Göran Wahnström

QUANTUM STRUCTURE
Lecture notes

Göteborg, 29 November 2021

Quantum structure

To determine the quantum structure, the wavefunction and the corresponding energies, is a computationally intensive task for many-particle systems in quantum mechanics. Here we will consider the electronic structure problem, a set of interacting electrons in an electrostatic field generated by stationary nuclei. We will introduce a few different methods and apply them to the most simple non-trivial problem, the ground-state energy for the helium atom. The emphasis will be on stochastic quantum Monte Carlo techniques, which are aiming at accurate solutions for many-electron systems.

Contents

1 The structure of matter

We can image a piece of matter as a collection of interacting atoms. This collection of particles may be in the gas phase (molecules, clusters, ...) or in a condensed phase (crystalline or amorphous solids, liquids, ...) and it could be either homogeneous or heterogeneous as molecules in solution and adsorbates on surfaces. However, at the atomic scale, we can unambiguously describe all these systems as a set of atomic nuclei and electrons interacting solely via electrostatic forces [1, 2]. Formally, we can write the Hamiltonian of such a system in the following general form

$$
\begin{aligned}
\mathcal{H} = & -\sum_{i=1}^{N} \frac{\hbar^2}{2m_e} \nabla_i^2 - \sum_{I=1}^{P} \frac{\hbar^2}{2M_I} \nabla_I^2 + \frac{1}{4\pi\epsilon_0} \frac{1}{2} \sum_{i=1}^{N} \sum_{j\neq i}^{N} \frac{e^2}{|\boldsymbol{r}_i - \boldsymbol{r}_j|} \\
& - \frac{1}{4\pi\epsilon_0} \sum_{I=1}^{P} \sum_{i=1}^{N} \frac{Z_I e^2}{|\boldsymbol{R}_I - \boldsymbol{r}_i|} + \frac{1}{4\pi\epsilon_0} \frac{1}{2} \sum_{I=1}^{P} \sum_{J\neq I}^{P} \frac{Z_I Z_J e^2}{|\boldsymbol{R}_I - \boldsymbol{R}_J|}
\end{aligned}
\tag{1}
$$

where $\boldsymbol{r}_i$ is the electronic coordinate for electron i and $\boldsymbol{R}_I$ the nuclear coordinate for nucleus I. $Z_I e$ and M_I are the the nuclear charges and masses, respectively, and m_e the electronic mass. The first two terms represent the kinetic energy for the electrons and nuclei, respectively, the third term the Coulomb repulsion between the electrons and the fourth term the Coulomb attraction between the electrons and nuclei. Finally, the last term contains the Coulomb repulsion between the nuclei.

1.1 The electronic structure problem

The Hamiltonian in Eqn (1) and the solution of the corresponding Schrödinger equation represents a formidable problem in computational physics. One therefore tries to simplify the problem. An important basic approximation that is often made is the so called *adiabatic* or *Born-Oppenheimer approximation*. It is based on the fact that the nuclei are much heavier than the electrons (the mass of a proton is 1836 times as large as the electron mass) and hence in a classical picture the nuclei move much more slowly than the electrons. The latter will then be able to essentially instantaneously adjust themselves to the current configuration of the nuclei and the problem is separated into an electronic structure problem and a nuclear motion problem.

This separation leads to a Hamiltonian for the electrons in the field generated by a static configuration of the nuclei, the Born-Oppenheimer Hamiltonian,

$$
\mathcal{H}_{BO} = -\sum_{i=1}^{N} \frac{\hbar^2}{2m_e} \nabla_i^2 + \frac{1}{4\pi\epsilon_0} \frac{1}{2} \sum_{i=1}^{N} \sum_{j\neq i}^{N} \frac{e^2}{|\boldsymbol{r}_i - \boldsymbol{r}_j|} - \frac{1}{4\pi\epsilon_0} \sum_{I=1}^{P} \sum_{i=1}^{N} \frac{Z_I e^2}{|\boldsymbol{R}_I - \boldsymbol{r}_i|}
\tag{2}
$$

The corresponding time-independent Schrödinger equation reads

$$\mathcal{H}_{BO} \ \Psi_n(\{\hat{\boldsymbol{r}}\}; \{\boldsymbol{R}\}) = E_n(\{\boldsymbol{R}\}) \ \Psi_n(\{\hat{\boldsymbol{r}}\}; \{\boldsymbol{R}\}) \tag{3}$$

where we have used the notation $\{\hat{\boldsymbol{r}}\}$ for the set of the N electronic coordinates and $\{\boldsymbol{R}\}$ for the set of P nuclear coordinates. We have also indicated with the *hat*-notation that the electronic coordinates are operators while the nuclear coordinates are ordinary variables, they are treated on a classical footing. The nuclear coordinates only enter as an external potential in the Born-Oppenheimer Hamiltonian. Still, the *electronic structure problem*, defined by the Hamiltonian in Eqn (2), is very demanding for many-electron systems. The main obstacle is the two-body nature of the Coulomb interaction that makes the Schrödinger equation not separable. The statistics also has to be taken into account. The electrons are fermions and therefore, the wave-function $\Psi(\{\hat{\boldsymbol{r}}\}; \{\boldsymbol{R}\})$ has to be anti-symmetric with respect to interchange of two electronic coordinates. The electronic coordinate is described by both its position coordinate and its spin coordinate.

1.2   The nuclear motion problem

The Born-Oppenheimer approximation also leads to a separate Schrödinger equation for the nuclei in which the electronic energy enters as a potential. Often one makes the further approximation that the motion of the nuclei can be treated using classical mechanics [1]. In that case the time-dependent Schrödinger equation simplifies to Newton's equation of motion

$$M_I \frac{d^2}{dt^2} \boldsymbol{R}_I(t) = -\nabla_I V_{\text{pot}}(\{\boldsymbol{R}\}) \ , \ I = 1, \dots, P \tag{4}$$

where the interaction potential is given by

$$V_{\text{pot}}(\{\boldsymbol{R}\}) = E_0(\{\boldsymbol{R}\}) + \frac{1}{4\pi\epsilon_0} \frac{1}{2} \sum_{I=1}^{P} \sum_{J\neq I}^{P} \frac{Z_I Z_J e^2}{|\boldsymbol{R}_I - \boldsymbol{R}_J|} \tag{5}$$

The first term $E_0(\{\boldsymbol{R}\})$ is the electronic energy defined as an eigenvalue in Eqn (3) and the second term is the nuclear interaction, present in Eqn (1) but left out in Eqn (2). We have here assumed that the nuclear motion occurs on the ground-state potential energy surface for the electronic degrees of freedom, $E_0(\{\boldsymbol{R}\})$.

Numerical solution of Eqn (4) for a many-particle system is denoted *molecular dynamics (MD) simulation*. If the interaction potential $V_{\text{pot}}(\{\boldsymbol{R}\})$ is obtained by solving the electronic structure problem, the eigenvalue problem in Eqn (3), it is denoted *first-principles* or *ab-initio molecular dynamics*. This is computationally very demanding. Often one tries to construct some simplified model for $V_{\text{pot}}(\{\boldsymbol{R}\})$, a parametrized force-field. The technique

then becomes much more efficient and the dynamics of large many-particle systems can be investigated. This is an important technique in computational physics and it is widely used. This technique was introduced and discussed in the lecture notes "Molecular Dynamics".

2   Model system - Helium

To illustrative various electronic structure calculation methods we will use the helium atom, a two-electron system. This can viewed as the most simple non-trivial system but it can be used to illustrative important computational methods as Hartree-Fock, Variational and Diffusion Monte Carlo. In atomic scale calculations it is very convenient to use atomic units. We will denote these by a.u., and they are defined in Sec. A.

The helium atom consists of two electrons and a nucleus with two protons. The Born-Oppenheimer Hamiltonian in Eqn (2) then takes the form

$$\mathcal{H}_{He} = -\frac{1}{2}(\nabla_1^2 + \nabla_2^2) - \frac{2}{r_1} - \frac{2}{r_2} + \frac{1}{r_{12}} \tag{6}$$

in atomic units, where $r_{12} = |\boldsymbol{r}_1 - \boldsymbol{r}_2|$, and where we have assumed the nucleus to be located at origo. If we neglect the inter-electron repulsion, the term $1/r_{12}$-term in the Hamiltonian, the problem is separable and easily solved. With the inter-electron Coulomb repulsion included, it becomes much more difficult to solve.

Furthermore, the electrons are fermions and the wave-function has to be anti-symmetric with respect to interchange of the two electronic coordinates. These are described by both position coordinates $\boldsymbol{r}_i$ and spin coordinates s_i. The ground-state for helium is a singlet and the spin part takes care of the antisymmetry and hence, for helium in its ground state we have the symmetry relation

$$\Psi(\boldsymbol{r}_1, \boldsymbol{r}_2) = \Psi(\boldsymbol{r}_2, \boldsymbol{r}_1)$$

3 Independent electron approximation

We will first consider the case where the two electrons in helium are independent of each other and the electrostatic repulsion between them are absent. The Hamiltonian for helium in Eqn (6) can then be written as

$$\mathcal{H}_{He} = \mathcal{H}_1 + \mathcal{H}_2 \tag{7}$$

where

$$\mathcal{H}_i = -\frac{1}{2}\nabla_i^2 - \frac{Z}{r_i} \quad \text{with } Z = 2 \quad \text{and } i = 1,2 \tag{8}$$

The ground-state wave-function and energy for the one-electron Hamiltonian $\mathcal{H}_i$ can be obtained analytically,

$$\phi(\boldsymbol{r_i}) = \frac{1}{\sqrt{4\pi}}Z^{3/2}2e^{-Zr_i} \quad , \quad E = -\frac{Z^2}{2} \tag{9}$$

The ground-state wave-function for helium can thus be written as

$$\Psi(\boldsymbol{r}_1, \boldsymbol{r}_2) = \phi(\boldsymbol{r}_1)\phi(\boldsymbol{r}_2) \tag{10}$$

and

$$\mathcal{H}_{He}\Psi = (\mathcal{H}_1 + \mathcal{H}_2)\Psi = \phi(\boldsymbol{r}_2)\mathcal{H}_1\phi(\boldsymbol{r}_1) + \phi(\boldsymbol{r}_1)\mathcal{H}_2\phi(\boldsymbol{r}_2) = (E_1 + E_2)\Psi$$

and the ground-state energy for helium (Z=2) becomes equal to

$$E = E_1 + E_2 = -4.0 \text{ a.u.} \tag{11}$$

This crude estimate for the ground-state energy for helium, E=-4.0 a.u.= -108.8 eV differs significantly from the experimental determined value $E = $ -2.903 a.u. = 79.0 eV. This demonstrates that the neglected electron-electron repulsion term makes a large contribution to the helium ground-state energy.

4 Central-field approximation

We will now consider the full Hamiltonian for helium

$$\mathcal{H}_{He} = -\frac{1}{2}\left(\nabla_1^2 + \nabla_2^2\right) - \frac{2}{r_1} - \frac{2}{r_2} + \frac{1}{r_{12}}$$

In the Independent electron approximation (see Sec. 3) we obtained the wave-function

$$\Psi(\mathbf{r}_1, \mathbf{r}_2) = \phi(\mathbf{r}_1)\phi(\mathbf{r}_2) = \frac{Z^3}{\pi}e^{-Z(r_1+r_2)} \tag{12}$$

If we neglect the inter-electron repulsion, the $1/r_{12}$-term in the Hamiltonian $\mathcal{H}_{He}$, this wave-function would have been exact with $Z=2$. If the inter-electron repulsion is included it will only be approximate. However, we can view it as a trial wave-function $\Psi_T(\mathbf{r}_1, \mathbf{r}_2)$ and make use of the variational theorem (see App. B). The expectation value of $\mathcal{H}_{He}$ would then give an upper limit for the true energy. Using the wave-function in Eqn (12) the expectation value can be calculated [3, 4]

$$\begin{aligned}
E &= \langle \mathcal{H}_{He} \rangle = \int d\mathbf{r}_1 \int d\mathbf{r}_2 \psi^*(\mathbf{r}_1, \mathbf{r}_2)\; \mathcal{H}_{He}\; \psi(\mathbf{r}_1, \mathbf{r}_2) \\
&= \frac{Z^2}{2} + \frac{Z^2}{2} - 2Z - 2Z + \frac{5}{8}Z = Z^2 - \frac{27}{8}Z \tag{13}
\end{aligned}$$

Using $Z = 2$ we get E = -2.75 a.u.. This is considerably closer to the experimental value compared with the result when the inter-electron repulsion term is neglected, the Independent electron approximation.

We can improve on this estimate by using also Z as a variational parameter. This corresponds to that the electron effectively does not feel the total charge from the nucleus. The other electron will screen the potential and reduce the effective charge from 2 to say $(2 - \sigma)$

$$V_{CF}(r) = -\frac{2 - \sigma}{r} \tag{14}$$

where $0 < \sigma < 1$. Using the wave-function in Eqn (12) with $Z = (2 - \sigma)$ we get

$$E = (2 - \sigma)^2 - \frac{27}{8}(2 - \sigma) = \sigma^2 - \frac{5}{8}\sigma - \frac{11}{4} \tag{15}$$

This has a minimum

$$E = -\left(\frac{27}{16}\right)^2 = -2.8477 \text{ a.u.}$$

for $\sigma = 5/16$. This value for E becomes even closer to the experimental value, E = -2.903 a.u..

The one-electron density is defined by

$$n(\boldsymbol{r}) = |\phi(\boldsymbol{r})|^2 \tag{16}$$

For the one-electron wave-function in Eqn (9) it is given by

$$n(\boldsymbol{r}) = \frac{1}{4\pi} Z^3 4 e^{-2Zr} \tag{17}$$

This is spherical symmetric. We can then ask for the probability to find the electron at a distance r from origo. This is equal to

$$\rho(r) = 4\pi r^2 \frac{1}{4\pi} Z^3 4 e^{-2Zr} = Z^3 4 r^2 e^{-2Zr} \tag{18}$$

In Fig. 1 we show $\rho(r)$ both for $Z = 2$ and for the optimal value $Z = 27/16$. The result is also compared with the corresponding density obtained from the Hartree method (see Sec. 5).

8

5   Hartree-Fock method

In using the central-field approximation in Sec. 4 for helium the total wave-function was approximated by a product of two one-electron wave-functions or orbitals $\phi(\boldsymbol{r})$

$$\Psi(\boldsymbol{r}_1, \boldsymbol{r}_2) = \phi_1(\boldsymbol{r}_1)\phi_2(\boldsymbol{r}_2) \tag{19}$$

This type of approximation simplifies the general many-electron problem considerably. It is denoted the *independent-electron approximation* and it was first used by Hartree 1928 to describe many-electron atoms [2, 3, 4]. It is the basis for the Hartree and Hartree-Fock methods in quantum mechanics.

Consider again the helium atom

$$\mathcal{H}_{He} = -\frac{1}{2}\left(\nabla_1^2 + \nabla_2^2\right) - \frac{2}{r_1} - \frac{2}{r_2} + \frac{1}{r_{12}}$$

Let now this Hamiltonian act on the ansatz in Eqn (19). We then obtain the Schrödinger equation

$$\left[-\frac{1}{2}\left(\nabla_1^2 + \nabla_2^2\right) - \frac{2}{r_1} - \frac{2}{r_2} + \frac{1}{r_{12}}\right]\phi_1(\boldsymbol{r}_1)\phi_2(\boldsymbol{r}_2) = E\ \phi_1(\boldsymbol{r}_1)\phi_2(\boldsymbol{r}_2) \tag{20}$$

We assume that the one-electron wave-functions $\phi_i(\boldsymbol{r})$ are normalized to unity,

$$\int d\boldsymbol{r}\ |\phi_i(\boldsymbol{r})|^2 = 1\ \ ,\ i = 1,2 \tag{21}$$

If we now multiply Eqn (20) from the left with $\phi_2^*(\boldsymbol{r}_2)$ and then integrating over $\boldsymbol{r}_2$ we get the eigenvalue equation

$$\left[-\frac{1}{2}\nabla_1^2 - \frac{2}{r_1} + \int d\boldsymbol{r}_2\frac{|\phi_2(\boldsymbol{r}_2)|^2}{r_{12}}\right]\phi_1(\boldsymbol{r}_1) = \epsilon_1\phi_1(\boldsymbol{r}_1) \tag{22}$$

where two integrals over $\boldsymbol{r}_2$ yielding constant values (i.e. not dependent on $\boldsymbol{r}_1$) have been adsorbed into ϵ_1. We can also multiply from the left with $\phi_1^*(\boldsymbol{r}_1)$ and then integrating over $\boldsymbol{r}_1$. The same equation is then obtained, but now for electron number 2

$$\left[-\frac{1}{2}\nabla_2^2 - \frac{2}{r_2} + \int d\boldsymbol{r}_1\frac{|\phi_1(\boldsymbol{r}_1)|^2}{r_{12}}\right]\phi_2(\boldsymbol{r}_2) = \epsilon_2\phi_2(\boldsymbol{r}_2) \tag{23}$$

If we now multiply Eqn (20) from left with $\phi_1^*(\boldsymbol{r}_1)\phi_2^*(\boldsymbol{r}_2)$ and integrate over $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$ we obtain the expression

$$E = \epsilon_1 + \epsilon_2 - \int d\boldsymbol{r}_1 d\boldsymbol{r}_2\frac{|\phi_1(\boldsymbol{r}_1)|^2\ |\phi_2(\boldsymbol{r}_2)|^2}{r_{12}} \tag{24}$$

for the energy of helium.

The effective Hamiltonian acting on the orbital for electron 1 in Eqn (22) has the form of a one-electron Hamiltonian with an effective potential

$$V_{eff}(\mathbf{r}_1) = -\frac{2}{r_1} + \int d\mathbf{r}_2 \frac{|\phi_2(\mathbf{r}_2)|^2}{r_{12}} \tag{25}$$

The first term is the Coulomb attraction from the nucleus while the second term, called the Hartree potential,

$$V_H(\boldsymbol{r}_1) \equiv \int d\boldsymbol{r}_2 \frac{|\phi_2(\boldsymbol{r}_2)|^2}{r_{12}} \tag{26}$$

describes the Coulomb repulsion with the other electron, generated through its charge distribution. This implies that the interaction with the other electron has only been taken into account in an approximate way. It is not the actual position of $\boldsymbol{r}_2$ that determines the wave-function for electron 1, but the average charge distribution of electron 2. This approach bears much relation to the *mean field approach* in statistical mechanics and it is the key consequence introduced by the independent electron ansatz in Eqn (19). We also notice that the effect of the Hartree potential is to screen the bare potential from the positive nucleus. The Hartree potential provides an explicit expression for the phenomenological screening parameter σ introduced in Eqn (14).

The two electrons in helium are identical and hence $\phi_1(\mathbf{r}) = \phi_2(\mathbf{r}) \equiv \phi(\mathbf{r})$ and $\epsilon_1 = \epsilon_2 \equiv \epsilon$. The *Hartree theory* for helium can therefore be summarized as:

The Hartree method for helium in atomic units. The Hartree energy for the ground state is obtained by solving the one-electron Schrödinger equation

$$\left[-\frac{1}{2}\nabla^2 - \frac{2}{r} + V_H(\boldsymbol{r}) \right] \phi(\boldsymbol{r}) = \epsilon\phi(\boldsymbol{r}) \tag{27}$$

with the Hartree potential $V_H(\boldsymbol{r})$ given by

$$V_H(\boldsymbol{r}) = \int d\boldsymbol{r}' \frac{n(\boldsymbol{r}')}{|\boldsymbol{r} - \boldsymbol{r}'|} \tag{28}$$

and where $n(\boldsymbol{r}) = |\phi(\boldsymbol{r})|^2$ is the one-electron density. The energy is then given by

$$E = 2\epsilon - \int d\boldsymbol{r} V_H(\boldsymbol{r})n(\boldsymbol{r}) \tag{29}$$

An important aspect for numerical computations of the Hartree energy is that the Hartree potential in Eqn (28) requires the knowledge of the orbital, which is not known before we have solved the one-electron problem

in Eqn (27). However, Eqn (27) contains the unknown Hartree potential and this, therefore, forms a *self-consistency* problem, which has to be solved in an iterative manner. Furthermore, the energy in Eqn (29) is not equal to the sum of the two orbital energies. This is due to that the inter-electron interaction is included twice, both in Eqn (22) and in Eqn (23). To account for this double counting a contribution from the inter-electron interaction is therefore subtracted in the expression for the energy in Eqn (29).

The total wave-function for a many-electron system should be antisymmetric in the electron coordinates. For helium in its groundstate the spin part takes care of the antisymmetry and the ansatz in Eqn (19) is correct. However in general it is crucial to take the antisymmetry into account. That was done in a systematic way by Fock and Slater 1930. This generalisation is known as the *Hartree-Fock theory*. An ansatz for the total wavefunction is made in terms of products of one-electron orbitals. The variational method is used to determine the optimal orbitals with respect to the energy. The obtained ground-state energy can therefore be viewed as the "best" estimate, i.e. the lowest energy, within the independent electron description. For helium one obtains the energy E = -2.8616 a.u. by solving the Eqs (27)-(29) numerically.
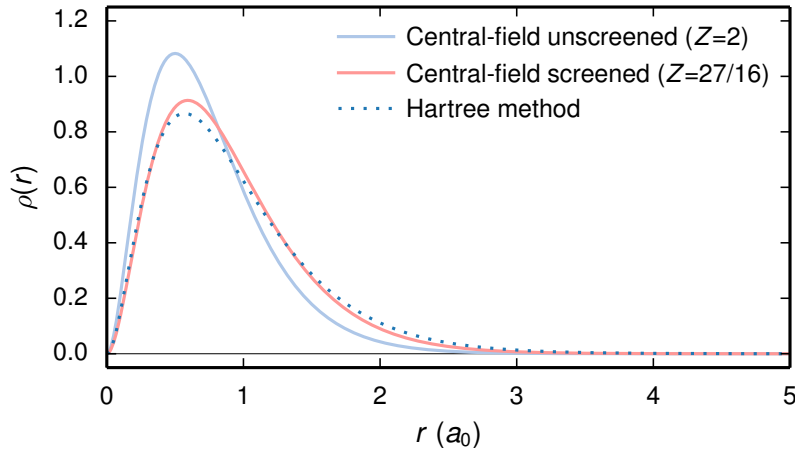

Figure 1: The probability to find an electron at a distance r from origo. The result from three different approximations are shown.

6 Variational Monte Carlo

In the Hartree and Hartree-Fock methods the many-electron wave-function is approximated by products of one-electron wave-functions, orbitals. This implies that the interaction between the electrons has only been taken into account in an approximate way, through self-consistently evaluated orbitals. It is a type of mean-field approach. What is left out is called *correlation*. In the *variational Monte Carlo method* [2, 5, 6, 7] these effects can be included by using approximate many-electron wave-functions. In the general case this implies high-dimensional integrals which can be evaluated using the Metropolis Monte Carlo method. Here, we will again consider the helium atom and we will assume the correlated product

$$\Psi(\boldsymbol{r}_1, \boldsymbol{r}_2) = \phi_1(\boldsymbol{r}_1)\phi_2(\boldsymbol{r}_2)f(r_{12}) \tag{30}$$

instead of the ansatz in Eqn (19).

6.1 Importance sampling

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

$$E[\Psi_\mathsf{T}] = \frac{\int d\mathcal{R}\,\Psi_\mathsf{T}^*(\mathcal{R})\mathcal{H}\Psi_\mathsf{T}(\mathcal{R})}{\int d\mathcal{R}\,\Psi_\mathsf{T}^*(\mathcal{R})\Psi_\mathsf{T}(\mathcal{R})} \geq E_0 \tag{31}$$

is always larger or equal to the ground state energy E_0. Only if $\Psi_\mathsf{T}(\mathcal{R})$ is equal to the ground state wave function $\Phi_0(\mathcal{R})$, the expectation value is equal to E_0. Here, $\mathcal{R}$ denotes the combined coordinates of N electrons, $\mathcal{R} = (\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_N)$, and $\Psi_\mathsf{T}(\mathcal{R}) = \Psi_\mathsf{T}(\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_N)$. If the system contains many electrons the calculation of the expectation value of the energy involves integrals over many degrees of freedom. The high-dimensional integrals may be evaluated using Monte Carlo based techniques and, hence, the method is called *Variational Monte Carlo* [2].

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 importance sampling technique in which the sampling of points in configuration space is increased in the regions where the wave function assumes appreciable values. We can rewrite Eqn (31) as

$$E[\Psi_\mathsf{T}] = \int d\mathcal{R}\, E_L(\mathcal{R})P(\mathcal{R}) \tag{32}$$

where

$$E_L(\mathcal{R}) = \frac{\mathcal{H}\Psi_\mathsf{T}(\mathcal{R})}{\Psi_\mathsf{T}(\mathcal{R})} \tag{33}$$

is called the *local energy* and

$$P(\mathcal{R}) = \frac{|\Psi_\mathsf{T}(\mathcal{R})|^2}{\int d\mathcal{R} \, |\Psi_\mathsf{T}(\mathcal{R})|^2} \tag{34}$$

is a normalized probability distribution, 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 electrons. If $\Psi_\mathsf{T}(\mathcal{R})$ is close to the ground state wave function, the local energy will depend weakly on $\mathcal{R}$, and if $\Psi_\mathsf{T}(\mathcal{R}) \equiv \Phi_0(\mathcal{R})$ 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 a set of configurations $\mathcal{R}_i$ according to the probability distribution $P(\mathcal{R})$. The energy is approximated as

$$E = \langle E_L(\mathcal{R}) \rangle \simeq \frac{1}{N} \sum_{i=1}^{N} E_L(\mathcal{R}_i) \tag{35}$$

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

$$\sigma^2 = \left\langle [E_L(\mathcal{R}) - \langle E_L(\mathcal{R}) \rangle]^2 \right\rangle \tag{36}$$

6.2   Stochastic gradient optimization

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

$$E(\boldsymbol{\alpha}) = \frac{\int d\mathcal{R} \Psi_\mathsf{T}^*(\mathcal{R}) \mathcal{H} \Psi_\mathsf{T}(\mathcal{R})}{\int d\mathcal{R} \Psi_\mathsf{T}^*(\mathcal{R}) \Psi_\mathsf{T}(\mathcal{R})} \tag{37}$$

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 (37) we obtain the following expression in terms of the analytical derivative of the trial wave function

$$\nabla_{\boldsymbol{\alpha}} E(\boldsymbol{\alpha}) = 2 \left[\langle E_L(\mathcal{R}) \nabla_{\boldsymbol{\alpha}} \ln \Psi_\mathsf{T}(\mathcal{R}) \rangle - \langle E_L(\mathcal{R}) \rangle \langle \nabla_{\boldsymbol{\alpha}} \ln \Psi_\mathsf{T}(\mathcal{R}) \rangle \right] \tag{38}$$

where we have assumed a real trial wave function, $\Psi_\mathsf{T}(\mathcal{R}) = \Psi_\mathsf{T}^*(\mathcal{R})$. The averages in eqn (38) can be esimated in the usual way in the Metropolis method.

The gradient can then be used in the stochastic gradient technique [9], a damped steepest descent method, according to

$$\boldsymbol{\alpha}_{p+1} = \boldsymbol{\alpha}_p - \gamma_p \nabla_{\boldsymbol{\alpha}} E(\boldsymbol{\alpha}_p) \tag{39}$$

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

$$\gamma_p = Ap^{-\beta} \tag{40}$$

where the exponent β should be in the range $0.5 < \beta \leq 1$ [9, 10] and A is related to the inverse of the Hessian. By iterating, $\boldsymbol{\alpha}_p$ may approach the correct value [9, 10] 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 [11] and often the Hessian is computed to increase the efficiency [11].

6.3 Guided sampling

A shortcoming with the Metropolis importance sampling technique is that many attempted moves will be rejected. The moves are proposed without any knowledge of which regions that are important. An alternative strategy is to directly guide the random walker into the important regions. This is called *smarter or force-biased Monte Carlo*.

The technique can be derived by considering the Smoluchowski equation

$$\frac{\partial}{\partial t}\rho(\mathcal{R},\tau) = \frac{1}{2}\frac{\partial}{\partial \mathcal{R}}\left[\frac{\partial}{\partial \mathcal{R}} - F(\mathcal{R})\right]\rho(\mathcal{R},\tau) \tag{41}$$

The equilibrium, or stationary, distribution is given by

$$\rho(\mathcal{R},\tau \to \infty) \equiv \rho^{st}(\mathcal{R}) = C\exp\left[-V(\mathcal{R})\right] \tag{42}$$

where C is a normalization constant and

$$F(\mathcal{R}) = \frac{1}{\rho^{st}(\mathcal{R})}\frac{\partial \rho^{st}(\mathcal{R})}{\partial \mathcal{R}} = -\frac{\partial}{\partial \mathcal{R}}V(\mathcal{R}) \tag{43}$$

is the force. In the present case we would like to generate points according to the probability distribution in Eqn (34). We therefore select

$$\rho^{st}(\mathcal{R}) = \frac{|\Psi_{\mathsf{T}}(\mathcal{R})|^2}{\int d\mathcal{R}\,|\Psi_{\mathsf{T}}(\mathcal{R})|^2} \tag{44}$$

for the stationary distribution. The corresponding force will then be given by

$$F(\mathcal{R}) = \frac{2}{\Psi_{\mathsf{T}}(\mathcal{R})}\frac{\partial \Psi_{\mathsf{T}}(\mathcal{R})}{\partial \mathcal{R}} \tag{45}$$

where a real wave function is assumed, $\Psi_{\mathsf{T}}(\mathcal{R}) = \Psi_{\mathsf{T}}^*(\mathcal{R})$. The equation can be solved by generating random walkers according to

$$\mathcal{R}_{n+1} = \mathcal{R}_n + \frac{\Delta\tau}{2}F(\mathcal{R}_n) + \sqrt{\Delta\tau}\,\mathcal{G}_n \tag{46}$$

where $\mathcal{G}_n$ is a Gaussian random number with zero mean, unit variance and uncorrelated in time $\langle \mathcal{G}_n \mathcal{G}_{n'} \rangle = \delta_{n,n'}$.

Using a finite step $\Delta\tau$ we have introduced a time-step error of the order $(\Delta\tau)^2$. However, it is possible to eliminate this error by combining the above algorithm with a Metropolis acceptance/rejection step. Consider a transition from $\mathcal{R}' = \mathcal{R}_n$ to $\mathcal{R} = \mathcal{R}_{n+1}$. This is given by the probability

$$\omega_{\mathcal{R}\leftarrow\mathcal{R}'} \equiv \omega_{\mathcal{R}\mathcal{R}'} = G(\mathcal{R}, \mathcal{R}'; \Delta\tau) \propto \exp\left\{ -\frac{1}{2\Delta\tau} \left[\mathcal{R} - \mathcal{R}' - \frac{\Delta\tau}{2} F(\mathcal{R}') \right]^2 \right\}$$
(47)

This is not symmetric in $\mathcal{R}$ and $\mathcal{R}'$ as F depends only on $\mathcal{R}'$. Therefore, we have to use the generalized Metropolis algorithm. We then view $\omega_{\mathcal{R}\mathcal{R}'}$ as a trial step probability and introduce the acceptance probability $\alpha_{\mathcal{R}\mathcal{R}'}$. Detailed balance

$$\alpha_{\mathcal{R}\mathcal{R}'}\ \omega_{\mathcal{R}\mathcal{R}'}\ \rho^{st}(\mathcal{R}') = \alpha_{\mathcal{R}'\mathcal{R}}\ \omega_{\mathcal{R}'\mathcal{R}}\ \rho^{st}(\mathcal{R})$$
(48)

is then satisfied if we use

$$\alpha_{\mathcal{R}\leftarrow\mathcal{R}'} \equiv \alpha_{\mathcal{R}\mathcal{R}'} = \min\left[1, \frac{\omega_{\mathcal{R}'\mathcal{R}}\ \rho^{st}(\mathcal{R})}{\omega_{\mathcal{R}\mathcal{R}'}\ \rho^{st}(\mathcal{R}')} \right]$$

As usual, if the step is not accepted the old configuration has to be counted once more. We notice that the quotient $\omega_{\mathcal{R}'\mathcal{R}}/\omega_{\mathcal{R}\mathcal{R}'}$ is approximately equal to the ratio $\rho^{st}(\mathcal{R}')/\rho^{st}(\mathcal{R})$ when the stationary condition, equilibrium, has been reached. The acceptance ratio is close to 1 and the method is very efficient. The Metropolis acceptance/rejection step is merely a correction for the time step discretisation error made in the steeping procedure in Eqn (46).

6.4 Trial wave functions

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 trial wave function for a many-electron system is often written on the form

$$\Psi_{\mathsf{T}}(\boldsymbol{x}_1, \dots, \boldsymbol{x}_N) = \Psi_{AS}(\boldsymbol{x}_1, \dots, \boldsymbol{x}_N) \exp\left[\frac{1}{2} \sum_i^N \sum_{i\neq j}^N f(r_{ij}) \right]$$
(49)

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 most commonly used form for $f(r)$ is the Padé-Jastrow function [5]

$$f(r) = \frac{a_1 r + a_2 r + \dots}{1 + b_1 r + b_2 r + \dots}$$

The trial wave function should also fulfill the boundary conditions, the so called cusp conditions [5]. They can be written on the form (in atomic units)

$$\lim_{r_{ij} \to 0} \frac{1}{\Psi_\mathsf{T}} \frac{\partial \Psi_\mathsf{T}}{\partial r_{ij}} = \frac{1}{4} \quad \text{if } s_i = s_j \tag{50}$$

$$\lim_{r_{ij} \to 0} \frac{1}{\Psi_\mathsf{T}} \frac{\partial \Psi_\mathsf{T}}{\partial r_{ij}} = \frac{1}{2} \quad \text{if } s_i \neq s_j \tag{51}$$

$$\lim_{r_{iI} \to 0} \frac{1}{\Psi_\mathsf{T}} \frac{\partial \Psi_\mathsf{T}}{\partial r_{iI}} = -Z_I \tag{52}$$

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

6.5 The helium atom

In Hartree method the wave-function for helium was approximated by a product of two one-electron orbitals

$$\Psi_\mathsf{T}(\boldsymbol{r}_1, \boldsymbol{r}_2) = \phi(r_1)\phi(r_2)$$

This implies that the joint probability density

$$n_2(\boldsymbol{r}_1, \boldsymbol{r}_2) \equiv |\Psi_\mathsf{T}(\boldsymbol{r}_1, \boldsymbol{r}_2)|^2$$

could be written as a product of two one-electron densities. The probability $n_2(\boldsymbol{r}_1, \boldsymbol{r}_2)$ to find an electron at $\boldsymbol{r}_1$ and another one at $\boldsymbol{r}_2$ is uncorrelated,

$$n_2(\boldsymbol{r}_1, \boldsymbol{r}_2) = n(\boldsymbol{r}_1)n(\boldsymbol{r}_2)$$

To introduce correlation effects we now write the trial wave-function on the form

$$\Psi_\mathsf{T}(\boldsymbol{r}_1, \boldsymbol{r}_2) = \phi(r_1)\phi(r_2)f(r_{12}) \tag{53}$$

where $f(r_{12})$ depends on both coordinates and introduces correlation effects. We make the ansatz

$$\phi(r) = \exp\left[-\alpha_1 r\right] \tag{54}$$

$$f(r) = \exp\left[\frac{\alpha_2 r}{1 + \alpha_3 r}\right] \tag{55}$$

The correlation function $f(r)$ is of the Padé-Jastrow type to low order. The cusp conditions imply that $\alpha_1 = 2$ and $\alpha_2 = 1/2$, i.e.

$$\Psi_\mathsf{T}(\boldsymbol{r}_1, \boldsymbol{r}_2) = \exp\left[-2r_1\right] \, \exp\left[-2r_2\right] \, \exp\left[\frac{r_{12}}{2(1 + \alpha r_{12})}\right] \tag{56}$$

16

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

$$E_L(\boldsymbol{r}_1, \boldsymbol{r}_2) = -4 + \frac{(\hat{\boldsymbol{r}}_1 - \hat{\boldsymbol{r}}_2) \cdot (\boldsymbol{r}_1 - \boldsymbol{r}_2)}{r_{12}(1 + \alpha r_{12})^2}$$
$$-\frac{1}{r_{12}(1 + \alpha r_{12})^3} - \frac{1}{4(1 + \alpha r_{12})^4} + \frac{1}{r_{12}} \tag{57}$$

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

7 Diffusion Monte Carlo

In the variational Monte Carlo method the accuracy is limited by the form of the trial many-electron wavefunction. Here we will now remove this restriction. The form of the many-electron wavefunction will be automatically refined and will approach the exact ground-state wavefunction. This will be done by mapping the Schrödinger equation into a diffusion-reaction process which will be solved by simulating the time-evolution of a set of random walkers. The method is usually denoted the *Diffusion Monte Carlo* (DMC) method [2, 5] and has been applied to various systems [6, 7, 8].

7.1 Imaginary time Schrödinger equation

Consider the time-dependent Schrödinger equation

$$i\frac{\partial \Psi(t)}{\partial t} = \mathcal{H}\Psi(t)$$

with the solution

$$\Psi(t) = e^{-i\mathcal{H}t}\Psi(0) = \sum_{n=0}^{\infty} c_n \Phi_n e^{-iE_n t}$$

where Φ_n and E_n are the exact energy eigenstates and eigenvalues, respectively, and

$$c_n = \langle \Phi_n | \Psi(0) \rangle$$

The exact energy eigenstates and eigenvalues are not known. We now make a formal transformation to imaginary time

$$t \to -i\tau$$

and we get

$$\frac{\partial \Psi(\tau)}{\partial \tau} = -\mathcal{H}\Psi(\tau) \tag{58}$$

with the solution

$$\Psi(\tau) = e^{-\mathcal{H}\tau}\Psi(0) \tag{59}$$

We shift the energy scale

$$\mathcal{H} \to \mathcal{H} - E_\mathsf{T} \tag{60}$$

and

$$\Psi(\tau) = e^{-(\mathcal{H}-E_\mathsf{T})\tau}\Psi(0) = \sum_{n=0}^{\infty} c_n \Phi_n e^{-(E_n - E_\mathsf{T})\tau} \tag{61}$$

For $\tau \to \infty$ we then obtain:

- if $E_\mathsf{T} < E_0$
  $\Psi(\tau \to \infty)$ decays exponentially

- if $E_T > E_0$

  $\Psi(\tau \to \infty)$ diverges exponentially

- if $E_T = E_0$

  $\Psi(\tau \to \infty) \to c_0 \Phi_0$

  the wavefunction converges to the ground-state wavefunction Φ_0, up to a constant factor c_0, provided the initial wavefunction $\Psi(0)$ has a numerically significant overlap with Φ_0 (i.e. $c_0 \neq 0$)

This forms the basis for the Diffusion Monte Carlo (DMC) method. An initial value of the reference energy E_T is chosen and during the simulation E_T will be adjusted in such away that it approaches the unknown ground state energy E_0.

7.2 Diffusion Monte Carlo

Consider now a single electron moving in a one-dimensional potential $V(x)$. The generalisation to a many-electron system is straightforward. We will represent the wavefunction for the electron with a set of random walkers. Initially the number of random walkers is set to N_0 with positions

$$x^{(j)} \;\; ; \;\; j = 1, \ldots, N_0$$

The number of walkers will fluctuate along the random walk and this will be used to adjust the reference energy E_T such that it will approach the correct ground state energy E_0.

The Hamiltonian for the electron is given by

$$\mathcal{H} = -\frac{1}{2}\frac{\partial^2}{\partial x^2} + [V(x) - E_T] \tag{62}$$

where the energy has been shifted with E_T. The Schrödinger equation in imaginary time is given by

$$\frac{\partial}{\partial \tau}\Psi(x,\tau) = -[\mathcal{K} + \mathcal{V}]\Psi(x,\tau) \tag{63}$$

with

$$\begin{aligned}
\mathcal{K} &= -\frac{1}{2}\frac{\partial^2}{\partial x^2} \\
\mathcal{V} &= V(x) - E_T
\end{aligned}$$

and has the formal solution

$$\Psi(x,\tau) = e^{-\mathcal{L}\tau}\Psi(x,0) \;\; ; \;\; \mathcal{L} = \mathcal{K} + \mathcal{V} \tag{64}$$

It is instructive to introduce the Green's function

$$G(x, x_0; \tau) = \langle x | e^{-\mathcal{L}\tau} | x_0 \rangle \tag{65}$$

and the wavefunction at time τ is then given by

$$\Psi(x, \tau) = \int dx_0 \; G(x, x_0; \tau)\Psi(x_0, 0)$$

where $\Psi(x_0, 0)$ is the wavefunction at the initial time. Next we divide the time evolution into M small time steps with size $\Delta\tau$

$$
\begin{aligned}
G(x, x_0; \tau) \;\; = \;\; & \int dx_1 \ldots dx_{M-1} \langle x|e^{-\mathcal{L}\Delta\tau}|x_{M-1}\rangle \langle x_{M-1}|e^{-\mathcal{L}\Delta\tau}|x_{M-2}\rangle \\
\ldots \;\; & \langle x_2|e^{-\mathcal{L}\Delta\tau}|x_1\rangle \langle x_1|e^{-\mathcal{L}\Delta\tau}|x_0\rangle
\end{aligned}
$$

where $\tau = M\Delta\tau$. We have to define how to determine the short-time propagator

$$G(x, x'; \Delta\tau) \equiv \langle x|e^{-\mathcal{L}\Delta\tau}|x'\rangle = \langle x|e^{-(\mathcal{L}_R+\mathcal{L}_D)\Delta\tau}|x'\rangle \tag{66}$$

The operators $\mathcal{L}_R \equiv \mathcal{V}$ and $\mathcal{L}_D \equiv \mathcal{K}$ do not commute but for sufficiently small $\Delta\tau$ we may write

$$
\begin{aligned}
e^{-\mathcal{L}\Delta\tau} \;\; = \;\; & e^{-(\mathcal{L}_R+\mathcal{L}_D)\Delta\tau} \\
\simeq \;\; & e^{-\mathcal{L}_R\Delta\tau}e^{-\mathcal{L}_D\Delta\tau} + \mathcal{O}(\Delta\tau^2)
\end{aligned}
\tag{67}
$$

and the short-time propagator can be approximated by the product of a diffusive part G_D and a reactive part G_R according to

$$G(x, x'; \Delta\tau) = \int dx'' G_R(x, x''; \Delta\tau) \; G_D(x'', x'; \Delta\tau)$$

Diffusive part The Green's function

$$G_D(x, x'; \Delta\tau) = \langle x|e^{-\mathcal{L}_D\Delta\tau}|x'\rangle = \langle x|e^{-\mathcal{K}\Delta\tau}|x'\rangle \tag{68}$$

describes the spreading of the wavefunction due to the kinetic energy. We have

$$
\begin{aligned}
\langle x|e^{-\mathcal{K}\Delta\tau}|x'\rangle \;\; = \;\; & \int dk \int dk' \langle x|k\rangle \langle k|e^{-\mathcal{K}\Delta\tau}|k'\rangle \langle k'|x'\rangle \\
= \;\; & \int dk \langle x|k\rangle e^{\frac{(ik)^2}{2}\Delta\tau} \langle k|x'\rangle = \int \frac{dk}{2\pi} e^{ik(x-x')} e^{-\frac{k^2\Delta\tau}{2}} \\
= \;\; & \frac{1}{\sqrt{2\pi\Delta\tau}} e^{-\frac{(x-x')^2}{2\Delta\tau}}
\end{aligned}
$$

The diffusive part can be taking into account by displacing each random walker according to

$$x = x' + \sqrt{\Delta\tau} \; \mathcal{G} \tag{69}$$

where $\mathcal{G}$ is a Gaussian random number with zero mean and unit variance (cf. Sec. Diffusion equation in the Lecture notes *Brownian Dynamics*).

Reactive part The Green's function for the reactive part

$$G_R(x, x'; \Delta\tau) = \langle x|e^{-\mathcal{L}_R \Delta\tau}|x'\rangle = e^{-(V(x)-E_\mathsf{T})\Delta\tau}\,\delta(x - x') \qquad (70)$$

introduces a weight factor

$$W(x) = e^{-(V(x)-E_\mathsf{T})\Delta\tau} \qquad (71)$$

The accumulated product of the weight factors $W(x_i), i = 1, \ldots$ can be added to each random walker. However, in many cases a more numerically efficient method is to use a branching method, a birth-and-death simulation process. In order to implement the branching correctly, we must make an *integer* number of copies of the snapshot that is equal on average to the real number $W(x)$. To implement this one can simply replace the random walker at x with m number of random walkers, where

$$m = \text{int}[W(x) + \mathcal{U}] \qquad (72)$$

Here, int$[\ldots]$ denotes the integer part and $\mathcal{U}$ is a uniform random number on [0,1]. If $W(x) > 1$ new walkers may be created and if $W(x) < 1$ the walker may be killed. This is repeated for each random walker at each time step.

Adjustment of E_T We would like the distribution of the random walkers to approach the ground-state wavefunction. This can be achieved if the reference energy E_T is adjusted to the unknown ground state energy E_0. To obtain the equilibrium distribution, the ground state distribution, the number of random walkers should approach a constant value. However as a result of the birth-and-death process the number of walkers are changing. To obtain a constant number of walkers we have to adjust the reference energy E_T. When E_T is approaching the ground state energy E_0 the number of walkers will not change and the ground state distribution has been obtained.

Denote the number of random walkers at step i with N_i. The initial number of random walkers is N_0. At iteration $i+1$ the reference energy can be set to

$$E_\mathsf{T}(i + 1) = E_\mathsf{T}(i) - \gamma \ln \frac{N_i}{N_0}$$

In this equation $E_\mathsf{T}(i + 1)$ is determined from knowledge of the previous value $E_\mathsf{T}(i)$. A better way is to replace the instantaneous value $E_\mathsf{T}(i)$ by the accumulated average value $\langle E_\mathsf{T} \rangle (i)$

$$\langle E_\mathsf{T} \rangle (i) = \frac{1}{i - i_0} \sum_{i'=i_0+1}^{i} E_\mathsf{T}(i')$$

where i_0 is an initial number of steps that are disregarded, the "equilibrium steps", and hence

$$E_\mathsf{T}(i + 1) = \langle E_\mathsf{T} \rangle (i) - \gamma \ln \frac{N_i}{N_0}$$

The accumulated average value can be determined iteratively as

$$\langle E_\mathsf{T}\rangle\,(i+1) = \frac{1}{i-i_0+1}E_\mathsf{T}(i+1) + \frac{i-i_0}{i-i_0+1}\,\langle E_\mathsf{T}\rangle\,(i)$$

The damping parameter γ is of the order one or less and can be adjusted during the simulation. The random walkers approaches the equilibrium distribution and the ground-state energy E_0 can then be obtained from an average along the random walk.

7.3   Diffusion Monte Carlo with importance sampling

Ordinary diffusion Monte Carlo becomes quite inefficient if the potential $V(x)$ varies strongly. In particular, serious problems occur if the potential is unbounded, which is quite common in real applications. Importance sampling techniques have therefore been introduced to make the sampling more efficient. Consider the function

$$f(x,\tau) \equiv \Psi_\mathsf{T}(x)\Psi(x,\tau) \tag{73}$$

where $\Psi_\mathsf{T}(x)$ is some trial wavefunction which models the real wavefunction $\Psi(x)$ in a reasonable way. As in variational Monte Carlo we define a local energy

$$E_L(x) \equiv \frac{\mathcal{H}\Psi_\mathsf{T}(x)}{\Psi_\mathsf{T}(x)} = -\frac{1}{2}\frac{1}{\Psi_\mathsf{T}(x)}\frac{d^2\Psi_\mathsf{T}(x)}{dx^2} + V(x) \tag{74}$$

The Schrödinger equation in imaginary time reads

$$\frac{\partial}{\partial\tau}\Psi(x,\tau) = \frac{1}{2}\frac{\partial^2}{\partial x^2}\Psi(x,\tau) - (V(x)-E_\mathsf{T})\Psi(x,\tau)$$

Multiply with $\Psi_\mathsf{T}(x)$ from the left and use the expression for the local energy

$$\begin{aligned}
\frac{\partial}{\partial\tau}f(x,\tau) &= \frac{1}{2}\Psi_\mathsf{T}(x)\frac{\partial^2\Psi(x,\tau)}{\partial x^2} - (V(x)-E_\mathsf{T})f(x,\tau)\\
&= \frac{1}{2}\left[\Psi_\mathsf{T}(x)\frac{\partial^2\Psi(x,\tau)}{\partial x^2} - \Psi(x,\tau)\frac{d^2\Psi_\mathsf{T}(x)}{dx^2}\right] - [E_L(x)-E_\mathsf{T}]f(x,\tau)
\end{aligned}$$

This can be rewritten as

$$\frac{\partial}{\partial\tau}f(x,\tau) = -\left[-\frac{1}{2}\frac{\partial^2}{\partial x^2} + \frac{\partial}{\partial x}v_F(x) + [E_L(x)-E_\mathsf{T}]\right]f(x,\tau) \tag{75}$$

where

$$v_F(x) = \frac{1}{\Psi_\mathsf{T}(x)}\frac{d}{dx}\Psi_\mathsf{T}(x) \tag{76}$$

This corresponds to a diffusion equation with a "drift velocity" $v_F(x)$ (cf. section Diffusion equation in the Lecture notes *Brownian dynamics*). The

drift velocity will push the walkers into regions which are more important. The number of walkers will also fluctuate considerably less due to that in general $|E_L(x) - E_\mathsf{T}| \ll |V(x) - E_\mathsf{T}|$.

The solution for $f(x, \tau)$ can formally be written as

$$f(x, \tau) = e^{-(\mathcal{L}_R + \mathcal{L}_D + \mathcal{L}_F)\tau} f(x, 0) \tag{77}$$

As before, we divide the time evolution into small time-steps with size $\Delta\tau$ and then approximate the corresponding short time propagator according to

$$\begin{aligned} e^{-\mathcal{L}\Delta\tau} &= e^{-(\mathcal{L}_R + \mathcal{L}_D + \mathcal{L}_F)\Delta\tau} \\ &\simeq e^{-\mathcal{L}_R\Delta\tau} e^{-\mathcal{L}_D\Delta\tau} e^{-\mathcal{L}_F\Delta\tau} + \mathcal{O}(\Delta\tau^2) \end{aligned} \tag{78}$$

A more accurate approximation can be obtained using the more symmetric decomposition

$$\begin{aligned} e^{-\mathcal{L}\Delta\tau} &= e^{-(\mathcal{L}_R + \mathcal{L}_D + \mathcal{L}_F)\Delta\tau} \\ &\simeq e^{-\mathcal{L}_F\Delta\tau/2} e^{-\mathcal{L}_D\Delta\tau/2} e^{-\mathcal{L}_R\Delta\tau} e^{-\mathcal{L}_D\Delta\tau/2} e^{-\mathcal{L}_F\Delta\tau/2} + \mathcal{O}(\Delta\tau^3) \end{aligned} \tag{79}$$

Diffusive part The Green's function for the diffusive part is the same as before

$$G_D(x, x'; \Delta\tau) = \langle x | e^{-\mathcal{L}_D\Delta\tau} | x' \rangle \tag{80}$$

with

$$x = x' + \sqrt{\Delta\tau}\, \mathcal{G} \tag{81}$$

where $\mathcal{G}$ is a Gaussian random number with zero mean and unit variance.

Drift part The Green's function for the drift part

$$G_F(x, x'; \Delta\tau) = \langle x | e^{-\mathcal{L}_F\Delta\tau} | x' \rangle \tag{82}$$

describes the movement due to the drift velocity $v_F(x)$. We have

$$\langle x | e^{-\mathcal{L}_F\Delta\tau} | x' \rangle = \delta(x - x(\Delta\tau))$$

where $x(\Delta\tau)$ is the solution to

$$\frac{d}{dt} x(t) = v_F(x(t)) \quad, \quad \text{with} \ \ x(0) = x'$$

An approximation, accurate to first order in $\Delta\tau$, is

$$x(\Delta\tau) = v_F(x')\Delta\tau + x' \tag{83}$$

If second order accuracy is needed the following expression can be used [12]

$$x(\Delta\tau) = v_F(x_{1/2})\Delta\tau + x' \quad, \quad \text{with} \ \ x_{1/2} = v_F(x')\Delta\tau/2 + x' \tag{84}$$

Reactive part The Green's function for the reactive part is now given by

$$G_F(x, x'; \Delta\tau) = \langle x|e^{-\mathcal{L}_R \Delta\tau}|x'\rangle = e^{-(E_L(x) - E_\mathsf{T})\Delta\tau} \delta(x - x') \qquad (85)$$

and

$$W(x) \equiv e^{-(E_L(x) - E_\mathsf{T})\Delta\tau} \qquad (86)$$

and can be evaluated in the same way as before. However in this case the weightfactor $W(x)$ will fluctuate considerably less.

Adjustment of E_T The number of walkers can be adjusted in the same way as before,

$$E_\mathsf{T}(i + 1) = \langle E_\mathsf{T}\rangle(i) - \gamma \ln \frac{N_i}{N_0}$$

and the ground-state energy E_0 can then be obtained from an average along the random walk.

Time-step error The DMC method contains a finite error due to that $\Delta\tau$ is finite, not equal to zero. To obtain the exact value one can solve for several time-steps $\Delta\tau$ and then extrapolate to $\Delta\tau = 0$. Using the decomposition in Eqn (78) the error for the ground-state energy will scale linearly with $\Delta\tau$ and Eqn (83) can be used. If the decomposition in Eqn (79) is used the error will scale quadratically, provided that the drift term also scales quadratically. That is not the case if Eqn (83) is used but Eqn (84) will restore quadratic convergence.

Techniques where the exact Green's function is sampled have also been developed [5, 13]. In that case the finite time step error in DMC is eliminated and the exact value is obtained without any extrapolation. These techniques are more involved compared with DMC and are usually denoted *Green's Function Monte Carlo* (GFMC).

A   Atomic units

Atomic units (a.u.) are obtained by putting $m = e = \hbar = 4\pi\epsilon_0 = 1$, where m is the mass of the electron, e is the (absolute value) of the charge of the electron, and ϵ_0 is the vacuum permittivity.

In these units the length unit becomes the Bohr radius

$$a_0 = \frac{4\pi\epsilon_0\hbar^2}{me^2} = 0.529\text{Å}$$

and the energy unit the Hartree energy

$$E_H = \frac{me^4}{(4\pi\epsilon_0\hbar)^2} = 27.2\text{eV}$$

B The variational theorem

The variational method in quantum mechanics is an important technique to facilitate numerical solutions of the time-independent Schrödinger equation for many-electron systems, such as the electronic structure problem in Eqn (3) [3, 4]. We will frequently make use of this method.

Consider a system described by the Hamiltonian $\mathcal{H}$, which is known. Denote the corresponding exact, but unknown, energy eigenfunctions and eigenvalues with Φ_n and E_n, respectively, i.e.

$$\mathcal{H}\Phi_n(\mathcal{R}) = E_n\Phi_n(\mathcal{R}) \ .$$

Here, we use the notation $\mathcal{R} = (\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N)$ for the collective positions of all electrons. Construct a trial many-electron wave function $\Psi_{\mathsf{T}}(\mathcal{R})$. The corresponding expectation value for the energy E can be viewed as a functional of the trial wave function $\Psi_{\mathsf{T}}(\mathcal{R})$ and is given by

$$E\left[\Psi_{\mathsf{T}}\right] = \frac{\int d\mathcal{R}\,\Psi_{\mathsf{T}}^*(\mathcal{R})\mathcal{H}\Psi_{\mathsf{T}}(\mathcal{R})}{\int d\mathcal{R}\,\Psi_{\mathsf{T}}^*(\mathcal{R})\Psi_{\mathsf{T}}(\mathcal{R})}$$

By formally expanding $\Psi_{\mathsf{T}}(\mathcal{R})$ in terms of the exact eigenfunctions

$$\Psi_{\mathsf{T}}(\mathcal{R}) = \sum_n c_n\Phi_n(\mathcal{R})$$

we get

$$E\left[\Psi_{\mathsf{T}}\right] = \frac{\sum_n E_n\,|c_n|^2}{\sum_n |c_n|^2}$$

If we now subtract E_0, the exact unknown ground-state energy, from both sides, we have

$$E\left[\Psi_{\mathsf{T}}\right] - E_0 = \frac{\sum_n(E_n - E_0)\,|c_n|^2}{\sum_n |c_n|^2}$$

Since $E_n \geq E_0$ for all n the right-hand side is non-negative and we have derived *the variational theorem*

$$E\left[\Psi_{\mathsf{T}}\right] \geq E_0 \tag{87}$$

i.e. the expectation value for the energy $E\left[\Psi_{\mathsf{T}}\right]$ using an arbitrary wavefunction $\Psi_{\mathsf{T}}(\mathcal{R})$ is always larger or equal to the exact ground state energy E_0.

Eqn (87), the variational theorem, constitutes the basis for the *Rayleigh-Ritz variational method* for the approximate calculation of E_0. This methods consists in evaluating the quantity $E\left[\Psi_{\mathsf{T}}\right]$ by using a trial wave function $\Psi_{\mathsf{T}}(\mathcal{R})$ which depends on a certain number of *variational parameters*, and

then to minimize $E[\Psi_{\mathsf{T}}]$ with respect to these parameters in order to obtain the best approximation of E_0 allowed by the form chosen for $\Psi_{\mathsf{T}}(\mathcal{R})$. The result constitutes an upper limit for the exact ground state energy E_0, and it is likely that the result will be close if the form of the trial wave function resembles well the exact ground state wave function $\Phi_0(\mathcal{R})$.

References

[1] J. Kohanoff, *Electronic Structure Calculations for Solids and Molecules*, Cambridge Univ. Press, 2006.

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

[3] L. I. Schiff, *Quantum mechanics*, 3rd ed., McGraw-Hill, 1955.

[4] B. H. Brandsen and C. J. Joachain, *Physics of Atoms and Molecules*, 2nd ed., Prentice Hall, 2003.

[5] B. L. Hammond, W. A. Lester Jr., P. J. Reynolds, *Monte Carlo Methods in Ab Initio Quantum Chemistry*, World Scientific, 1994.

[6] W. H. C. Foulkes, L. Mitas, R. J. Needs, and G. Rajagopal, Rev. Mod. Phys. **73**, 33 (2001).

[7] R. J. Needs, M. D.Towler, N. D. Drummond, and P. Lopez Rios, J. Phys.: Condens. Matter **22**, 023201 (2010).

[8] R. Q. Hood, P. R. C. Kent, and F. A. Reboredo, Phys. Rev. B **85**, 134109 (2012).

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

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

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

[12] S. A. Chin, Phys. Rev. A **42** 69961 (1990).

[13] M. H. Kalos, Phys. Rev. **128**, 1791 (1962).