

Chapter 6

Monte Carlo Methods in Quantum Physics

Quantum Monte Carlo (QMC) methods are widely used nowadays for systems with large number of particles. In main distinction of these simulation techniques is in their stochastic nature, which is not the case for other simulation methods, such as molecular dynamics. For more on the topic please refer to [14] and [17].

These methods are flexible and at the same time accurate for many-body systems. There are two most common types of QMC, first is variational Monte Carlo method (VMC) and projector Monte Carlo method (PMC). For PMC methods differ by a form of projector, e.i. exponential or linear. The most common choice is an exponential one also sometimes called as imaginary time propagator.

The essence of VMC method is to choose a trial wave function, depending on some parameters and then for this particular trial wave function to find expectation values of operators and optimize the parameters.

At the same time PMC methods also use trial wave functions. In general PMC is a stochastic application of power method to compute expectation values for eigenstate corresponding to largest absolute eigenvalue. The main advantage of the PMC methods compared to the common matrix diagonalization algorithms, for example iterative and deterministic Lanczos algorithm, is that for stochastic method we do not need to store the whole vector, but just some randomly sampled elements of the vector [31].

In this thesis we are implementing a Coupled Cluster Quantum Monte Carlo (CC-QMC) method which has much in common with one of the PMC methods a diffusion Monte Carlo (DMC) method. Below we present an overview of some of these methods.

6.1 Variational Monte Carlo

The expectation value of energy is expressed as following:

$$\langle E \rangle = \frac{\int d\mathbf{R} \Psi^*(\mathbf{R}) H(\mathbf{R}) \Psi(\mathbf{R})}{\int d\mathbf{R} \Psi^*(\mathbf{R}) \Psi(\mathbf{R})}, \quad (6.1)$$

where $\Psi(\mathbf{R})$ is a wave-function of the system and $H(\mathbf{R})$ is a Hamiltonian. \mathbf{R} stands for the spatial coordinates of all the particles. For the system containing many particles both wave-function and Hamiltonian depend on all coordinates of all the particles. In this case integral (6.1) becomes multidimensional. Additional challenge is a wave-function which is unknown. For the VMC algorithm one can replace it with some kind of parametrized function, usually referred to as trial wave function. If the trial wave-function is chosen to take into account as much physics of the system as possible, the value of $\langle E \rangle$ converged to ground state energy $\langle E_0 \rangle$.

The main steps of VMC method are the following:

- Choice of a trial wave-function which depends on variational parameters.
- Evaluation of the expectation value of the Hamiltonian.
- Minimization algorithm in space of variational parameters.
- Monte Carlo calculations with optimal parameters and error estimation.

6.2 Diffusion Monte Carlo

The DMC method employ the imaginary time propagator:

$$P(\tau) = e^{\tau(E_T \hat{1} - \hat{H})}, \quad (6.2)$$

here $\tau = it$ is imaginary time.

This method is called diffusion one can do a mapping between a classical diffusion equation and an imaginary-time Schrödinger equation. Let's consider the following form of Schrödinger equation:

$$\frac{\partial}{\partial \tau} |\Psi(\mathbf{x}, \tau)\rangle = \frac{1}{2} \nabla^2 |\Psi(\mathbf{x}, \tau)\rangle + (E_T - V(\mathbf{x})) |\Psi(\mathbf{x}, \tau)\rangle, \quad (6.3)$$

This equation is very similar to a classical equation:

$$\frac{\partial C}{\partial t} = D \frac{\partial^2 C}{\partial t^2} - kC, \quad (6.4)$$

here D is diffusion constant, C is concentration of particles and k proportionality coefficient describing source or sink of particles. First part on the right hand side of this equation describe a classical diffusion process and the second part is just a first order rate equation. The meaning of k is the following - particles can be destroyed or created proportional to k .

If we compare two equations together (6.3) and (6.4) they are not really alike, due to imaginary time and a wave function. However similarities between this two allow us to use some techniques normally used for classical simulations to imaginary-time Schrödinger equation. Monte Carlo methods are capable to simulate both diffusion

and rate, so one can just build up a method that combine this two processes into one.

In this thesis we do not aim to dig into the DMC method and derive any equations. The main idea of the DMC is to represent the wave function as a population of random walkers. The trial wave function obtained for the VMC can be used as input for DMC method to improve the result. The DMC is one of the most accurate methods for the many body problems so far.

6.3 Stochastic Coupled Cluster Theory

In chapter 5 we have introduced a coupled cluster method. However, the main problem with exact coupled cluster computations is with system size. Scaling starts with the sixth order of N (here N represents system size) that makes even small systems very costly in computations. Coupled Cluster Monte Carlo (CCMC) is a new method which has been first described by Alex Thom in [30]. The method has been developed similarly to Full Configuration Interaction Quantum Monte Carlo (FCIQMC) of Alavi *et al.* in [6], [9] and [5].

In order to explain how we get to CCMC algorithm we first present a brief description of FCIQMC.

6.3.1 Full Configuration Interaction Quantum Monte Carlo

The CI wave function is represented by the following expansion:

$$|\Psi_{\text{CI}}\rangle = C_0 |D_0\rangle + \sum_{ai} C_i^a |D_i^a\rangle + \sum_{i<j, a<b} C_{ij}^{ab} |D_{ij}^{ab}\rangle + \dots, \quad (6.5)$$

here $|D_0\rangle^1$ is HF determinant, $|D_i^a\rangle$ and $|D_{ij}^{ab}\rangle$ are excited determinants.

In order to determine coefficients C we need to solve a set of projection equations.

$$\langle D_0 | \hat{H} - E | \Psi_{\text{CI}} \rangle = 0 \quad (6.6)$$

$$\langle D_i^a | \hat{H} - E | \Psi_{\text{CI}} \rangle = 0 \quad (6.7)$$

$$\langle D_{ij}^{ab} | \hat{H} - E | \Psi_{\text{CI}} \rangle = 0 \quad (6.8)$$

$$\langle D_{ijk}^{abc} | \hat{H} - E | \Psi_{\text{CI}} \rangle = 0, \quad (6.9)$$

and so on.

Instead of performing this process iteratively we now need to re-express equations. As we already mentioned in section 6.2 while describing DMC this can be done using the imaginary time propagator. In this case operator is given by:

$$\hat{P} = 1 - \tau(\hat{H} - E). \quad (6.10)$$

¹we use $|D_0\rangle$ for $|\Psi_0\rangle$, $|D_i^a\rangle$ for $|\Psi_i^a\rangle$ and so on to distinguish between deterministic and stochastic methods.

It projects CI wave function to itself. Using this approach we will get a new set of equations, formulated as:

$$C_{\mathbf{I}} - \tau \langle D_{\mathbf{I}} | \hat{H} - E | D_{\mathbf{I}} \rangle C_{\mathbf{I}} - \tau \sum_{\mathbf{J} \rightarrow \mathbf{I}, \mathbf{J} \neq \mathbf{I}} \langle D_{\mathbf{I}} | \hat{H} | D_{\mathbf{J}} \rangle C_{\mathbf{J}} = C_{\mathbf{I}}, \quad (6.11)$$

here \mathbf{I}, \mathbf{J} are some generic indexes and $\mathbf{J} \rightarrow \mathbf{I}$ means that this two indexes are connected (corresponding matrix element of Hamiltonian is different from zero). As in DMC we may study two processes independently, namely:

$$C_{\mathbf{I}} - \tau \langle D_{\mathbf{I}} | \hat{H} - E | D_{\mathbf{I}} \rangle C_{\mathbf{I}} \rightarrow C_{\mathbf{I}} \quad (6.12)$$

$$C_{\mathbf{I}} - \tau \sum_{\mathbf{J} \rightarrow \mathbf{I}, \mathbf{J} \neq \mathbf{I}} \langle D_{\mathbf{I}} | \hat{H} | D_{\mathbf{J}} \rangle C_{\mathbf{J}} \rightarrow C_{\mathbf{I}} \quad (6.13)$$

$$(6.14)$$

After this we need to discretize the coefficients. To simulate this stochastically we should create a population of walkers in the SD space. Each walker has a sign and belong to any determinant. Number of walkers on each specific determinant, e.i. $D_{\mathbf{I}}$ is proportional to coefficient $C_{\mathbf{I}}$. The population dynamics in this case should obey equations (6.13) and (6.14). In order to run the simulation we need to consider three processes:

1. **Spawning.** As soon as walker can be positioned on any determinant, we consider random walker at determinant $D_{\mathbf{J}}$. After this we randomly pick up a connected determinant $D_{\mathbf{I}}$ (assuming $|H_{\mathbf{IJ}}| \neq 0$). The probability to spawn a walker from $\mathbf{J} \rightarrow \mathbf{I}$ is proportional to $\tau |H_{\mathbf{IJ}}|$ and sign $-\text{sgn}(H_{\mathbf{IJ}})$. The spawning process is corresponding to equation (6.14) and can be formulated as $-\tau H_{\mathbf{IJ}} C_{\mathbf{J}} + C_{\mathbf{I}} \rightarrow C_{\mathbf{I}}$.
2. **Birth/Death.** Here we simulate creation or destruction of the already existing walker at some determinant, for example $D_{\mathbf{I}}$. This process corresponds to the equation (6.13) and the probability is now proportional to $-\tau(H_{\mathbf{II}} - E)$ and sign distinguish between either walker is created or destroyed. Negative sign corresponds to death and positive sign corresponds to birth.
3. **Annihilation.** All pairs of walkers with opposite sign are removed, if they belong to the same determinant.

The population control is carried out by varying of the remaining parameter, so-called shift E . This parameter can influence the population dynamics in a following way - if the value of E exceeds the lowest eigenvalue, then the population of walkers will increase and vice versa.

6.3.2 Coupled Cluster Quantum Monte Carlo

Following the same procedure as for FCIQMC we present a CC wave function in a similar way as we did it for CI in equation (6.5). The CC wave function is represented by the following expansion:

$$|\Psi_{CC}\rangle = N_0 e^{\frac{\hat{T}}{N_0}} |D_0\rangle, \quad (6.15)$$

where N_0 determines the normalization of wave function and the equation for cluster operator \hat{T} is presented in description of deterministic CC method in Chapter 5 by equation (5.3).

Let's write it in a more detailed way using a cluster operator equation:

$$|\Psi_{CC}\rangle = N_0 |D_0\rangle + \sum_{\mathbf{I}} |D_0\rangle t_{\mathbf{I}} \hat{c}_{\mathbf{I}} + \frac{1}{2N_0} \sum_{\mathbf{I}\mathbf{J}} |D_0\rangle t_{\mathbf{I}} t_{\mathbf{J}} \hat{c}_{\mathbf{I}} \hat{c}_{\mathbf{J}} |D_0\rangle + \dots, \quad (6.16)$$

where indexes \mathbf{I}, \mathbf{J} are generic indexes for an excitation.² We use this kind of notations to make equations more alike those for FCIQMC.

After this we can present the CC equations:

$$\langle D_0 | \hat{H} - E | \Psi_{CC} \rangle = 0, \quad (6.17)$$

$$\langle D_i^a | \hat{H} - E | \Psi_{CC} \rangle = 0, \quad (6.18)$$

$$\langle D_{ij}^{ab} | \hat{H} - E | \Psi_{CC} \rangle = 0, \quad (6.19)$$

$$(6.20)$$

and so on.

However we are not following the traditional practice here and not rewrite (6.17) to create iterative pattern. We are going to consider the following equation instead:

$$\langle D_0 | 1 - \delta\tau(\hat{H} - E) | \Psi_{CC} \rangle = \langle D_{\mathbf{m}} | \Psi_{CC} \rangle, \quad (6.21)$$

here $\delta\tau$ is just a small positive number. Here we must make a remark regarding the value of $\delta\tau$. It must satisfy the following inequality:

$$\delta\tau \leq \frac{2}{E_{max} - E}, \quad (6.22)$$

here E_{max} is the largest eigenvalue of \hat{H} , E should be close to the ground state energy. It is common to take E equal to reference energy in the beginning of simulation. Let's now consider a projection on a single excited determinant $D_{\mathbf{m}}$:

$$\langle D_{\mathbf{m}} | e^{\hat{T}} D_0 \rangle = t_{\mathbf{m}} + \mathcal{O}(\hat{T}^2), \quad (6.23)$$

²Indexing is different from the deterministic CC theory, because we are using notations by Thom and Spencer from [29]. In this notation each cluster has just one generic index, regardless of level of the excitation for example, $\hat{c}_a^\dagger \hat{c}_i$ is just written as $\hat{c}_{\mathbf{I}}$.

here we do not consider contributions from the higher terms. After doing this we obtain:

$$t_{\mathbf{m}} - \delta\tau \langle D_{\mathbf{m}} | (\hat{H} - E) | \Psi_{CC} \rangle = t_{\mathbf{m}}. \quad (6.24)$$

Now we are ready to rewrite the equation (6.17) in an iterative way:

$$t_{\{\mathbf{m}\}}(\tau) - \delta\tau \langle D_{\mathbf{i}} | (\hat{H} - E) | \Psi_{CC}(\tau) \rangle = t_{\mathbf{m}}(\tau + \delta\tau). \quad (6.25)$$

Second term in the equation above is not easy to evaluate. The solution here is to do it stochastically by sampling this term. In order to do this we should consider two stochastic processes: sampling of the wave function and sampling of the action of Hamiltonian.

Sampling of the wave function is achieved by a random selection of all possible clusters and after this is done collapsing them into an excited determinant.

Sampling of the action of Hamiltonian is done after this. Once we get an excited determinant $D_{\mathbf{n}}$ we select another excited determinant $D_{\mathbf{m}}$ a single or double excitation of this chosen determinant and compute the corresponding $t_{\mathbf{m}}$. However this is not a very efficient algorithm as we have to sample a large number of Monte Carlo cycles. This comes from the fact that many clusters have small amplitudes. A better way is to follow the same discretization procedure as have been done in the FCIQMC method. Here we follow Thom in his article from 2010 [30].

We are aiming to create a population of walkers in order to discretize the amplitudes. In case of CCQMC walkers are called "excips" and the corresponding operators are named "excitors"³.

6.3.3 CCQMC algorithm

In section 6.3.1 we have already presented the main steps for the FCIQMC. For the CCQMC method we have a very similar algorithm:

1. **Spawning.** After selecting a number of clusters and collapsing them into an excited determinant $D_{\mathbf{n}}$ one should pick a connected determinant $D_{\mathbf{m}}$. A new excip is created with the probability proportional to $|\delta\tau H_{\mathbf{nm}}|$.
2. **Birth/Death.** Once we selected a determinant it is associated with a population of excips living on it (might be zero as we have composite clusters). This excip can die with probability $-\tau(H_{\mathbf{nm}} - E)$. "Die" in this case means that we should generate a new excip with opposite sign. This part is different from FCIQMC.
3. **Annihilation.** This part is same as in FCIQMC. Annihilation means that all pairs of walkers with opposite sign are removed, if they belong to the same determinant.

³Here one should be careful with terminology because it can differ from one article to another. Imaginary particles for stochastic approach to solve the Schrödinger equation were first introduced by Andersen article [1] and were called "psips". In this thesis we follow the Thom and Spencer in [29] and name operators as "excitors" and particles as "excips". In some articles both particles and operators are addressed to as "excitors".

6.3.4 Normalization

Our description uses intermediate normalization as an assumption and it requires a normalization constant to be introduced.

$$|\Psi_{CC}\rangle = N_0 e^{\hat{T}} |D_0\rangle. \quad (6.26)$$

here N_0 is normalization constant.

Such normalization allows to produce fractional populations on excitors, needed for convergence. Here we say population on reference, implying that N_0 are excips as well. However this is not the actual case, because they are not "real" excips of the cluster, just some parameter used for normalization.

6.3.5 Sampling probabilities

We have to sample a wave function. In this case a wave function is approximated by the exponential expansion in terms of cluster operator. We will conduct sampling by selecting a particular cluster from the expansion. Here we need to introduce some definitions.

Definition 13. Cluster size.

Size of a cluster is defined by number of excips that are used to collapse the wave function. Reference determinant is considered as a cluster of size 0.

The probability to choose a cluster of certain size is given by

$$p_{\text{size}} = \frac{1}{2^{s+1}}. \quad (6.27)$$

here s is size of a cluster. Each cluster has it's amplitude A .

After we have selected a cluster of size s , we need another probability - a probability to select a combination of excitors e .

$$p_{\text{clust}}(e|s) = s! \prod_{i=1}^s \frac{|N_i|}{N_{ex}}. \quad (6.28)$$

here $|N_i|$ is a population on excitor i , module here is needed because population of each excitor is stored with sing and can in some cases be negative.

Generation probabilities

We need to compute and asses the generation probability for any excited determinant in order to compute the spawning probability. Here we present the generation probability to obtain a double excitation for HEG, meaning we have two excited

determinants $|\Phi_m\rangle$ and $|\Phi_n\rangle$ connected through a double excitation such as $|\Phi_m\rangle = a_r^\dagger a_s^\dagger a_q a_p |\Phi_n\rangle$. the value we need to compute is the following:

$$p_{excit}(n, m) = p_{excit}(r, s|p, q)p_{excit}(p, q), \quad (6.29)$$

here $p_{excit}(p, q)$ is a probability to pick two occupied orbitals and $p_{excit}(r, s|p, q)$ is a probability to pick up two excited orbitals, after we have already chosen the occupied ones.

For the system with number of electrons equal N_e the probability to pick up two of them is given by:

$$p_{excit}(p, q) = \binom{N}{2}^{-1} = \frac{2}{N_e(N_e - 1)}. \quad (6.30)$$

The conditional probability need a bit more care:

$$p_{excit}(r, s|p, q) = p_{excit}(r|s, p, q)p_{excit}(s|p, q) + p_{excit}(s|r, p, q)p_{excit}(r|p, q). \quad (6.31)$$

For HEG the values $p_{excit}(r|s, p, q)$ and $p_{excit}(s|r, p, q)$ are equal to one, because there is only one possibility to pick up the second excited orbital if all the other orbitals are already known. The formula is then simplifies to:

$$p_{excit}(r, s|p, q) = p_{excit}(s|p, q) + p_{excit}(r|p, q). \quad (6.32)$$

This values only depend on the number of electrons and the size of the basis set, that is why we are able to compute them before running the algorithm:

$$p_{excit}(s|p, q) = \frac{1}{N_{virt}}, \quad (6.33)$$

here N_{virt} is a number of all possible combinations of virtual states that can be chosen for a known pair of occupied states. In case of excitation from reference this value is always the same for a given pair (p,q). If we spawn from already excited determinants (not from reference) we should check if the sates already existing in the parent determinant are not counted in the N_{virt} . If it is the case we should subtract them to avoid double counting. This should be done on the fly, as soon as it is too costly to recalculate such value for all possible excited determinants.

For the single excitation $|\Phi_m\rangle = a_r^\dagger a_p |\Phi_n\rangle$. the probability is given by:

$$p_{excit}(r, p) = p_{excit}(r|p)p_{excit}(p), \quad (6.34)$$

here $p_{excit}(p)$ is a probability to select one occupied orbital and is given by:

$$p_{excit}(p) = \frac{1}{N_e}. \quad (6.35)$$

The $p_{excit}(r|p)$ is a probability to select the virtual one after this. However for the HEG we do not need the single excitations, so we haven't implement this.

Spawn and death probabilities

Combining these probabilities together we get:

$$p_{select}(e) = \frac{s! \prod_{i=1}^s |N_i|}{2^{s+1} N_{ex}}. \quad (6.36)$$

It is also important that we use same number of excips at each time step:

$$p_{sel} = N_0 + N_{ex}. \quad (6.37)$$

Spawn probability is:

$$p_{total} = \frac{|\delta\tau H_{nm} A|}{p_{select} p_{sel} p_{clust} p_{excit}}. \quad (6.38)$$

Death probability is:

$$p_{death} = \frac{|\delta\tau (H_{mm} - S) A|}{p_{select} p_{sel} p_{clust}}. \quad (6.39)$$

6.3.6 Population Control

Population control is needed to sustain a stable simulation. If not done properly we would just have excips dying or rapidly growing. There are many ways the population control can be done. For example we can update the "shift" E in the following way:

$$E = \langle E_0 \rangle + \alpha \frac{N_w^{(0)}}{N_w}, \quad (6.40)$$

here $\langle E_0 \rangle$ lowest eigenvalue at time τ , α is small damping parameter, $N_w^{(0)}$ the population one want to reach and N_w total number of random walker at time τ .

However this method introduce bias into the simulation. Another method to carry out population control is to update the shift as:

$$E(\tau) = E(\tau + A\delta\tau) - \frac{\gamma}{A\delta\tau} \ln \left(\frac{N_{ex}(\tau)}{N_{ex}(\tau + A\delta\tau)} \right), \quad (6.41)$$

here γ is a small damping parameter, A is number of steps when we update the E .

6.3.7 Energy Estimation for CCQMC

For this method we have two different and independent options to estimate energy. First is to use the shift updated for under the population control as energy estimator. Second option is to sample a projection energy as:

$$E_{proj} = \frac{\langle D_0 | \hat{H} | \Psi_{CC} \rangle}{\langle D_0 | \Psi_{CC} \rangle} \quad (6.42)$$

Part II

Implementation and Results