Department of Applied Physics, Chalmers, Göteborg
Göran Wahnström, Dan Fors, 2015-11-24

H2b Variational Monte Carlo

Here you are asked to determine the ground state energy for the helium atom using the variational Monte Carlo technique. It is an application of the Metropolis algorithm. Atomic units ($\hbar = e = m_e = 4\pi\epsilon_0 = 1$) are used and denoted a.u.. The problem can be solved using `MATLAB` but substantial computational time is required to obtain accurate results.

The variational Monte Carlo method

Consider a system described by the Hamiltonian $\mathcal{H}$. The variation theorem then states that for some arbitrary (trial) wave function Ψ_T the expectation value of the energy

$$E[\Psi_T] = \frac{\langle \Psi_T \mid \mathcal{H} \mid \Psi_T \rangle}{\langle \Psi_T \mid \Psi_T \rangle} \geq E_0 \tag{1}$$

is always larger or equal to the ground state energy E_0. Only if Ψ_T is equal to the ground state wave function Ψ_0, the expectation value is equal to E_0.

If the system contains many particles the calculation of the expectation value of the energy involves integrals over many degrees of freedom. For a system of N interaction particles in 3 dimensions the corresponding integrals

$$E[\Psi_T] = \frac{\int d\mathcal{R}\Psi_T^*(\mathcal{R})\mathcal{H}\Psi_T(\mathcal{R})}{\int d\mathcal{R}\Psi_T^*(\mathcal{R})\Psi_T(\mathcal{R})} \tag{2}$$

become $3N$ dimensional. Here, $\mathcal{R}$ denotes the combined coordinate of all N particles, $\mathcal{R} = (\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_N)$, and $\Psi_T(\mathcal{R}) = \Psi_T(\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_N)$. The high-dimesional integrals may be evaluated using Monte Carlo based techniques and, hence, the method is called *Variational Monte Carlo* [1].

In realistic systems the many-particle wave function assumes very small values in large part of the configuration space, so a straightforward procedure using homogeneously distributed random points in configuration space is bound to fail. This suggests that it might be efficient to use the Metropolis algorithm in which the sampling of points in configuration space is pushed towards those regions where the wave function assumes appreciable values. We can rewrite eqn (2) as

$$E[\Psi_T] = \int d\mathcal{R}E_L(\mathcal{R})\rho(\mathcal{R}) \tag{3}$$

where

$$E_L(\mathcal{R}) = \frac{\mathcal{H}\Psi_T(\mathcal{R})}{\Psi_T(\mathcal{R})} \tag{4}$$

is called the *local energy* and

$$\rho(\mathcal{R}) = \frac{\mid \Psi_T(\mathcal{R}) \mid^2}{\int d\mathcal{R} \mid \Psi_T(\mathcal{R}) \mid^2} \tag{5}$$

is a normalized probability distribution for the coordinate $\mathcal{R}$, which can be used as the weight function in the Metropolis algorithm. The local energy $E_L(\mathcal{R})$ is a function that depends on the coordinates for all the particles. If $\Psi_T(\mathcal{R})$ is close to the ground state wave function, the local energy will depend weakly on $\mathcal{R}$, and if $\Psi_T(\mathcal{R}) \equiv \Psi_0(\mathcal{R})$ it is constant and equal to the ground state energy E_0.

We can now compute the expectation value of the energy using the Metropolis algorithm and sample configurations $\mathcal{R}_i$ according to the probability distribution $\rho(\mathcal{R})$. The energy is approximated as

$$E = \langle E_L(\mathcal{R}) \rangle \simeq \frac{1}{N} \sum_{i=1}^{N} E_L(\mathcal{R}_i) \tag{6}$$

and an estimate of the error is obtained by determining the variance

$$\sigma^2 = \left\langle [E_L(\mathcal{R}) - \langle E_L(\mathcal{R}) \rangle]^2 \right\rangle \tag{7}$$

In the variational method one then tries to optimize the wave function $\Psi_T(\mathcal{R})$ in order to minimize the energy. If $\Psi_T(\mathcal{R})$ is parametrized with a set of S parameters $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_S)$, the expectation value becomes an ordinary function of $\boldsymbol{\alpha}$

$$E(\boldsymbol{\alpha}) = \frac{\langle \Psi_T \mid \mathcal{H} \mid \Psi_T \rangle}{\langle \Psi_T \mid \Psi_T \rangle} \tag{8}$$

More efficient minimization techniques are available if the gradient of $E(\boldsymbol{\alpha})$ is known. In the present case, with stochastic fluctuations, accurate values for the gradient are difficult to obtain with straightforward finite differencing. However, by differenting eqn (8) we obtain the following expression in terms of the analytical derivative of the trial wave function

$$\nabla_{\boldsymbol{\alpha}} E(\boldsymbol{\alpha}) = 2 \left[\langle E_L(\mathcal{R}) \nabla_{\boldsymbol{\alpha}} \ln \Psi_T(\mathcal{R}) \rangle - \langle E_L(\mathcal{R}) \rangle \langle \nabla_{\boldsymbol{\alpha}} \ln \Psi_T(\mathcal{R}) \rangle \right] \tag{9}$$

where we have assumed a real trial wave function, $\Psi_T(\mathcal{R}) = \Psi_T^*(\mathcal{R})$. The averages in eqn (9) can be esimated in the usual way in the Metropolis method.

The gradient can then be used in the stochastic gradient technique [2], a damped steepest descent method, according to

$$\boldsymbol{\alpha}_{p+1} = \boldsymbol{\alpha}_p - \gamma_p \nabla_{\boldsymbol{\alpha}} E(\boldsymbol{\alpha}_p) \tag{10}$$

where γ_p is a scaling factor. It is an iterative method and p is used to denote the iteration number. A suitable form for the scaling factor is

$$\gamma_p = A p^{-\beta} \tag{11}$$

where the exponent β should be in the range $0.5 < \beta \leq 1$ [2, 3] and A is related to the inverse of the Hessian. By iterating, $\boldsymbol{\alpha}_p$ may approach the correct value [2, 3] and the scaling factor γ_p has then an important role in averaging out the stochastic fluctuations of the gradient $\nabla_{\boldsymbol{\alpha}} E(\boldsymbol{\alpha}_p)$.

Other techniques minimize the variance instead of the energy [4] and often the Hessian is computed to increase the efficiency [4].

Trial wave function

The trial wave function should approximate the true ground state wave function as closely as possible, in particular, it should fulfill the symmetry conditions. The standard trial wave function for a many-electron system is hence written on the form

$$\Psi_T(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N) = \Psi_{AS}(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N) \exp\left[\frac{1}{2}\sum_{i,j=1}^{N} f(r_{ij})\right] \tag{12}$$

where $\boldsymbol{x}_i = (\boldsymbol{r}_i, s_i)$ denotes the position $\boldsymbol{r}_i$ and spin s_i of electron i. Ψ_{AS} is the Slater determinant composed of spin-orbitals and and $f(r_{ij})$ is a function which contains two-particle correlation effects and $r_{ij} = |\boldsymbol{r}_j - \boldsymbol{r}_i|$. The trial wave function should also fulfill the boundary conditions, the so called cusp conditions. They can be written on the form (in atomic units)

$$\lim_{r_{ij}\to 0} \frac{1}{\Psi_T}\frac{\partial \Psi_T}{\partial r_{ij}} = \frac{1}{4} \quad \text{if } s_i = s_j \tag{13}$$

$$\lim_{r_{ij}\to 0} \frac{1}{\Psi_T}\frac{\partial \Psi_T}{\partial r_{ij}} = \frac{1}{2} \quad \text{if } s_i \neq s_j \tag{14}$$

$$\lim_{r_{in}\to 0} \frac{1}{\Psi_T}\frac{\partial \Psi_T}{\partial r_{in}} = -Z_n \tag{15}$$

where $\boldsymbol{R}_n$ and Z_n denote the position and atomic number for nucleus n, respectively, and $r_{in} = |\boldsymbol{R}_n - \boldsymbol{r}_i|$. These conditions imply that the leading singularity in Coulomb interaction, when two particles come close together, cancels when evaluating the local energy E_L.

The Helium atom

The Born-Oppenheimer Hamiltonian for the many-electron system reads (in atomic units)

$$\mathcal{H} = -\frac{1}{2}\sum_i \nabla_i^2 - \sum_n \sum_i \frac{Z_n}{|\boldsymbol{R}_n - \boldsymbol{r}_i|} + \frac{1}{2}\sum_i \sum_{j(\neq i)} \frac{1}{|\boldsymbol{r}_j - \boldsymbol{r}_i|} \tag{16}$$

which for the helium atom (with the nucleus in origo) simplifies to

$$\mathcal{H} = -\frac{1}{2}(\nabla_1^2 + \nabla_2^2) - \frac{2}{r_1} - \frac{2}{r_2} + \frac{1}{r_{12}} \tag{17}$$

A very crude solution can be obtained by totally neglect the Coulomb interaction $1/r_{12}$ between the electrons. Each electron then experiences an unscreened nuclear potential with charge $Z = 2$. The corresponding (normalized) wave function can be written on the form

$$\Psi(\boldsymbol{r}_1, \boldsymbol{r}_2) = 2^{1/2}\varphi(r_1)\varphi(r_2) \tag{18}$$

with

$$\varphi(r) = \frac{2^{5/2}}{\sqrt{4\pi}} \exp\left[-2r\right] \tag{19}$$

and $E_0 = -4$ a.u..

We will now take into account the Coulomb interaction between the two electrons. This introduces correlation effects. We write the trial wavefunction for the ground state on the standard form

$$\Psi_T(\boldsymbol{r}_1, \boldsymbol{r}_2) = \phi(r_1)\phi(r_2)f(r_{12}) \tag{20}$$

We make the ansatz

$$\phi(r) = \exp\left[-\alpha_1 r\right] \tag{21}$$

$$f(r) = \exp\left[\frac{\alpha_2 r}{1+\alpha_3 r}\right] \tag{22}$$

The cusp conditions imply that $\alpha_1 = 2$ and $\alpha_2 = 1/2$, i.e.

$$\Psi_T(\boldsymbol{r}_1, \boldsymbol{r}_2) = \exp\left[-2r_1\right]\ \exp\left[-2r_2\right]\ \exp\left[\frac{r_{12}}{2(1+\alpha r_{12})}\right] \tag{23}$$

and we are left with only one variational parameter $\alpha = \alpha_3$. It controls the distance over which the trial wave function correlates the two electrons. After some algebra the following expression for the local energy can be derived

$$\begin{aligned} E_L(\boldsymbol{r}_1, \boldsymbol{r}_2) = -4 + \frac{(\hat{\boldsymbol{r}}_1 - \hat{\boldsymbol{r}}_2)\cdot(\boldsymbol{r}_1 - \boldsymbol{r}_2)}{r_{12}(1+\alpha r_{12})^2} \\ -\frac{1}{r_{12}(1+\alpha r_{12})^3} - \frac{1}{4(1+\alpha r_{12})^4} + \frac{1}{r_{12}} \end{aligned} \tag{24}$$

where $\hat{\boldsymbol{r}}$ denote a unit vector along $\boldsymbol{r}$.

Task

1. Consider the helium atom. Implement the Metropolis algorithm and use the ansatz in eqn (23) for the trial wave function. Use the value $\alpha = 0.1$.

   You have to define how to generate a new configuration. The simpliest is to choose one of the six coordinates $(\boldsymbol{r}_1, \boldsymbol{r}_2)$ at random and displace

it symmetrically. You can also choose to displace all six coordinates at each step. Introduce a parameter d to control the size of the symmetric displacement(s). Determine an appropriate value for d by monitoring the acceptance-rejection ratio.

To convince yourself that the program is working correctly, it is instructive to investigate the distribution of the sampled points. You can then compare with the model in eqn (18) for which

$$P(r) = 2^5 r^2 \exp\left[-4r\right]$$

In the model in eqn (18) the Coloumb interaction is neglected. Neglect the Coloumb interaction in your simulation and plot the probability density $P(r)$ to find an electron at the distance r from the nucleus. Do you obtain the above the result?

If you now add the Coloumb interaction and plot the probability density $P(r)$, do you then obtain a deviation from the above result. Notice, to clearly identify the deviation you have to perform sufficiently number of iterations. Could you explain qualitatively the deviation?

Determine also the ground state energy by averaging the local energy along the Markov chain. You should obtain about $E_0 = -3$ a.u.. The model without Coloumb interaction gives $E_0 = -4$ a.u.. The Coloumb interaction is, however, not weak and amounts to about the energy 1 a.u.. Do you obtain a reasonable number? (4p)

2. In a Monte Carlo calculation using the Metropolis algorithm a large number of steps, $N_{tot} = N_{eq} + N$, are generated. The first N_{eq} steps is discarded, the system has to equilibrate, and the subsequent N steps are used to compute different quantities. The configurations are correlated and the statistical inefficiency, the correlation length, has to be determined in order to be able to make a proper estimate of the error.

   Consider again the case $\alpha = 0.1$ and determine the local energy at each step. Investigate the number of steps N_{eq} required to equilibrate the system by monitoring the variation of the local energy. When N_{eq} is reached no direct drift of the local energy should be observable.

   Determine next the statistical inefficiency using both the correlation function method and block averaging. (3p)

3. You should now have a correct working program. Perform simulations for several different α-values in the range $0.05 < \alpha < 0.25$. Determine the energy with error bars. For the latter, you may use the statistical inefficency you obtained above, i.e. you may neglect the (weak) α-dependence of the statistical inefficiency.

You may also perform several independent calculations (for the same α-value) to verify that you get results independent on the starting configuration. In some cases one can get stuck in local minima, favourable regions in configuration space, and the sampling of configuration space will not be correct. Independent runs are then a useful test. The information from the different runs can also be used to reduce the error bars.

What optimum value do you obtain for the energy? What is the corresponding α-value? (4p)

4. Variational Monte Carlo is an optimization method. It is possible to adjust the parameter α directly in the simulation. In the case of a large set of variational parameters $\boldsymbol{\alpha}_p$ this becomes essential.

   Perform such a study using the algorithm in eqn (10). For γ_p use the expression in eqn (11) with A=1 and investigate a few different values for the exponent β. Do you converge to the same α value? Is your result for α consistent with what you obtained above? (4p)

5. You should now have obtained an optimum value for α. Perform a careful simulation using this α-value and compare your result with the experimental number -2.9033 a.u.. (1p)

References

[1] J. M. Thijssen, *Computational Physics* (2nd ed., Cambridge Univ. Press, 2007).

[2] A. Harju, B. Barbiellini, S. Siljamäki, R. M. Nieminen, G. Ortiz, Phys. Rev. Lett. **79**, 1173 (1997).

[3] A. Harju, S. Siljamäki, R. M. Nieminen, Phys. Rev. B **65**, 075309 (2002).

[4] C. J. Umrigar and C. Filippi, Phys. Rev. Lett. **94**, 150201 (2005).