

Computing the ground state energy of helium

Daniel Martin

School of Physics and Astronomy,
The University of Manchester

Fourth Year M.Phys Project Report

May 2007

This experiment was performed in collaboration with P. Hughes
and under the supervision of Prof. J. Forshaw.

Abstract

Computer programs have been written to simulate various systems using Monte Carlo methods. The Ising ferro-magnet has been successfully simulated and its behaviour investigated in two and three dimensions. Quantum Monte Carlo methods have been implemented and the complete ionisation energy of helium has been measured using these methods, and found to be -2.902 ± 0.001 Hartrees, consistent with experiment.

Contents

1	Introduction	2
2	Monte Carlo methods	2
2.1	Monte Carlo integration	2
2.1.1	Importance sampling	3
2.2	The Metropolis algorithm	4
2.2.1	The single walker optimisation	4
2.3	The Ising ferro-magnet	5
2.3.1	Method	6
2.3.2	Results	7
3	Quantum Monte Carlo methods	9
3.1	Variational Monte Carlo	9
3.1.1	Fokker-Planck variational Monte Carlo	11
3.2	Green's Function Diffusion Monte Carlo	13
3.2.1	The 1D quantum harmonic oscillator	14
3.3	Case Study: The H^2 molecule	15
3.3.1	The He atom	16
3.3.2	The H^2 molecule	16
4	Conclusion	18

1 Introduction

The goal of this project is to calculate the total ionisation energy of helium in its ground state. This has been done using Monte-Carlo methods.

Monte-Carlo methods, named after the famous casino-rich town in Monaco, are stochastic methods in that they involve an element of chance, using (pseudo-)random¹ numbers. Monte Carlo methods can be very efficient for studying systems with many dimensions.

The Ising ferro-magnet has been studied as an introduction to Monte Carlo methods and the Metropolis algorithm. Then quantum Monte Carlo methods were tested and used to simulate the helium atom and the hydrogen molecule. A diffusion Monte Carlo method was used to measure the exact ground-states of the systems.

The diffusion Monte Carlo method used in this report is the Green's function Monte Carlo, with a guide function. However other methods can be used, such as one based on the Green's function that does not use a guide function. Also path integrals can be used in place of the Green's function to form path integral (diffusion) Monte Carlo methods.

This report begins by introducing Monte Carlo methods, and the Metropolis algorithm. It then explains how the Ising ferro-magnet was studied and gives graphs of the physical properties, measured by the method. It goes on to cover different quantum Monte Carlo methods. First detailing the standard variational Monte Carlo method, before covering a more efficient version based on the Fokker-Planck equation. Finally the Green's function diffusion Monte Carlo method is discussed. Each method is applied to the one-dimensional quantum harmonic oscillator allowing the methods to be easily compared. The methods are then applied to study the helium atom and the hydrogen molecule.

2 Monte Carlo methods

Monte Carlo methods, are simply methods that rely on 'random' numbers.² In this project we are interested in Monte Carlo methods from a statistical physics stand-point. These methods use random numbers to provide artificial dynamics. They are often used for simulating systems with many dimensions as for these systems Monte Carlo methods are more efficient than dynamics simulations. They are therefore well suited to simulating many-particle systems; two particles in a three-dimensional space may be considered to be a six-dimensional system, for example. The easiest way to understand the methods used in this project is to first consider Monte Carlo integration.

2.1 Monte Carlo integration

Imagine we wish to solve the integral:

$$I = \int_0^1 [f(x)] dx \quad (1)$$

Where $f(x)$ is some arbitrary function.

¹Computers are incapable of producing truly random numbers.

²In practice one cannot generate truly random numbers on a computer and instead must settle for 'pseudo-random' numbers.

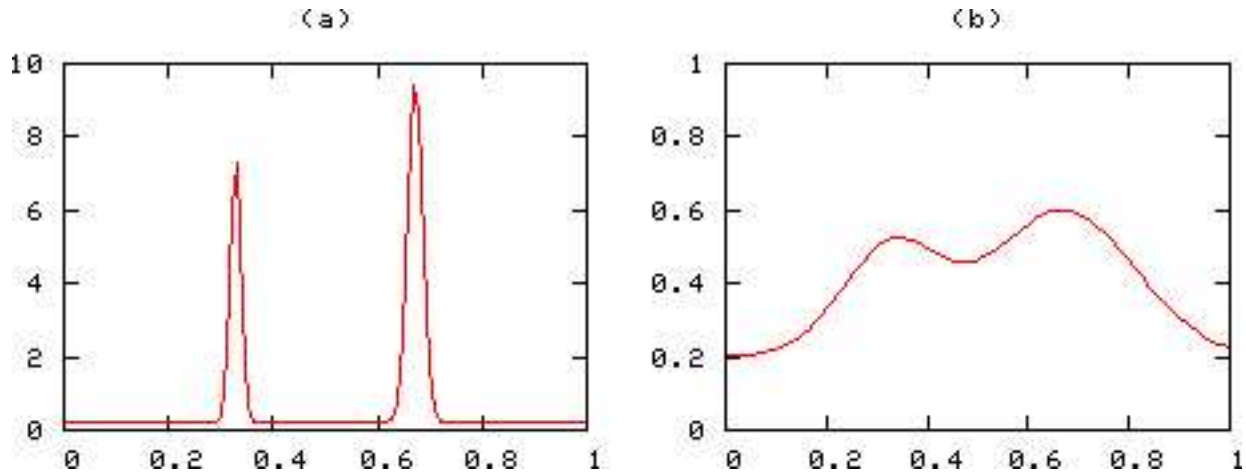


Figure 1: Monte Carlo integration (with crude sampling) will lead to greater errors on the function shown in graph (a) compared to graph (b).

We can approximate this integral using the following formula:

$$I \approx \tilde{I} = \frac{1}{N} \sum_{i=1}^N [f(x_i)] = \langle f \rangle \quad (2)$$

Where x_i are random numbers between 0 and 1, distributed uniformly.

The variance of the result is given by:

$$\sigma_{\tilde{I}}^2 = \frac{1}{N} \sigma_f^2 = \frac{1}{N} [\langle f^2 \rangle - \langle f \rangle^2] \quad (3)$$

and so the uncertainty on the estimate of I decreases as $N^{-1/2}$. This means that for this one dimensional problem, Monte Carlo integration is less efficient than standard quadrature methods that use equally distributed sampling points rather than random ones.³ However, we can apply this Monte Carlo method to multi-dimensional integrals and (unlike standard quadrature methods) retain the same uncertainty scaling. This means that Monte Carlo integration is efficient for evaluating many-dimensional integrals.

It should be noted that Monte Carlo integration is not susceptible to correlations in the pseudo-random number generator. In fact it is possible to use quasi-random numbers in place of pseudo-random ones.

2.1.1 Importance sampling

For many functions, the contribution to the integral from different regions of the function varies greatly (for example (a) in figure 1). Because standard Monte Carlo integration samples the function homogeneously, the regions which contribute most to the integral can be sampled poorly, leading to large statistical errors. We can reduce these errors by redefining (1) as:

$$I = \int_0^1 \rho(x) \left[\frac{f(x)}{\rho(x)} \right] dx \quad (4)$$

³Further discussion of this as well as Monte Carlo integration in general, can be found in section 10.2 of [1]:p272-275.

Where $\rho(x)$ is some arbitrary function.

If we now require that $\rho(x)$ is normalised across the interval, ($\int_0^1 \rho(x)dx = 1$) then (2) can be redefined as:

$$I \approx \tilde{I} = \frac{1}{N} \sum_{i=1}^N \left[\frac{f(x_i)}{\rho(x_i)} \right] \quad (5)$$

Where x_i are random numbers between 0 and 1, distributed according to $\rho(x)$.

If we choose $\rho(x)$ to have a similar shape to $f(x)$, then f/ρ will not vary much across the interval and (5) will yield a more accurate estimate of I than (2). This method is known as *importance sampling Monte Carlo*. However, generating random points according to an arbitrary distribution, $\rho(x)$, is a non-trivial exercise. In this project methods based on the Metropolis algorithm have been used.

2.2 The Metropolis algorithm

The algorithm of Metropolis et al.^[2] provides a way to generate random points, \mathbf{r}_i , according to an arbitrary distribution, $\rho(\mathbf{r})$. A single point is generated using a random walk through the space. Starting from a random initial position, the walker takes many *Metropolis steps* until it has *thermalised* to the distribution. The number of steps required for this must be ascertained by the simulator.

A Metropolis step is taken by first taking a trial step from \mathbf{r} to \mathbf{r}' . For example \mathbf{r}' may be obtained by moving from \mathbf{r} to a random point in a hyper-cube, centred on \mathbf{r} . (Again, the optimum side length of the hyper-cube must be determined by the simulator.) The trial step is then accepted or rejected according to:

$$R(\mathbf{r}, \mathbf{r}') = \frac{\rho(\mathbf{r}')}{\rho(\mathbf{r})} \quad (6)$$

If $R \geq 1$, then the step is accepted, otherwise the step is accepted with probability R .⁴

So if we require N points distributed according to $\rho(\mathbf{r})$, N walkers can simply be thermalised using the Metropolis algorithm. However, it can take a great many steps to thermalise a walker. Therefore we may look for a less computationally expensive approach.

2.2.1 The single walker optimisation

Once a walker has been thermalised, it will remain thermalised when undergoing further Metropolis steps. Therefore, rather than thermalising many walkers we can generate new points by thermalising a single walker and taking extra Metropolis steps. However, the new points will be correlated, as they have been generated as a part of a ‘Markov chain’.⁵ Strongly correlated points have the effect of artificially reducing the size of the statistical errors, causing us to over-estimate the accuracy of our simulation. Therefore it is necessary to reduce the correlation between the points that are used. This can be achieved by increasing the number of Metropolis steps between accepted points, or *anti-correlatory steps*. Because the number of anti-correlatory steps required to make the correlation between accepted points negligible is much lower than the number of steps required

⁴This can be easily implemented by comparing R to a uniform random number between 0 & 1. If R is larger then the step is accepted.

⁵Further discussion of Markov chains can be found in section 10.3 of [1]:p275-280.

to thermalise a walker, the *thermalisation steps*, this scheme provides a substantial computational saving at the inconvenience of requiring the simulator to determine the number of anti-correlatory steps required.

In practice there is another problem with this scheme. Imagine a distribution, $\rho(\mathbf{r})$, with two very favourable regions, separated by a very unfavourable region. (Graph (a) of figure 1 is a good example.) In this case a walker can effectively become ‘trapped’ inside one of the favourable regions as the chance of the walker passing through the unfavourable region is very small. This effect means that the use of a single walker can lead to the simulation failing to sample the entire distribution. We can overcome this issue through a compromise. Let’s say that we require N points. Rather than thermalising N walkers each yielding a single point, or thermalising a single walker yielding N points, we can use G walkers each generating N/G points. If $N \gg G$, then a substantial computational saving is still achieved whilst successfully sampling the entire distribution; the walkers are unlikely to all become trapped in the same region.

2.3 The Ising ferro-magnet

In order to demonstrate these Monte Carlo methods, they have been applied to the Ising ferro-magnet.

The Ising model of a ferro-magnet consists of a grid of spins, each of which are restricted to pointing either ‘up’ or ‘down’, with a magnitude of 1. The spins can interact with one another (usually only interactions between nearest neighbours are considered) and also with an external magnetic field (of course we are only interested in the projection of such a field into the same axis as the spins).

The *configuration space* is a space which contains all possible states or configurations of the systems. The configuration space for the Ising model with N spins, is N -dimensional. However these dimensions are discrete rather than continuous, each dimension taking either the value ‘up’ (+1) or ‘down’ (−1). We wish to ‘integrate’ over the entire configuration space in order to measure the physical properties of the model.⁶ Specifically we measure the magnetisation, M , the magnetic susceptibility, χ , the energy, E and the specific heat for constant external field strength, C_B .

$$M = \sum_{\mathbf{S}} \rho(\mathbf{S}) \left[\sum_i S_i \right] \quad (7)$$

$$\chi = \sum_{\mathbf{S}} \rho(\mathbf{S}) \left[\sum_i S_i \right]^2 - M^2 \quad (8)$$

$$E = \sum_{\mathbf{S}} \rho(\mathbf{S}) H(\mathbf{S}) \quad (9)$$

$$C_B = \sum_{\mathbf{S}} \rho(\mathbf{S}) H^2(\mathbf{S}) - E^2 \quad (10)$$

Where \mathbf{S} is the configuration vector, S_i is the value of the spin at grid point, i , $H(\mathbf{S})$ is the Hamiltonian and $\rho(\mathbf{S})$ is the distribution.

The Hamiltonian, ($H\mathbf{S}$), is:

$$H(\mathbf{S}) = -J \sum_{\langle \alpha, \beta \rangle} [S_\alpha S_\beta] - B \sum_{\alpha} [S_\alpha] \quad (11)$$

⁶As the dimension of the configuration space are discrete, the integrals become summations.

Where $\sum_{\langle\alpha,\beta\rangle}$ represents a sum over the interacting spin pairs,⁷ J is the interaction strength, and B is the external field strength.

The canonical ensemble was used to describe the system and so the distribution, $\rho(\mathbf{S})$, is:

$$\rho(\mathbf{S}) = \frac{e^{-H(\mathbf{S})}}{Z} \quad (12)$$

Where the partition function, Z is:

$$Z = \sum_{\mathbf{S}} e^{-H(\mathbf{S})} \quad (13)$$

It should be noted that the $1/k_B T$ factor has been absorbed into J and B .

It is easy to see that $\rho(\mathbf{S})$ is normalised:

$$\sum_{\mathbf{S}} \rho(\mathbf{S}) = \frac{1}{Z} \sum_{\mathbf{S}} e^{-H(\mathbf{S})} = \frac{Z}{Z} = 1 \quad (14)$$

Therefore a sensible strategy for solving equations (7) to (10) is importance sampling Monte Carlo as described in section 2.1.1.⁸ In this case we can estimate (9), the energy, for example, using the following equation:

$$E \approx \tilde{E} = \frac{1}{N} \sum_{i=1}^N H(\mathbf{S}_i) \quad (15)$$

Where \mathbf{S}_i are random configurations distributed according to the distribution, $\rho(\mathbf{S})$.

Similar equations can be derived for the other physical properties. The statistical errors on the estimates can be found using equation (3).

2.3.1 Method

The Ising ferro-magnet was studied in two and three dimensions, with the external magnetic field, B , set to zero. Only the interactions between nearest neighbours were considered and also periodic boundary conditions were used, giving the system the topology of a torus.

Configurations were generated using the the Metropolis algorithm. Perhaps the most intuitive way to achieve this would be to choose randomly whether or not to flip each spin on the lattice, and then accept or reject the new trial configuration using the Metropolis algorithm. However this leads to a trial configuration that is very different to the previous one, increasing the chance of rejection. A more efficient method is to attempt to flip only one spin at a time, ie. moving the walker in only one dimension at a time. The algorithm is then implemented by performing sweeps over the lattice, performing a Metropolis step for each spin.

In this scheme it is possible to optimise the Metropolis algorithm. As explained in section 2.2 the Metropolis algorithm accepts or rejects trial steps depending on the ratio of the distribution (or weight) functions (see equation (6)). Substituting in our distribution function (12) we find:

$$R(\mathbf{S}, \mathbf{S}') = \frac{\rho(\mathbf{S}')}{\rho(\mathbf{S})} = e^{-H(\mathbf{S}') + H(\mathbf{S})} \quad (16)$$

If we then substitute in the Hamiltonian, $H(\mathbf{S})$, we find:

$$R(\mathbf{S}, \mathbf{S}') = \exp(-2S_{i,j} [J(S_{i+1,j} + S_{i-1,j} + S_{i,j+1} + S_{i,j-1}) + B]) \quad (17)$$

⁷Care must be taken not to double count.

⁸Direct evaluation is not a realistic option. As for even a small lattice there are many possible configurations, eg. a 10×10 lattice has $2^{100} \approx 10^{30}$ configurations.

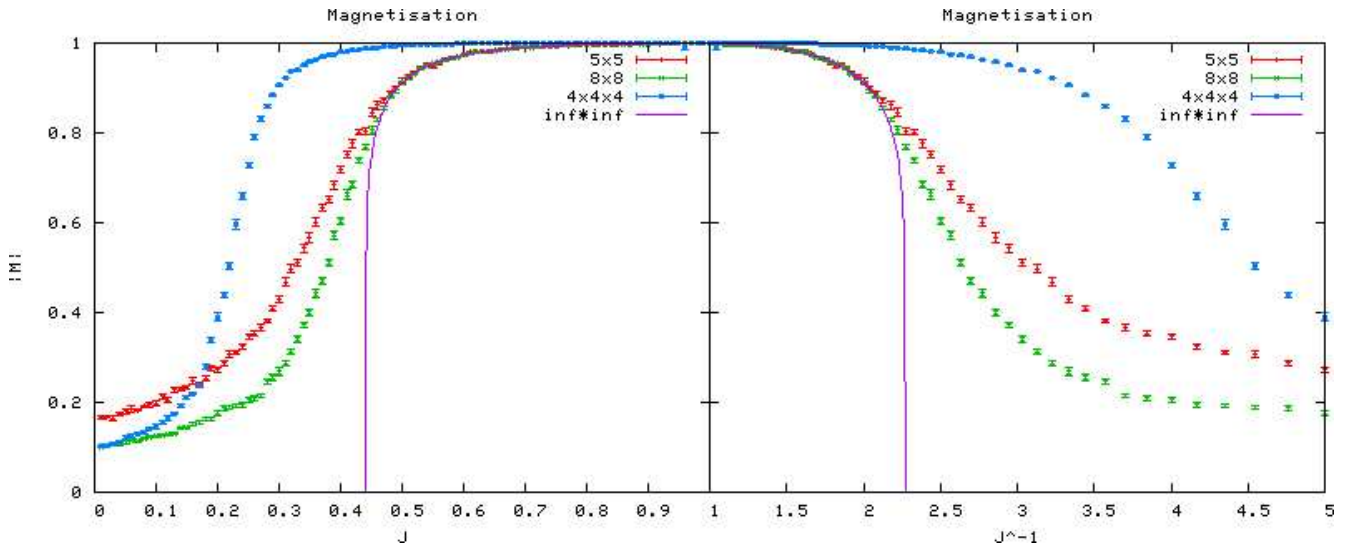


Figure 2: The modulus of magnetisation per spin, for various lattices, in the Ising ferro-magnet simulation.

Where $S_{i,j}$ is the spin attempting to flip.

This means that $R(\mathcal{S}, \mathcal{S}')$ can only take ten values.⁹ These values can easily be calculated at the start of the simulation and stored in a table to save the computational expense of recalculating the distribution function every time we attempt a spin flip.

2.3.2 Results

Lattices of differing sizes and topology were simulated. The smallest, composed of 25 spins, was a 2D lattice (5×5). Two 64 spin lattices were simulated, one of which was 2D (8×8) the other 3D ($4 \times 4 \times 4$). The external magnetic field, B was set to zero, and the interaction strength, J , was varied from 0 to 1.

Three of the four graphs are shown with the Onsager solution, for an infinite square lattice. Each graph is shown in two forms, the form on the left shows the behaviour when changing the interaction strength at constant temperature, whilst the other shows the behaviour when changing the temperature at constant interaction strength.

Figure 2 shows the modulus of the magnetisation per spin, of the lattices. The critical point is clearly seen at $J = 0.441$, on the Onsager solution. As $J \rightarrow 0$ the finite lattices do not reach zero magnetisation, unlike the Onsager solution. This is to be expected however as a random finite lattice will probably have some net magnetisation. The larger the lattice the closer the magnetisation will be to zero and this is seen in the results. As the interaction strength approaches zero, the effect of the topology of the lattice vanishes and so the two 64 spin lattices approach the same magnetisation for small J . The finite 2D lattices become magnetised gradually, compared to the infinite lattice which has a critical point at which magnetisation rises sharply. The 3D lattice gains magnetisation faster than its 2D counterpart. This also is to be expected due to the greater number of interactions per spin.

Figure 3 shows the magnetic susceptibility per spin, of the lattices. The 2D lattices peak just below the critical point of the Onsager solution. The larger the lattice, the higher the peak, and the closer the peak to the critical point. The 3D lattice peaks at a much lower value of J than its 2D counterpart; the peak itself also has a lower value.

⁹A more detailed discussion of this can be found in section 8.4 of [3]:p218

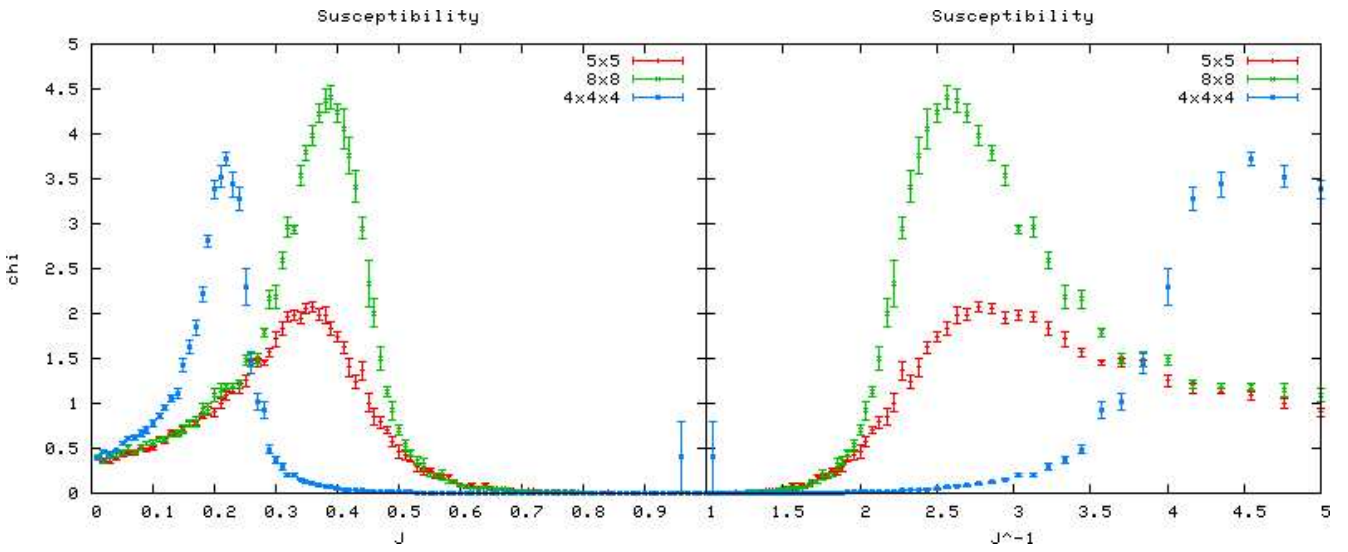


Figure 3: The magnetic susceptibility per spin, for various lattices, in the Ising ferro-magnet simulation.

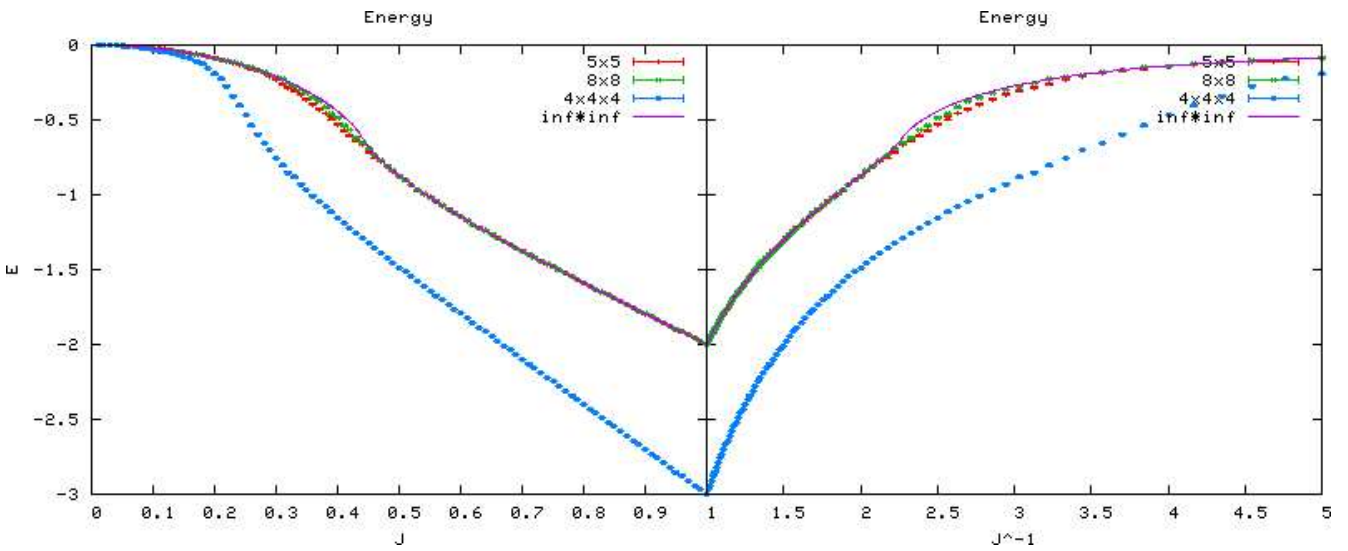


Figure 4: The energy per spin, for various lattices, in the Ising ferro-magnet simulation.

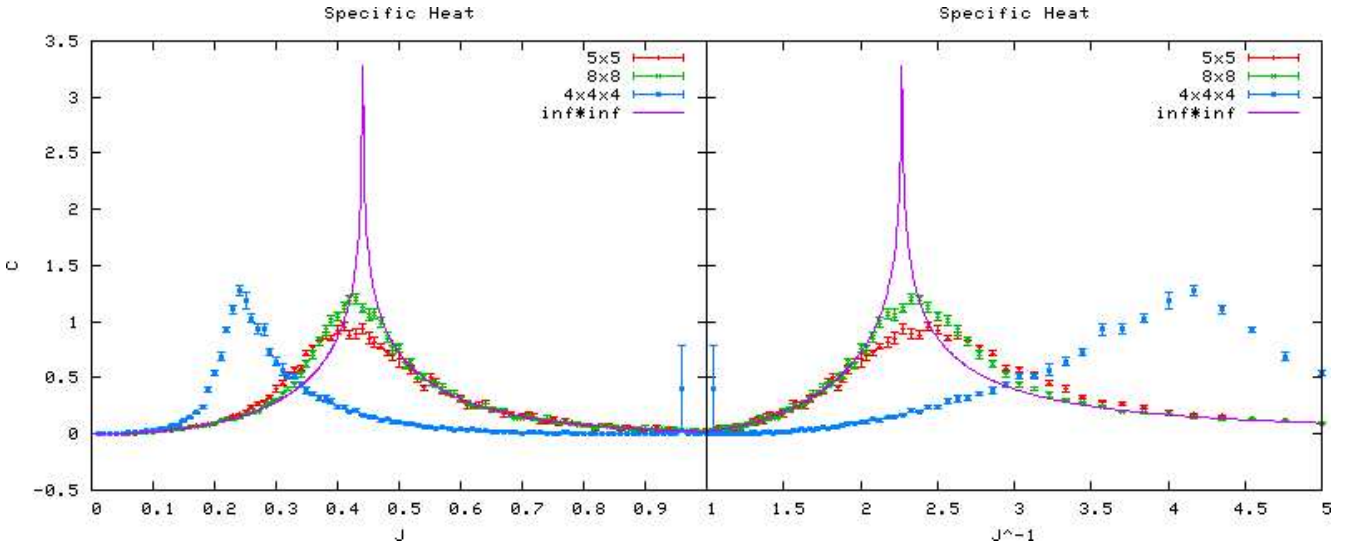


Figure 5: The specific heat per spin (for constant B), for various lattices, in the Ising ferro-magnet simulation.

Figure 4 shows the energy per spin, of the lattices. The behaviour of the 2D lattices is very close to the Onsager solution. Naturally the larger the lattice the closer its behaviour to that of the Onsager solution. The 3D lattice loses energy much faster as the interaction strength is increased (or as temperature is decreased).

Figure 5 shows the energy per spin (for constant B), of the lattices. The Onsager solution has a pole at the critical point. The finite 2D lattices do not have poles but rather reach a finite peak at a value of J just below the Onsager solution's critical point. Similar to the magnetic susceptibility, the larger the lattice the higher its peak, and the closer it is to the critical point. The 3D lattice also has a finite peak, at a similar height to its 2D counterpart, but at a much lower interaction strength.

3 Quantum Monte Carlo methods

This section covers Monte Carlo methods to calculate the ground-state energies of quantum systems.

3.1 Variational Monte Carlo

The variational energy given by:

$$E_v = \frac{\langle \psi_T(\mathbf{r}) | \hat{H} | \psi_T(\mathbf{r}) \rangle}{\langle \psi_T(\mathbf{r}) | \psi_T(\mathbf{r}) \rangle} \geq E_0 \quad (18)$$

Where E_v is the variational energy, \mathbf{r} is the configuration vector, $\psi_T(\mathbf{r})$ is any trial wave-function,¹⁰ \hat{H} is the Hamiltonian operator and E_0 is the ground-state energy.

provides us with an upper bounds on the ground-state energy. The closer the trial wave-function, $\psi_T(\mathbf{r})$, to the true ground-state, the lower the value of E_v , with the equality being satisfied only when the trial wave-function is the exact ground-state.

¹⁰The trial wave-function need not be normalised.

In general, equation (18) cannot be solved analytically, and so it must be solved through numerical methods. If we require that the trial wave-function, $\psi_T(\mathbf{r})$, is real,¹¹ we can rearrange (18) to:

$$E_v = \int \left[\frac{\psi_T^2(\mathbf{r})}{\langle \psi_T(\mathbf{r}) | \psi_T(\mathbf{r}) \rangle} E_L(\mathbf{r}) \right] d\mathbf{r} \quad (19)$$

Where $E_L(\mathbf{r})$ is called the *local energy* and is given by:

$$E_L(\mathbf{r}) = \frac{\hat{H}\psi_T(\mathbf{r})}{\psi_T(\mathbf{r})} \quad (20)$$

It is clear that if $\psi_T(\mathbf{r})$ is an energy eigenstate, then $E_L(\mathbf{r})$ does not depend on \mathbf{r} .

Once again a sensible strategy to solve equation (19) is importance sampling Monte Carlo, using the Metropolis algorithm,¹² in which case the equation is estimated by:

$$E_v \approx \tilde{E}_v = \frac{1}{N} \sum_{i=0}^N E_L(\mathbf{r}_i) \quad (21)$$

Where r_i are random configurations distributed according to $\rho(\mathbf{r}) = \psi_T^2(\mathbf{r}) / \langle \psi_T(\mathbf{r}) | \psi_T(\mathbf{r}) \rangle$.

Again it is clear that $\rho(\mathbf{r})$ is normalised:

$$\int [(\rho(\mathbf{r}))] d\mathbf{r} = \frac{1}{\langle \psi_T(\mathbf{r}) | \psi_T(\mathbf{r}) \rangle} \int [\psi_T^2(\mathbf{r})] d\mathbf{r} = \frac{\langle \psi_T(\mathbf{r}) | \psi_T(\mathbf{r}) \rangle}{\langle \psi_T(\mathbf{r}) | \psi_T(\mathbf{r}) \rangle} = 1 \quad (22)$$

If we plug $\rho(\mathbf{r})$ into (6) we see that the denominator cancels and we can therefore take $\rho(\mathbf{r}) = \psi_T^2(\mathbf{r})$.

The 1D quantum harmonic oscillator

In order to test the VMC method it was first applied to a very simple system, the one-dimensional quantum harmonic oscillator, which has the potential:

$$V(x) = \frac{1}{2}m\omega^2x^2 \quad (23)$$

Where m is the mass, ω is the angular frequency and x is the position and also the configuration of the system.

The ground state wave-function can be found analytically, and the unnormalised form is:

$$\psi_0(x) = \exp\left(\frac{-x^2}{2a^2}\right) \quad (24)$$

Where $a = \sqrt{\hbar/m\omega}$.

Finally the ground state energy is:

$$E_0 = \frac{\hbar\omega}{2} \quad (25)$$

In order to test the VMC method, $\hbar = m = \omega = 1$ is set and the following family of trial wave-functions used:

$$\psi_T(x) = \exp\left(\frac{-x^2e}{2a^2}\right) \quad (26)$$

¹¹As is the case for bosonic systems.

¹²The name of the entire method we have described is the *Variational Monte Carlo (VMC)*.

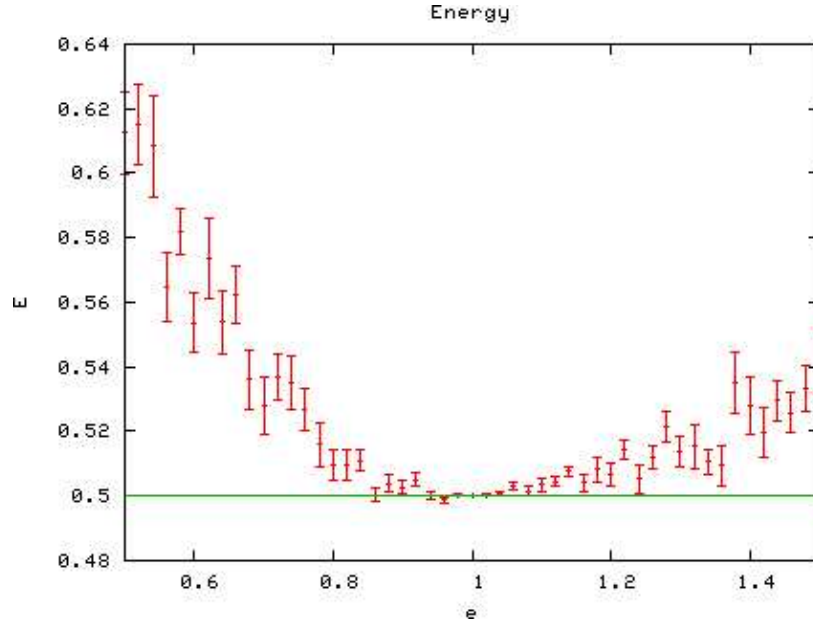


Figure 6: Energy of the 1D-harmonic oscillator from the standard VMC method. The greater the quality of the trial wave function, the closer to the result to the ground energy, and the smaller the error bars.

Where e is an artificial *variational parameter* on the width of the Gaussian wave-function.

This family includes the ground-state (when $e = 1$). Substituting (26) into (20) we find:

$$E_L(x) = \frac{e}{2}\hbar\omega + \frac{m\omega^2 x^2}{2} (1 - e^2) \quad (27)$$

which for $e = 1$ gives E_0 independent of x , as expected.

Results

Figure 6 shows the results of a run of the VMC method. The artificial variational parameter, e , is varied from 0.5 to 1.5. When $e = 1$ the trial wave-function is the exact ground-state and so the VMC method gives the exact ground energy with zero statistical errors. As e is varied away from 1, the result for the energy increases as well as the error on the result.

3.1.1 Fokker-Planck variational Monte Carlo

In imaginary time, $\tau = it$ and with zero potential, $V(\mathbf{r})$, the time-dependent Schrödinger equation (TDSE) becomes the Fokker-Planck equation (28).

$$\frac{\partial \rho(\mathbf{r}, t)}{\partial t} = \frac{1}{2} \nabla [\nabla - \mathbf{F}(\mathbf{r})] \rho(\mathbf{r}, t) \quad (28)$$

Where, $\hbar = 1$ (atomic units).

It is possible to formulate a VMC method using the Fokker-Planck equation.¹³ This scheme samples the distribution, $\rho(\mathbf{r}) = \psi_T^2(\mathbf{r})$, by simulating a diffusion process. The time-evolution operator is split using a 1st order approximation,¹⁴ and so small, discrete time steps are used.

¹³A more detailed explanation of this can be found in sections 12.2.4 and 12.2.5 of [1]:p320-328

¹⁴A 2nd order splitting could be used but has no benefit on the accuracy of the simulation.[1]:p362

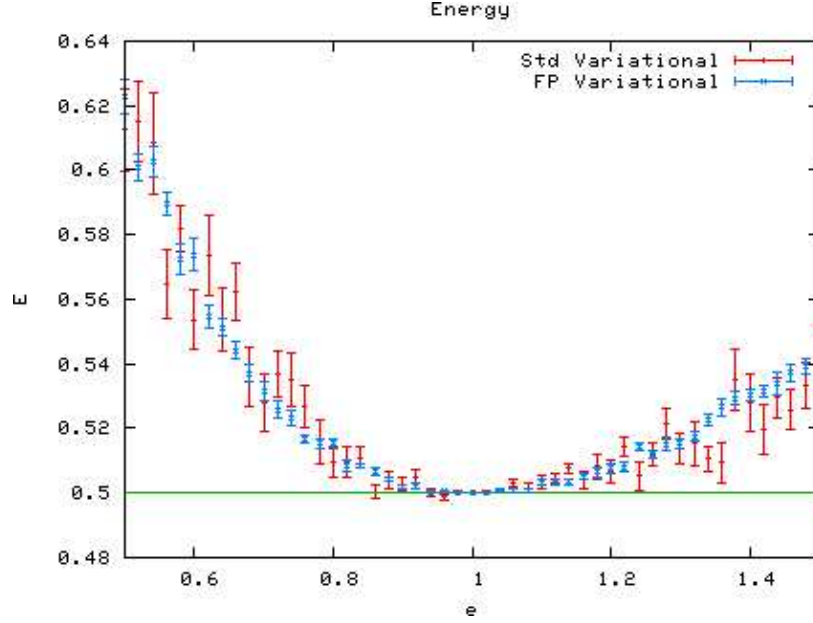


Figure 7: Energy of the 1D-harmonic oscillator. The Fokker-Planck VMC method is consistent with, and more efficient than, the standard VMC method.

Rather than moving walkers randomly in a hyper-cube as described in section 2.2, walkers are shifted from \mathbf{r} to $\mathbf{r}' = \mathbf{r} + \mathbf{F}(\mathbf{r})\Delta\tau/2 + \eta\sqrt{\Delta\tau}$. The ‘drift term’, $\mathbf{F}(\mathbf{r})$, is taken to be:

$$\mathbf{F}(\mathbf{r}) = \frac{2\nabla\psi_T(\mathbf{r})}{\psi_T(\mathbf{r})} \quad (29)$$

Unlike the standard VMC method, the transition probability of a walker moving from \mathbf{r} to \mathbf{r}' in a discrete imaginary-time step, $\Delta\tau$, is given by the Green’s function:

$$G(\mathbf{r}, \mathbf{r}'; \Delta\tau) = \frac{1}{(2\pi\Delta\tau)^{D/2}} \exp\left(-\frac{[\mathbf{r}' - \mathbf{r} - \Delta\tau\mathbf{F}(\mathbf{r})/2]^2}{2\Delta\tau}\right) + O(\Delta\tau^2) \quad (30)$$

Where D is the number of dimensions of the system.

Equation (30) is not symmetric under the interchange of \mathbf{r} and \mathbf{r}' and therefore if the Metropolis algorithm is to be used,¹⁵ the generalised form, known as the Metropolis-Hastings algorithm,^[4] must be used in order to guarantee detailed balance. This means that (6) now becomes:

$$R(\mathbf{r}, \mathbf{r}') = \frac{T_{\Delta\tau}(\mathbf{r}' \rightarrow \mathbf{r})\rho(\mathbf{r}')}{T_{\Delta\tau}(\mathbf{r} \rightarrow \mathbf{r}')\rho(\mathbf{r})} = \frac{G(\mathbf{r}', \mathbf{r}; \Delta\tau)\psi_T^2(\mathbf{r}')}{G(\mathbf{r}, \mathbf{r}'; \Delta\tau)\psi_T^2(\mathbf{r})} \quad (31)$$

The 1D quantum harmonic oscillator

The Fokker-Planck Variational Monte Carlo (*FP-VMC*) method was tested on the one-dimensional quantum harmonic oscillator, allowing the method to be compared to the standard VMC method.

Results

Figure 7 shows the results of a run of the FP-VMC method of the one-dimensional quantum harmonic oscillator. It was created from 10 walkers, each generating 1,000 configurations. The walkers

¹⁵This is not required, but importance sampling will correct for the discrete time step.

underwent 100,000 thermalisation steps and used 300 steps between accepted configurations. The imaginary step size was set to $\Delta\tau = 0.01$.

The results are consistent with the results of the standard VMC method (originally from figure 6), however they have smaller statistical errors, due in part to the greater number of configurations use. Despite using 10 times more configurations than the standard VMC run, the FP-VMC run was faster to execute.

3.2 Green's Function Diffusion Monte Carlo

Projector or *Diffusion Monte Carlo* (DMC) methods provide a way to find the exact ground-state energy of a quantum system, even when the ground-state wave-function is not known. The method described in this section is called *Green's function Diffusion Monte Carlo with a guide function* (GF-DMC).¹⁶ In section 3.1.1 we mentioned that in imaginary time and with no potential, the TDSE becomes (28), the Fokker-Planck equation. This time the potential, $V(\mathbf{r})$, is not set to zero. Taking $\rho(\mathbf{r}, \tau) = \psi_T(\mathbf{r})\Psi(\mathbf{r}, \tau)$ to find the modified Fokker-Planck equation:

$$\frac{\partial \rho}{\partial \tau} = \frac{1}{2} \nabla (\nabla - \mathbf{F}(\mathbf{r})) \rho(\mathbf{r}, \tau) - (E_L(\mathbf{r}) - E_T) \rho(\mathbf{r}, \tau) \quad (32)$$

Where E_T is an artificial term used to normalise the Green's function.

This equation describes a diffusion process with a 'potential'. The Green's function takes a similar form to before, (30):

$$G(\mathbf{r}, \mathbf{r}'; \Delta\tau) = e^{(-\Delta\tau[E_L(\mathbf{r}) - E_T])} \frac{1}{(2\pi\Delta\tau)^{D/2}} \exp\left(\frac{-[\mathbf{r}' - \mathbf{r} - \Delta\tau\mathbf{F}(\mathbf{r})/2]^2}{2\Delta\tau}\right) + O(\Delta\tau^2) \quad (33)$$

The Green's function, (33), is now a small-time approximation of the imaginary time operator.

$$|\Psi(\mathbf{r}, t)\rangle = e^{-i\hat{H}t/\hbar} |\psi(\mathbf{r}, 0)\rangle \quad (34)$$

$$\begin{aligned} |\Psi(\mathbf{r}, \tau)\rangle &= e^{-\hat{H}\tau/\hbar} |\psi(\mathbf{r}, 0)\rangle \\ &= \sum_{n=0}^{\infty} c_n e^{-E_n\tau/\hbar} |\psi(\mathbf{r}, 0)\rangle \end{aligned} \quad (35)$$

It can be seen from (35) that as we evolve the distribution in imaginary time all the energy eigenstates decay. However, it is the ground-state energy that decays the slowest. Therefore the GF-DMC method gives the ground-state energy, unless the trial wave-function is orthogonal to the true ground-state.

The most efficient way to implement the first term of (33) is in a 'branching process'. So one step in the GF-DMC algorithm is composed of a diffusion step and a branching step. This means the diffusion step is simply the Metropolis step taken in exactly the same way as described in section 3.1.1 for the FP-VMC, if the step is accepted then the branching step is performed. The branching step works by creating or destroying walkers. It calculates, $q = \exp(-\Delta\tau([E_L(\mathbf{r}') + E_L(\mathbf{r})]/2 - E_T))$, if $q < 1$ then the walker is destroyed with probability, $1 - q$, (and survives with probability, q) otherwise if $q > 1$ then either $[q]$ or $[q - 1]$ new walkers are created at the current

¹⁶A version that does not use a guide function is described in section 12.3.1 of [1]:p238-331. Alternatively [3]:p225-229 and [1]:p341-352 describe *Path Integral Diffusion Monte Carlo* methods.

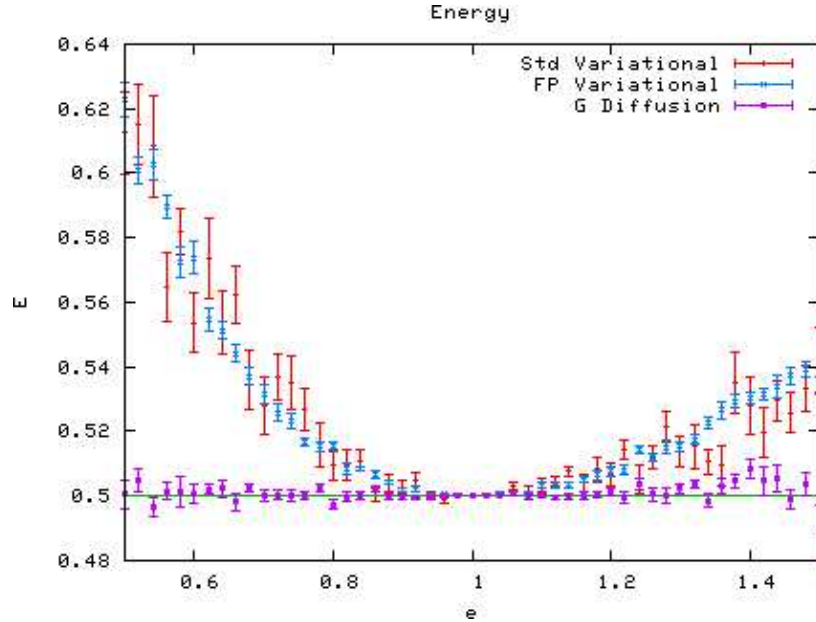


Figure 8: Energy of the 1D-harmonic oscillator. Regardless of the trial wave-function, the DMC method gives the ground state, although the quality of the trial wave-function has a strong effect on the errors.

walker’s position (with probabilities, $q - [q]$ and $1 - q + [q]$ respectively), where $[q]$ is q truncated to an integer.¹⁷

As previously mentioned, E_T is an artificial parameter used to normalise the Green’s function. In terms of the branching process its job is to hold the number of walkers in the simulation roughly constant. In order to do this it must be adjusted at every time-step, using the following equation:

$$E_T = E_0 + \alpha \cdot \ln \left(\frac{\tilde{N}}{N} \right) \quad (36)$$

Where N is the current number of walkers, \tilde{N} is the ‘target’ number of walkers, α is some small parameter and E_0 is an estimate of the ground-state.¹⁸

As the simulation progresses, E_T approaches the ground-state energy.

3.2.1 The 1D quantum harmonic oscillator

The one-dimensional quantum harmonic oscillator was used once again in order to test the GF-DMC method.

Results

Figure 8 shows the results of a run of the GF-DMC method of the one-dimensional quantum harmonic oscillator. It was created by first executing the FP-VMC method, using 10 walkers, each generating 1,000 configurations. The walkers underwent 100,000 thermalisation steps and used 300 steps between accepted configurations. The imaginary-time step size was set to $\Delta\tau = 0.01$.

¹⁷So for example, if $q = 2.34$, then $[q] = 2$.

¹⁸A sensible choice for E_0 would be the previous value of E_T , as E_T tends towards the ground-state energy. For the first time-step, E_0 could be taken from the variational energy.

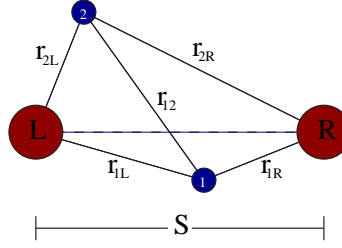


Figure 9: Vectors used in calculations for simulating H_2 and He.

Then all 10,000 configurations became walkers in the GF-DMC and underwent 500 imaginary-time steps.

Unlike the previous VMC runs, the GF-DMC method gives the ground-state energy regardless of the variational parameter in the trial-wavefunction. However, it can be seen that the quality of the trial wave-functions affects the statistical errors on the GF-DMC results, as the errors vanish around $e = 1$.

3.3 Case Study: The hydrogen molecule

The Hydrogen molecule can be simulated using the quantum Monte Carlo methods discussed in this report. We take the Born-Oppenheimer approximation and hold the protons in fixed positions whilst simulating the bond made by the electrons. Atomic units are used.

The potential governing the inter-proton separation is given by:

$$U(S) = \frac{e^2}{S} + E_0(S) \quad (37)$$

Where the first term is the Coulomb repulsion of the protons, e is the electronic charge, S is the inter-proton separation and $E_0(S)$ is the ground-state energy eigenvalue of the two electron system.

$E_0(S)$ is found by solving the Schrödinger equation for the two electrons:

$$\begin{aligned} \hat{H}\psi(\mathbf{r}_1, \mathbf{r}_2; S) &= \frac{-\hbar^2}{2m} \sum_{i=1}^2 \nabla_i^2 \psi(\mathbf{r}_i, \mathbf{r}_2; S) + V(\mathbf{r}_1, \mathbf{r}_2; S)\psi(\mathbf{r}_1, \mathbf{r}_2; S) \\ &= E_0(S)\psi(\mathbf{r}_1, \mathbf{r}_2; S) \end{aligned} \quad (38)$$

Where m is the mass of an electron and \mathbf{r}_i is the position of the i^{th} electron.

Considering only the Coulomb force in the potential, $V(\mathbf{r}_1, \mathbf{r}_2; S)$:

$$V(\mathbf{R}) = V(\mathbf{r}_1, \mathbf{r}_2; S) = e^2 \left[\frac{1}{r_{12}} - \frac{1}{r_{1L}} - \frac{1}{r_{1R}} - \frac{1}{r_{2L}} - \frac{1}{r_{2R}} \right] \quad (39)$$

Where r_{ij} are the distances between particles, as labeled in figure 9, and \mathbf{R} is the 6 dimensional configuration vector, encompassing the two 3 dimensional particle positions \mathbf{r}_1 and \mathbf{r}_2 .

The following family of trial wave-functions were chosen:

$$\psi_T(\mathbf{R}) = \psi_T(\mathbf{r}_1, \mathbf{r}_2) = \phi(\mathbf{r}_1)\phi(\mathbf{r}_2)f(r_{12}) \quad (40)$$

$$\phi(\mathbf{r}_i) = e^{-r_{iL}/a} + e^{-r_{iR}/a} \quad (41)$$

$$f(r) = \exp\left(\frac{r}{\alpha(1 + \beta r)}\right) \quad (42)$$

Where $\phi(\mathbf{r}_i)$ is an independent-particle wave-function, and $f(r_{12})$ is the term that deals with the correlation between the two electrons due to their Coulomb repulsion. a , α and β are variational parameters.

Because the wave-function is required to meet the Coulomb ‘cusp’ conditions, it is found that $\alpha = 2a_0$ (where a_0 is the Bohr radius), and that a satisfies the following transcendental equation:¹⁹

$$a = \frac{a_0}{1 + e^{-S/a}} \quad (43)$$

This can easily be solved numerically, to the required degree of accuracy, using the Newton-Raphson method, leaving only a single variational parameter, β .

After a great deal of algebra, the local energy is found to be:

$$E_L(\mathbf{R}) = -\frac{1}{2} \sum_{i=1}^2 \left[\frac{2}{r_{12}\alpha(1 + \beta r_{12})^3} + \frac{1}{a^2} + \frac{1}{\alpha^2(1 + \beta r_{12})^4} - \frac{2(e^{-r_{iL}/a}/r_{iL} + e^{-r_{iR}/a}/r_{iR})}{a(e^{-r_{iL}/a} + e^{-r_{iR}/a})} + \frac{(-1)^i \cdot 2(e^{-r_{iL}/a} \mathbf{r}_{iL} \cdot \mathbf{r}_{12} + e^{-r_{iR}/a} \mathbf{r}_{iR} \cdot \mathbf{r}_{12})}{a\alpha(1 + \beta r_{12})^2(e^{-r_{iL}/a} + e^{-r_{iR}/a})} \right] + V(\mathbf{R}) \quad (44)$$

3.3.1 The helium atom

The helium atom is also governed by the equations stated above. It is simply the special case where $S = 0$, ie. the protons are ‘on top of each other’. (Because we are holding the protons fixed and are only considering the coulomb interaction, we do not need to concern ourselves with the neutrons.)

Because the value of S for helium is known, it provides a convenient system for which the value of β can be determined, whilst at the same time calculating the complete ionisation energy of a system, whos energy is known from experiment.

Figure 10, shows the results of a FP-VMC run, generated using 10 walkers, each producing 3,000 configurations. The imaginary-time step was set to 0.0025, and the walkers underwent 100,000 thermalising steps, and took 600 steps between accepted configurations. A parabola was fitted to the points around the minimum ($0.11 \leq \beta \leq 0.30$), which gave the minimum to be at $\beta = 0.195 \pm 0.005$. The fit gave a reduced chi-squared of 1.25.

Figure 10, also shows results from the GF-DMC. To which horizontal lines are fitted. The first run had the imaginary-time step set to 0.005, and 800 imaginary-time steps taken. This gives the result of -2.9001 ± 0.0004 Hartrees, with a reduced chi-squared of 1.79. The second run had the imaginary-time step set to 0.0025, and 800 imaginary-time steps taken. This gives the result of -2.902 ± 0.001 Hartrees, with a reduced chi-squared of 0.41. The exact value of the complete ionisation of helium is 2.903 Hartrees.²⁰ The high value of the first GF-DMC run is probably due to the larger imaginary-time step.

3.3.2 The hydrogen molecule

Once β has been determined to a good degree of accuracy, the inter-proton separation, S , can be varied to find the potential, $U(S)$, as described by (37).

Figure 11 displays the results of varying S . Fitting a parabola to the minimum gave the location as 1.43 ± 0.01 Bohr radii, at this value the complete ionisation energy is, -1.166 ± 0.001 Hartrees, although the fit had a reduced chi-squared of 2.01. The form of the fit is certainly close to that found in other works.^[5]

¹⁹More detailed discussion can be found in [3]:p223-225

²⁰_{[1]:p318}

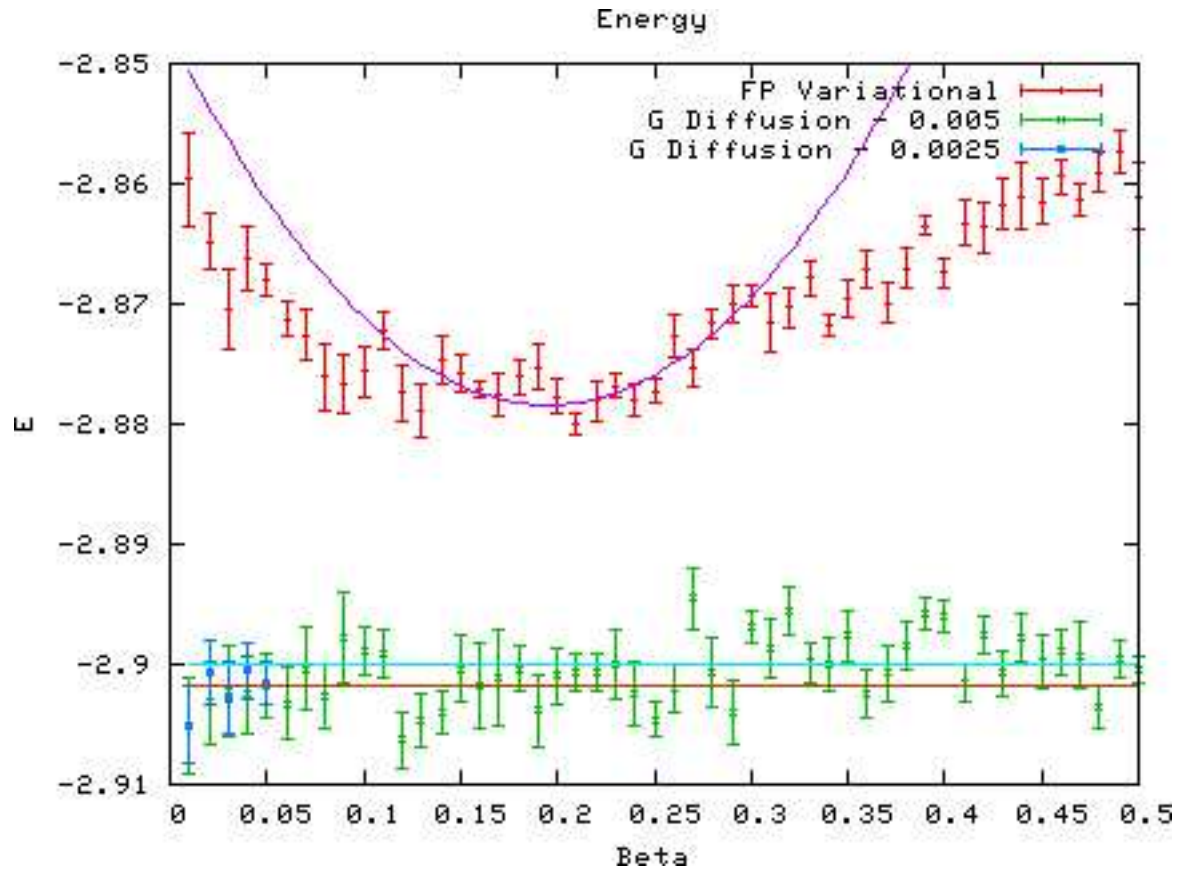


Figure 10: Complete ionisation energy of He atom results, against the value of the variational parameter, β . The FP-VMC run gives us the optimal value of β , the GF-DMC run, gives a value for the complete ionisation energy.

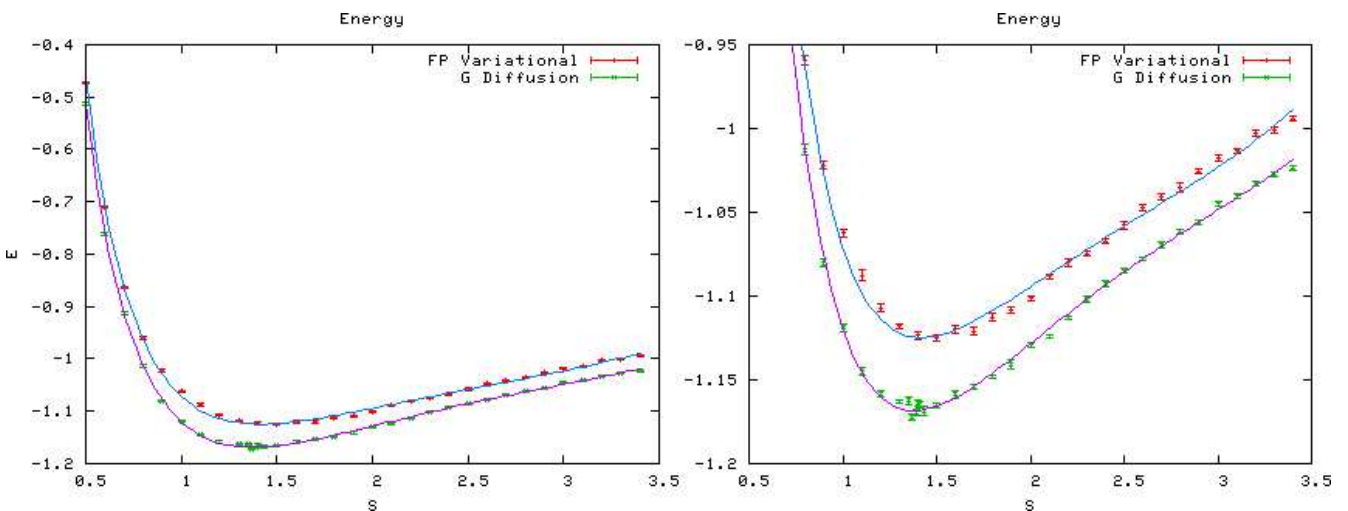


Figure 11: Complete ionisation energy of H₂ molecule results, against the inter-proton separation, S .

4 Conclusion

Computer programs have been written to use Monte Carlo methods to simulate various systems. The Ising ferro-magnet has been successfully simulated in two and three dimensions and its behaviour reported. Various quantum monte carlo methods have been implemented and investigated, each one having been tested on the simple quantum harmonic oscillator system. The helium atom has been simulated and its complete ionisation energy was measured to be -2.902 ± 0.001 Hartrees, consistent with experimental data. The hydrogen molecule was simulated and found to have a complete ionisation energy of -1.166 ± 0.001 Hartrees, with an inter-proton separation of 1.43 ± 0.01 Bohr radii. However doubt is cast over the accuracy of these figures given that the reduced chi-squared of the fit was 2.01.

The project could be extended in many ways. For instance the more sophisticated trial wavefunction could be used for the hydrogen molecule in order to increase the accuracy of the hydrogen and helium results. The effects of the discrete imaginary-time step could be investigated, as well as the effect of α in (36). Larger extensions would be to implement a path integral Monte Carlo method, and investigate the difference with the Green's function Monte Carlo already implemented in this project. Alternatively the fixed-node approximation could be implemented and fermionic systems such as the lithium atom could be investigated.

The author's personal contribution to the project included writing the programs from scratch and fixing its 'bugs', using the programs to generate data, creating graphs from the data, and studying the graphs, whilst he relied on his colleague, P. Hughes, to carry out background research, to handle much the mathematics involved, and to make and plot further graphs. This project has been interesting as it uses such a conceptually simple view of quantum mechanics and because it has provided the foundation for performing more complex simulations. The author would like to express his thanks to Prof F. Forshaw for his guidance over the course of this project.

References

- [1] J.M. Thijssen, "Computational physics", Cambridge University Press (1999)
- [2] N. Metropolis, A.E. Rosenbluth, M.N. Rosenbluth, A.H. Teller, E. Teller, "Equation of state calculations by fast computing machines", J. Chem. Phys. 21, (1953) p1087-1092
- [3] S.E. Koonin, D.C. Meredith, "Computational physics", Addison-Wesley Publishing (1990)
- [4] W.K. Hastings, "Monte Carlo Sampling Methods Using Markov Chains and Their Applications", Biometrika, 57, 1 (1970) p97-109
- [5] Ko, Wing Ho "Demonstrating quantum Monte Carlo methods through the study of hydrogen molecule",
<http://www.physics.uiuc.edu/education/undergraduate/REU/2004%20REU/Final%20Papers/wingko.pdf> (Accessed 14th May 2007)
- [6] J. Forshaw, private correspondence