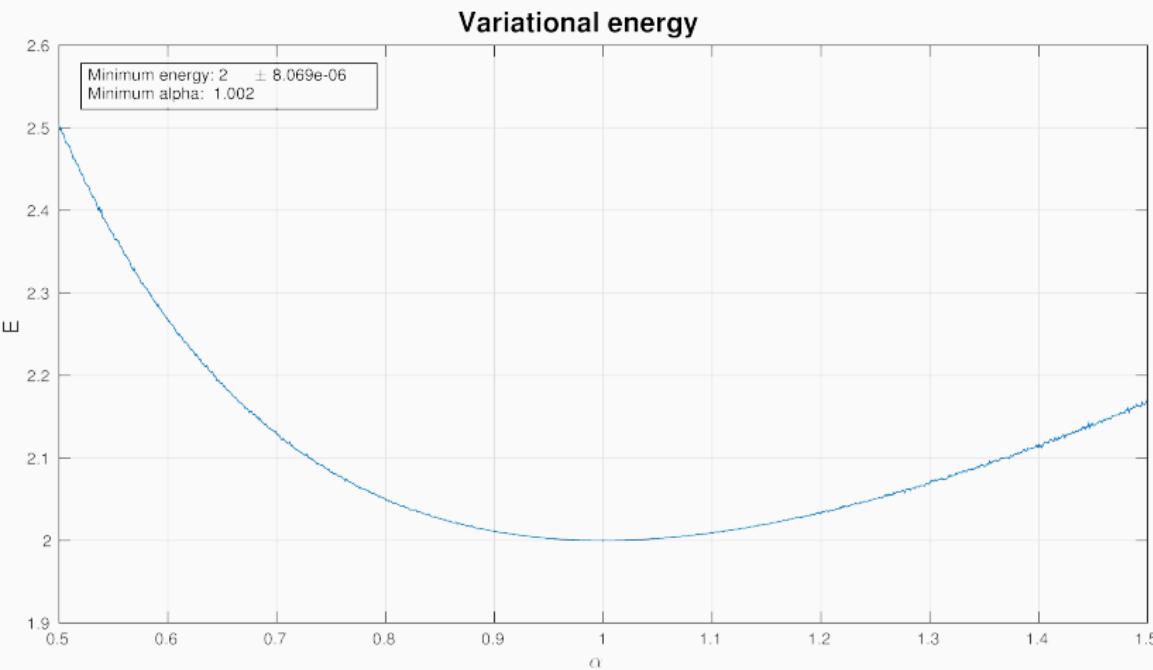
For  $\alpha = 1$  we expect to obtain a variational energy **exactly** equal to the ground state energy  $E_{\text{gs}} = 2\omega$ .

Since we know the exact wave-function, the problem is trivial.

We model our trial wave-function by adding a variational parameter  $\alpha$ .

For  $\alpha = 1$  we expect to obtain a variational energy **exactly** equal to the ground state energy  $E_{\text{gs}} = 2\omega$ .

In other words, for  $\alpha = 1$  the upper bound given by the variational principle coincides with the *true* ground state energy.



The variational energy versus the variational parameter  $\alpha$ . The settings used are: brute force sampling with step length 2, no Jastrow factor, no parallelization, 1000 variations of  $\alpha$  around 1 with step 0.001,  $1 \times 10^7$  Monte Carlo steps. Acceptance ratio varies from 40 to 60 %.

The full Hamiltonian is

$$\hat{H} = -\frac{1}{2}\nabla_1^2 - \frac{1}{2}\nabla_2^2 + \frac{1}{2}\omega r_1^2 + \frac{1}{2}\omega r_2^2 + \frac{1}{|\vec{r}_1 - \vec{r}_2|}. \quad (14)$$

**Problem:** the term

$$\frac{1}{|\vec{r}_1 - \vec{r}_2|} \quad (15)$$

could be a division by zero.

**Solution:** add, in the wave-function, a factor that cancels the divergence in the Hamiltonian.

The Hamiltonian of the full system, that includes the electron-electron repulsion, has an extra term that diverges when the two particles are near to each other.

The divergence can be canceled if we are smart in the choice of the wave-function, by adding a term that compensates the singularity.

The extra factor is usually modeled as

$$\exp\left(\frac{ar}{1+\beta r}\right) \quad (16)$$

that is called the **Padé-Jastrow** factor.

$r$  is the inter-particle distance,  $a$  is a parameter that depends on the spin, while  $\beta$  is a *variational parameter*.

For  $\omega = 1$  we get the upper bound  $E = 3$ , that is compatible with the *actual* ground state energy.

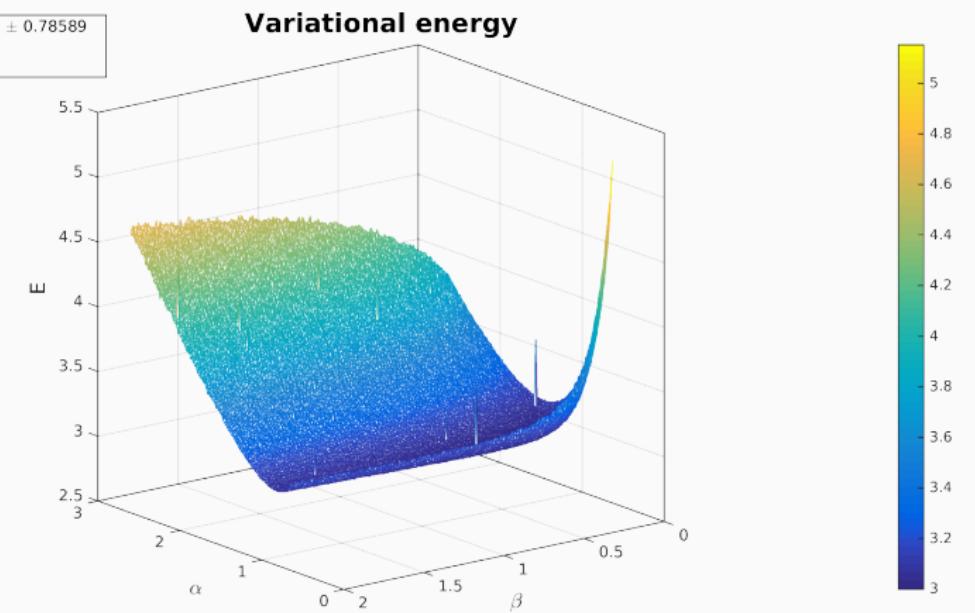
An extra factor to the non-interacting wave-function is added, that has the form **READ THE SLIDE**.

In two dimensions

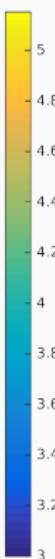
$$a = \begin{cases} 1 & \text{anti-parallel spin} \\ 1/3 & \text{parallel spin} \end{cases} \quad (17)$$

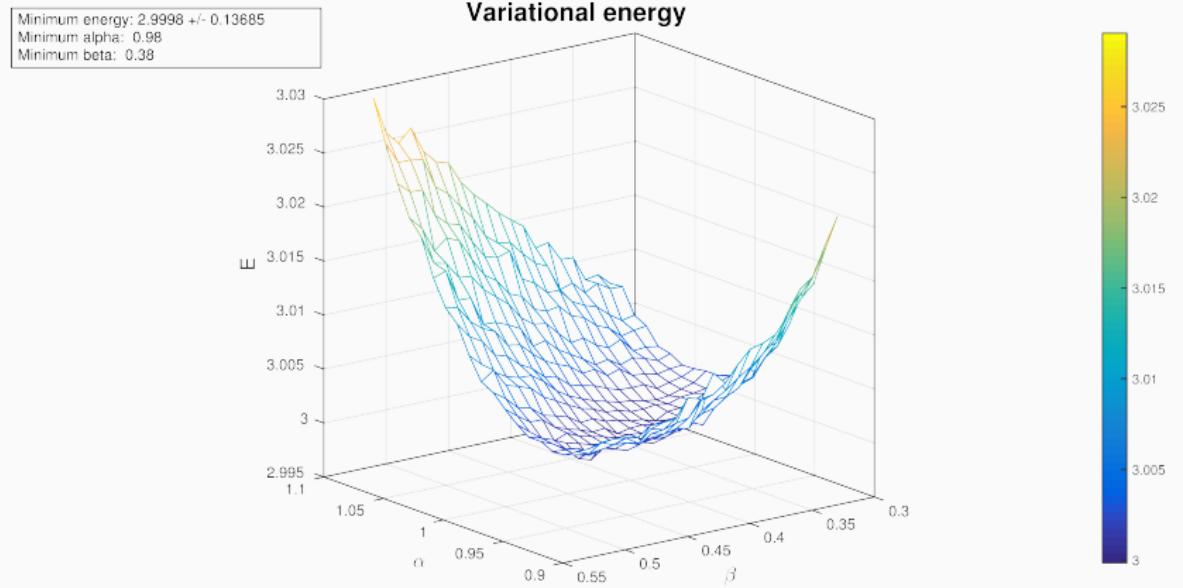
The calculated upper bound coincides with the ground state energy calculated by Taut theoretically.

Minimum energy:  $2.9985 \pm 0.78589$   
Minimum alpha: 1.02  
Minimum beta: 0.36



The variational energy versus the variational parameters  $\alpha$  and  $\beta$ . The settings used are: brute force sampling with step length 2, Jastrow factor, no parallelization, 200 variations of  $\alpha$  and  $\beta$  with step 0.01,  $1 \times 10^5$  Monte Carlo steps. Acceptance ratio varies from 45 to 55 %.





The variational energy versus the variational parameters  $\alpha$  and  $\beta$ . The settings used are: brute force sampling with step length 2, Jastrow factor, no parallelization, 20 variations of  $\alpha$  and  $\beta$  with step 0.01,  $1 \times 10^6$  Monte Carlo steps. Acceptance ratio varies from 45 to 55 %. The relative distance is about 1.64 in natural units.

*Idea:* take multiple measurements for each point in order to lower the error.

If we take  $n$  measurements of the same quantity, its error falls like

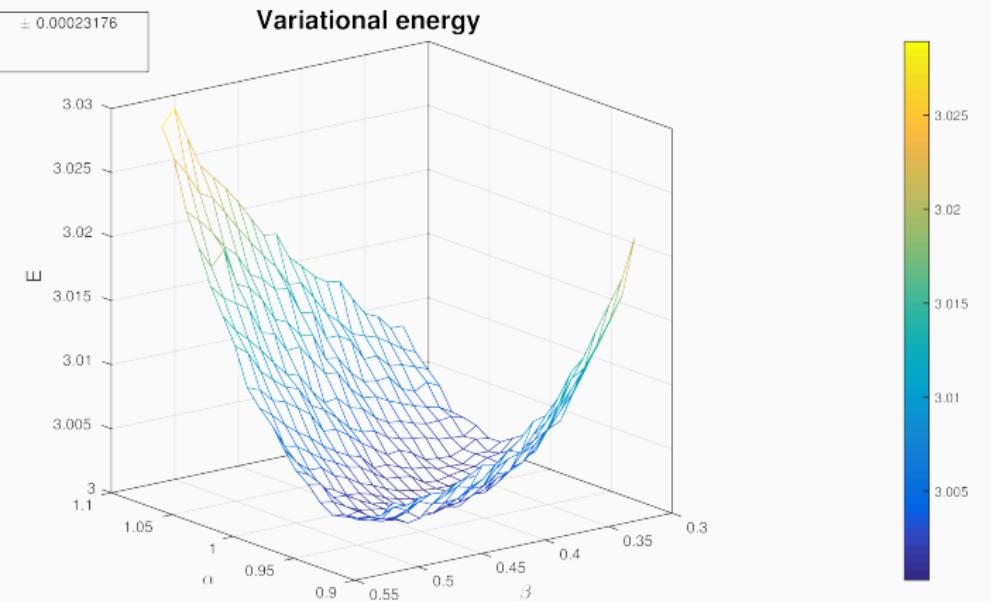
$$\frac{1}{\sqrt{n}}$$

We obtain more precise energy values even with less Monte Carlo steps.

Despite the fact that we have very good values, you see that the error is still quite large.

We can improve it by taking multiple measurements at each point of the grid, and then take the mean. If we do that, the error falls like the  $1/\sqrt{n}$  and we can obtain more precise energy values even with less Monte Carlo steps.

Minimum energy: 3.0003 ± 0.00023176  
Minimum alpha: 0.99  
Minimum beta: 0.4



The variational energy versus the variational parameters  $\alpha$  and  $\beta$ . The settings used are: brute force sampling with step length 2, Jastrow factor, parallelization (8 threads), 20 variations of  $\alpha$  and  $\beta$  with step 0.01,  $2 \times 10^5$  Monte Carlo steps. Acceptance ratio varies from 45 to 55 %.

*Idea:* implement importance sampling in order to waste less points.

*Problem:* if the time-step is too small and we don't use enough Monte Carlo steps, we have bad data.

*Reason:* the walker remains near its starting point and will not cover completely the integration space.

*Solution:*

- For a given time-step, increase the number of Monte Carlo cycles until the acceptance ratio is slightly below 100 %.
- For a given number of Monte Carlo cycles, increase the time-step until the acceptance ratio is slightly below 100 %.

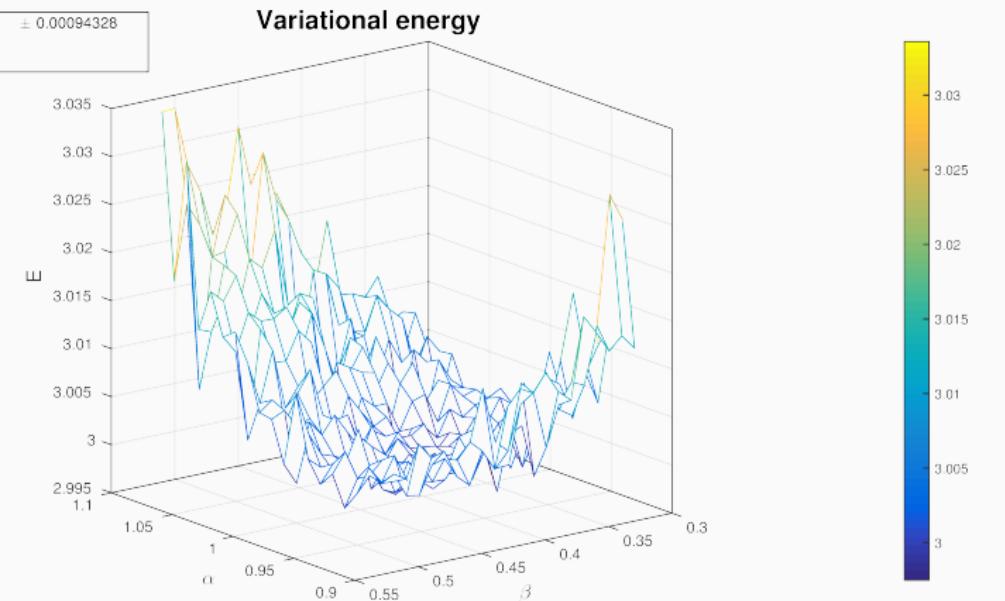
A further improvement would be to implement the importance sampling, as explained in the previous section.

The only thing we have to care about is the amount of the time-step. [READ THE SLIDE](#).

The solution is to test different values for the time-step and use the one that gives an acceptance ratio that is slightly below 100 %.

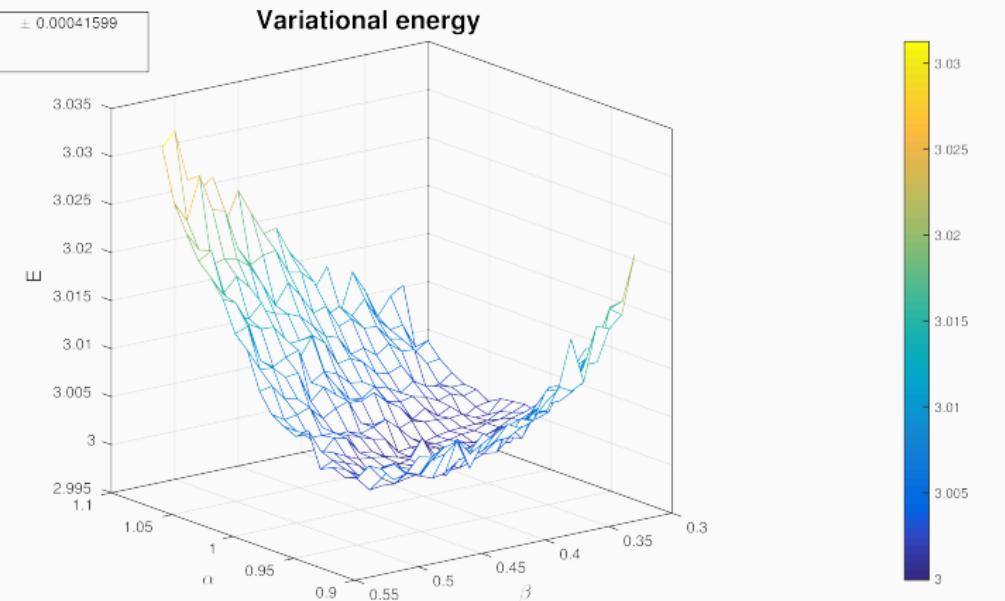
$\omega = 1$ 

Minimum energy: 2.9975 ± 0.00094328  
Minimum alpha: 0.95  
Minimum beta: 0.39



The variational energy versus the variational parameters  $\alpha$  and  $\beta$ . The settings used are: importance sampling with  $\Delta t = 0.001$ , Jastrow factor, parallelization (8 threads), 20 variations of  $\alpha$  and  $\beta$  with step 0.01,  $2 \times 10^5$  Monte Carlo steps. Acceptance ratio is about 99.999 %.

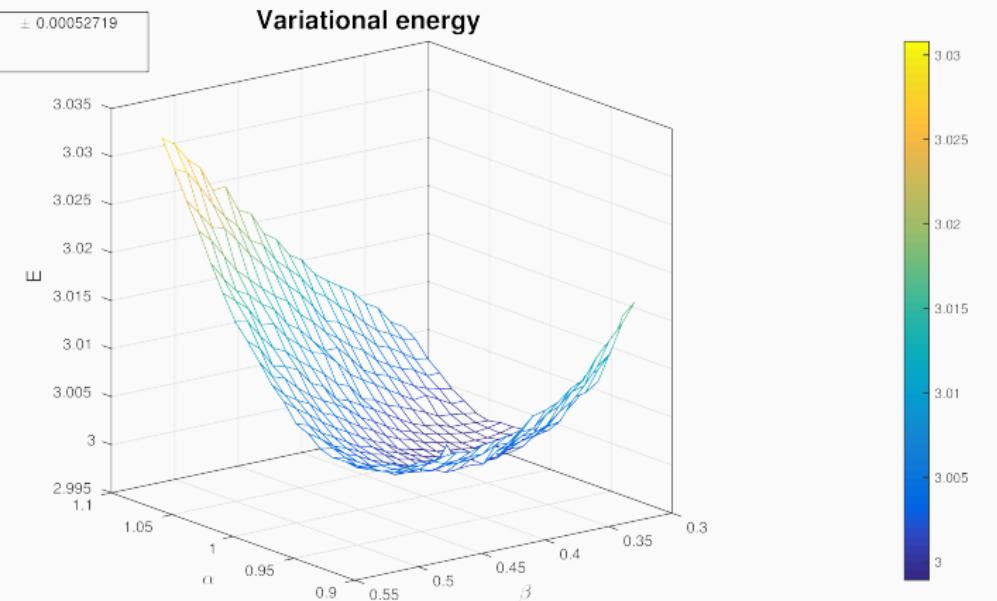
Minimum energy: 2.9999 ± 0.00041599  
Minimum alpha: 1.01  
Minimum beta: 0.39



The variational energy versus the variational parameters  $\alpha$  and  $\beta$ . The settings used are: importance sampling with  $\Delta t = 0.01$ , Jastrow factor, parallelization (8 threads), 20 variations of  $\alpha$  and  $\beta$  with step 0.01,  $2 \times 10^5$  Monte Carlo steps. Acceptance ratio is about 99.999 %.

$\omega = 1$ 

Minimum energy: 2.9989 ± 0.00052719  
Minimum alpha: 1  
Minimum beta: 0.36



The variational energy versus the variational parameters  $\alpha$  and  $\beta$ . The settings used are: importance sampling with  $\Delta t = 0.1$ , Jastrow factor, parallelization (8 threads), 20 variations of  $\alpha$  and  $\beta$  with step 0.01,  $2 \times 10^5$  Monte Carlo steps. Acceptance ratio is about 99.793 %.

## THE 6-ELECTRONS SYSTEM

---

The trial wave-function is like the one we used for 2 electrons, but this time is more general.

For 2 electrons, it was

$$\psi_T(\vec{r}_1, \vec{r}_2) = \phi(\vec{r}_1)\phi(\vec{r}_2) \exp\left(\frac{ar_{12}}{1 + \beta r_{12}}\right) \quad (18)$$

where  $\phi$  is the single-particle spatial wave-function for  $(n_x, n_y) = (0, 0)$ , that is equal for both electron 1 and electron 2, and  $r_{12}$  is the inter-particle distance.

Now, we can extend the previous reasoning to the 6-electrons case, trying to discuss it in the most possible general way.

Recall that, for 2-electrons, the wave-function had this form. That is, it was the product among the single-particle wave-functions and a repulsion factor called the Jastrow factor.

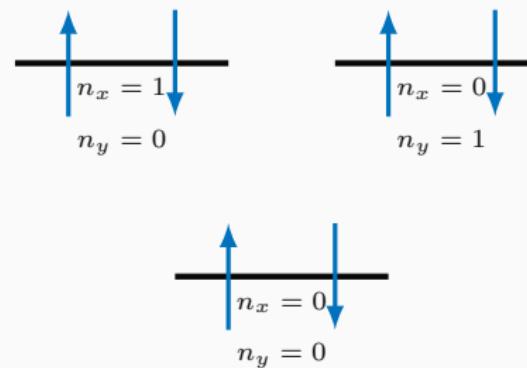
For  $N$  particles the generalization is immediate, as you can see. This matrix  $S$  is named the Slater matrix.

For  $N$  electrons, the generalization is straightforward:

$$\psi_T(\vec{r}_1, \dots, \vec{r}_N) = |S| \prod_{i < j}^N \exp\left(\frac{ar_{ij}}{1 + \beta r_{ij}}\right) \quad (19)$$

where  $|S|$  is the **Slater determinant**.

$$|S| = \begin{vmatrix} \phi_1(\vec{r}_1) & \phi_2(\vec{r}_1) & \dots & \phi_N(\vec{r}_1) \\ \phi_1(\vec{r}_2) & \phi_2(\vec{r}_2) & \dots & \phi_N(\vec{r}_2) \\ \vdots & & & \vdots \\ \phi_1(\vec{r}_N) & \phi_2(\vec{r}_N) & \dots & \phi_N(\vec{r}_N) \end{vmatrix} \quad (20)$$



**Smart thing:** for spin-independent Hamiltonians, the Slater determinant can be split in a product of **two** Slater determinants, one for the single-particle orbitals with spin up and the other for single-particle orbitals with spin down.

Now, let's take a look at the configuration of the system. You see that we are free to label the particles in whatever way we want; for example, we can say that (**INDICATE ALL SPIN UP**) this is particle 1, this is particle 2 and this is particle 3, and the others are numbered accordingly. If we do that, we can split the Slater determinant into two smaller determinants; one for spin-up particles, and one for spin-down particles.

**SHOW NEXT SLIDE WITHOUT COMMENTS.**

We can arbitrarily choose that the first  $N/2$  particles are spin-up and the other half are spin-down. This way, our trial wave-function can be written as

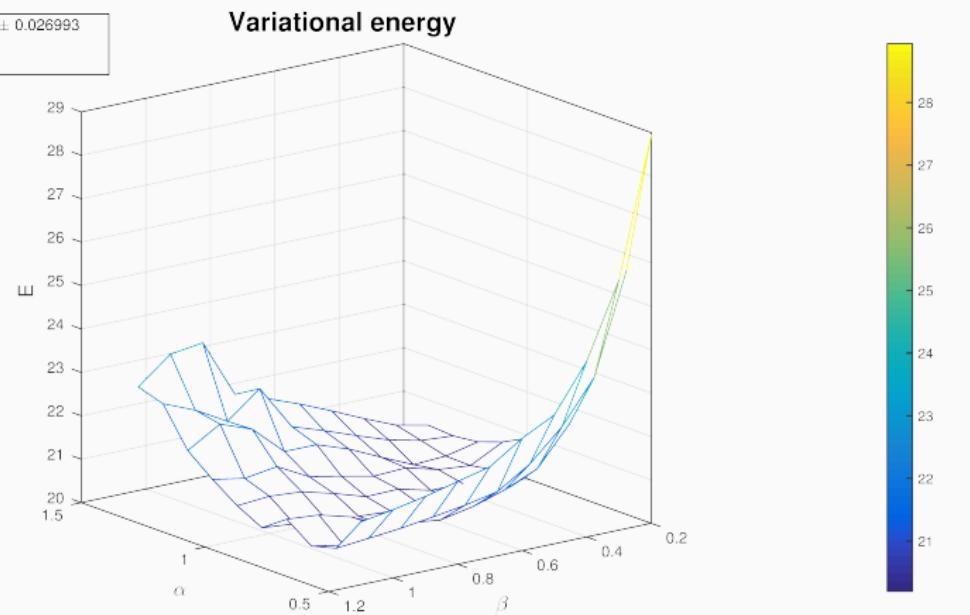
$$\psi_T(\vec{r}_1, \dots, \vec{r}_N) = |S^\uparrow||S^\downarrow|J, \quad (21)$$

where (and similarly for  $|S^\downarrow|$ )

$$S^\uparrow = \begin{vmatrix} \phi_1(\vec{r}_1) & \phi_2(\vec{r}_1) & \dots & \phi_{N/2}(\vec{r}_1) \\ \phi_1(\vec{r}_2) & \phi_2(\vec{r}_2) & \dots & \phi_{N/2}(\vec{r}_2) \\ \vdots & & & \vdots \\ \phi_1(\vec{r}_{N/2}) & \phi_2(\vec{r}_{N/2}) & \dots & \phi_{N/2}(\vec{r}_{N/2}) \end{vmatrix} \quad (22)$$

and  $J$  is the Jastrow factor.

Minimum energy: 20.1947  $\pm$  0.026993  
Minimum alpha: 0.9  
Minimum beta: 0.6



The variational energy versus the variational parameters  $\alpha$  and  $\beta$ . The settings used are: importance sampling with  $\Delta t = 0.1$ , Jastrow factor, parallelization (8 threads), 10 variations of  $\alpha$  and  $\beta$  with step 0.1,  $2 \times 10^5$  Monte Carlo steps.  $\omega = 1.00$ .

Our result is in very good accordance with the one calculated by DMC, that is  $20.1597(2)$  a.u..

## ANALYTICAL DERIVATIVES

---

*Idea:* speed up the calculation of the acceptance ratio by performing analytical derivatives.

Remember that

$$\psi_T = |S^\uparrow||S^\downarrow|J \quad (23)$$

and

$$\vec{F}_i = 2\frac{1}{\psi_T}\nabla_i\psi_T, \quad E_L = \frac{1}{2}\sum_i\frac{1}{\psi_T}\nabla_i^2\psi_T + \sum_i V_i. \quad (24)$$

The most time-consuming part of the Metropolis-Hastings algorithm is the computation of the acceptance ratio, the kinetic energy and the quantum force, that involve the calculation of several first and second derivatives. A way to – possibly – improve our program is to use analytical derivatives. Note that this is possible in very few cases – like this one, so we will take advantage of it.

Remember that... [READ](#)

Doing the calculations

$$\frac{1}{\psi_T} \nabla_i \psi_T = \frac{\nabla_i |S^\alpha|}{|S^\alpha|} + \frac{\nabla_i J}{J} \quad (25)$$

and

$$\frac{1}{\psi_T} \nabla_i^2 \psi_T = \frac{\nabla_i^2 |S^\alpha|}{|S^\alpha|} + \frac{\nabla_i^2 J}{J} + 2 \frac{\nabla_i |S^\alpha|}{|S^\alpha|} \frac{\nabla_i J}{J}, \quad (26)$$

where  $\alpha$  is the spin relative to particle  $i$ .

Further simplifications give

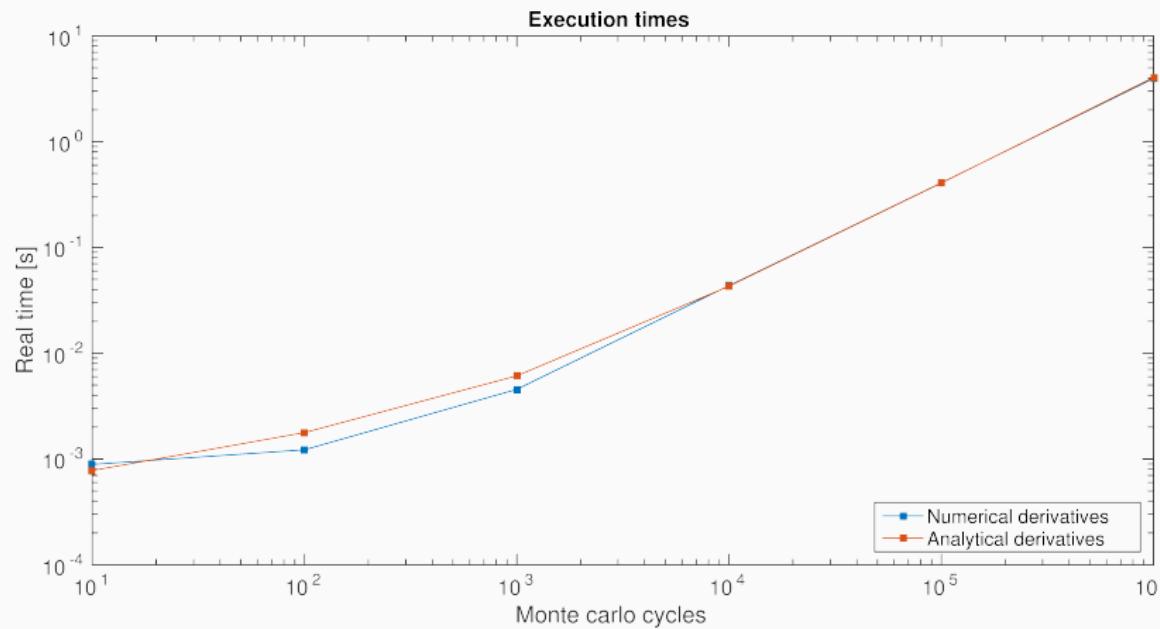
$$\frac{\nabla_i J}{J} = \sum_{k \neq i=1}^N \frac{a_{ik}}{r_{ik}} \frac{\vec{r}_i - \vec{r}_k}{(1 + \beta r_{ik})^2} \quad (27)$$

$$\frac{\nabla_i^2 J}{J} = \left| \frac{\nabla_i J}{J} \right|^2 + \sum_{k \neq i=1}^N a_{ik} \frac{(d-3)(\beta r_{ik} + 1) + 2}{r_{ik}(\beta r_{ik} + 1)^3} \quad (28)$$

$$\frac{\nabla_i |S|}{|S|} = \sum_k \nabla_i \phi_k(\vec{r}_i^{\text{new}}) (S_{ki}^{\text{new}})^{-1} \quad (29)$$

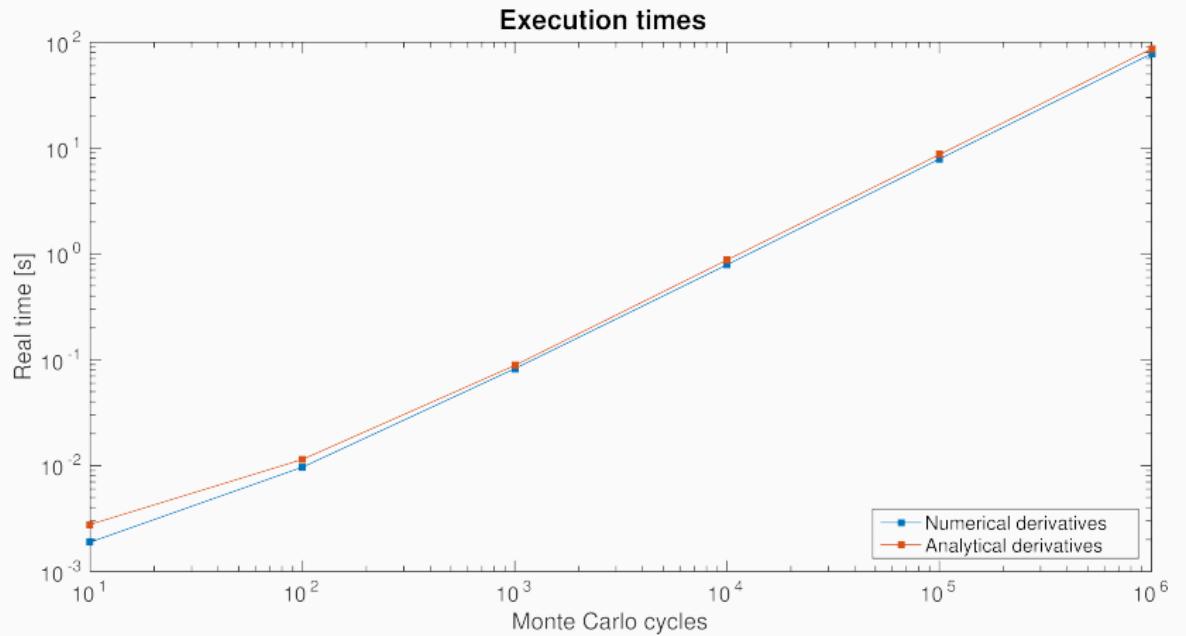
$$\frac{\nabla_i^2 |S|}{|S|} = \sum_k \nabla_i^2 \phi_k(\vec{r}_i^{\text{new}}) (S_{ki}^{\text{new}})^{-1} \quad (30)$$

## EXECUTION TIMES – 2 ELECTRONS



Execution times for 2-electrons and a single pair of variational parameters  $(\alpha, \beta)$ . The GNU/Linux system tool `time` was used to take the measurements.

## EXECUTION TIMES – 6 ELECTRONS



Execution times for 6-electrons and a single pair of variational parameters  $(\alpha, \beta)$ . The GNU/Linux system tool `time` was used to take the measurements.

As you see, we don't have great improvements; on the contrary, we obtain a code that is even a little bit slower than the numerical one for the 6-electrons case! However, there are some things that have to be pointed out. First of all we decided to perform a numerical inversion of a matrix, that is really time-consuming (it's a  $\mathcal{O}(n^3)$  operation). Secondly there is no optimization of the acceptance ratio, that can be further simplified . Lastly, the benefits of numerical derivatives could not appear until one reaches 12 or even 20 particles; this is because the computational complexity of a determinant of a matrix is  $\mathcal{O}(n^3)$ , that for  $3 \times 3$  matrices (as the 6-electrons case) is still acceptable. The struggle in calculating big determinants all the time is more evident for higher numbers of particles; for the 12-particles case the Slater determinant is 8 times slower than the 6-particles case, while for the 20-particles case is approximately 37 times slower.

- We don't really have great improvements.
- We are performing a *numerical inversion of a matrix* ( $\sim \mathcal{O}(n^3)$ ).
- There is no optimization of the acceptance ratio.
- The determinant of a matrix has a computational complexity of  $\sim \mathcal{O}(n^3)$ , so we might not see improvements up to 12-20 particles.

## CONCLUSION

---

- The VMC technique is an easy-to-implement tool that gives an upper bound to the ground state energy of a system.
- Although is used as a starting point for DMC calculations, if we choose our trial wave-function in a smart way this technique gives an upper bound that is surprisingly compatible with the actual ground-state energy.
- For  $N$  interacting electrons in an attractive potential, a good approximation of the actual wave-function is

$$\psi_T(\vec{r}_1, \dots, \vec{r}_N) = |S^\uparrow||S^\downarrow|J$$

- For a quantum dot, the calculations can be further simplified by analytical means.
- In general, such a simplification is not possible; what one does is to optimize the code and to parallelize it as much as possible via OpenMP/MPI and possibly GPU parallelization.
- **Thanks for your attention!**

QUESTIONS?