ABSTRACT

GUO, SHI. Quantum Monte Carlo Studies of Ultra-cold Molecules and Rashba Interactions. (Under the direction of Lubos Mitas.)

Quantum Monte Carlo method(QMC) has proved to be powerful in predicting material properties. We have applied QMC method to solve different electronic structure problems including ultra-cold molecules, molecules with heavy elements and electron gas with presence of spin-orbit interaction.

The ultracold molecules are of great interest to experiments. Two kinds of ultracold molecules: LiSr and KRb, are studied in this thesis. Both binding energy curve and dipole moment curve are calculated as a function of bond length. We use Hartree Fock(HF), Configuration Interaction(CI) and the Quantum Monte Carlo methods and compare all the results. We assess the many-body effects on the calculated quantities and the calculation shows the dipole moment is quite sensitive to the treatment of correlation.

The dissociation energy of Thorium halide molecules($ThCl_n$ and $ThBr_n$) is studied using QMC and the behaviour of the molecules with different number of halides is analyzed. We also compare our results with density functional theory calculations and experimental data. The comparison shows a reasonable agreement for $ThCl_n$ case, however, $ThBr_n$ case remains problematic.

Furthermore, we carry out a development of a new quantum Monte Carlo method for calculations with Hamiltonians that contain spin-dependent operators. This is crucial in particular for treatment of spin-orbit effects in systems with heavy atoms or effective spin-orbit couplings such as the Rashba interaction. We introduce a continuous spin representation and develop a new spin-sampling procedure in which the spin becomes a dynamic variable instead of being restricted to a static label. Due to the presence of the spin-orbit term in the Hamiltonian, the wavefunction becomes complex. For the purpose of the treatment of complex wavefunctions, fixed-phase approximation is applied. The energies of electron gas with Rashba interactions are calculated at different densities. The result is very promising and comparable to other independent calculations in the literature.

Quantum Monte Carlo Studies of Ultra-cold Molecules and Rashba Interactions

by
Shi Guo

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Physics

Raleigh, North Carolina

2014

APPROVED BY:

Christopher Roland	Marco Buongiorno Nardelli
Donald Brenner	Lubos Mitas Chair of Advisory Committee

DEDICATION

To my parents.

BIOGRAPHY

I was born in 1984, in the city of Changchun, Jilin province, China. I obtained my bachelor's degree in physics at Shanghai Jiaotong University, Shanghai, China. I came to the USA and began my Ph.D study in Department of Physics, North Carolina State University in 2007.

ACKNOWLEDGEMENTS

I would like to thank my advisor, Lubos Mitas for his support. I can not complete my research without his advice and help. I would like to thank Prof. Marco Buongiorno Nardelli, Prof. Christopher Roland and Donald Brenner for being my committee members and evaluating my work. I would also like to thank Department of Physics, North Carolina State University, especially to Prof. Harald Ade and Prof. Chueng Ji for being the director of graduate programs.

I also would like to thank all of my past and current group members: Jindrich Kolorenc, Michal Bajdich, Shuming Hu, Xin Li, Rene Derian and Minyi Zhu. They provided me great knowledge and advice, and offered a lot of help throughout my research. Especially I would like to appreciate Shuming Hu. I had many discussions with him during which he clarified me many confusions and addressed many issues for me.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

Introduction

Quantum mechanics have been successful in interpreting the properties of condensed matter systems at microscale. It describes the nature of fundamental particles in our daily world: atoms, molecules, electrons, etc. The Schrödinger equation, proposed in 1925, provide us an exact mathematical formulation of non-relativistic quantum physics . With the solution to the Schrödinger equation, we are able to know and predict many properties of quantum systems. However, while physicists are often proud of being able to precisely describe how a system behaves quantitatively, solving the equations underlying the theory is a great challenge, especially when the size of the system becomes large. Thanks to the tremendous development in computer technology, many calculations which used to take immense time to perform, can be quickly done on modern computers. But there still exist many unsolved problems because they suffer difficulties beyond the computational efficiency. Theoretical physicists aim to exactly solve problems relying on the equations from first principles and many people have been devoting their lives to developing various kinds of methods. Quantum Monte Carlo, presented in this thesis, is one of the methods to simulate quantum systems. Generally, the term "Monte Carlo" refers to a class of computer algorithms to compute the results based on random sampling. The word "quantum" indicates we are using the Monte Carlo technique to solve problems of quantum systems based on the knowledge of quantum mechanics. In this thesis, we will focus on electronic structure problems, a particular area of the quantum system problems.

The greatest difficulty in the electronic structure problems is the treatment of electron correlation. Many well-developed methods can not take the correlation effect into account without introducing approximations. Quantum Monte Carlo(QMC), however,

has the capability to incorporate the electron correlation directly by applying the many-body wavefunction, which will be explained in Chapter 3. This leads to highly accurate results. The milestone of the QMC calculation was the description of homogeneous electron gas in 1980[1], which provided an accurate local density approximations used in density functional theory calculations. Over the past two decades, QMC have become an excellent method to treat systems ranging from analytical models to real materials including electron gas, atoms, molecules, and solids. Before we introduce the QMC method, in this chapter, we start with some fundamental knowledge of quantum mechanics followed by some traditional well-developed methods of solving electronic structure problems.

1.1 Many-body Schrödinger Equation

We consider the non-relativistic time-independent Schrödinger equation

$$H\Phi(\mathbf{R},\mathbf{r}) = E\Phi(\mathbf{R},\mathbf{r}) \tag{1.1}$$

H is the Hamiltonian of the system, $\Phi(\mathbf{R},\mathbf{r})$ is the wavefunction with $\mathbf{R}$ being the positions of nuclei $\mathbf{r}$ being the positions of electrons and E is the energy eigenvalue. For a system which contains electrons and nuclei, we can write the non-relativistic Hamiltonian H as the following:

$$\begin{aligned} H = &-\sum_i \frac{\hbar^2}{2m_i}\bigtriangledown^2_{r_i} - \sum_I \frac{\hbar^2}{2M_I}\bigtriangledown^2_{R_I} - \sum_{i,I}\frac{1}{4\pi\epsilon_0}\frac{Z_I e^2}{|\mathbf{r}_i-\mathbf{R}_I|} + \sum_{i<j}\frac{1}{4\pi\epsilon_0}\frac{e^2}{|\mathbf{r}_i-\mathbf{r}_j|} \\ &+ \sum_{I<J}\frac{1}{4\pi\epsilon_0}\frac{Z_I Z_J e^2}{|\mathbf{R}_I-\mathbf{R}_J|} \end{aligned} \tag{1.2}$$

m_i is the mass of the ith electron, and M_I is the mass of the Ith nuclei. For convenience, we write the equation using atomic units by setting $e=1$, $4\pi\epsilon_0=1$, $m=1$, $\hbar=1$, then the Hamiltonian becomes

$$\begin{aligned} H &= -\sum_i \frac{1}{2}\bigtriangledown^2_{r_i} - \sum_I \frac{1}{2M_I}\bigtriangledown^2_{R_I} - \sum_{i,I}\frac{Z_I}{|\mathbf{r}_i-\mathbf{R}_I|} + \sum_{i<j}\frac{1}{|\mathbf{r}_i-\mathbf{r}_j|} + \sum_{I<J}\frac{Z_I Z_J}{|\mathbf{R}_I-\mathbf{R}_J|} \\ &= T_i + T_I + V_{iI} + V_{ij} + V_{IJ} \end{aligned} \tag{1.3}$$

The indices i and j refer to the electrons, and I and J refer to the nuclei. The first two terms are the kinetic energy operator of all the electrons and nuclei, respectively. The third term describes the attractive Coulomb interaction between the electrons and nuclei. The fourth term describes the repulsive interaction between the electrons, and the last term is the repulsive interaction between the nuclei. This is a quite complicated equation, some approximation has to be made before any further treatment. In fact, consider the mass of the nuclear is 1836 times larger than that of the electron, it is eligible to ignore the motion of the nuclei. Therefore, the kinetic energies of the nuclei are assumed to be zero, and the potential energy between the interactions of the nuclei can be added to the total energy as a constant. This approximation is called the *Born-Oppenheimer approximation*. Under this approximation, we can separate the whole Schrödinger equation into two parts: the electron part and the nuclear part, and then solve each of them separately.

The electron part of the Schrödinger equation is

$$(T_i + V_{ij} + V_{iI} + V_{IJ})\Phi_{e,n}(\mathbf{r}, \mathbf{R}) = E_{e,n}(\mathbf{R})\Phi_{e,n}(\mathbf{r}, \mathbf{R}) \tag{1.4}$$

where $\Phi_{e,n}(\mathbf{r}, \mathbf{R})$ is the wavefunction of electron given the positions of all the nuclei fixed. $E_{e,n}(\mathbf{R})$ is the total eigenenergy of electrons which depends on the positions of nuclei: $\mathbf{R}$. This term is also called as Born-Oppenheimer *potential energy surface*(PBS). Since V_{IJ} are constant values, we usually add it to our total energy at the end of the calculation. After we solve the electron part, the term $E_{e,n}(\mathbf{R})$ enters the equation for the nuclei as a constant potential energy. The part of the equation corresponding the nuclei with the wavefunction $\Phi_n^{'}(\mathbf{R})$ can be written as

$$(T_I + E_{e,n})\Phi_n^{'}(\mathbf{R}) = E_n\Phi_n^{'}(\mathbf{R}) \tag{1.5}$$

The focus of the thesis is to solve the electron part of the Schrödinger equation under Born-Oppenheimer approximation

$$(-\frac{1}{2}\sum_i \bigtriangledown_i^2 - \sum_{i,I}\frac{Z_I}{|\mathbf{r_i} - \mathbf{R_I}|} + \sum_{i<j}\frac{1}{|\mathbf{r_i} - \mathbf{r_j}|} + \sum_{I<J}\frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|})\Phi_{e,n}(\mathbf{r}) = E_{e,n}\Phi_{e,n}(\mathbf{r}) \tag{1.6}$$

where $\mathbf{R}_I$ are fixed as discussed before. $\bigtriangledown_i$ means taking the derivatives with respect to the coordinates of the ith electron.

1.2 The Hartree-Fock Method

In this section we introduce an approach to solving many-body Schrödinger equation: Hartree-Fock method[2, 3, 4]. We remember electrons are fermions. Therefore, the wavefunction must be antisymmetric. If we define a particle-exchange operator P_{ij}, then for a give N-particle state Ψ, we write:

$$P_{ij}\Psi(\mathbf{r}_1,\sigma_1,\mathbf{r}_2,\sigma_2,...,\mathbf{r}_i,\sigma_i,...,\mathbf{r}_j,\sigma_j...\mathbf{r}_N,\sigma_N) = \Psi(\mathbf{r}_1,\sigma_1,\mathbf{r}_2,\sigma_2,...,\mathbf{r}_j,\sigma_j...,\mathbf{r}_i,\sigma_i,...,\mathbf{r}_N,\sigma_N) \tag{1.7}$$

For fermions, we must have

$$P_{ij}\Psi = -\Psi \tag{1.8}$$

The idea of the Hartree-Fock approximation is to build a variational many-body wavefunction that satisfies antisymmetry condition using independent one-electron states. So this approximation can be viewed as independent-electron approximation. The variational wavefunction Ψ is what is called Slater determinant and it is written as follows:

$$\Psi(\mathbf{r}_1,\mathbf{r}_2,...,\mathbf{r}_N) = \frac{1}{\sqrt{N!}}\begin{vmatrix} \psi_1(\mathbf{r}_1,\sigma_1) & \psi_1(\mathbf{r}_2,\sigma_2) & ... & \psi_1(\mathbf{r}_N,\sigma_N) \\ \psi_2(\mathbf{r}_1,\sigma_1) & \psi_2(\mathbf{r}_2,\sigma_2) & ... & \psi_2(\mathbf{r}_N,\sigma_N) \\ . & . & & . \\ . & . & & . \\ . & . & & . \\ \psi_N(\mathbf{r}_1,\sigma_1) & \psi_N(\mathbf{r}_2,\sigma_2) & & \psi_N(\mathbf{r}_N,\sigma_N) \end{vmatrix} \tag{1.9}$$

Where $\psi_i(\mathbf{r}_j)$ are the ith one-electron orbitals for the jth electron. All the one-electron orbitals ψ_i need to be optimized to minimize the total energy and we will discuss how to do it in the next step. Note that the Slater determinant is antisymmetric, which means, when we exchange the ith and jth particle, we are exchanging the ith column and the jth colomn of the Slater determinant, so the result is we get a minus sign as shown in Eq. 1.8. Therefore the Pauli exclusion principle is well satisfied here. If the ith particle and the jth particle happen to be in the same state, we have exactly the same two rows in the determinant, so the wavefunction vanishes.

1.2.1 The Hatree-Fock Equation

We now derive the Hartree-Fock equation here. For convenience, we write our Hamiltonian for N electrons in the following way:

$$H = \sum_{i}^{N} h(i) + \sum_{i,j;i<j} g(i,j) \tag{1.10}$$

where $h(i)$ is a one-electron operator, $g(i,j)$ is a two-electron operator. It is easy to see that from Eq. 1.6

$$h(i) = -\frac{1}{2} \bigtriangledown_i^2 - \sum_I \frac{Z_I}{|\mathbf{r}_i - \mathbf{R}_I|} \tag{1.11}$$

$$g(i,j) = \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} \tag{1.12}$$

Give the Slater determinant above as our wavefunction, we can calculate the expectation value of $h(i)$ and $g(i,j)$

$$< \Psi | \sum_i h(i) | \Psi >= \frac{N!}{N!} \sum_m < \psi_m | h | \psi_m >= \sum_m \int d\mathbf{r} \psi_m^* h(\mathbf{r}) \psi_m \tag{1.13}$$

$$< \Psi | \sum_{i,j} g(i,j) | \Psi >= \sum_{m,n} < \psi_m \psi_n | g | \psi_m \psi_n > - \sum_{m,n} < \psi_m \psi_n | g | \psi_n \psi_m > \tag{1.14}$$

where

$$< \psi_m \psi_n | g | \psi_m \psi_n >= \int d\mathbf{r}_1 d\mathbf{r}_2 \psi_m^*(\mathbf{r}_1) \psi_n^*(\mathbf{r}_2) \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} \psi_m(\mathbf{r}_1) \psi_n(\mathbf{r}_2) \tag{1.15}$$

So the expectation value of the total energy is

$$E = \sum_m < \psi_m | h | \psi_m > + \frac{1}{2} \sum_{mn} [< \psi_m \psi_n | g | \psi_m \psi_n > - < \psi_m \psi_n | g | \psi_m \psi_n >] \tag{1.16}$$

The first term is total kinetic energy. We define operators J and K

$$J_m(\mathbf{r}) \psi(\mathbf{r}) = \int \psi_m^*(\mathbf{r}') \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} \psi_m(\mathbf{r}') \psi(\mathbf{r}) d\mathbf{r} \tag{1.17}$$

$$K_m(\mathbf{r}) \psi(\mathbf{r}) = \int \psi_m^*(\mathbf{r}') \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} \psi(\mathbf{r}') \psi_m(\mathbf{r}) d\mathbf{r} \tag{1.18}$$

$$F = \sum_m h_m + \sum_m J_m - \sum_m K_m = h + J - K \tag{1.19}$$

with $h = \sum h_m$, $J = \sum J_m$,$K = \sum K_m$. And

$$E = \sum_m < \psi_m | h + \frac{1}{2}(J - K) | \psi_m > \tag{1.20}$$

We call the operator J the Coulomb operator and K the exchange operator. And F is the Fock operator. The corresponding expectation values for J and K are called the Hartree energy and the exchange energy. The Hartree energy is the classical Coulomb energy while the exchange is purely quantum-mechanical with no classical counterpart. The variational total energy E is a function of Ψ, with constraint that

$$< \psi_m | \psi_n >= \delta_{mn} \tag{1.21}$$

Based on Lagrange multiplier theorem, this problem can be solved by self-consistency procedure(SCF). When the total energy reaches the minimum, the one-electron state satisfies the Fock equation.

$$F\psi_k = \epsilon_k \psi_k \tag{1.22}$$

ϵ_k is the eigenvalue of the Fock operator. For most of the problems in quantum chemistry, we express our one-electron state as linear combination of basis sets, which we will explain in the next section. The one-electron state can be written by

$$\psi_k(r) = \sum_s C_{sk} \phi_s(r) \tag{1.23}$$

The matrix form of the Fock equation(known as *Roothaan equation*) is

$$\mathbf{FC_k} = \epsilon_k \mathbf{SC_k} \tag{1.24}$$

$S_{ij} =< \phi_i | \phi_j >$ is the overlap matrix. Finding the correct ψ_k becomes finding the eigenvalue ϵ_k and eigenvector C_k of the Fock matrix: $F_{ij} =< \phi_i | F | \phi_j >=< \phi_i | h | \phi_j > + \frac{1}{2} < \phi_i | J - K | \phi_j >$. The second term depends on the density matrix $P_{ij} = 2 \sum_k C_{ik} C_{jk}^*$. So to carry out the SCF calculation, we first need to guess the density matrix, then evaluate Fock matrix, diagonalize and find the eigenvalue(ϵ) and eigenvector($\mathbf{C_k}$). Then the new density matrix is constructed, we diagonalize the Fock matrix again. We repeat the

procedure until we reach the convergence which means the result difference between the two consecutive steps is smaller than threshold value.

1.2.2 Gaussian Basis Sets

In the calculation presented in this thesis, two forms of one-electron orbitals are used: planewave for periodic systems and linear combinations of Gaussian basis sets for atomic and molecular systems, where each molecular orbital is a linear combination of atomic orbitals(LCAO) ϕ_n

$$\psi(\mathbf{r}) = \sum_n C_n \phi_n(\mathbf{r}) \tag{1.25}$$

Each atomic orbital is represented by a linear combination of basis sets:

$$\phi_n(\mathbf{r}) = \sum_{lm} N_{lm} (\sum_g R_{lmg}(r)) r^l Y_{lm}(\mathbf{r}) \tag{1.26}$$

where N_{lm} is the normalization factor, Y_{lm} is spherical harmonic functions. So the whole basis set consists of both radial part and angular part, and is a sum over all different angular momentum channels. In each channel, the radial part of the basis is a linear combination of several Gaussian functions: $R_{lmg}(\alpha, r)$

$$R_{lmg}(r) = C_{lmg} exp(-\alpha_{lmg}(r - r_0)^2) \tag{1.27}$$

where r_0 is the center of the nuclear, α_{lmg} are constant numbers, which can be obtained by optimization. In our calculations, we took the atomic basis from basis set exchange website. Theoretically, both Slater-type functions and Gaussian-type function can be used for radial part. The advantage of using Gaussian functions as our basis sets is when we need to calculate the expectation value of any operator the integrals involved are quite easy to evaluate. But the Gaussian functions do not have the correct asymptotic behavior called cusp condition when $r \to 0$. The cusp condition will be discussed elaborately in the following Section 3.1.3. Usually, one may need many Gaussian functions in order to correctly represent the asymptotic behavior of the radial part of the orbitals. Therefore, there are many parameters and the calculation is quite time-consuming. To reduce the computational cost, we use contracted basis set by taking the linear combination of

several functions, for each angular momentum channel

$$R_{lmg}(r) = C_{lmg}(\sum_{h} D_{lmgh} exp(-\alpha_{lmgh} r^2)) \tag{1.28}$$

where the coefficients D_{lmgh} are frozen in the calculation, only C_{lmg} is optimized in the calculation. Making the coefficients frozen reduces the variational freedom, but for most cases, it is enough to predict the chemical properties.

Up to here, we see Hartree Fock approximation nicely explains the Coulomb interaction and the exchange interaction. However, Hartree Fock fails to describe the correlation between the electrons with different spins. This will be addressed in the next few sections.

1.3 Post-Hartree-Fock methods

In many cases, calculations using Hartree-Fock approximation lead to unsatisfactory results. This is because of the missing correlations in the Hartree Fock approximation. For a given system, the difference between the exact total energy and the Hartree-Fock total energy is called *correlation energy*. So the main goal of the post Hatree Fock methods is to capture the correlation energy beyond the Hartree-Fock level. A basic fact is for a given basis, all the orbitals in the Hartree-Fock solution form a complete orthogonal set. We can write our variational wavefunction as a linear combination of many configuration state functions(CSF) which satisfy the requirement of symmetry. And each CSF could be one or several determinants built up by the Hartree-Fock orbitals. The form of the wavefunction is

$$\Psi = \Psi_0 + \sum_{i} \sum_{p} c_i^p \Psi_i^p + \sum_{ij} \sum_{pq} c_{ij}^{pq} \Psi_{ij}^{pq} + \sum_{ijk} \sum_{pqr} c_{ijk}^{pqr} \Psi_{ijk}^{pqr} + ... \tag{1.29}$$

The first term Ψ_0 is the Slater determinant in Hartree-Fock solution. Ψ_i^p means a new CSF by replacing the ith orbital in Ψ_0 with the pth orbitals, which is unoccupied. This is single excitation. Ψ_{ij}^{pq} is replacing both ith and jth obtials with pth and qth, which is double excitation. We can build the expansion as large as possible to obtain the correlation effect. A self-consistent-field procedure is carried out similar to what is described in Hartree-Fock to minimize the total energy and determine all the weight(c_i^p,c_{ij}^{pq},c_{ijk}^{pqr}) of CSFs. The same minimization technique used in Hartree-Fock can be directly applied

here. Using Lagrange multiplier method, minimizing the energy is equal to solving the following equation:

$$HC = SCE \tag{1.30}$$

H is the CI matrix, of which the elements are the Hamiltonian evaluated between the two configuration state function. $H_{ij} =< \Psi_i|H|\Psi_j >$, where Ψ_i andΨ_j are CSFs. C is the vector containing all the coefficients in Eq. 1.29. $S =< \Psi_i|\Psi_j >$ is the overlap matrix. This method is called the *configuration interaction*(CI). We can go one step forward by varying both coefficients of the CSFs and coefficients of each orbital to minimize the total energy. This is the *Multi-Configuration Self-Consistent-Field calculation*. Theoretically, the CI method can be exact, but in practice, the result of CI can be sensitive to several factors. Here we want to highlight a few relevant ones. First, one needs to be careful to choose excitation levels. Based on Brillouin's theorem, we know

$$< \Psi_i^p|H|\Psi >= 0 \tag{1.31}$$

This means single excitations do not contribute anything, so double excitation must be added for the purpose of gaining the correlation energy. For single and double CI, the computational cost scales as N^6. It has been shown including triple and quadruple can have a nontrivial effect on the total energy. For calculating those other quantities other than energy, choice of excitation levels are crucial to obtain a reasonable result. We will see this in Chapter 4 when we want to calculate the dipole moment of molecules. In some cases, it is possible to include all the possible excitation states and this is called full CI calculation. For large systems, full CI is almost formidable. We must truncate some of the CSFs and keep only the most important ones. Unfortunately, such a truncated CI is not size-consistent. The total energy does not scale linearly with the size of the system. For example, consider the binding energy of a molecule(AB) made up of two small pieces: A and B. Using doubly CI, one can take the difference between the total energy of AB(E_{AB}) and sum of total energy E_A and E_B to get the binding energy. Another way is take the difference between E_{AB} and total energy of molecule AB when the distance between A and B are extremely large(We call it $A....B$) therefore the binding effect can be neglected. These two approaches provide different results. This is because using CI double excitation to treat two separate atoms A and B is not equivalent to $A....B$. In fact, treating A and B using double excitation is equivalent of treating $A....B$ by quadruple excitation.

Second, the result is highly dependent on basis sets. Especially for the full CI calculation, the result is solely determined by the basis. Big basis sets are necessary to make the calculation converge for a complex system. Up to know, highly accurate CI calculations are only available in small systems.

Some limitations in CI methods can be overcome by using *Coupled-Cluster* method. This methods is size-consistent, and is almost one of the most exact method for molecular systems.

1.4 Density Functional Theory

The methods introduced in the previous section all focus on how to build the wavefunction. The limitations of these methods make them difficult to apply in extended systems such as solids. Now we introduce another method: Density Functional Theory(DFT)[5, 6, 7]. The fundamental idea of DFT is any property of a physical system can be considered as some functional of the electron density. And this is stated in the Hohenberg-Kohn theorem as follows:

Theorem 1 The energy of a system are uniquely determined by the electron density $n_0(\mathbf{r})$.

Theorem 2 The energy functional $E[n_0(\mathbf{r})]$ is defined and the electron density of the ground state minimizes the energy functional.

Kohn and Sham developed the framework called Kohn-Sham approach, which formulates the theory. This approach assumes that the interacting system can be replaced by a non-interacting system by constructing an effective Hamiltonian. Therefore, solving a many-body problem becomes solving a one-body problem with an effective external potential. The real Hamiltonian is replaced by an effective Hamiltonian H_{eff} containing the external potential V_{eff} satisfying

$$H_{eff}\psi_k(\mathbf{r}) = [-\frac{1}{2}\bigtriangledown^2 + V_{ext}]\psi_k(\mathbf{r}) = E\psi_k(\mathbf{r}) \tag{1.32}$$

Suppose $\psi_k(\mathbf{r})$ is the solution to this effective Hamiltonian, the one-electron orbital. And the electron density is

$$n(\mathbf{r}) = \sum_{k=1}^{N} f_k|\psi_k(\mathbf{r})|^2 \tag{1.33}$$

the sum here is over all the occupied orbitals. Since each orbital can hold up to two electrons due to spin-degeneracy, f_k is the occupation number, which can be 1 or 2.

Given the non-interacting system described by Eq. 1.32, the total kinetic energy is

$$T_s[n] = \sum_{i=1}^{N} < \psi_k(\mathbf{r})| - \frac{1}{2} \bigtriangledown_i^2 \, |\psi_k(\mathbf{r}) > \tag{1.34}$$

The classical Coulomb energy(Hartree energy) is

$$H_e = \frac{1}{2} \int \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}' \tag{1.35}$$

The electron-nuclei interaction is

$$V_{eI} = \sum_I \int \frac{Z_I n(\mathbf{r})}{|\mathbf{r} - \mathbf{R}_I|} d\mathbf{r} \tag{1.36}$$

We define $T[n]$ as the true kinetic energy of the system, $V_{ee}[n]$ as the true electron-electron interaction The true total energy of interacting system is

$$E = T[n] + V_{ee}[n] + V_{eI} \tag{1.37}$$

If we calculate the kinetic energy as a non-interacting system and Coulomb energy as a Hartree energy, then the total energy becomes

$$E_s = T_s[n] + H_e[n] + V_{eI} \tag{1.38}$$

The energy component we miss in E_s is $E - E_s$, which we call the exchange-correlation energy E_{xc}.

$$E_{xc} = T[n] - T_s[n] + V_{ee}[n] - H_e[n] \tag{1.39}$$

So

$$E = \sum_{i=1}^{N} \int \psi_i^*(\mathbf{r})[-\frac{1}{2} \bigtriangledown_i^2] \psi_i(\mathbf{r}) d\mathbf{r} + \frac{1}{2} \int \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}' + \sum_I \int \frac{Z_I n(\mathbf{r})}{|\mathbf{r} - R_I|} d\mathbf{r} + E_{xc}[n(\mathbf{r})] \tag{1.40}$$

Using Lagrange multiplier, variation of E subject to $< \psi_k | \psi_l >= \delta_{kl}$ yields

$$[-\frac{1}{2} \bigtriangledown^2 - \sum_I \frac{Z_I}{|\mathbf{r} - \mathbf{R}_I|} + \int d^3\mathbf{r}' n(\mathbf{r}') \frac{1}{|\mathbf{r} - \mathbf{r}'|} + V_{xc}[n(\mathbf{r})](\mathbf{r})]\psi_k(\mathbf{r}) = \epsilon_k \psi_k(\mathbf{r}) \quad (1.41)$$

Comparing Eq. 1.40 to Eq. 1.16, we have the same energy component as what we have in Hartree-Fock, including kinetic, electron-nuclei interaction energy and Hartree energy. The fourth term is new, and this describes the exchange and correlation energy together. This is the essence of density function theory. For a given electron, the interaction effect from other electrons is treated by an averaged external potential as a functional of the density. Like the solutions to HF and CI, the DFT solution is also self-consistent. We first choose the density $n(\mathbf{r})$ by some initial guess, construct the exchange-correlation potential using the density, solve Eq. 1.41 and calculate the density from the solution. Then we construct the exchange-correlation potential again and solve the Schrödinger equation iteratively until the change in the density in two consecutive loops is smaller than the chosen threshold.

1.4.1 Local Density Approximation

The exact exchange-correlation energy is never known. After the DFT method came out, various forms of the exchange-correlation functionals have been proposed and developed. In local density approximation(LDA), the exchange-correlation energy is an integral of the exchange-correlation density of homogeneous electron gas.

$$E_{xc} = \int d^3\mathbf{r} \epsilon_{xc}[n(\mathbf{r})] n(\mathbf{r}) \quad (1.42)$$

$\epsilon_{xc}[n]$ is the exchange-correlation energy per particle of an homogeneous electron gas, thus it contains both exchange part(ϵ_x) and correlation part (ϵ_c). The exchange part can be derived analytically using homogeneous electron gas model, yielding the following form:

$$\epsilon_x[n(r)] = Const. * n^{\frac{1}{3}}(r) \quad (1.43)$$

The total energy of homogeneous electron gas is calculated with great accuracy by quantum Monte Carlo methods, which we will discuss in the following chapters.

The correlation functional is developed by fitting to the data of quantum Monte Carlo

results[8, 1]. Perdew and Wang proposed an expression[9]

$$\epsilon_c^{LDA}(r_s,\xi) = \epsilon_c(r_s,0) + a_c(r_s)\frac{f(\xi)}{f''(\xi)}(1-\xi^4) + [\epsilon(r_s,1) - \epsilon(r_s,0)]f(\xi)\xi^4 \tag{1.44}$$

where

$$r_s = [\frac{3}{4\pi}(n(r)^{-1})]^{\frac{1}{3}} \tag{1.45}$$

$$f(\xi) = [(1+\xi)^{4/3} + (1-\xi)^{4/3} - 2]/(2^{4/3} - 2) \tag{1.46}$$

$\epsilon_c(r_s,0)$,$\epsilon_c(r_s,1)$,$a_c(r_s)$ are determined by the empirical formula. ξ is spin-polarization.

$$\begin{aligned} & G(r_s, A, \alpha_1, \beta_1, \beta_2, \beta_3, \beta_4) \\ = \quad & -2A(1+\alpha_1 r_s) ln[1 + \frac{1}{2A(\beta_1 r_s^{1/2} + \beta_2 r_s + \beta_3 r_s^{3/2} + \beta_4 r_s^2)}] \end{aligned} \tag{1.47}$$

Since many systems are spin-polarized, the spin-up and the spin-down electrons are treated separately. Then, the exchange-correlation functional is a functional of spin-up electron density and spin-down electron density. The method is *local spin density approximation*(LSDA).

$$E_{xc}[n^{up}(r), n^{down}(\mathbf{r})] = \int d^3\mathbf{r} n(\mathbf{r}) \epsilon_{xc}(n^{up}(\mathbf{r}), n^{down}(\mathbf{r})) \tag{1.48}$$

n^{up} and n^{down} denote the spin-up and spin-down electron density.

1.4.2 Functionals Beyond LDA

LDA(LSDA) is very accurate for homogeneous systems, but does not work well for inhomogeneous systems. In order to improve the performance of the functionals, people introduced other type of functionals not only depending on the electron density, but also the gradient of the density.

$$E_{xc}^{GGA}[n^{up}, n^{down}] = \int d^3\mathbf{r} n(\mathbf{r}) \epsilon_{xc}(n^{up}, n^{down}, |\bigtriangledown n^{up}|, |\bigtriangledown n^{down}|) \tag{1.49}$$

The functional is called Generalized-Gradient Approximations(GGA). The most widely used forms of the functionals is PBE[10].

Another improvement is called hybrid functional. Comparing DFT to Hartree Fock,

the exchange part in Hartree Fock is exact, but this effect is treated approximately in LDA and GGA. The hybrid functional can take advantage of the exact exchange in Hartree Fock and mix the HF exchange with exchange-correlation functional. The most popular hybrid functional is B3LYP.

$$E_{xc}^{B3LYP} = E_{xc}^{LDA} + a_0(E_x^{HF} - E_x^{LDA}) + a_x(E_x^{GGA} - E_x^{LDA}) + a_c(E_c^{GGA} - E_c^{LDA}) \quad (1.50)$$

where $a_0 = 0.2$,$a_x = 0.72$, and $a_c = 0.81$. E_x^{GGA} is Becke 88[11], E_c^{GGA} is correlation functional of LEE, YANG and PARR [12]. E_C^{LDA} is the VMN functional. Those parameters are fitted to the experimental data. B3LYP functional is more often used for chemical systems, such as atoms and molecule. For solid, B3LYP performs worse than GGA.

DFT is popular due to its low computational cost to offer a reasonable accuracy. We now discuss some drawbacks of DFT. The hartree energy artificially add a self-interaction part, this term is canceled by the exchange part in the Hartree-Fock, but it is not the case in DFT. Some methods such as self-interaction correction and (LDA+U) are developed, but the problem has not been completely settled. The results of DFT calculations are highly-dependent on the choice of functionals. Different functionals lead to different results. Unlike Hartree-Fock and CI, which guarantee the energy is an upper bound of the true energy, the DFT result can be either above or lower than the exact value. Improvement can only be possible when there is experimental data available. Therefore DFT results are hard to use for prediction.

Chapter 2

Quantum Monte Carlo Method

The Quantum Monte Carlo(QMC) method discussed here differs from all the methods mentioned in the previous chapter dramatically. This method is able to treat many-body effect directly by applying many-body wavefunctions, instead of considering the electron-electron interaction in a "mean field", which is an effective interaction created by other electrons formulated by an effective Hamiltonian. The QMC method does not introduce too many approximations in the Hamiltonian but keeps its original form as much as possible. Since the many-body wavefunction used in QMC can explicitly describe the electron correlation and the Hamiltonian is the real Hamiltonian with only some reasonable approximations, QMC can make a big improvement having a much better accuracy and acceptable computational cost. Like other methods, QMC suffers some fundamental difficulties as well, especially for large and complex systems. Here we first review the classical Monte Carlo method, and go through the necessary technique of mathematics and statistics, then elaborate the major methods of QMC.

2.1 Classical Monte Carlo Method

Monte Carlo algorithm is a class of computer algorithm which computes the result by repeated random sampling. One application of Monte Carlo method is solving integrals. Consider a simple integral, which integrates a function $f(x)$ from a to b

$$I = \int_a^b f(x)dx \tag{2.1}$$

There are many ways to solve this based on standard numerical methods. For example, we may use rectangular rule, trapezoid rule, Simpson rule, etc. All of these methods have something in common. We first choose equally spaced grid on x axis. Suppose the number of grids is N, each grid has length $\frac{b-a}{N}$. Then evaluate the integrand at each x_i. The integral is a sum

$$I = \frac{b-a}{N} \sum_{i=1}^{N} w_i f(x_i) \tag{2.2}$$

where $f(x_i)$ is the value of integrand at x_i, w_i is some weight which differs by each method. These methods attempt to use a fixed number of points x_i to approximate all the possible values x_i over the range $[a, b]$. This approximation generates an error term which depends on the size of the grid $h = \frac{b-a}{N}$. The error term is usually proportional to h^k and N^{-k}. k is an positive integer associated to each method. It may vary depending on the method chosen.

Monte Carlo methods do this job in an alternative way. We don't choose all the fixed x_is, instead, randomly generate N random variables that are uniformly distributed on $[a, b]$. If N is sufficiently large, the integral is given by Eq. 2.2 with all w_i are equal to 1.

$$I = \frac{b-a}{N} \sum_{i=1}^{N} f(x_i) \tag{2.3}$$

This can be proved by the Law of Large numbers. That is, the average of all the $f(x_i)$ converges to the expectation of $f(x_i)$ almost surely for sufficiently large N.

$$\lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} f(x_i) = E(f(x)) = \int_a^b f(x) \frac{1}{b-a} dx = \frac{1}{b-a} I \tag{2.4}$$

Then this equation can be rearranged to Eq. 2.3

The error term of Monte Carlo integration can be easily calculated by Central Limit Theorem.The variance is proportional to $1/N$, so the error term is $N^{-1/2}$. Comparing this to the error in the standard numerical ways, which is N^{-k}, we see the numerical ways would be more favorable since k is always greater than 1, making N^{-k} smaller $N^{-1/2}$.

The power of Monte Carlo integration manifests when one wants to evaluate an integral in higher dimensions. For a multidimensional integral, it is easy to prove that the standard numerical integration gives the error term $N^{-k/d}$. But for Monte Carlo integra-

tion, the central limit theorem is valid in the same way regardless of the dimension, so the error term remains the same as long as we keep the same number of samples. When $d > 2k$, the error term in Monte Carlo integration becomes smaller, which makes Monte Carlo integration more efficient in higher dimensions.

Most of the time we need to calculate the expectation of function composed of a random variable. For example, we want to evaluate

$$I = \int f(x)p(x)dx \tag{2.5}$$

This is the expectation of $f(x)$ given the probability density function of x is $p(x)$. Monte Carlo integration is able to handle this. Here we don't generate a uniformly distributed random variables but rather a set of random variables x_i distributed according to $p(x)$.The Monte Carlo integral becomes

$$I = \frac{1}{N}\sum_{i=1}^{N} f(x_i) \tag{2.6}$$

The sample variance of $f(x_i)$ is a well-defined quantity

$$\sigma_f^2 = \frac{1}{N-1}\sum_{n=1}^{N}[f(x_i) - \frac{1}{N}\sum_{j=1}^{N} f(x_j)]^2 \tag{2.7}$$

The variance of the integral I when N is large is given by central limit theorem, which is $\frac{\sigma_f^2}{N}$, therefore the standard error is $\pm\sigma_f/\sqrt{N}$.

2.2 Metropolis Algorithm

The Monte Carlo integration needs to generate random numbers of which the probability density function is $p(x)$. This is easy to do by an inverse transform if we already had uniformly distributed random variables in hand generated by any random number generator. Suppose random numbers u_i is uniformly distributed, then $x = F^{-1}(u)$ is a random variable with cumulative density function $F(x)$. But if F is a complicated multivariate function, it is impossible to follow this approach because finding the inverse function may not be trivial. The Metropolis algorithm[13] comes here as a very useful tool. The purpose of the Metropolis algorithm is to generate a state of random sample point asymptotically according to a given distribution $P(x)$ using Markov Chain.

The algorithm works in 3 three steps:
(1)Generate random sample points(walkers) $X_n = \{X_i : 1 \leq i \leq n\}$.
(2)For each sample point, make a trial move according to any convenient choice of distribution function $T(\Delta x = X_i^{'} - X_i)$. So T here can be considered as a transition probability function from X_i to $X_i^{'}$, and we also write it as $T(X_i \to X_i^{'})$. Usually we choose T as a uniform distribution function or normal distribution function.
(3)For each move, calculate the acceptance probability.

$$A(X_i \to X_i^{'}) = Min(1, \frac{T(X_i^{'} \to X_i)P(X_i{'})}{T(X_i \to X_i^{'})P(X_i))}) \tag{2.8}$$

$P(x)$ is the desired distribution function. If the acceptance probability is 1, then we accept the move. If not, we compare the ratio term $\frac{T(X_i^{'} \to X_i)P(X_i{'})}{T(X_i \to X_i^{'})P(X_i))}$ with another uniformly distributed random variable from 0 to 1. If the ratio term is bigger, then we accept the move otherwise reject it.

The path that each sample point takes is a random walk. It is easy to see this random walk generated by Metropolis algorithm is a Markov Chain. In order to make the Metropolis algorithm converge to the desired distribution, we have to prove the Markov Chain is stationary. A nice property of Metropolis algorithm which can guarantee the stationary Markov Chain is the detailed balanced condition. It means when the equilibrium has reached, there would be no net flow of walkers. Suppose the density of the walkers at position X and $X^{'}$ are $P(X)$ and $P(X^{'})$. The detailed balanced condition states the probability of a walker going from X to $X^{'}$ must be equal to the probability of the reverse way. Mathematically, we can write

$$P(X)T(X \to X^{'})A(X \to X^{'}) = P(X^{'})T(X^{'} \to X)A(X^{'} \to X) \tag{2.9}$$

Rearranging the equation gives

$$\frac{P(X)}{P(X')} = \frac{A(X^{'} \to X)T(X^{'} \to X)}{A(X \to X')T(X \to X')} \tag{2.10}$$

The detailed balanced condition is not necessary for the distribution to be stationary, but one can rigorously show it is a sufficient condition. We now prove Eq. 2.10 is true

from the definition of the acceptance probability in Eq. 2.8. In Eq. 2.8,if

$$1 > \frac{T(X^{'} \to X)P(X')}{T(X \to X')P(X)} \tag{2.11}$$

then

$$\frac{A(X^{'} \to X)T(X^{'} \to X)}{A(X \to X')T(X \to X')} = \frac{T(X^{'} \to X)}{\frac{T(X' \to X)P(X')}{T(X \to X')P(X)}T(X \to X')} = \frac{P(X)}{P(X')} \tag{2.12}$$

otherwise

$$1 < \frac{T(X^{'} \to X)P(X')}{T(X \to X')P(X)} \tag{2.13}$$

then

$$\frac{A(X^{'} \to X)T(X^{'} \to X)}{A(X \to X')T(X \to X')} = \frac{\frac{T(X \to X')P(X)}{T(X' \to X)P(X')}T(X^{'} \to X)}{T(X \to X')} = \frac{P(X)}{P(X')} \tag{2.14}$$

So the Eq. 2.10 holds for both cases. Therefore Metropolis algorithm satisfies the detailed balance condition, which implies a stationary Markov chain.

2.3 Variational Monte Carlo

Variational Monte Carlo(VMC) is a method to evaluate the expectation value of a given wavefunction based on the combination of variational theorem and Monte Carlo sampling. Here we first show the variational theorem in quantum mechanics and then show how we do the calculation of VMC.

Variational theorem is what the word "variational" in variational Monte Carlo means. In VMC, the wavefunction we use is a trial wave function Ψ_T, which is the best available wavefunction provided by some other methods (Hartree Fock, CI or DFT). We evaluate the quantity

$$E_V = \frac{\langle \Psi_T | H | \Psi_T \rangle}{\langle \Psi_T | \Psi_T \rangle} \tag{2.15}$$

If the true ground state is E_0, then the variational theorem states

$$E_V \geq E_0 \tag{2.16}$$

We prove this by first expanding the trial wavefunction Ψ_T using the eigenstates of the

Hamiltonian.

$$\Psi_T = \sum_k \langle \Psi_k | \Psi_T \rangle \Psi_k \tag{2.17}$$

Since Ψ_k is the eigenstate, so

$$H\Psi_k = E_k \Psi_k \tag{2.18}$$

$$\begin{aligned} E_V &= \frac{\sum |\langle \Psi_k | \Psi_T \rangle|^2 E_k}{\sum |\langle \Psi_k | \Psi_T \rangle|^2} \\ &\geq \frac{\sum |\langle \Psi_k | \Psi_T \rangle|^2 E_0}{\sum |\langle \Psi_k | \Psi_T \rangle|^2} = E_0 \end{aligned}$$

We calculate E_V in Eq.2.15 using Monte Carlo integral. We rewrite the variational energy:

$$E_V = \frac{\int |\Psi_T(R)|^2 [\Psi_T(R)^{-1} H \Psi_T(R)] dR}{\int |\Psi_T(R)|^2 dR} \tag{2.19}$$

Then define the local energy

$$E_{local} = \Psi_T(R)^{-1} H \Psi_T(R) \tag{2.20}$$

We can consider E_V as the expectation of E_{local}, and the corresponding probability density function of E_{local} is

$$p(R) = \frac{|\Psi_T(R)|^2}{\int |\Psi_T(R)|^2 dR} \tag{2.21}$$

The variational Monte Carlo algorithm is summarized as follows:
(1)Generate a set of walkers that are distributed according to $|\Psi_T(R)|^2$ using Metropolis algorithm.
(2)For each walker, evaluate the local energy based on Eq. 2.20.
(3)Average the local energies at all walkers to get the expectation value. If there are M walkers totally, then the expectation value is

$$\langle E \rangle = \frac{1}{M} \sum_{i=1}^{M} (E_{local})_i \tag{2.22}$$

The variational nature of this method gives us the upper bound of the ground-state energy. For other operators other than the Hamiltonian, the expectation can be calculated

using this method, but the result is not variational. This disadvantage of VMC is the result is highly dependent on the choice of trial wavefunction. This obstacle is overcome in the method introduced in the next section.

2.4 Diffusion Monte Carlo Method

2.4.1 Diffusion Equations and Green's Functions

In the following sections, we will be introducing Diffusion Monte Carlo(DMC) method. The model used in DMC is based on the diffusion process. We start with the diffusion equation:

$$\frac{\partial \rho(x,t)}{\partial t} = D\frac{\partial^2 \rho(x,t)}{\partial x^2} = D\hat{L}\rho(x,t) \tag{2.23}$$

where $\rho(x,t)$ is the density of the diffusing material as function of space and time. The diffusion equation describes how the density evolves in time during this diffusion process. D is the diffusion constant.

The solution can be found easily if we follow the routine of solving a partial differential equation. But we try to solve the problem using Green's function, then implement the Green's function using Monte Carlo algorithm. Given an initial distribution $\rho(x,0)$, we can write down the solution using the Dirac notation

$$|\rho,t\rangle = exp(Dt\hat{L})|\rho,0\rangle \tag{2.24}$$

then

$$\langle x|\rho,t\rangle = \rho(x,t) = \int dx^{'}\langle x|exp(Dt\hat{L})|x^{'}\rangle\langle x^{'}|\rho,0\rangle \tag{2.25}$$

We define the Green's function G

$$G(x^{'},x;t) = \langle x|exp(Dt\hat{L})|x^{'}\rangle \tag{2.26}$$

and $\rho(x,t)$ can be written as:

$$\rho(x,t) = \int dx^{'}G(x^{'},x;t)\rho(x^{'},0) \tag{2.27}$$

From Eq.2.23 and 2.27, we can see

$$\frac{\partial}{\partial t}\int dx^{'}G(x^{'},x;t)\rho(x,0) = D\frac{\partial^2}{\partial x^2}\int dx^{'}G(x^{'},x;t)\rho(x,0) \tag{2.28}$$

Here we assume the differentiation operator and the integration operator are interchangeable. This can be rigorously proved by Fubini's theorem. We find out the Green's function $G(x,x^{'};t)$ satisfies the same equation as stated in Eq.2.23.

$$\frac{\partial G(x^{'},x;t)}{\partial t} = D\frac{\partial^2 G(x^{'},x;t)}{\partial x^2} \tag{2.29}$$

In order to have a unique solution, we need an initial condition. When $t = 0$

$$G(x^{'},x;0) = \delta(x - x^{'}) \tag{2.30}$$

To make the derivation easier, we write the Green's function using the momentum operator $\hat{p}^2 = -\frac{\partial^2}{\partial x^2}$ and change the position representation to the momentum representation:

$$\begin{aligned} G(x^{'},x;t) &= \langle x^{'}|exp(-Dt\frac{\partial^2}{\partial x^2})|x\rangle \\ &= \int dp\langle x^{'}|p\rangle exp(-Dt\frac{\partial^2}{\partial x^2})\langle p|x\rangle \\ &= \int_{-\infty}^{\infty} dp\frac{1}{\sqrt{2\pi}}e^{ipx^{'}}exp(-Dtp^2)\frac{1}{\sqrt{2\pi}}e^{-ipx} \\ &= \frac{1}{\sqrt{4\pi Dt}}exp[-(x-x^{'})^2/4Dt] \end{aligned} \tag{2.31}$$

This gives us the explicit form of Green's function. We can implement the diffusion process using random walk based on the Green's function. The state space is the position of sample points(walkers). Then probability density of walkers as a function of its position is

$$\rho(x,0) = \sum_i \delta(x - x_i) \tag{2.32}$$

Using this form, we write Eq. 2.27 as

$$\begin{aligned}\rho(x,t) &= \int dx^{'} G(x^{'},x;t)\rho(x^{'},0) \\ &= \int dx^{'} \frac{1}{\sqrt{4\pi Dt}} exp(-(x-x^{'})^2/4Dt) \sum_i \delta(x^{'}-x_i^{'}) \\ &= \sum_i \frac{1}{\sqrt{4\pi Dt}} exp(-(x-x_i^{'})^2/4Dt) \qquad (2.33)\end{aligned}$$

$$(2.34)$$

So the new distribution is a sum of normal distribution centered at each $x_i^{'}$ with variance $\sqrt{2Dt}$. By repeating this procedure we construct a Markov process with transition probability given by the Green's function. We use random walk to simulate this Markov process in discrete time. The movement of the walkers in the state space is given by:

$$x(t+\Delta t) = x(t) + \eta\sqrt{2D\Delta t} \qquad (2.35)$$

where η is a Gaussian random variable.

The equation we use in DMC is rather more complicated than a simple diffusion equation. When $D = 1/2$, consider the equation

$$\frac{\partial\rho(x,t)}{\partial t} = (\frac{1}{2}\frac{\partial^2}{\partial x^2} - V(x))\rho(x,t) \qquad (2.36)$$

Using the same method, we write the solution as

$$|\rho, t >= exp[-t(T+V)]|\rho, 0 > \qquad (2.37)$$

T is the kinetic operator $T = \frac{p^2}{2} = -1/2(\partial^2/\partial x^2)$. And $V(x)$ is the potential operator. Since the operators T and V do not commute, a precise way to evaluate the Green's function is based on Campbell-Baker-Hausdorff(CBH) commutators. For a time interval Δt

$$e^{-\Delta t(T+V)} = e^{-\Delta tT} e^{-\Delta tV} e^{-\Delta tCBH(T,V)} \qquad (2.38)$$

where $CBH(T,V) = \frac{1}{2}[T,V] + \frac{1}{12}[T,[T,V]] + ...$ If Δt is small, we may introduce an

approximation and ignore the CBH commutator, then

$$e^{-\Delta t(T+V)=e^{-\Delta tT}e^{-\Delta tV}} + O(\Delta t)^2 \tag{2.39}$$

The Green's function is the matrix element of the exponential operator in the position representation. After we made this approximation, the Green's function is written as

$$G(x^{'}, x; \Delta t) = \langle x|e^{-\Delta tT}e^{-\Delta tV}|x^{'}\rangle \tag{2.40}$$

From what we did above, the Green's function involved with kinetic operator is done by setting the diffusion constant $D = 1/2$.

$$G_T(x^{'}, x; \Delta t) = \frac{1}{\sqrt{2\pi\Delta t}}e^{-(x-x^{'})^2/(2\Delta t)} \tag{2.41}$$

Since the potential operator V commute with position operator. The Green's function involved with potential part is trivial. We can write down the full Green's function under the approximation given by Eq. 2.39 as

$$G(x^{'}, x; \Delta t) = G_T(x^{'}, x; t)e^{-\Delta tV(x^{'})} + (\Delta t)^2 \tag{2.42}$$

But there is a serious problem here. The term involving $V(x^{'})$ in the expression fails to satisfy the normalization condition, which prevents us interpreting the Green's function as a transition probability and then construct Markov Chain. However, we can modify Eq. 2.36 by adding another term E_T:

$$\frac{\partial\rho(x,t)}{\partial t} = \frac{1}{2}\frac{\partial^2}{\partial x^2} - (V(x) - E_T)\rho(x,t) \tag{2.43}$$

E_T here is chosen to keep the Green's function normalized. We will discuss how to choose E_T in the next section. Furthermore, there exists a more precise approximation to deal with the potential energy in the Green's function. The expression is

$$e^{-\Delta t(T+V)} = e^{-\Delta t\hat{V}/2}e^{-\Delta t\hat{T}}e^{-\Delta t\hat{V}/2} + (\Delta t)^3 \tag{2.44}$$

A more complex form of diffusion equation appearing in DMC is diffusion with drift:

$$\frac{\partial \rho(x,t)}{\partial t} = \frac{1}{2}\frac{\partial}{\partial x}[\frac{\partial}{\partial x} - F(x)]\rho(x,t) \tag{2.45}$$

This is Fokker-Planck equation, $F(x)$ is defined as

$$F(x) = \frac{1}{\rho(x)}\frac{d\rho(x)}{dx} \tag{2.46}$$

Following the same way of deriving the Green's function,we obtain the following expression:

$$G(x^{'}, x; \Delta t) = \frac{1}{\sqrt{2\pi\Delta t}} e^{-[x-x^{'}-F(x)\Delta t/2]^2/2\Delta t} \tag{2.47}$$

This Green's function can be implemented using random walk similar to Eq. 2.35 but adding a drift term:

$$x(t+\Delta t) = x(t) + \Delta t F(x)/2 + \eta\sqrt{\Delta t} \tag{2.48}$$

So far all that we discussed above is in one dimension. In three dimension, the Green's function for a simple diffusion process in Eq. 2.23 with $D = 1/2$ is

$$G_{3N}(\mathbf{R}^{'}, \mathbf{R}; \Delta t) = \frac{1}{(2\pi\Delta t)^{3N/2}} e^{-(\mathbf{R}-\mathbf{R}^{'})^2/(2\Delta t)} \tag{2.49}$$

The Green's function with a drift term is

$$G_{3N}(\mathbf{R}^{'}, \mathbf{R}; \Delta t) = \frac{1}{(2\pi\Delta t)^{3N/2}} e^{-(\mathbf{R}-\mathbf{R}^{'}-\Delta t F(\mathbf{R}^{'})/2)^2/(2\Delta t)} \tag{2.50}$$

2.4.2 Diffusion Monte Carlo Algorithm

The diffusion Monte Carlo can also be called as projection Monte Carlo. To understand the meaning of projection, we consider again the simple diffusion equation plus a potential term $V(x)$, as we see in Eq. 2.36. We expand the operator of Green's function with the adjusted term E_T in terms of the eigenstate of the Hamiltonian. We obtain

$$e^{-t(\hat{H}-E_T)} = \sum_n |\Phi_n\rangle e^{-t(E_n-E_t)} \langle\Phi_n| \tag{2.51}$$

E_n is the eigenenergy. It is straightforward to see for large t, the ground state energy E_G dominates in the sum $E_G - E_T$. Therefore we can interpret the Green's function operator as a projection operator on the ground state. We discussed how to implement the Green's function associated with the kinetic operator using a random walk. Now we need to treat the potential part. The potential part of the Green's function is $exp[-\Delta t(V(\mathbf{R}^{'}) - E_T)]$ or $exp[-\Delta t(V(\mathbf{R}) + V(\mathbf{R}^{'}) - 2E_T)/2]$. We can interpret this as a weight w being assigned to each walker, thus we introduce a process called branching process. The branching process states if $w < 1$, then the walker survives with probability w. If $w \geq 1$, the walker continue to evolve in space, and a new walker is created at the same position with probability $w - 1$. In the previous section, we introduced E_T in order to keep the normalization of the Green's function and construct the Markov chain. We now need to find a way to calculate E_T. We understand the fact that keeping the normalization of the Green's function means to keep the total probability density conserved. When we use a set of walkers to represent this process, it is equivalent to having the total number of walkers conserved, preventing a fast growing or decreasing population. Suppose the target number is M, and at a certain step the total number is $M^{'}$. If we multiply a factor $M/M^{'}$ in the Green's function, then the normalization is satisfied. We can achieve this by changing the E_T by $ln(M/M^{'})$. So we update E_T based on

$$E_T^{new} = E_T^{old} + Cln(M/M^{'}) \tag{2.52}$$

C can be chosen to control how fast the number of walkers approach the target number.

2.4.3 Guiding Function and Importance Sampling

When we directly apply the diffusion Monte Carlo algorithm to solve Eq. 2.36, we might encounter some problems. First, the approximation used in Eq. 2.39 requires a very small timestep to be taken. This can be very inefficient. Second, all the walkers are free to diffuse around. If our potential has singularities and the walker is in the close proximity of the singular point, the branching process will generate too many copies of the walker. This causes a biased sample of the distribution function and leads to a very slow convergence and large statistical errors. To overcome this, we choose a better sample distribution

function as our guiding function

$$f(\mathbf{R},t) = \Phi(\mathbf{R},t)\Psi_T(\mathbf{R}) \tag{2.53}$$

Where $\Psi_T(\mathbf{R})$ is some trial function which is the best guess of the true wavefunction we can find. From Eq. 2.36, it is easy to see that $f(\mathbf{R},t)$ satisfies a Fokker-Planck type of equation

$$\frac{\partial f(\mathbf{R},t)}{\partial t} = \frac{1}{2}\bigtriangledown^2_{\mathbf{R}} f(\mathbf{R},t) - \bigtriangledown \cdot [v_D(\mathbf{R})f(\mathbf{R},t)] - [E_L(\mathbf{R}) - E_T]f(\mathbf{R},t) \tag{2.54}$$

Here $v_D(\mathbf{R})$ is the drift velocity defined as

$$v_D(\mathbf{R}) = \bigtriangledown ln|\Psi_T(\mathbf{R})| = \Psi_T(\mathbf{R})^{-1} \bigtriangledown \Psi_T(\mathbf{R}) \tag{2.55}$$

And $E_L(\mathbf{R})$ is

$$E_L(\mathbf{R}) = \frac{H\Psi_T(\mathbf{R})}{\Psi_T(\mathbf{R})} \tag{2.56}$$

Using the trial function, the DMC algorithm becomes a combination of diffusion with drift and DMC with a potential term, which is implemented using branching process. The Green's function is written as a multiplication of two parts:

$$G(\mathbf{R}^{'},\mathbf{R};t) \approx G_d(\mathbf{R}^{'},\mathbf{R};t)G_b(\mathbf{R}^{'},\mathbf{R};t) \tag{2.57}$$

The diffusion part $G_d(\mathbf{R}^{'},\mathbf{R};t)$ is

$$G_d(\mathbf{R}^{'},\mathbf{R};t) = (2\pi t)^{-3N/2} exp[-\frac{(\mathbf{R}-\mathbf{R}^{'}-tv_D(\mathbf{R}^{'}))^2}{2t}] \tag{2.58}$$

The branching part G_b is

$$G_b(\mathbf{R}^{'},\mathbf{R};t) = exp\{-t[E_L(\mathbf{R}) + E_L(\mathbf{R}^{'}) - 2E_T]/2\} \tag{2.59}$$

The method above is called importance sampling. We notice the Green's function change in both diffusion part and branching part. The improvements we make are:
(1)In the diffusion part, the drift term we add can push the walkers going to the favorable region where the probability density is higher therefore we can sample that region more

often.
(2)In the branching part, the weights of walkers depends on the local energy instead of the potential energy only. Usually the local energy is a smoother function than the potential function so we avoid encountering singularities.

The last point we want to address here is the modification in the Green's function. In order to have better accuracy, smaller time steps are usually used to reduce the errors, but can lead to low efficiency. So we introduce a remedy which is the acceptance-rejection procedure in order to guarantee the detailed balance condition. The probability of acceptance is

$$P(\mathbf{R}^{'} \rightarrow \mathbf{R}) = min[1, \frac{G_d(\mathbf{R} \rightarrow \mathbf{R}^{'}, t)\Psi_T^2(\mathbf{R})}{G_d(\mathbf{R}^{'} \rightarrow \mathbf{R}, t)\Psi_T^2(\mathbf{R}^{'})}] \tag{2.60}$$

2.4.4 Expectation Values in DMC

In DMC, we calculate the expectation value of the energy by averaging the local energy at all sample points:

$$< E_{DMC} >= \frac{1}{N}\sum_{i=1}^{N} E_L(\mathbf{R}_i). \tag{2.61}$$

By using importance sampling, we can write $< E_{DMC} >$ as

$$\begin{aligned} < E_{DMC} > &= \frac{\int d\mathbf{R} f(\mathbf{R}) E_L(\mathbf{R})}{\int d\mathbf{R} f(\mathbf{R})} && (2.62)\\ &= \frac{\int d\mathbf{R}\Phi(\mathbf{R}) H\Psi_T(\mathbf{R})}{\int d\mathbf{R}\Phi(\mathbf{R})\Psi_T(\mathbf{R})} && (2.63)\\ &= E_0. && (2.64) \end{aligned}$$

It is clear to see the expectation value of ground state energy in DMC is equal to the true ground state energy.

2.4.5 Fixed-Node Approximation

We mentioned in the last section that $f(\mathbf{R}, t) = \Phi(\mathbf{R}, t)\Psi_T(\mathbf{R})$ represents the distribution function of the sample, so it is meaningful only when it take positive values. For bosons, the wavefunction of the ground state is guaranteed positive everywhere, however, this is not the case for fermions. The approximation used in the calculation to address this

problem is fixed-node approximation. And this is not the only available one, but it is the simplest one and used throughout the calculations presented in this thesis. In fixed-node approximation, we enforce $\Phi(\mathbf{R}, t)$ to have the same sign as $\Psi_T(\mathbf{R})$ everywhere. Therefore, every time the walker moves, we check if the sign of $\Psi_T(\mathbf{R})$ changes or not. We only accept the move which does not change the sign of $\Psi_T(\mathbf{R})$. In fact, adding this constraint is equivalent to changing the real Hamiltonian to the effective Hamiltonian. Consider the region where $\Psi_T > 0$, the effective Hamiltonian under fixed-node approximation(H^{FN}) can be written as follow:

$$H^{FN} = H + V_\infty \tag{2.65}$$

where $V_\infty(R) = 0$ if $\Psi_T(R) > 0$ and $V_\infty = \infty$ if $\Psi_T \leq 0$.
If $\Psi_T < 0$, then

$$H^{FN} = H + V_\infty \tag{2.66}$$

where $V_\infty(\mathbf{R}) = 0$ if $\Psi_T(\mathbf{R}) < 0$ and $V_\infty = \infty$ if $\Psi_T \geq 0$.
The infinite potential is added to prohibit any walker from crossing the regions where the wavefunction changes the sign. It can be rigorously proved that the expectation value of the energy under the fixed-node Hamiltonian (H^{FN}) is a strictly upper bound of the one in the true Hamiltonian. The detailed proof can be found in the Appendix of Reference [14].

2.4.6 Fixed-Phase Approximation

The fixed-node approximation is good enough if the trial wavefunction used is real. This is usually the case in the calculation for atomic and molecular systems. As long as the Hamiltonian of the systems has time-reversal symmetry, one can always use real wavefunction. However, many of the interesting properties of condensed matter physics involves breaking of time-reversal symmetry. Fox example, if the Hamiltonian contains magnetic field, it would be necessary to employ complex wavefunctions. An equivalent approach to tackle the fermion sign problem when the wavefunction is complex is the fixed-phase approximation.

We first write our complex trial wavefunction as $|\Psi_T(\mathbf{R})|exp(i\phi_T(\mathbf{R}))$, where $|\Psi_T(\mathbf{R})|$ is the modulus of the trial wavefunction, $\phi_T(\mathbf{R})$ is the phase of the wavefunction. We take the simplest Hamiltonian which contains only the kinetic energy operator and potential

operator and apply it to the wavefunction.

$$[\sum_i -\frac{1}{2}\bigtriangledown_i^2 + \sum_{i<j}\frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}]|\Psi_T(\mathbf{R})|exp(i\phi_T(\mathbf{R})) = E|\Psi_T(\mathbf{R})|exp(i\phi_T(\mathbf{R})) \tag{2.67}$$

where i denotes the ith electron. Using a little math, we can reorganize the equation above, and split into the following two equations, one being the real part, the other being the imaginary part.

$$[\sum_i \frac{1}{2}\bigtriangledown_i^2 + \sum_{i<j}\frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} + \frac{1}{2}\sum_i(\bigtriangledown_i\phi_T(\mathbf{R}))^2]|\Psi_T(\mathbf{R})| = E|\Psi_T(\mathbf{R})| \tag{2.68}$$

$$\sum_i \bigtriangledown_i \cdot [|\Psi_T(\mathbf{R})|^2 \bigtriangledown_i \phi_T(\mathbf{R})] = 0 \tag{2.69}$$

The first equation corresponds to the real part, which means we map the fermion problem into a bosonic one, of which the wavefunction is positive everywhere. Note the third term in the bracket on the left hand side of the equation is never an operator. It is a constant factor that only depends on the phase of the trial wavefunction. The second equation corresponds to the imaginary part. Since any eigenenergy must be real, the right hand side must be equal to nothing but 0.

The idea of the fixed-phase approximation is to keep the phase of the wavefunction the same as the phase of the trial wavefunction and solve the exactly bosonic problem for $|\Phi_T(\mathbf{R})|$ using DMC. It can also be shown that the expectation of the energy under the fixed-phase approximation is an upper bound of the true energy.

2.4.7 Summary of DMC Algorithm

We summarize the DMC algorithm:
1.Generate a set of walkers from some initial distribution. This is done in VMC. Usually we run VMC calculation first and take the distribution of the walkers generated by VMC to run DMC.
2.Repeat the following iterative steps:
(1)Evaluate drift velocity v_D of each walker.
(2)Move each walker to the proposed position according to

$$\mathbf{R} = \mathbf{R}^{'} + \Delta t v_D(\mathbf{R}^{'}) + \chi \tag{2.70}$$

χ is a 3N-dimensional random vector drawn from a normal distribution with mean 0 and variance Δt

(3)Check whether the walker has crossed the nodal surface by checking whether the sign of wavefunction has changed from $\mathbf{R}^{'}$ to $\mathbf{R}$.

(4)Accept the step according the acceptance ratio in Eq. 2.60.

(5)Calculate the weight of each walker

$$w = exp[-t(E_L(\mathbf{R}) + E_L(\mathbf{R}^{'}) - 2E_T)/2] \tag{2.71}$$

and branch each walkers by calculating the copies of each walker. The number of copies is

$$N_{new} = [\eta + w] \tag{2.72}$$

[ ] denotes the integer part, and η is a uniform random variable on the interval [0,1].

(6)Calculate the expectation values of interest by averaging the local energies of all the walkers.

(7)Adjust E_T to control the population of walkers.

3.After step 2 is repeated till the error bar of the quantity of interest is small enough.

2.4.8 Summary of Error Sources in DMC

Finally, we summarize the error sources in DMC.

1. Time step errors.

This is because of the approximation made in Eq. 2.39. We can control the error by taking smaller timestep and using acceptance/rejection step, and finally extrapolate the result to zero time step.

2. Population control bias

This bias exists because we artificially introduce E_T to control the total population of walkers. The choice of E_T is empirical in Eq. 2.52. We can improve the result by using larger number of population.

3. Fixed-node errors

This is clearly explained in the fixed-node approximation subsection. There is no way to control the error except choosing a better trial wavefunction. However, the error is quite small. For most chemical systems, it only accounts for 5% of the correlation energy.

4. Finite size errors

This error exists when we need to calculate the extensive properties of systems with infinite size. However, our simulation is possible only for finite systems. We will elaborate this problem in Chapter 6.

Chapter 3

Trial Wave Functions and Pseudopotentials in Quantum Monte Carlo

In this chapter we will be discussing how to build trial wavefunctions and how to treat systems with heavy atoms using pseudopotentials. The quality of trial wavefunctions is the most important factor that affects the accuracy of QMC calculations. A good trial wavefunction leads to a very efficient calculation and much smaller variance of those quantities we want to evaluate. Here we start with the properties of the many-body wavefunctions and then discuss how to build a good trial wavefunction.

3.1 Trial Wave Functions

3.1.1 Slater Determinant

Electrons are fermions, therefore the wavefunctions must satisfy the anti-symmetry conditions. The most common form of the anti-symmetric wavefunction is expressed in terms

of Slater determinants, which was discussed in the Hartree Fock Approximation.

$$\Psi(r_1, r_2, ..., r_N, \sigma_1, \sigma_2, ..., \sigma_N) = \begin{vmatrix} \psi_1(r_1, \sigma_1) & \psi_1(r_2, \sigma_2) & ... & \psi_1(r_N, \sigma_N) \\ \psi_2(r_1, \sigma_1) & \psi_2(r_2, \sigma_2) & ... & \psi_2(r_N, \sigma_N) \\ . & . & & . \\ . & . & & . \\ . & . & & . \\ \psi_N(r_1, \sigma_1) & \psi_N(r_2, \sigma_2) & & \psi_N(r_N, \sigma_N) \end{vmatrix}, \quad (3.1)$$

where the σ_i is the spin variables, and ψ_i is one-electron orbital which we can take from HF or DFT calculations. For each orbital, we have to specify both the spatial coordinates and the spin of each electron. In many cases, if the Hamiltonian does not contain spin operator explicitly, the Hamiltonian commutes with the total spin operator S^2 and the z-component of spin operator S_z. The spin becomes a constant of motion so we can do spin assignment to each electron. Once the spins are assigned, they never change over time. Suppose the first $N^{\uparrow}$ electrons are spin-up, and $N^{\downarrow} = N - N^{\uparrow}$ are spin-down. Now we can remove the spin variable and replace the original Slater determinant Ψ with $\Psi^{'}$ as a multiplication of two determinants: one is for spin-up electrons, and the other is for spin-down electrons.

$$\begin{aligned} & \Psi^{'}(r_1, r_2, ..., r_n, \sigma_1, \sigma_2, ..., \sigma_n) \\ = \; & \Psi^{'}(r_1, \uparrow, ...r_{N_\uparrow}, \uparrow r_{N_\uparrow+1} \downarrow, ..., r_N, \downarrow) \\ = \; & \frac{1}{\sqrt{N_\uparrow!}} \begin{vmatrix} \psi_1(r_1) & \psi_1(r_2) & ... & \psi_1(r_{N_\uparrow}) \\ \psi_2(r_1) & \psi_2(r_2) & ... & \psi_2(r_{N_\uparrow}) \\ . & . & & . \\ . & . & & . \\ . & . & & . \\ \psi_{N_\uparrow}(r_1) & \psi_{N_\uparrow}(r_2) & ... & \psi_{N_\uparrow}(r_{N_\uparrow}) \end{vmatrix} \frac{1}{\sqrt{N_\downarrow!}} \begin{vmatrix} \psi_{N_\uparrow+1}(r_{N_\uparrow+1}) & ... & \psi_{N_\uparrow+1}(r_N) \\ \psi_{N_\uparrow+2}(r_{N_\uparrow+1}) & ... & \psi_{N_\uparrow+2}(r_N) \\ . & ... & . \\ . & ... & . \\ . & ... & . \\ \psi_N(r_{N_\uparrow+1}) & & \psi_N(r_N) \end{vmatrix}. \end{aligned}$$

It can be rigourously shown that the expectation of an operator($\hat{O}$) which does not contain spin operator calculated using the multiplication of the two determinants is the same as the one calculated using one single slater determinant[15].

$$\langle \hat{O} \rangle = \frac{\langle \Psi^{'} | \hat{O} | \Psi^{'} \rangle}{\langle \Psi^{'} | \hat{O} | \Psi^{'} \rangle} = \frac{\langle \Psi | \hat{O} | \Psi \rangle}{\langle \Psi | \hat{O} | \Psi \rangle}. \quad (3.2)$$

3.1.2 Multireference Determinants

The final result of DMC calculation is strongly affected by the quality of the wavefunction. By applying the fixed-node approximation, we enforce the projected wavefunction to have the same nodes as the trial wavefunction. Therefore, the node of the trial wavefunction has a direct consequence on the accuracy of the result. Our goal is to have a better trial wavefunction with a better node. We can do this by applying a linear expansion of Slater determinants taken from Configuration Interaction. The form of CI expansion was discussed in Section 1.3. It is written as the following

$$\Psi = \Psi_0 + \sum_i \sum_p c_i^p \Psi_i^p + \sum_{ij} \sum_{pq} c_{ij}^{pq} \Psi_{ij}^{pq} + \sum_{ijk} \sum_{pqr} c_{ijk}^{pqr} \Psi_{ijk}^{pqr} + ... \tag{3.3}$$

Using a large expansion of Slater determinant leads to a very expensive calculation, so this is only applicable to some small chemical systems. Other methods, although not used in the calculations presented in this thesis, also exisit to improve the node of the wavefuction, such as Pfaffian pairing determinant or Backflow determinant for different systems.

3.1.3 Cusp Conditions

In QMC method, our local energy is defined in Chapter 2 as

$$\frac{H\Psi(\mathbf{R})}{\Psi(\mathbf{R})} = \frac{(-\frac{1}{2}\sum_i \bigtriangledown_i^2 + \sum_{i<j} V_{ij})\Psi(\mathbf{R})}{\Psi(\mathbf{R})}, \tag{3.4}$$

where $\mathbf{R}$ is a $3N$ dimensional vector representing the coordinates of N electrons in the system. To be able to sample the local energy, we need the local energy to be finite otherwise the result may have large fluctuations and may be meaningless. However, the potential part, either the electron-ion potential or the electron- electron potential, diverges when distance of the electron and ion or the electron and electron is zero. This fact enforces the divergence in the kinetic energy part has to exactly cancel the divergence in the potential part. We introduce the cusp condition in our variational trial wavefunction in order to satisfy this property.

Considering two electrons with mass m_1 and m_2, we define the total mass and reduced mass

$$M = m_1 + m_2; \tag{3.5}$$

$$\mu = \frac{m_1 m_2}{m_1 + m_2}; \tag{3.6}$$

$$\overline{\mathbf{r}} = \frac{m_1 \mathbf{r_1} + m_2 \mathbf{r_2}}{m_1 + m_2}; \tag{3.7}$$

$$\mathbf{r} = \mathbf{r_1} - \mathbf{r_2}. \tag{3.8}$$

The local energy is

$$E_L = -\frac{1}{2M}\frac{\bigtriangledown^2 \Psi(\overline{\mathbf{r}})}{\Psi(\overline{\mathbf{r}})} - \frac{1}{2\mu}\frac{\bigtriangledown^2 \Psi(\mathbf{r})}{\Psi(\mathbf{r})} + V(\mathbf{r}). \tag{3.9}$$

Because the potential is spherically symmetrical, for convenience we write the equation using spherical coordinates. We know the potential between the ith and jth particle is $V(r) = \frac{q_i q_j}{r}$, with $r = |\mathbf{r}_i - \mathbf{r}_j|$. So

$$E_L = -\frac{1}{2M}\frac{\bigtriangledown^2 \Psi(\overline{\mathbf{r}})}{\Psi(\overline{\mathbf{r}})} - \frac{1}{2\mu}\frac{1}{\Psi(r)}\frac{\partial^2 \Psi(r)}{\partial r^2} - \frac{1}{\mu}\frac{1}{r}\frac{1}{\Psi(r)}\frac{\partial \Psi(r)}{\partial r} + \frac{l(l+1)}{r^2} + \frac{q_i q_j}{r}. \tag{3.10}$$

For the simplest case, assume $l = 0$. It is easy to note the second term and the last term in Eq. 3.10 diverges when $r = 0$. Thus when the local energy takes a finite value, these two terms must be cancelled by each other.

$$-\frac{1}{\mu}\frac{1}{r}\frac{1}{\Psi(r)}\frac{\partial \Psi(r)}{\partial r} + \frac{q_i q_j}{r} = 0. \tag{3.11}$$

This leads to

$$\left.\frac{\partial \Psi(r)}{\partial r}\right|_{r=0} = \mu q_i q_j \Psi(r=0). \tag{3.12}$$

This is the general form of the cusp condition. We now discuss the specific forms of the cusp conditions in different circumstances.

(1) Electron-nucleus cusp

If $l = 0$, then it is straightforward. Since the mass of the nucleus is sufficiently larger than the mass of electron. μ is chosen to be 1. We can take $q_i = Z$, which is the charge of the nucleus, and $q_j = -1$. Then electron-nucleus cusp condition is

$$\left.\frac{1}{\Psi}\frac{\partial \Psi}{\partial r}\right|_{r=0} = -Z. \tag{3.13}$$

For $l > 0$, the radial part of the wavefunction can be written in the form $r^l\rho^r$ and ρ is finite at $r = 0$. Using the same way of analysis, the electron-nucleus cusp condition is

$$\frac{1}{\rho(r)}\frac{d\rho(r)}{dr} = -\frac{Z}{l+1}. \tag{3.14}$$

(2) Electron-electron cusp

For a system containing two electrons, the reduced mass is $\frac{1}{2}$. When they have the parallel spin, the ground state is s state with $l = 0$. Therefore the cusp condition is

$$\frac{1}{\rho(r)}\frac{d\rho(r)}{dr} = \frac{1}{2(l+1)} = \frac{1}{2}. \tag{3.15}$$

If they have anti-parallel spins, then the lowest state is p state with $l = 1$, so the cusp condition is

$$\frac{1}{\rho(r)}\frac{d\rho(r)}{dr} = \frac{1}{2(l+1)} = \frac{1}{4}. \tag{3.16}$$

In QMC, we introduce factor called Jastrow factor $e^{-u(r_{ij})}$ to enforce the cusp condition of the wavefunction. The wavefunction Ψ is

$$\Psi(\mathbf{R}) = e^{-u(r_{ij})}f(\mathbf{R}) = e^{-u(r_{ij})}f(\mathbf{R}^{(ij)}, r_{ij}, \overline{\mathbf{r}}), \tag{3.17}$$

where the superscript (ij) means the coordinates of all other electrons except the ith and jth electrons. And $\mathbf{r}_{ij} = \mathbf{r}_i - \mathbf{r}_j$, $\overline{\mathbf{r}} = \frac{1}{2}(\mathbf{r}_i + \mathbf{r}_j)$. We want to see how the Jastrow factor should satisfy the cusp condition. Now we only focus on the ith and jth electrons. The kinetic operator can be written as

$$-\frac{1}{2}\bigtriangledown^2 = -\bigtriangledown^2_{\mathbf{r}_{ij}} - \frac{1}{4}\bigtriangledown^2_{\overline{\mathbf{r}}}, \tag{3.18}$$

Since we are only interested in the situation when the ith electron and jth electron come very close, we do not care the term of $\bigtriangledown_{\overline{\mathbf{r}}}$. So we derive our local energy, which is

$$\begin{aligned}
\frac{1}{\Psi}H\Psi &= \frac{1}{e^{-u(r_{ij})}f(\mathbf{R})}(-\bigtriangledown^2_{\mathbf{r_{ij}}} + \frac{1}{r_{ij}})e^{-u(r_{ij})f(\mathbf{R})} \\
&= \frac{\partial^2 u}{\partial r_{ij}^2} + \frac{2}{r_{ij}}\frac{\partial u}{\partial r_{ij}} - (\frac{\partial u}{\partial r_{ij}})^2 + 2\frac{\partial u}{\partial r_{ij}}\hat{\mathbf{r}}_{ij}\cdot\frac{\bigtriangledown f}{f} - \frac{\bigtriangledown^2 f}{f} + \frac{1}{r_{ij}}.
\end{aligned}$$

There are two possibilities. The first one is when the spins of the two electrons are anti-parallel, so $f(r_{ij})$ is a finite value at $r_{ij} = 0$. The diverging terms are $\frac{2u'}{r_{ij}}$ and $\frac{1}{r_{ij}}$. Therefore, the cusp condition is

$$\left.\frac{du}{dr_{ij}}\right|_{r=0} = -\frac{1}{2}. \tag{3.19}$$

If the spins are parallel, then $f(r_{ij}) = 0$ when $r_{ij} = 0$. $f(r_{ij})$ should be an odd function of r_{ij}. When r_{ij} goes to zero, we can expand $f(r_{ij})$ up to the second order

$$f(r) = \bigtriangledown f(r)\Big|_{r=0} \cdot r. \tag{3.20}$$

This makes the term $2u'\hat{r}\cdot\frac{\bigtriangledown f}{f}$ also diverge. Let $a = \bigtriangledown f(r)|_{r=0}$. So the sum of the diverging terms are

$$\frac{2u'}{r} + \frac{2u'}{a\cdot r}\hat{r}\cdot a + \frac{1}{r} = \frac{4u'+1}{r}. \tag{3.21}$$

Then the cusp condition is

$$\left.\frac{du}{dr_{ij}}\right|_{r=0} = -\frac{1}{4}. \tag{3.22}$$

This is the same as shown in Eq. 3.15 and 3.16.

3.1.4 Slater-Jastrow Wavefunction

The most popular form of the wavefunction in QMC is called the Slater-Jastrow wavefunction. Most of the calculations presented in the thesis use Slater-Jastrow wavefunction. The Slater-Jastrow wavefunction is a product of Slater part and Jastrow part.

$$\Psi(\mathbf{R}) = Slater(\mathbf{R})exp(J(\mathbf{R})). \tag{3.23}$$

The Slater part can be a single determinant or a linear combination of determinants as discussed in the previous section. The Jastrow factor here is used to include the correlation effects between the particles. The choice of Jastrow factor has a lot of freedom as long as it satisfies the well-known conditions. Generally, the correlation factor $J(\mathbf{R})$ is

$$J(R) = \chi(r_i) + u(r_{ij}) + w(r_{ijI}). \tag{3.24}$$

$\chi(r_i)$ is the one-body term describing the electron-nucleus correlation with electron-nucleus cusp conditions. $u(r_{ij})$ is the two-body term describing the electron-electron correlation with electron-electron cusp conditions. Since the u_{ij} term is dependent on the relative distance between the two electrons, it is able to catch the correlation beyond the one-particle effect. For uniform system like electron gas, having the u_{ij} term is enough. However for nonuniform systems, this factor changes the electron density by pushing the electrons from the high-density regions to low-density regions. This negative effect is not very good as our Slater determinant usually provides a very accurate one-particle density. The $\chi(r_i)$ term can compensate for the loss by adjusting the electron density near the nucleus. For non-uniform systems like atomic and molecular systems, both $\chi(r_i)$ and $u(r_{ij})$ are needed as they work together to describe the correlation without decreasing the quality of the density. For more accurate calculations, it is necessary to include a three-body term which depends on the position of two electrons and one nucleus. [16]

3.1.5 Basis Functions of Jastrow Factor

In QMC calculations, the terms in the Jastrow factor $\chi(r_i)$, $u(r_{ij})$ and $w(r_{ijI})$ are expansions of the chosen basis functions. For example, the one-body term is expressed as

$$\chi(r_i) = \sum_k a_k f_k(i). \tag{3.25}$$

f_k is the basis function. Expressions for the two-body and three-body terms are similar. The basis function we use in the cusp conditions are

$$f(\rho) = c(\frac{\rho - \rho^2 + \frac{1}{3}\rho^3}{1 + \gamma(\rho - \rho^2 + \frac{1}{3}\rho^3)} - \frac{1}{\gamma + 3}), \tag{3.26}$$

where $\rho = r/r_{cut}$ and r_{cut} gives the cutoff distance. For $r > r_{cut}$, $f(\rho) = 0$. we enforce the cusp condition by choosing correct c values. In electron-nucleus cusp, $c = -\frac{Z}{l+1}$. In electron-electron cusp, $c = \frac{1}{2}$ for spin-unlike electrons, and $c = \frac{1}{4}$ for spin-like electrons. γ affects the curvature of the function and can be optimized. The cusp functions for electrons with antiparallel spin and parallel spin are plotted in Fig.3.1.

The Polynomial Pade functions are also used in the Jastrow factor. It can smooth out the wavefunction and give better variational energies and small fluctuations. Since the introduction of electron-electron cusp condition may have a negative effect on one-

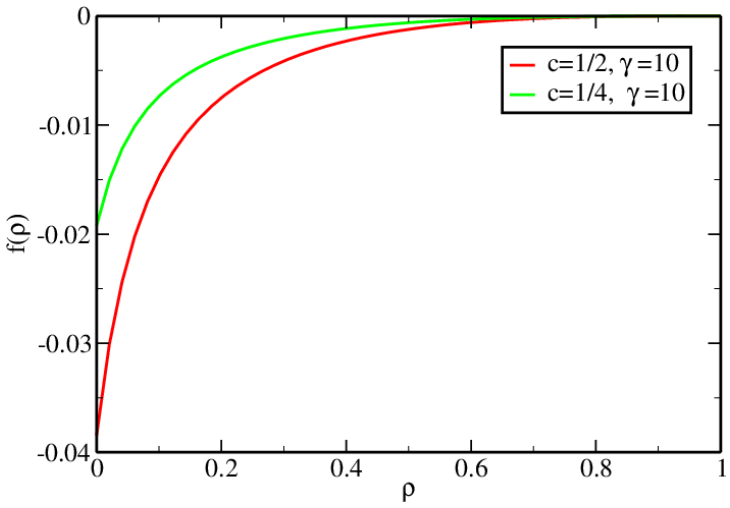

Figure 3.1: Basis function for cusp conditons

particle density, the Polynomial Pade functions can compensate for the negative effect and also preserve the cusp conditions. The form of Polynomial Pade functions is

$$f_{pade}(\rho) = \frac{1 - \rho^2(6 - 8\rho + 3\rho^2)}{1 + \gamma\rho^2(6 - 8\rho + 3\rho^2)}. \tag{3.27}$$

ρ is defined in the same way as in cusp functions. Note that when $\rho = 0$, the derivative

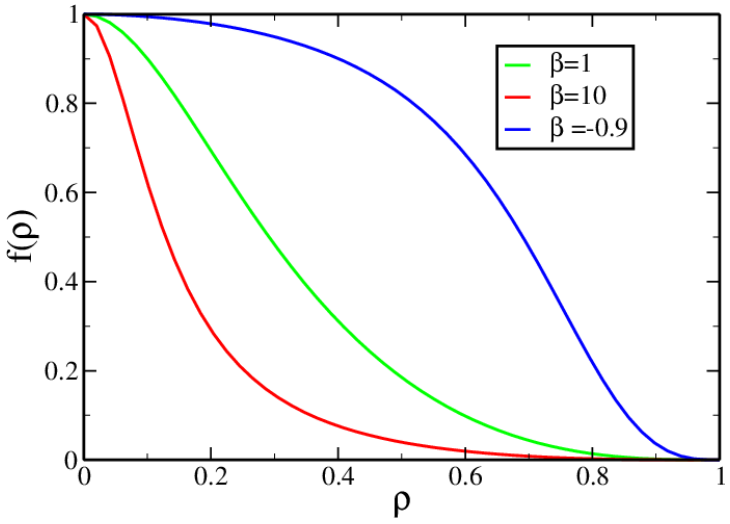

Figure 3.2: Basis for Polynomial Pade functions

$f^{'}(\rho)$ is zero, so the function does not change the cusp condition. When $r = r_{cut}$, the function approaches to zero, so the wavefunction in the region $r > r_{cut}$ is unchanged.

Usually, we use an expansion of the polynomial Pade functions in the Jastrow factor. For example, the piece of the one-body electron-nucleus functions is

$$\chi(r) = \sum_i c_i f_{pade}(r), \tag{3.28}$$

where c_is are the coefficient to be optimized. Several functions with different curvatures(different β) are employed so they provide more variational freedom.

3.1.6 Effect of Jastrow Factor on QMC Calculations

In variational Monte Carlo calculations, the effect of adding Jastrow factor is clearly visible. It brings down the total energy. This can be understood that the Jastrow factor reduces the probability of the two electrons coming too close to each other. When the electrons move further away from each other, they have less repulsive Coulomb energy and move slowly. Another positive effect of Jastrow factor is to reduce the fluctuations in energy because as electrons are further away from each other the wavefunction becomes smoother. In diffusion Monte Carlo calculations, since the Jastrow factor remains positive everywhere, it does not change the structure of the trial wavefunction nodes. Therefore, in principle, if the running time is long enough, the added Jastrow factor does not change the outcome of DMC calculations. The only thing that the Jastrow factor can do is to reduce the fluctuations. However, for most systems with large atoms, the pseudopotential is needed, then by using some approximations we can generate errors. The pseudopotential will be discussed in the following sessions.

3.2 Wavefunction Optimization

In Jastrow wavefunction, the basis sets of cusp and Pade functions contain several variational parameters. Suppose our sample points $\mathbf{R}$ are fixed, our variational energy E_V is dependent on the parameters in Jastrow factor.

$$E_V(\mathbf{c}) = \frac{\int |\Psi_T(\mathbf{c})|^2 E_L(\mathbf{c}) d\mathbf{R}}{\int |\Psi_T(\mathbf{c})|^2 d\mathbf{R}}, \tag{3.29}$$

where $\mathbf{c}$ is the set of parameters in Jastrow. We want those parameters chosen in such a way that they provide better variational energies and small fluctuations. This is achieved by optimizing the wavefunction. There are several approaches to do this. One can optimize the wavefunction by optimizing the variance or the variational energy, or we may choose a mixed function to optimize the variance and energy together.

If we choose to minimize the variance of the energy, then the target function is

$$f_{min}\{\mathbf{c}\} = \sigma^2(\mathbf{c}) = \frac{\int \Psi_T^2(\mathbf{c})[E_L(\mathbf{c}) - E_V(\mathbf{c})]^2 d\mathbf{R}}{\int \Psi_T^2(\mathbf{c}) d\mathbf{R}}. \tag{3.30}$$

Using the target function is good but not that convenient. Whenever the parameter set $\mathbf{c}$ changes, we need to generate a new distribution function $\Psi_T^2(\mathbf{c})$. This can be avoided if we write our target function $f_{min}(\mathbf{c})$ in the following way

$$f_{min}\{\mathbf{c}\} = \sigma^2(\mathbf{c}) = \frac{\int \Psi_T^2(\mathbf{c}_0) w(\mathbf{c})[E_L(\mathbf{c}) - E_V(\mathbf{c})]^2 d\mathbf{R}}{\int \Psi_T^2(\mathbf{c}_0) w(\mathbf{c}) d\mathbf{R}}, \tag{3.31}$$

where $\mathbf{c}_0$ is the initial value of parameters, and $w(\mathbf{c}) = \frac{\Phi_T^2(\mathbf{c})}{\Phi_T^2(\mathbf{c}_0)}$. This is the correlated sampling.

Another way is to minimize the total energy in Eq. 3.29. In the calculation presented throughout this thesis, a mixed procedure is employed to both reduce the fluctuations and have a better variational energy. The target function we use here in the mixed procedure is

$$f_{min}(\mathbf{c}) = weight * E_V(\mathbf{c}) + (1 - weight) * \sigma^2(\mathbf{c}), \tag{3.32}$$

where the value weight is here is 0.95.

Several numerical methods exist to optimize the target function. Here we choose the Levenberg-Marquardt method which is a combination of steep descent method and Newton method. To do this, we need to compute the Hession matrix $\mathbf{H}$ and the gradient vector $\mathbf{g}$ of the target function with respect to the parameter set $\{\mathbf{c}\}$. We first start with some initial value $\mathbf{c}^{(0)}$, and update $\{\mathbf{c}\}$ by

$$\mathbf{c}^{(i+1)} = \mathbf{c}^{(i)} - (\mathbf{H} + \mu\mathbf{I})^{-1}\mathbf{g}, \tag{3.33}$$

where μ is the damping parameter we need to determine. μ has to be positive so $\mathbf{H} + \mu\mathbf{I}$ is a positive definite matrix so f_{min} is descending. If $\mu = 0$, then the method is reduced to

Quasi-Newton Method. If μ is chosen to be really large, the diagonal elements in $\mathbf{H}+\mu\mathbf{I}$ are dominant, then the method becomes the steep descent method. At each step, we compute the following

$$r = \frac{f_{min}^{(i)} - f_{min}^{(i+1)}}{\frac{1}{2}(\mathbf{c^{(i+1)}} - \mathbf{c^{(i)}})(\mathbf{c^{(i+1)}} - \mathbf{c^{(i)}} - \mathbf{g})}. \tag{3.34}$$

Then if $r > 0$,

$$\mu^{(i+1)} = \mu^{(i)} max[\frac{1}{3}, 1-(2r-1)^3] \tag{3.35}$$

otherwise

$$\mu^{(i+1)} = 2^n \mu^{(i)} \tag{3.36}$$

The details of this method is discussed in Ref. [17]

It is straightforward to calculate the gradient and the Hessian of the target function for optimizing variance and energy[18] [19]. In optimizing the variance, the ith component of gradient vector of the target function is

$$g_i = 2[< E_L^{(i)}(E_L - E_V) > + < \frac{\Psi_T^{(i)}}{\Psi_T} E_L^2 > - < \frac{\Psi_T^{(i)}}{\Psi_T} >< E_L^2 > -2E_V < \frac{\Psi_T^{(i)}}{\Psi}(E_L - E_V) >], \tag{3.37}$$

where $E_L^{(i)}$, $\Psi_T^{(i)}$ is the partial derivative with respect to c_i.

$$H_{ij} = 2 < (E_L^{(i)}) - E_V^{(i)})(E_L^{(j)} - E_V^{(j)}) > \tag{3.38}$$

In minimizing the energy, the gradient components are

$$g_i = 2 < \frac{\Psi_T^{(i)}}{\Psi_T}(E_L - E_V) > \tag{3.39}$$

and the Hessian components are

$$H_{ij} = 2[< (\frac{\Psi_T^{(ij)}}{\Psi_T} + \frac{\Psi_T^{(i)}\Psi_T^{(j)}}{\Psi_T^2})(E_L - E_V) > - < \frac{\Psi_T^{(i)}}{\Psi_T} E_V^{(j)} > - < \frac{\Psi_T^{(j)}}{\Psi_T} E_V^{(i)} > + < \frac{\Psi_T^{(i)}}{\Psi} E_L^{(j)} >], \tag{3.40}$$

where the superscript denotes components of the gradient or Hessian.

3.3 Pseudopotentials

3.3.1 The need of Pseudopotentials

In quantum Monte Carlo calculations, the computational cost increases rapidly with the size of the systems. It has been estimated the scaling of the calculation with Z ranging from $Z^{5.5}$ [20]to $Z^{6.5}$[21]. Due to this large scaling, many application of QMC methods to large systems are almost impossible. On the other hand, in the atomic or molecular systems, the deep Coulomb potential near the nuclei gives rise to large oscillations of the wavefunction, leading large fluctuations of the local energy. This problem can be alleviated by using a more accurate trial wavefunction but may still be present. Moreover, for large systems a very good quality of trial wavefunction is difficult to obtain. Fortunately, most of the properties that we are interested in do not depend too much on the core electrons(electrons from the inner shell of the atom), since the core states tend to be rigid and do not affect many properties of electronic structure in the valence space, such as chemical bonds, excitations, band gaps. So we introduce effective potential called pseudopotentials which remove the existing potential in the region of the core electron orbitals and replace with an effective potential to mimic the effects of the core electrons.

3.3.2 The Construction of Pseudopotentials

The construction of pseudopotentials is not trivial. The pseudopotential must satisfy several conditions. The most important one is that outside the region of core electrons, the states of valence electrons should be identical to those states as if the core electrons are present. There exist many schemes to construct pseudopotentials used in ab initio calculations such as norm-conserving pseudopotentials and ultra-soft pseudopotentials [22], [23], [24], [25]. In the calculations presented in this thesis, we will apply energy-consistent pseudopotentials. This requires the eigenenergies should match exactly the all-electron energies of a number of different configurations for the atom under consideration. Since the atomic system has spherical symmetry, electrons in the valence space with different angular momentum l could be treated differently, which means a l-dependent pseudopotential. In QMC calculation, the pseudopotential for a given atom consists of a local part universal to all the angular momenta and a nonlocal part which is dependent on angular momentum. For a specific atom in the core region, the pseudopotential takes

the form

$$V_{ps}(r) = V_{loc}(r) + \hat{V}_{nl}(r), \tag{3.41}$$

where $V_{loc}(r)$ is the local part, which usually includes the potential from the ion. $\hat{V}_{nl}(r)$ is the nonlocal part. The nonlocal part can be written as

$$V_{nl}(r) = \sum_{l,m} V_l(r)|lm >< lm|, \tag{3.42}$$

where $V_l(r)$ is a function of r, which is l-dependent. More precisely, it should depend on quantum number j, and $j = l \pm \frac{1}{2}$, so we have two terms: $V_{l+1/2}$ and $V_{l-1/2}$. But if we don't take fully relativistic effect into account, we can simply take the average of $V_{l+1/2}$ and $V_{l-1/2}$, and fold both terms into a single term $V_l(r)$. $|lm >< lm|$ is a projection operator. When this nonlocal operator acts on a function, it works as

$$\hat{V}_{nl}(r)f(\mathbf{r}) = \sum_{l,m} V_{nl}(r)Y_{lm}(\Omega) \int Y^*_{lm}(\Omega^{'})f(\mathbf{r}^{'})d\Omega^{'}, \tag{3.43}$$

where Y_{lm} is spherical harmonic functions.

3.3.3 Pseudopotential in VMC and DMC

In VMC, the contribution of the nonlocal pseudopotential operator to the local energy can be written as

$$\begin{aligned} E_{nl} &= \frac{\hat{V}_{nl}\Psi_T}{\Psi_T} \\ &= \sum_l V_{nl}(r_i) \sum_{m=-l}^{l} Y_{lm}(\Omega_{r_i}) \int Y^*_{lm}(\Omega_{r_i^{'}}) \\ &\quad \times \frac{\Psi_T(r_1, ..., r_{i-1}, r_i^{'}, r_{i+1}, ..., r_N)}{\Psi_T(r_1, ...r_{i-1}, r_i, r_{i+1,...,}r_N)} d\Omega_{r_i^{'}}, \end{aligned} \tag{3.44}$$

because

$$\sum_{m=-l}^{l} Y_{lm}(\Omega_{r_i})Y^*_{lm}(\Omega_{r_i^{'}}) = \frac{2l+1}{4\pi} P_l[cos(\theta_i^{'})], \tag{3.45}$$

where P_l is the Legendre polynomial,

$$E_{nl} = \sum_l V_{nl}(r_i) \frac{2l+1}{4\pi} \int P_l[cos(\theta_i^{'})] \times \frac{\Psi_T(r_1, ..., r_{i-1}, r_i^{'}, r_{i+1}, ..., r_N)}{\Psi_T(r_1, ...r_{i-1}, r_i, r_{i+1,...,r_N})} d\Omega_{r_i^{'}}. \quad (3.46)$$

This involves an integration over the surface of the sphere centered at $r^{'}$. This integral can be calculated numerically by choosing certain number of grid points on the surface of the sphere. It is proved that for the maximum $l = 3$ it is sufficient to choose 6 grid points for reasonable accuracy[26]. Higher angular momentum requires more grid points.

In DMC, the case is more complex since the nonlocal operator brings trouble when we want to evaluate the Green's function. One way to overcome this is to apply localization approximation. We rewrite our diffusion-drift equation

$$\partial_t f = \frac{1}{2} \bigtriangledown^2 - \bigtriangledown \cdot (\mathbf{v}_D f) - \frac{(\hat{H} - E_T)\Psi_T}{\Psi_T} f + (\frac{\hat{V}_{nl}\Psi_T}{\Psi_T} - \frac{\hat{V}_{nl}\Phi}{\Phi}) f. \quad (3.47)$$

In the localization approximation, the last term in curly brackets is ignored. This generates an error and one can prove the error is proportional to $(\Psi_T - \Phi_0)^2$[26]

Chapter 4

Study of Dipole Moments of LiSr and KRb Molecules

sections of this chapter also appeared in

Study of dipole moments of LiSr and KRb molecules by quantum Monte Carlo methods

Shi Guo , Michal Bajdich , Lubos Mitas, Peter J. Reynolds

Molecular Physics, Volume 111, Issue 12-13, Pages 1744-1752,

Heteronuclear dimers are of significant interest to experiments seeking to exploit ultracold polar molecules in a number of novel ways including precision measurement, quantum computing, and quantum simulation. We calculate highly accurate Born-Oppenheimer total energies and electric dipole moments as a function of internuclear separation for two such dimers, LiSr and KRb. We apply fully-correlated, high-accuracy quantum Monte Carlo methods for evaluating these molecular properties in a many-body framework. We use small-core effective potentials combined with multi-reference Slater-Jastrow trial wave functions to provide accurate nodes for the fixed-node diffusion Monte Carlo method. For reference and comparison, we calculate the same properties with Hartree-Fock and with restricted Configuration Interaction methods, and carefully assess the impact of the recovered many-body correlations on the calculated quantities. For LiSr we find a highly nonlinear dipole moment curve, which may make this molecule's dipole moment tunable through vibrational state control.

4.1 Introduction

Motivated both by the desire for deeper understanding of basic physics and of numerous potential applications, there has been a great deal of work over the last few years devoted to cooling small molecules [27] and molecular ions [28]. This follows up on a few decades of exciting progress in cooling and trapping of atoms and atomic ions. Molecules, of course, are far more challenging, due to their complex internal structure, particularly their many internal degrees of freedom, such as vibrational and rotational levels, in addition to fine and hyperfine structure. A number of approaches have been pursued, with varying degrees of success. While making molecules cold has been achieved in numerous ways, reaching the ultra-cold regime, and particularly the quantum degenerate regime, has been limited.

The most general methods of cooling molecules use buffer gases and supersonic expansion [29, 30, 31], or velocity filtering [32] techniques. Another direct approach is Stark decceleration [33]. On the other hand, direct laser cooling of molecules, in analogy to the very successful approach for neutral atoms, was long deemed impractical due to the complex level structures. The simplest and most widespread method of direct laser cooling of atoms is Doppler cooling, where radiative forces originate from momentum transfer to atoms from a laser field, and subsequent spontaneous emission of slightly higher energy photons into random directions. Repeating this optical cycle tens of thousands of times

cools neutral atoms very quickly to the Doppler limit (which is mass dependent, but typically reaches sub-mK temperatures). For molecules, the problem with the conventional scheme is that excited states can radiatively decay out to a multitude of other states. This leads to decays that destroy any closed cycling transitions. Exciting population from all these states back to the starting point requires an impractically large number of lasers. Only recently has laser cooling of a special class of molecules been demonstrated [34, 35] based on insights from earlier work [36, 37].

Most successful at reaching into the ultra-cold regime are methods that create the molecules from previously cooled atoms. However, the number of molecules produced in this way is fairly small ($\approx 10^4$). Such methods include photoassociation [38] and magnetoassociation (exploiting Feshbach resonances) [39]. While the former has been a well-established technique to produce homonuclear dimers from cold atoms, its application to produce heteronuclear molecules from two different species of laser-cooled alkali atoms is more recent. Such diatomic molecules can be polar, and this is of particular interest (see, for example, reviews in [27, 30, 40, 41]).

The usefulness of polar molecules arises because they, unlike neutral atoms, interact via the characteristic long-range, anisotropic dipolar interaction, making them controllable by external electric fields. On the other hand, they are less strongly coupled to the environment than ions, making them less prone to decoherence [42]. This makes them attractive for quantum manipulation, such as for quantum information processing [43, 44, 45, 46, 47, 48], and precision measurements to test symmetries [49], [50], [51], [52] and the constancy of fundamental "constants" [53, 54, 55, 56, 57, 58].

Additional interest lies in alkali-alkali earth dimers such as LiSr, one of the molecules studied here. In contrast with alkali dimers, the unpaired spin makes them able to be manipulated with both external electric and magnetic fields. This provides an avenue for different physics to be explored, whether in fundamental tests, as qubits, in optical lattices where the competing interactions are important, or potentially even for use in atomic clocks where atoms such as Sr are already showing great potential [59].

Cooling, trapping and quantum manipulation of polar molecules will have an important impact on a diverse range of fields beyond just fundamental tests and quantum computing. One such example is in condensed matter physics, through quantum "emulation" or simulation of many-body quantum physics that is inaccessible to even today's (and any day's) high-performance computers, let alone to analytical solution (for ex-

ample [60, 61, 62]). Cold polar molecules trapped in optical lattices make an extremely interesting many-body system (with long-range dipole-dipole interactions) promising a rich variety of novel quantum phases, such as supersolids and topologically ordered states [61, 63]. One can imagine many other novel strongly-correlated quantum phases, particularly in reduced dimensions [64, 65, 66]. Local control of the density or orientation of polar molecules may allow one to create or simulate charge density waves or (pseudo-) spin-density waves, or even a random, glassy system.

Another and rapidly emerging area is ultra-cold chemistry [39, 40, 67, 68, 69]. In the regime where the de Broglie wavelength of the molecule becomes comparable, or even orders of magnitude larger than the molecule itself, the classical notion of a reaction coordinate is at odds with quantum mechanics. Coherent population transfer methods (e.g., STIRAP) have already produced a gas of fermionic $^{40}K^{87}Rb$ molecules with a temperature of a few hundred nano-Kelvin at a phase-space density getting very close to that needed to achieve degeneracy. In this regime one expects entirely new phenomena in looking at chemical reactions at ultralow energies [70]. In addition to just potential energy surfaces being relevant, the ultracold reaction rate is controlled by quantum statistics and quantum coherence effects, including tunneling. Moreover, quantum statistics only allows certain collisions, and tunneling leads to threshold laws. In this regime, quantum control also can take on new meaning, with resonance-mediated reactions and collective many-body effects becoming prominent. Polar molecules also provide a handle for control through application of relatively weak electric fields.

Among applications, cold molecules hold great promise for improving sensors of all sorts. So far, all demonstrated matter-wave interferometric sensors utilize atoms. Using dipolar molecules instead of atoms could lead to orders-of-magnitude improvements in the sensitivity of such sensors [71]. The advantage comes from the ability to guide the molecules with modest electric field gradients, thereby creating steeper confining potentials. Steeper potential gradients allow one to load waveguide structures more efficiently; this leads to a larger number of particles and better statistical sensitivity. Also because of the steeper potentials, molecules can be kept further away from the wires and guiding surfaces, thus improving performance, as these are major sources of decoherence. Moreover, because sensitivity (e.g., to rotation) is proportional to the area enclosed by an interferometer, a large-area storage ring (with electrostatic guiding) [72] becomes a possible interferometer.

Bi-alkali molecules produced thus far include RbCs [38, 73], KRb [74, 75, 76, 77], NaCs [78] and LiCs [79, 80]. Photoassociation and magnetoassociation can be combined with STIRAP to allow coherent population transfer into a low-energy bound state, and significantly enhances the rate of molecular production [74, 77, 81, 82].

The strength of the dipole moment in these diatomics is critical to their possible uses, ranging from their potential as qubits, their use in precision measurement experiments, the nature of their quantum phase transitions, as well as simply their ability to be manipulated.

This motivates our study, which is focused on applying first principles methods to two polar diatomics, one a bi-alkali, and one an alkali-alkali earth. Specifically, we use quantum Monte Carlo in combination with basis set methods to look carefully at KRb and LiSr, both in their ground states ($X^1\Sigma^+$ and $X^2\Sigma^+$, respectively).

4.2 Methods

Since the bonds in all these systems are weak, and consist of a mixture of covalent and van der Waals bonding mechanisms, first principles calculations tend to be quite involved and highly sensitive to basis sets, degree of correlation included, nature of approximations, and other factors. Density Functional Theory (DFT) for van der Waals systems is often less than satisfactory, typically showing a varying degree of overbinding without any obvious systematic trends. Several new DFT functionals with corrections for dispersion interactions have been proposed very recently; see for example, [83, 84]. However, thorough testing and benchmarking of these new DFT approaches is necessary before they can be reliably applied across a variety of systems. Dispersion interactions result from transient-induced polarizations between the interacting constituents, and are therefore subtle many-body effects which are difficult to capture in the functionals framework.

On the other hand, the powerful arsenal of basis-set correlated methods, such as Configuration Interaction (CI), have their own challenges for systems such as these. For example, the dipole moment can be very sensitive to the basis set, and convergence for weakly bonded systems can be very slow [85]. In addition, the dipole moment appears highly sensitive to the level of correlation used, especially for problems which require multi-reference treatment. This is well known, but a systematic study can be found in Ref. [85]. Our CI results here also demonstrate this. It is therefore important to explore

other types of methods to understand the impact of many-body effects more thoroughly. That is why we need the highly accurate alternative approach, which is quantum Monte Carlo (QMC) [86, 21]. QMC is very attractive since it is in principle exact; in practice due to approximations it has a residual weak sensitivity to the size of basis sets, and it captures the correlations at a level of 90-95% [86, 15, 87]. That is something that is quite difficult to achieve by correlated methods based on expansions in basis sets.

In fact, there are previous studies of molecular dipole moments calculated by QMC methods for a few molecular systems. Schautz and colleagues carried out QMC study of the carbon monoxide molecule and obtained a very good estimate for this non-trivial problem in which correlation reverses the sign of the dipole obtained at the Hartree-Fock level [88]. The dipole moment of the lithium hydride molecule was computed by fixed-node diffusion Monte Carlo (DMC) in a good agreement with experiment [89]. Recently, several transition metal monoxide molecules have also been calculated by QMC methods. Besides the binding energies and equilibrium bond lengths, the dipole moments were also calculated, all with reasonably good agreement to experiment [90]. In addition, transition dipole moments such as for the Li atom were calculated by QMC approaches some time ago [91].

For our calculations here, we employ the two most common QMC methods, variational and fixed-node diffusion Monte Carlo (VMC and DMC). The details of these two methods were discussed in Chapter 2.

The choice of the trialfunction is critical for the calculations because the quality of the trial wave function plays two roles. First, highly accurate trial functions lead to smaller statistical fluctutations and faster QMC sampling and convergence. Second, better trial functions typically go hand-in-hand with improvements of the nodal hypersurfaces, and hence result in smaller fixed-node bias. Clearly, we want to use the best trial function we can. Traditional methods, including correlated approaches, provide a convenient starting point in this respect. In addition, however, quantum Monte Carlo can employ explicitly correlated wave functions as well, without prohibitive computational cost. This enables us to reach beyond the limits of traditional approaches, even before beginning to project out the true (fixed node) ground state.

In our calculations, we have used the Slater-Jastrow trial wave functions which embody both the traditional computational chemistry starting point and explicit correlation in the two factors. The explicit form of the Slater-Jastrow trial wave function was dis-

cussed in Chapter 3. The orbitals for our Slater determinats are obtained from Hartree-Fock calculations, and our CI expansion is calculated using the GAMESS package. The Jastrow variational parameters and the CI expansion coefficients are optimized in VMC by minimizing a linear combination of the variational energy and the variance of the local energy, $[H\Psi_T]/\Psi_T$. The Jastrow variational parameters are optimized in VMC by minimizing a linear combination of the variational energy and the variance of the local energy, $[H\Psi_T]/\Psi_T$. The optimized trial function is then employed in our fixed-node DMC runs.

In diffusion Monte Carlo, expectation values of quantities which commute with the Hamiltonian are exact [86]; however, expectation values of non-commuting operators, $\hat{O}$, are estimated by mixed estimators, i.e., with the distribution $\Psi_T\Phi$, instead of Φ^2. One can correct for this in a number of ways [91, 92]. A commonly used approximation is

$$< \Phi|\hat{O}|\Phi > \approx 2 < \Phi|\hat{O}|\Psi_T > - < \Psi_T|\hat{O}|\Psi_T > .$$

Exact sampling of Φ^2, while computationally more complex and significantly more costly, is also possible [93, 15, 92]. Accuracy gained from exact sampling methods is limited by any approximations involved, both fundamental and technical. These include, e.g., the fixed-node restriction and the localization approximation in treating the nonlocal effective core potentials (see below). Since the optimization of the wave function is done stochastically on finite samples, it is difficult to eliminate a possible optimization bias reflecting differences in the localization error with different Jastrows. This is especially so in the case of weakly bonded systems such as LiSr (with bonding of the order of 0.006 a.u.). Therefore, with any likely improvements to be invisible, the results below are based on the mixed estimator, which provides a good balance between accuracy and efficiency of the calculations.

4.3 Calculational Details

The pseudopotential we ultimately used for the bulk of the calculations here removes small cores for the elements beyond the first row, with the resulting configurations of valence electrons in KRb given as K($3s^2 3p^6 4s^1$) and Rb($4s^2 4p^6 5s^1$). Gaussian basis sets used in the calculations were gradually increased in size until we observed saturation to about $\approx 0.001\ E_h$ at the self-consistent level; basis sets of 25s22p13d/[4s4p3d] were used for

both K and Rb. In the LiSr calculation, we kept all Li atom electrons in the valence space, while for Sr we eliminated 18 core electrons, so that the valence configuration became Sr($4s^2 4p^6 5s^2$). Quite extensive basis sets for Sr and Li, with sizes 7s6p5d1f/[7s4p2d1f] and 27s7p1d/[7s5p1d], provided saturated HF energies at the level of 0.001 E_h or better.

A proper description of the dipole moments requires high-quality correlated methods. This is true in particular for LiSr, which does not exhibit any binding at the HF level. (KRb, on the other hand, has a single σ bond, so that using a single-reference wave function, and CI with single and double excitations, was deemed to be appropriate in this case.) For LiSr we explored several correlation levels and sizes of active occupied and virtual spaces. In particular, we carried out singles and doubles (SD) with 5 active occupied and 70 virtual orbitals; SD and triples (SDT) with 5 active occupied and 55 virtual orbitals; SDT and quadruples (SDTQ) with 3 active occupied and 48 virtual orbitals. Excitations beyond doubles proved to be important for the LiSr dipole moment. Even with quadruples included, it remained difficult to be certain that the correlations were adequately described.

This is where QMC methods provide an important alternative, and a powerful option to probe for correlations beyond what is practical with CI. For QMC, the CI(SD) wave functions only serve as a starting point—as trial functions. They are truncated to only their most significant terms by imposing a cutoff on the weights of the configuration state functions. Different cutoff values from 0.05 to 0.01 were tested. It turned out that the optimal cutoff value is around 0.03-0.05. (Values smaller than this no longer decreased the QMC energy within the detectability of the error bars.) Using a cutoff of 0.05 for LiSr and 0.03 for KRb, we were able to improve the nodal surfaces (evidenced by lower QMC energy) with reasonable efficiency and statistical consistency. For completeness and to gain more confidence of our DMC results, we performed DMC calculations with a very large CI determinant expansion. The trial wavefunction was taken from the CI(SDT) expansion with a smaller cutoff value 0.025. The Jastrow factor has the simplest form including only cusp conditions. No optimization of Jastrow factor was used so that there are no optimization bias. We compare the results of several DMC calculations.

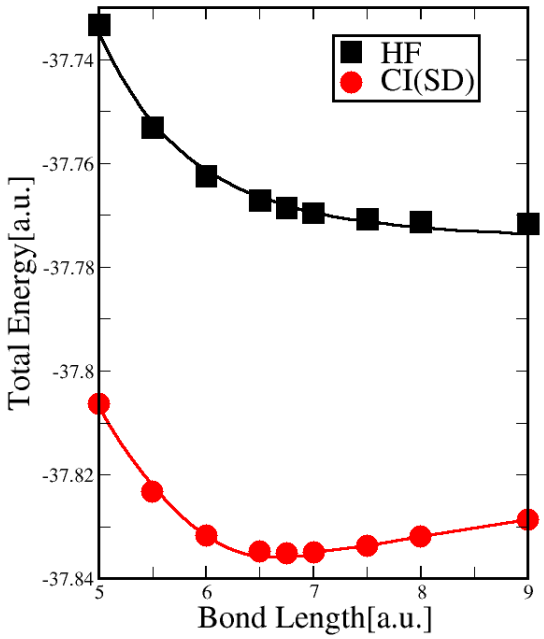

Figure 4.1: The binding energies of LiSr as a function of the bond length with HF and CI(SD) methods.

4.4 Results and Discussion

The energies that we obtain as a function of bond length are shown in for LiSr in Figure 4.1 and Figure 4.3, and for KRb in Figure 4.2 and Figure 4.4. We first compare our HF and CI(SD) calculations and then comprare the best CI calculation with the diffusion QMC (DMC) results. For LiSr, we show the QMC results with two different trial functions, DMC1 and DMC2. We see that the QMC correlation energy for both molecules is considerably better than that found by the CI method, and provides significant overall improvement of the potential energy surfaces.

As previously discussed, the dipole moment of LiSr is particularly sensitive—to pretty much everything (basis set size, level of correlation in the theory, nodes in the QMC trial function, etc.). We show in Figure 4.5 this sensitivity with respect to amount of correlation in the theory. As is clear, CI SDT, which is often sufficient, has not converged here. It would be difficult to know if even CI SDTQ is sufficient were it not for our QMC

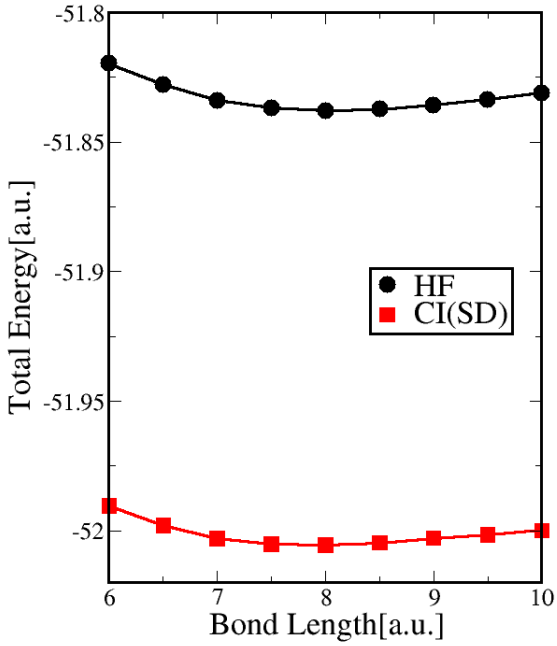

Figure 4.2: The binding energies of KRb as a function of the bond length with HF and CI(SD) mehtods

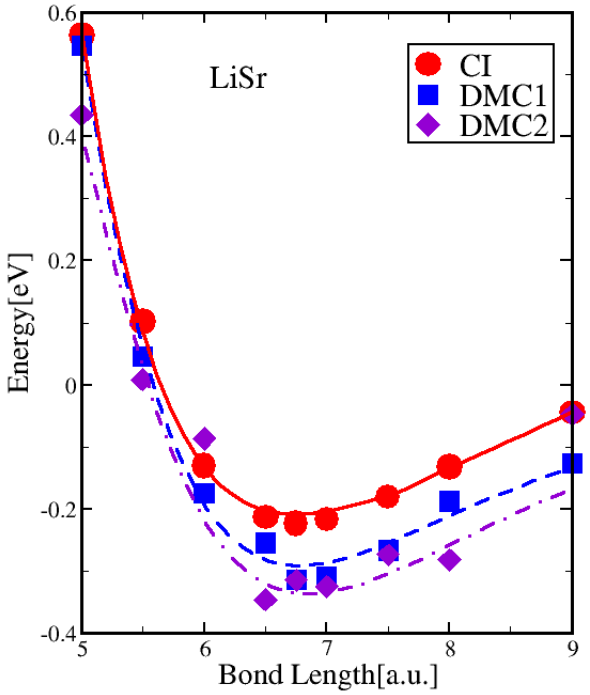

Figure 4.3: The binding energies of LiSr as a function of the bond length with different methods. Two optimized Jastrow DMC trial wave functions are used for LiSr. DMC1 denotes a single Slater determinant, while DMC2 is multi-reference with a cutoff weight of 0.05. The statistical error bars of the QMC results are approximately of the symbol size. The CI curves correspond to CI/SD

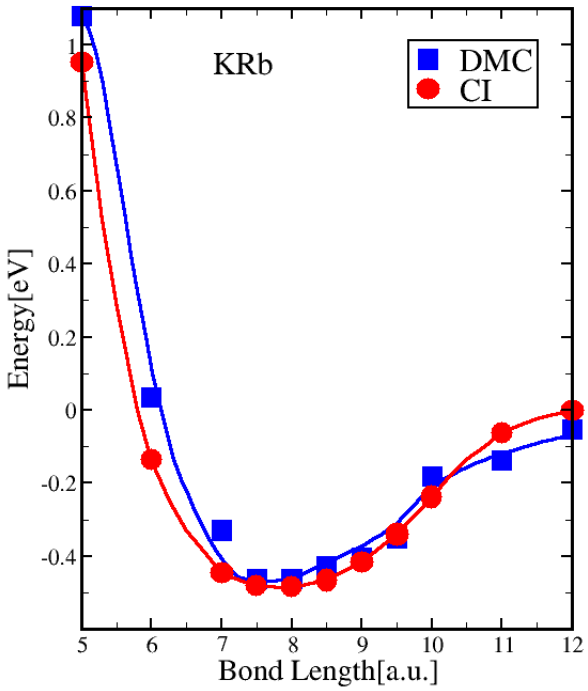

Figure 4.4: The binding energies of KRb as a function of the bond length with different methods. The DMC trial wave function is multi-reference, with a cutoff weight of 0.03, and again an optimized Jastrow. Here and in subsequent figures, the statistical error bars of the QMC results are approximately of the symbol size. The CI curves correspond to CI/SD.

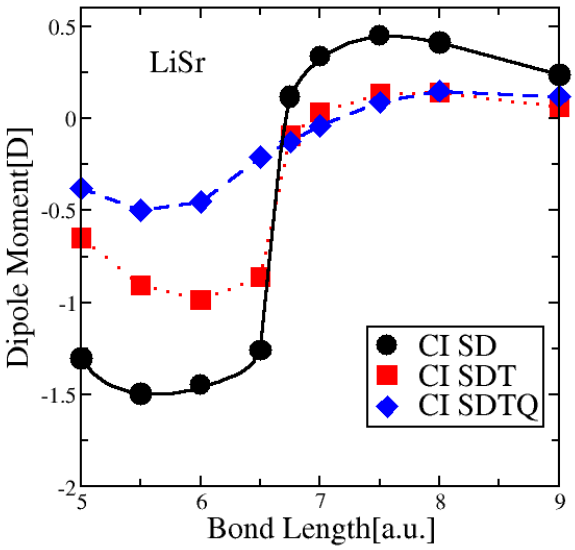

Figure 4.5: The dipole moment of LiSr calculated by the CI method with SD, SDT and SDTQ levels of correlation, as described in the text.

calculation. In Figure 4.6 we compare the CI SDTQ result for the dipole moment with our two QMC calculations with different trial functions. While the two QMC results are not identical, they vary immensely less than the different CI calculations. Moreover, the QMC with refined nodes from the multi-reference trial function gives almost the same results as CI SDTQ. This provides additional confidence in this result. This can be seen better in the expanded Figure 4.8, where we compare the dipole moment of LiSr as computed from only correlated methods, and compare our results against a recent result in the literature as well. Since the dipole moment does not commute with the Hamiltonian, by optimizing the Jastrow factor to obtian a lower variational energy, it is easy to generate some optimiation bias on the dipole moment. To assess how large the optimization bias is, in Figure 4.9, we compare the DMC result using large CI expansion but no optimization with the ones with optimized Jastrow. The two DMC results show very good agreement overall, indicating the optimization bias is very small in this case.

Figure 4.7 shows the dipole moment results for KRb. The bonding behavior, as well as the dipole moment of KRb, is rather straightforward to understand. There is a single

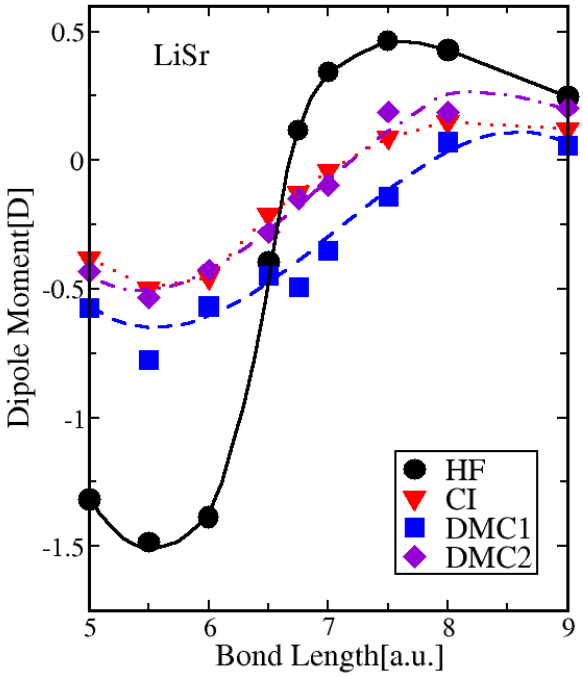

Figure 4.6: The dipole moment of LiSr as a function of internuclear distance, as obtained in DMC, HF and CI methods (CI/SDTQ). The trial wave functions, DMC, DMC1 and DMC2 are the same as described in Figure 4.3.

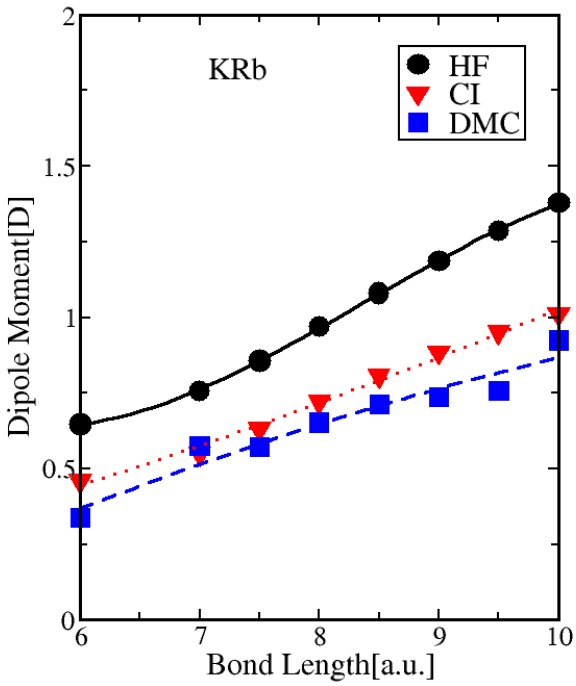

Figure 4.7: The dipole moment of KRb as a function of internuclear distance, as obtained in DMC, HF and CI methods (CI/SD). The trial wave functions, DMC, is the same as described in Figure 4.4

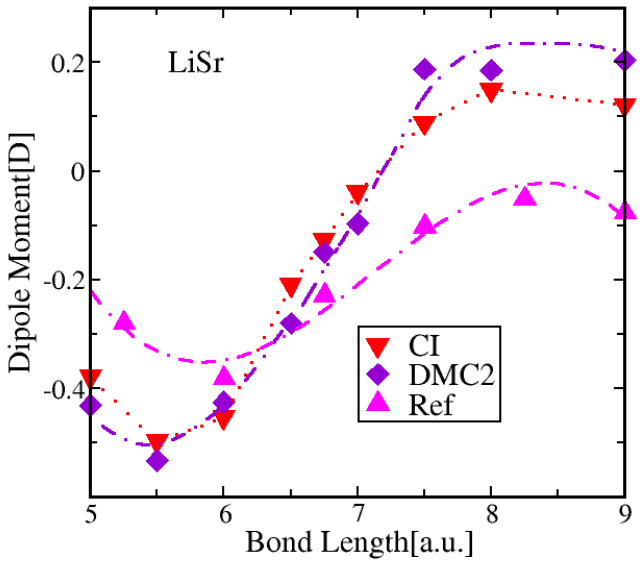

Figure 4.8: Dipole moment for LiSr from our best CI and QMC calculations compared with results from Ref.[94]. The pronounced nonlinearity allows switching of the value of d by state selection to a vibrational state near dissociation. See text.

σ bond formed from the K(4s) and Rb(5s) atomic states. The dipole moment monotonically increases with the bond length, corresponding to growth of the distance between the effective charges, and also some possible growing enhancement of the electron density in the region of the K atom. This trend is observed already at the HF level, indicating that it is driven by the one-particle self-consistent balance of the energy contributions. However, quantitatively, the HF dipole moment can be seen to be too large. This is because electron correlation decreases the charge polarization over the whole range of calculated interatomic distances. This can be equivalently understood as a result of the well-known HF bias towards larger ionicity of the bonds. The absence of correlation tends to make the exchange more prominent, which drives the electronic structure towards stronger charge polarization. It is interesting to observe that the dipole moments computed from both of our correlated methods appear to be rather close, although the degree of description of the many-body effects is very different, with the CI recovering only about 30% of the correlation energy compared to $\approx 90\%$ in QMC.

On the other hand, LiSr has a more complicated electronic structure due to the

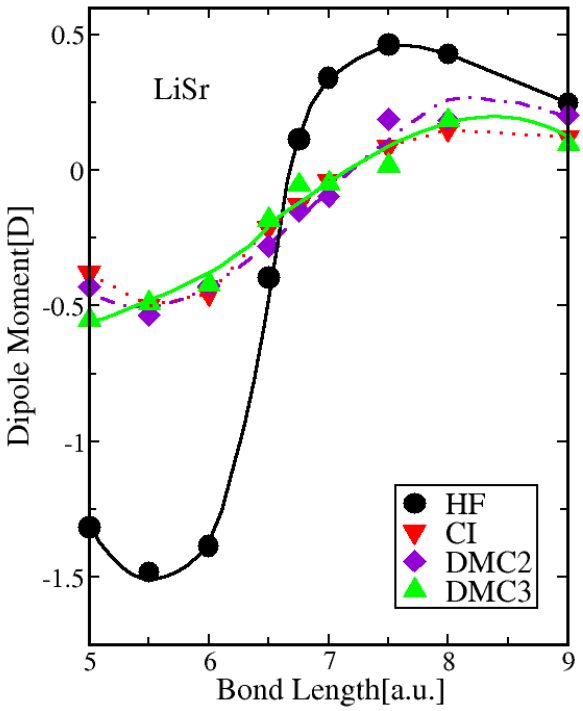

Figure 4.9: Dipole moment for LiSr from our HF, CI(SD), and different DMC methods. The trial wavefunction for DMC2 is the same as Figure 4.3. DMC3 uses CI(SDT) determinant expansion with cutoff 0.025 and no optimization of Jastrow.

presence of the two ($5s^2$) electrons in the outermost shell of the Sr atom, with the resulting open-shell $X^2\Sigma^+$ molecular ground state. The last valence electron occupies an anti-bonding σ level which is formed by the combination of Li(2s) and Sr(5s) atomic states. The overall bond is thus weaker, and this accounts for the lack of binding seen in the HF method in Figure 4.1. Even after including $\approx 90\%$ of the correlation energy via QMC, as we have done here, the binding is small, only of the order of ≈ 0.25 eV. Even more interestingly, the dipole moment changes sign as a function of separation! This behavior is visible in all three methods, implying that one can trace its origin to changes at the single-particle level. Indeed, by plotting the isosurfaces of the Hartree-Fock HOMO (antibonding σ) and LUMO orbitals (see Figure 4.10 and Figure 4.11) we see that as the bond length increases, the nature of the orbitals change very significantly. These changes lead to sizeable restructuring of the electron density already at the HF level, with the resulting sign change of the dipole moment as plotted in Figure 4.5,Figure 4.6, Figure 4.8, Figure 4.9.

The quality of the trial wave function and its nodal surface has a significantly more pronounced impact on the LiSr dipole moment result than on the KRb moment, and one cannot rely on a single configuration for the former. The inclusion of the most important configurations from the CI expansion improves the accuracy of the DMC estimations quite significantly. It is interesting, if somewhat unexpected, that our QMC results obtained with our best trial function, apart from the statistical noise, agree remarkably well with our most accurate CI calculations. This consistency is reassuring since the employed methods are largely independent, and capture the electron correlations in very different manners. The need to employ triples and quadruples in CI indicates that correlations beyond multi-reference configurations based on singles/doubles excitations are important. On the QMC side, the single configuration trial function already captures the main correction to the HF dipole moment, due to the correlations, as can be seen in Figure 4.6. However, for high accuracy results, the complicated orbital restructuring of the LiSr molecule clearly requires a multi-reference treatment which includes the dominant non-dynamical effects explicitly, thereby improving the nodal surface. This accounts for the difference seen in Figure 4.6 between DMC1 and DMC2. This observation is very much in line with previous QMC calculations of other systems with significant multi-reference or near-degeneracy effects.

The spectroscopic constants for these molecules have been computed, and are shown

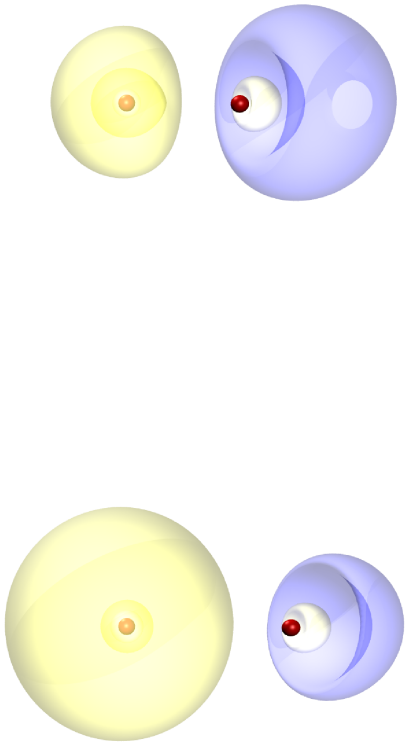

Figure 4.10: Isosufaces of the positive and negative lobes of the highest occupied molecular orbital (HOMO) of LiSr, for interatomic distances 5 a.u. (top) and 7.5 a.u. (bottom).

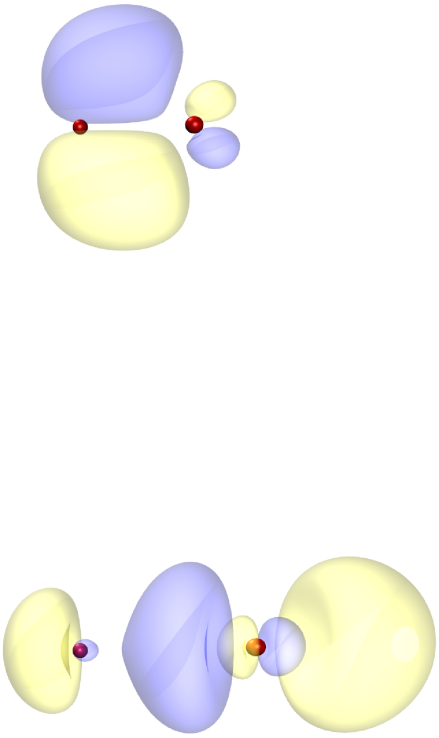

Figure 4.11: Isosurfaces of the positive and negative lobes of the lowest unoccupied molecular orbital (LUMO) of LiSr, for interatomic distances 5 a.u. (top) and 7.5 a.u. (bottom).

Table 4.1: Spectroscopic constants for the LiSr molecule: bond length R_e, potential well depth D_e, harmonic constant ω_e, and the ground state averaged dipole moment $\langle d \rangle$. The values in the row labeled DMC are calculated from the DMC2 data displayed in Figure 4.3 and Figure 4.6 Numbers in parentheses give the statistical uncertainty in the last significant figure.

	R_e (a.u.)	D_e (cm^{-1})	$\omega_e(cm^{-1})$	$\langle d \rangle$ (D)
DMC	6.80(5)	$2.7(3) \times 10^3$	167(7)	- 0.14(2)
Ref [95]	6.57	2.587×10^3	185	- 0.340
Ref [94]	6.71	2.401×10^3	184	- 0.244

in Tables 4.1 and 4.2, together with the most recently published calculations which used much less fully converged coupled cluster and CI methods. For LiSr we calculated the dipole moment also in the first vibrationally excited state, and we found that it is basically the same as in the ground state within the given error bar. This can be understood from the fact that our results show the dipole being approximately linear with bond length over the range $\approx$ 6.0 - 7.7 a.u., and this interval covers the spatial range of both the ground and the first vibrational states. Nevertheless, the nonlinearity of the LiSr dipole moment may be exploitable for vibrational dependent control of d. In particular, it appears that the highest bound vibrational states might have considerably smaller dipole moments, making it possible to largely switch off (and on) the dipole moment through tailored excitations.

The overall agreement between the calculational approaches, considering how small the quantities are, is very reasonable. The differences reflect the systematic biases of the approaches used, and clearly illustrate the challenge of describing weakly-bonded systems with high accuracy. Since these calculations reflect the state of the art of these computational methodologies, this shows what is currently feasible. It is clearly desireable to increase the accuracy further, in order to decrease the uncertainties. This can be accomplished with further development of the methods; in particular, for QMC approaches one would need much higher sensitivity in quantities which do not commute with the Hamiltonian so that the optimization of much larger multi-reference wave functions would be feasible. This is a promising direction for future explorations.

Table 4.2: Spectroscopic constants for the KRb molecule, defined as in 4.1. The values in the row labeled DMC correspond to the data displayed in Figure 4.4 and Figure 4.7.

	R_e (a.u.)	D_e (cm^{-1})	$\omega_e(cm^{-1})$	$\langle d \rangle$(D)
DMC	7.58(5)	$3.8(3) \times 10^3$	77(2)	0.58(2)
Ref [95]	7.63	4.199×10^3	76	0.615

4.5 Conclusion

We have carried out careful calculations of the dipole moments and potential energy curves for the KRb and LiSr molecules using Hartree-Fock, Configuration Interaction, and fixed-node diffusion quantum Monte Carlo methods. The calculations show significant effects of the electronic correlations on the magnitude of the dipole moments over the whole range of the investigated interatomic distances.

Although single determinant QMC already captures most of the correlation energy, the low-lying excitations need to be explicitly included into the trial function for the LiSr dimer, particularly for accurate computation of the dipole moment. The need is clear from the fact that the weights of some configurations beyond the reference HF configuration are significant. Comparisons of CI(SD) with CI(SDTQ) clearly illustrates the multi-reference nature of the ground state. The accuracy of our results is supported by the consistency between our most extensively correlated CI approach and our methodologically quite distinct QMC results.

In addition, the high percentage of the correlation energy recovered by QMC ($\approx 90\%$) captures most of the dynamical correlation, and therefore provides much more extensive insight into how correlations affect sensitive molecular properties.

Chapter 5

Dissociation Energy Study of Thorium Compounds

5.1 Motivation

Thorium element is a radioactive element in the last row of the periodic table with atomic number 90. Like uranium, thorium is also an important nuclei material which can be used as a nuclear fuel in reactors. Compared to uranium, thorium has its own advantages to be favorable because of its greater safety benefits. Due to the great interest, it is worthwhile to study the electronic structure properties using first-principles methods. However, the presence of the strongly localized f electrons adds tremendous difficulty to the calculations. Up to now, fairly number of the chemical compounds consisting of elements from the last row have been studied. For example, calculations on UF_n and Cl_n were done using DFT with hybrid functionals.[96] [97]. The bond dissociation energies they obtained agree very well to experiment data. To our best of knowledge, few calculations on thorium have been done using QMC. In this Chapter, we present the benchmark QMC calculations on several types of thorium compounds, and we compare the QMC result with both other theoretical calculations and experiment results.

5.2 Calculation Details

The molecular compounds we want to investigate are $ThCl_n$ and $ThBr_n$, where n=1,2,3,4. Th atoms has 90 electrons totally, so it is necessary to apply pseudopotential to elimi-

nate core electrons. For Th, we use large core Stuttgart pseudopotential removing 78 core electrons which include all d and f electrons and keep 12 valence electrons in the calculation. Gaussian basis sets are used in the calculation. Basis sets contraction 8s8p6d5f2g/[5s5p4d5f2g] is used for Th. For Cl and Br, we use energy-consistent pseudopotentials. Both have 7 electrons in the valance space. Contracted basis sets 11s11p2d1f/[3s3p2d1f] are used for both Cl and Br. We apply DFT method to optimize the geometry of each molecule using NWCHEM and GAMESS package. The exchange correlation functional we use is B3LYP, which proves to be successful for molecule calculations. It is a hybrid functional which mixes the exact Hartree Fock exchange and LDA and GGA functionals(see section 1.4.2). We then take the optimized geometry and run a sequence of QMC calculations(VMC, fixed-node DMC, see Chapter 2) and apply Slater-Jastrow wavefunction(see Chapter 3) to calculate the total energies, then evaluate the dissociation energies.

5.3 Results

We did the calculation for $ThCl_n$ and $ThBr_n$. The geometries of all the molecules is in Figure 5.1.

In the case of molecules with Cl element, we compare the QMC energies using different initial orbitals: Hartree Fock(HF) and DFT. The optimized bond lengths, HF energies, plus the DMC energies with error bars using both HF orbitals and DFT orbitals are listed in Table 5.1. In Table 5.2, we list the averaged binding energies in kcal/mol of $ThCl_n$ using the data in Table 5.1. The averaged binding energy is defined as $\Delta E = (E_{\mathrm{ThCl}_n} - E_{\mathrm{Th}} - n * E_{\mathrm{Cl}})/n$

Table 5.1: List of bond length[Angs], symmetry group, and total energies[a.u.] of each molecule

	$ThCl_4$	$ThCl_3$	$ThCl_2$	ThCl	Th	Cl
d	2.58	2.58	2.53	2.575	n/a	n/a
Symmetry	T_d	c_{3h}	c_{2v}	$c_{\infty v}$	n/a	n/a
HF	-94.845	-79.750	-65.034	-50.126	-35.233	-14.740
DMC(HF)	-96.264(4)	-81.100(4)	-65.933(2)	-50.765(2)	-35.610(2)	-14.967(1)
DMC(DFT)	-96.271(4)	-81.106(4)	-65.937(2)	-50.768(2)	-35.609(2)	-14.967(1)

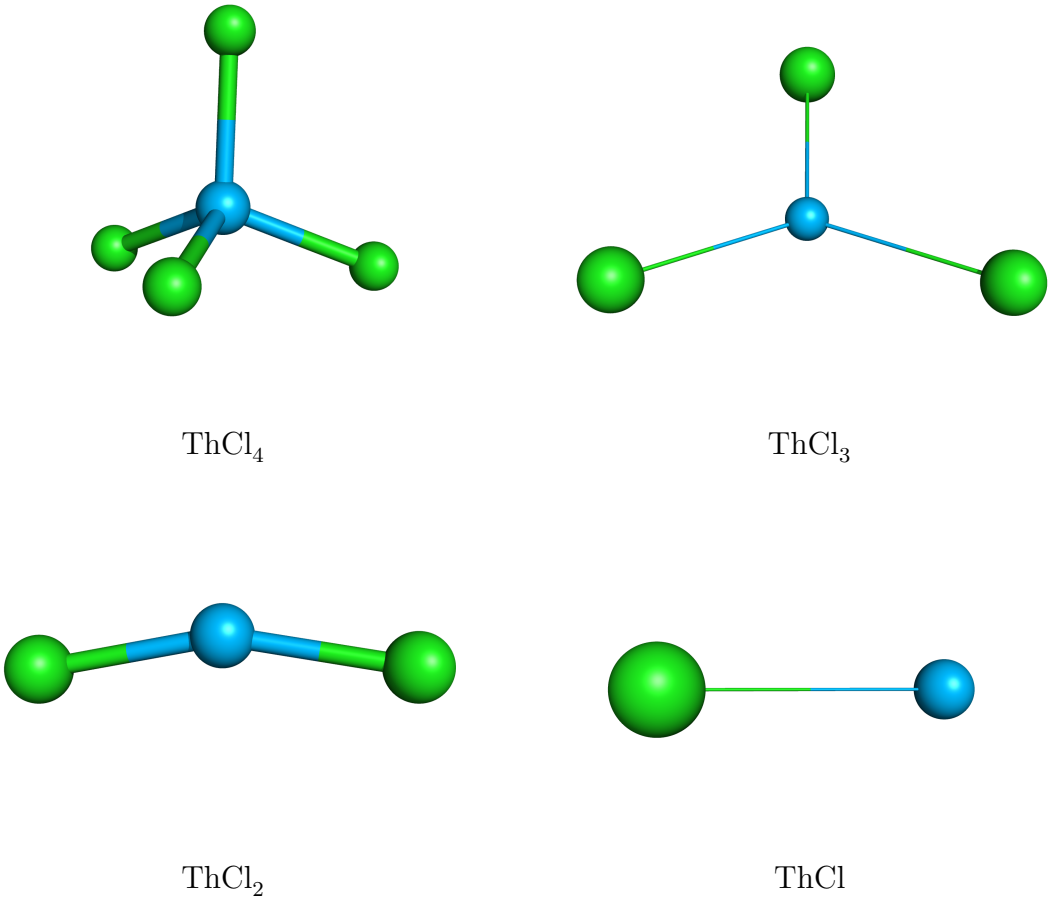

Figure 5.1: Geometries of each molecule

Table 5.2: List of averaged binding energies(kcal/mol) of $ThCl_n$ using DMC method

	$ThCl_4$	$ThCl_3$	$ThCl_2$	ThCl
DMC(HF)	123.05	122.9	122.3	117.9
DMC(DFT)	124.32	124.6	123.5	120.229

The electron configuration of Th atoms is [Rh] $6d^2 7s^2$. The 4 valance electrons $6d^2 7s^2$ form covalent bonds with at most 4 halogens, which are either Cl or Br. With different numbers of halogens and different orbital hybridization, the binding energies display a systematic change. In the Table 5.1, we see the QMC total energies calculated using DFT orbitals are all lower than those starting with HF orbitals. Although the energy difference are small, in QMC calculations DFT(B3LYP) orbitals give a better variational upper bound of the total energies and higher binding energies than HF orbitals. Comparing the averaged binding energies in Table 5.2, from ThCl to $ThCl_4$, we see a roughly increasing trend in averaged binding energies as in number of halogens increase, with ThCl being the lowest. This phenomenon is visible in QMC calculations using both HF and DFT orbitals.

Among those 4 molecules, ThCl is the simplest one. It is an open-shell structure with only sigma bond between the 7s electron of Th and 3p electron of Cl. With more halogens, the higher symmetry of geometry and the more complexity of the orbital hybridization (p and d orbitals in Th) systematically increase the value of the averaged binding energies. In order to compare our data with experimental data, we transform the data of averaged binding energies in Table 5.2 to bond dissociation energies(BDE). We define bond dissociation energy as

$$BDE_n = E_{ThX_n} - E_{ThX_{n-1}} - E_X \tag{5.1}$$

We plot the BDEs using our QMC result(using DFT orbitals) and the DFT reference together with the experimental data[98] in Figure 5.2. We clearly see both QMC and DFT give similar prediction, although different basis sets and pseudopotentials are used (DFT references shown here used a much poorer basis set than QMC). The differences between the QMC results and DFT references are quite reasonable. They are less than 5 kcal/mol which is within the resolution of the method. Since the correlation effects are treated quite differently in QMC and DFT, we conclude the correlation effect does not change the dissociation behavior. We also compare the data with experiments. Both DFT and QMC calculations agree well with the experimental data.

Cl and Br are both in the 7th column in the periodic table except that Br has one more electron shell than Cl. Therefore theoretically we expect they may have many similar chemical properties. With the similar core-valence partition, the prediction should be roughly the same. We want to test whether the additional electron shell in Br changes

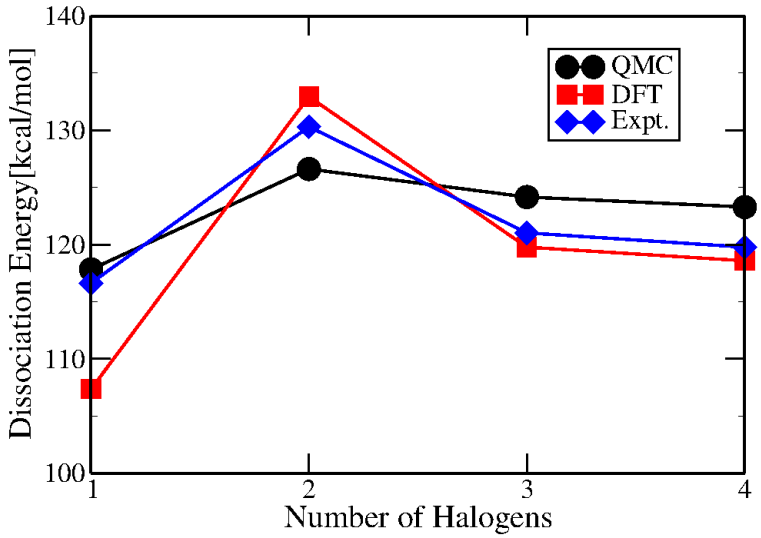

Figure 5.2: Bond dissociation energies for $ThCl_n$ molecules. QMC data uses DFT orbitals

the bonding behavior with Th, making bonding properties apart from Cl. We perform the calculations following the same procedure as what we did with Cl. The total energies of each molecule and averaged binding energies are listed in Table 5.3 and Table 5.4. We see the differences in total energies between HF trial wavefunction and DFT trial wavefunction. The differences appears larger for open shell molecules, as for $ThBr_3$ and ThBr, which indicates the larger bias from HF. It is verified in this calculation that the trend of the averaged bonding energies is similar to Cl cases in Table 5.2. ThBr has the lowest value. The values of all other molecules are above ThBr. It is also noticed that in terms of averaged binding energies the difference between ThBr and other 3 molecules ($ThBr_2$, $ThBr_3$, $ThBr_4$) is more or less 3 kcal/mol. And we have the same difference in the case with Cl. This is a strong evidence that Cl and Br behave similarly.

We again plot the BDEs in Figure 5.3 to compare our result with reference DFT calculation and experimental data[99]. There is a difference between QMC result and DFT result for all molecules except for ThBr. It is reasonable to claim the electron correlation accounts for this difference in a relatively small fraction. However, apart from $ThCl_n$ case, we have a big discrepancy with experimental data. The calculated values and the

Table 5.3: List of bond length[Angs] of each $ThBr_n$, symmetry, and total energies[a.u.]

	$ThBr_4$	$ThBr_3$	$ThBr_2$	ThBr	Br
d	2.68	2.71	2.74	2.74	
Symmetry	T_d	c_{3h}	$c2_v$	$c_{\infty v}$	
DMC(HF)	-89.599(4)	-76.090(3)	-62.599(3)	-49.085(2)	-13.320(1)
DMC(DFT)	-89.601(4)	76.104(3)	-62.602(3)	-49.096(3)	-13.320(1)

Table 5.4: List of averaged binding energies(kcal/mol) of $ThBr_n$, using DMC method

	$ThBr_4$	$ThBr_3$	$ThBr_2$	ThBr
DMC(HF)	111.23	109.20	109.93	97.65
DMC(DFT)	107.27	107.6	107.05	104.3

experimental values differ both qualitatively and quantitatively. The biggest discrepancy appears in the dissociation channel ThBr$-$ > Th + Br. In Ref [99], it is claimed that $ThBr_4$ has a distorted tetrahedron structure therefore its behavior is different. However, this doesn't show up in our calculation. Other DFT calculation does not show that either. The bias of the calculation can possibly come from the basis, ECP and method. QMC is a computational method which has the least dependence on the choice of the basis sets. The ECP is also tested in many other chemical environments. Since the same Th ECP has been used in ThCln cases and no big discrepancies occur, we test the ECP of Br by calculating the binding energy of HBr. The experimental shows the binding energy of HBr is about 3.8eV while both DFT and DMC results using this ECP are 3.9 - 4.0eV. So the ECP can not account for such larger errors in $ThBr_n$ case. Neither DFT nor QMC could generate such a big bias(as large as 20kcal/mol).

5.4 Conclusion

We used different ab initio methods(HF, DFT, QMC) and calculated the bond length, binding energies of $ThCl_n$ and $ThBr_n$ molecules. We compared our result to the experi-

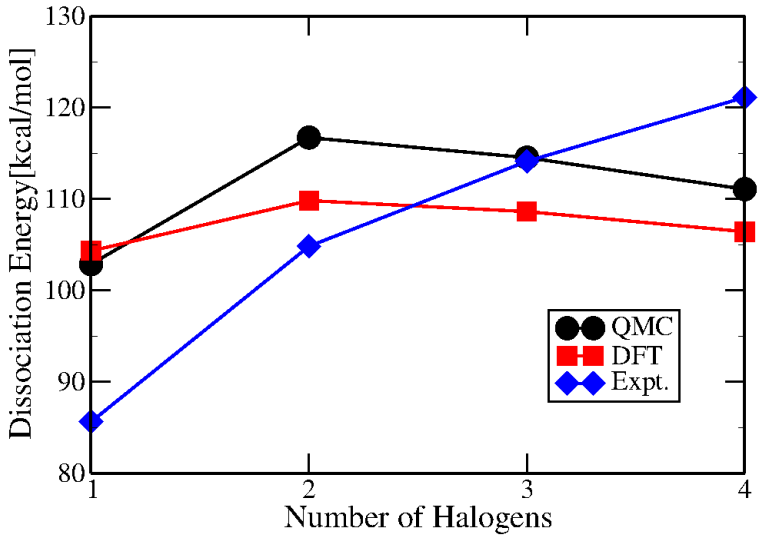

Figure 5.3: Bond dissociation energies for $ThBr_n$ molecules. QMC data uses DFT orbitals

mental data. In $ThCl_n$ cases, the QMC results agree with DFT results and experimental data. However, in $ThBr_n$ cases big differences are shown between the calculated results and experimental data, although there is an agreement between DFT results and QMC results.

Chapter 6

Spin-Orbit Interaction in Quantum Monte Carlo

Spin-orbit interaction is an interaction of particle's spin with its motion itself. In many systems, the spin-orbit interaction effect plays an important role and a careful treatment in the calculation is needed in order to correctly predict any properties of interest. Traditionally, in all QMC calculations(mentioned in the previous chapters) the spin of a particular electron is fixed by spin assignment. This treatment is not valid when we want to deal with the spin-orbit interaction since the operator of spin-orbit interaction is spin-dependent. Therefore, it becomes necessary to develop a novel approach which can handle spin business explicitly in QMC. In this chapter, we will introduce a new representation of electron spin, and then describe the QMC method with the treatment of spin, finally present an application of this method which is the calculation of two dimensional electron gas with Rashba interaction, discuss our result and make comparisons to other calculations in the literature.

6.1 Spin Basis and Wavefunction

In quantum mechanics, the spin operator is well defined using Pauli matrices,

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \qquad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{6.1}$$

The Pauli representation diagonalizes σ_z matrix with the eigenvalues ± 1 and each eigenvalue corresponds to one eigenvector, which is our spin basis:

$$\chi_z^{(+1)} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad \chi_z^{(-1)} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{6.2}$$

It is impossible to sample the spin in QMC if we use this basis because these are discrete functions and we are not able to make a smooth path of the random walk in the state space of the configuration. The way that we use to overcome this problem is to change the representation and introduce two continuous functions as our new spin basis set

$$\begin{aligned} \langle \omega | \chi^{\uparrow} \rangle &= \frac{1}{\sqrt{2\pi}} e^{i\omega} \\ \langle \omega | \chi^{\downarrow} \rangle &= \frac{1}{\sqrt{2\pi}} e^{-i\omega} \end{aligned} \tag{6.3}$$

Where $\omega \in [0, 2\pi)$. Here the $\uparrow$ and $\downarrow$ mean the z-component of the spin $s_z = 1/2$ and $s_z = -1/2$. Obviously, these two functions satisfy the orthogonality condition

$$\langle \chi^{\alpha} | \chi^{\beta} \rangle = \delta_{\alpha\beta} \tag{6.4}$$

An arbitrary quantum state which is a mixture of spin-up state and spin-down state can be written as

$$\chi(\omega) = a e^{i\omega} + b e^{-i\omega} \tag{6.5}$$

The spin operator, for example, is given as

$$s_z = 1/2(|\chi^{\uparrow}\rangle\langle\chi^{\uparrow}| - |\chi^{\downarrow}\rangle\langle\chi^{\downarrow}|) \tag{6.6}$$

When this acts on the state $\chi(\omega)$ it gives

$$\begin{aligned} \langle \omega | s_z | \chi \rangle &= \frac{1}{2} e^{i\omega} \int d\omega' (e^{i\omega'})^* \chi(\omega') - \frac{1}{2} e^{-i\omega} \int d\omega' (e^{-i\omega'})^* \chi(\omega') \\ &= \frac{1}{2} (a e^{i\omega} - b e^{-i\omega}) \end{aligned} \tag{6.7}$$

So it works as a projection operator. Note here the evaluation involves the integral of spin

from 0 to 2π, but the integral is actually trivial because what it does is exactly based on the orthogonality condition in Eq. 6.4. In traditional QMC calculations we build up the Slater determinant using one-electron orbitals, which we either calculate from other software packages or build it by hand. For the Hamiltonian which does not contain spin, we write the Slater determinant as product of two, one is for spin-up electrons and the other for spin-down electrons. However, this approach is not eligible when we take the spin-orbit interaction into account. Here, we are only allowed to use one Slater determinant and each one-electron orbital is a Dirac spinor. The one-electron orbital has the form

$$\psi(\mathbf{r},\omega) = a\phi^{\uparrow}(\mathbf{r})\frac{1}{\sqrt{2\pi}}e^{i\omega} + b\phi^{\downarrow}(\mathbf{r})\frac{1}{\sqrt{2\pi}}e^{-i\omega} \tag{6.8}$$

where $\phi^{\uparrow}$ is the spin-up component of the radial part, $\phi^{\downarrow}$ is the spin-down component of the radial part. The Slater determinant made up of those one-particle orbitals is

$$D(\mathbf{R},\mathbf{\Omega}) = det(\psi(\mathbf{r},\omega)) \tag{6.9}$$

6.2 Spin Dynamics in QMC

The Hamiltonian we want to solve is

$$H = T + V_C + V_{SO} \tag{6.10}$$

where T is the kinetic operator, V_C is the Coulomb operator. V_{SO} is the spin-orbit interaction term, which may have different forms for different systems. In our traditional QMC, the state space of the Markov chain is the position of all the configurations. We now need to increase one more degree of freedom and include spin variable ω in the configuration state space. In the Metropolis algorithm, the only additional step is that we propose a move for spin variable, and the rest of the algorithm works as usual. Therefore, our traditional VMC algorithm can be easily modified to include spin dynamics by slightly changing the Metropolis algorithm:

The algorithm works in 3 three steps:

(1)Generate random sample points(walkers) X_i and ω_i, where $X_i = \{X_i : 1 \leq i \leq n\}$, $\omega_i = \{\omega_i : 1 \leq i \leq n; 0 \leq \omega_i < 2\pi\}$. Here X_i denotes position of the ith configuration and ω_i denotes spin variable.

(2)For each sample point, make a trial move according to any convenient choice of distribution function $T(\delta x = X_i^{'} - X_i)$, $T(\delta\omega = \omega_i^{'} - \omega_i)$. The trial move of the spin variable is same as what we did for position variable, we choose T as a uniform distribution function or normal distribution function.
(3)For each move, calculate the acceptance probability.
Up to here, the Metropolis algorithm with spin is well-defined.

Since spin operator is a projection operator, an issue is how to evaluate the expectation value of spin-dependent operator. Similar to what we did when we evaluated the local energy of the pseudopotential operator, the local energy of any spin-dependent operator $\hat{O}(\mathbf{r}, s)$ is

$$E_{\hat{O}} = \frac{\hat{O}\Psi_T}{\Psi_T} \tag{6.11}$$

And the expectation value is the sum of the local energy over all the configurations. A key thing to notice here is our trial wavefunction Ψ_T is complex, so it is natural to have a complex local energy. But only the real part makes sense because any observable quantity has to be real. So when we average the energy of each walkers, we only keep the real part.

$$\langle\hat{O}\rangle = Re[\frac{1}{M}\sum_{i=1}^{M}\frac{\hat{O}\Psi_T(\mathbf{R}, \mathbf{\Omega})}{\Psi_T(\mathbf{R}, \mathbf{\Omega})}] \tag{6.12}$$

where M is the number of configurations.

Incorporating spin into DMC algorithm is more complex since we have to derive to Green's function for spin. We consider the basis we use for spin $e^{\pm i\omega}$ are the eigenstate for the Hamiltonian $H(\omega) = -\frac{1}{2}\frac{\partial^2}{\partial\omega^2} - \frac{1}{2}$

$$(-\frac{1}{2}\frac{\partial^2}{\partial\omega^2} - \frac{1}{2})e^{\pm i\omega} = (\frac{1}{2} - \frac{1}{2})e^{\pm i\omega} = 0 \tag{6.13}$$

We see here the first term of $H(\omega)$ works similarly as a kinetic operator. If we create an effective Hamiltonian

$$H_{eff}(\mathbf{R}, \mathbf{\Omega}) = H + H(\mathbf{\Omega}) \tag{6.14}$$

where H is the usual Hamiltonian, $\Omega = (\omega_1, \omega_2, ..., \omega_i...)$, $H(\mathbf{\Omega}) = \sum_i(-\frac{1}{2}\frac{\partial^2}{\partial\omega_i^2} - \frac{1}{2})$. In Eq. 6.13, we show that the additional term $H(\omega)$ contributes zero to the eigenenergy. The Green's function for spin is analogous to the one of a diffusion process that we derived

in Chapter 2.

$$G_S(\omega', \omega; \Delta t_\omega) = \frac{1}{\sqrt{2\pi\Delta t_\omega}} e^{-(\omega-\omega')^2/(2\Delta t_\omega)} \tag{6.15}$$

where t_ω means the time corresponding to the spin variable. The Green's function for the effective Hamiltonian H_{eff} is

$$\begin{aligned} G(\mathbf{R}', \mathbf{R}; \mathbf{\Omega}', \mathbf{\Omega}; \Delta t) &= \langle \mathbf{R}, \mathbf{\Omega} | e^{-\Delta t H - \Delta t_\omega H(\mathbf{\Omega})} | \mathbf{R}', \mathbf{\Omega}' \rangle \\ &= \langle \mathbf{R}, \mathbf{\Omega} | e^{-\Delta t (T + V_C + V_{SO}) - \Delta t_\omega H(\mathbf{\Omega})} | \mathbf{R}', \mathbf{\Omega}' \rangle \\ &\approx \langle \mathbf{R} | e^{-\Delta t (T + V_C + V_{SO})} | \mathbf{R}' \rangle \langle \mathbf{\Omega} | e^{-\Delta t_\omega H(\Omega)} | \mathbf{\Omega}' \rangle + O(\Delta t_\omega)^2 \end{aligned} \tag{6.16}$$

In DMC, the fixed-phase approximation(See Chapter 2) is applied to treat the complex wavefunctions.

6.3 Rashba Interaction in Two Dimensional Electron Gas

In this section, we will be discussing Rashba interaction in two dimensional electron gas. We first review some fundamental theory of electron gas, then we introduce the Rashba interaction, finally present and analyze the calculation results.

6.3.1 Basic Theory of Homogeneous Electron Gas(HEG)

Homogeneous electron gas, often called Jellium, is a model of interacting electrons placed in uniformly distributed positively charged background to ensure the system is charge-neutral. It is an ideal model, but it is very helpful for us to have a basic understanding of the scheme of electron interaction, especially, the electron correlation effect. We can use HEG to model the electron-electron interactions in the real system like solids. Without having atomic lattice and crystal structure to make the system very complex, we can solely focus on the quantum nature of electrons. We consider N non-interacting electron gas confined in a cubic box with sides of length L. The solution to the non-interacting

Hamiltonian which only contains sum of the kinetic operator is

$$\phi_{k\lambda}(r) = \frac{e^{i\mathbf{k}\cdot\mathbf{r}}}{\sqrt{V}}\chi_\lambda \tag{6.17}$$

where $V = L^3$ is the volume of the box, χ_λ is the spin functions in Pauli representation. The wavenumber k is determined by the periodic boundary conditions

$$k_i = \frac{2\pi n_i}{L} \quad i = x, y, z, \quad n_i = 0, \pm 1, \pm 2... \tag{6.18}$$

We consider the Hamiltonian of HEG.

$$H = H_e + H_b + H_{e-b} \tag{6.19}$$

where

$$H_e = \sum_{i=1}^{N} \frac{p_i^2}{2} + \sum_{i<j}^{N} \frac{e^{-\alpha|\mathbf{r_i}-\mathbf{r_j}|}}{|\mathbf{r_i}-\mathbf{r_j}|} \tag{6.20}$$

are the Hamiltonian for electrons,

$$H_b = \frac{1}{2}\int\int d^3r d^3r' \frac{n(\mathbf{r})n(\mathbf{r}')e^{-\alpha|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|} \tag{6.21}$$

is the energy of the positively charged background whose particle density is $n(\mathbf{r}) = N/V$, and

$$H_{e-b} = -\sum_{i=1}^{N}\int d^3r \frac{n(\mathbf{r})e^{-\alpha|\mathbf{r}-\mathbf{r_i}|}}{|\mathbf{r}-\mathbf{r_i}|} \tag{6.22}$$

is the interaction energy between the electrons and the positively charged background. Note the notation $\mathbf{r}$ denotes the position regarding the background, while $\mathbf{r_i}$ denotes electrons. To make the following mathematics meaningful at every step, we change to Coulomb potential to Yukawa potential which we can finally reduce to Coulomb potential by taking the limit as α approaches to 0. Because of the long-range properties of Coulomb interaction, all the three terms in Eq. 6.19 diverge in the thermodynamic limit when $N \to \infty, V \to \infty$, but we can ensure the density $n = N/V$ remains constant, and the bulk properties, for example E/N(energy per particle), are meaningful.

It is easy to work out the energy components of the terms H_b and H_{eb} in Eq. 6.19.

$$H_b = \frac{1}{2}\frac{N^2}{V}\frac{4\pi}{\alpha^2} \tag{6.23}$$

$$H_{e-b} = -\frac{N^2}{V}\frac{4\pi}{\alpha^2} \tag{6.24}$$

Thus the total Hamiltonian is reduced to

$$H = -\frac{1}{2}N^2V^{-1}4\pi\alpha^{-2} + H_e \tag{6.25}$$

This is what of physical interest. It can be rigorously shown that plane waves we used above is the solution to this Hamiltonian under Hartree-Fock approximation if we use Eq. 6.17 as our one-electron orbital to construct the Slater determinant $\Psi_T = det\{\phi_k(r)\}$. We here show the result of our interest directly. If the system is non-spin-polarized, that is, each k-wavenumber eigenstate is occupied by two electrons. Then the maximum wavenumber is determined by the density

$$2 \cdot \frac{4}{3}\pi k_F^3 \frac{V}{(2\pi)^3} = N \tag{6.26}$$

The factor of 2 accounts for the spin degeneracy. Thus

$$k_F^3 = \frac{3\pi^2 N}{V} \tag{6.27}$$

Define radius of the electron r_s which satisfies

$$V = \frac{4}{3}\pi r_s^3 N \tag{6.28}$$

The expectation value of the kinetic energy is

$$\langle\Psi_T|\sum_{i=1}^{N}\frac{p_i^2}{2}|\Psi_T\rangle = \frac{3}{5}\frac{k_F^2}{2}N = \frac{1}{2}\frac{N}{r_s^2}\frac{3}{5}(\frac{9\pi}{4})^{2/3} = \frac{1}{2}N\frac{2.21}{r_s^2} \tag{6.29}$$

The Coulomb energy is

$$\langle\Psi_T|\frac{1}{2}\sum_{i\neq j}^{N}\frac{e^{-\alpha|\mathbf{r_i}-\mathbf{r_j}|}}{|\mathbf{r_i}-\mathbf{r_j}|}|\Psi_T\rangle = \frac{1}{2}\frac{N^2}{V}\frac{4\pi}{\alpha^2} - \frac{1}{2}\frac{N}{V}\frac{4\pi}{\alpha^2} - \frac{1}{2}\frac{N}{r_s}(\frac{9\pi}{4})^{\frac{1}{3}}\frac{3}{2\pi} \tag{6.30}$$

It is clear that the first term here cancels the one in Eq.6.25. The second term vanishes when $L \to \infty$ and $\alpha \to 0$, but keeping $\alpha^{-1} \ll L$. Those two terms combined together are the Hartree energy, the classical component of Coulomb interaction. It corresponds to the interaction of electrons without exchanging the momentum. We also introduced this term in Chapter 1 when we were discussing the Hartree-Fock approximation. The last term is simplified to $-\frac{1}{2}N\frac{0.916}{r_s}$, which is the exchange energy corresponding to the interaction of electrons with exchanging the momentum. Thus the ground state energy per particle under Hartree-Fock approximation is

$$E/N = \frac{1}{2}[\frac{2.21}{r_s^2} - \frac{0.916}{r_s}] \tag{6.31}$$

which is a finite number. The first term is simply the same as the kinetic energy of free electrons. It is dominant when $r_s \to 0$, that is, the density is high. The second term is the exchange energy and it is negative. The exchange energy is the only term left due to Coulomb interaction under Hartree-Fock approximation, because as what we have shown, the Hartree part, is either canceled by the positive background charge or vanishes in the thermodynamic limit. The exchange becomes dominant when the density of electrons is low.

For 2D HEG, the same procedure can be followed and the ground state energy per particle is

$$E/N = \frac{1}{2}\frac{1}{r_s^2} - \frac{4\sqrt{2}}{3\pi r_s} = \frac{1}{2}(\frac{1}{r_s^2} - \frac{0.12}{r_s}) \tag{6.32}$$

6.3.2 Rashba Hamiltonian

Rashba interaction is a consequence of spin-orbit effect in 2 dimensional electron gas(2DEG). In the system, electrons are free to move in 2 dimensions($\hat{x} - \hat{y}$ plane) but is forbidden to move in the third direction($\hat{z}$). While the 3 dimensional electron gas is usually used to model the electron behavior in solid, 2DEG can be realized in semiconductor devices. For example, in MOSEFT(metal-oxide-semiconductor field effect transistor), when the tran-

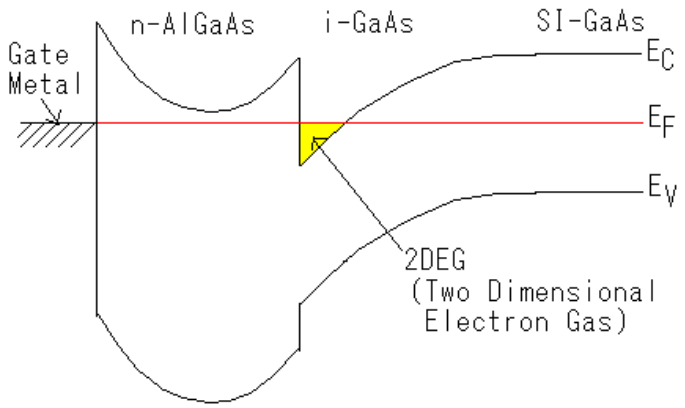

Figure 6.1: Band edge diagram of a basic HEMT. The picture is retrieved at http://en.wikipedia.org/wiki/File:
HighElectronMobilityTrasistor _Band_image.PNG

sistor is in the inversion mode, the electrons are confined between the semiconductor-oxide surface. 2DEG also can be found in high-electron-mobility-transistors(HEMTs) shown in Figure 6.1 and Figure 6.2.

If we add an electric field perpendicular to the $\hat{x} - \hat{y}$ plane, which is in $\hat{z}$ direction,

$$\mathbf{E} = E_0 \hat{z} \tag{6.33}$$

Due to relativistic corrections, based on the electrodynamics theory, an electron moving in an electric field experiences another effective magnetic field

$$\mathbf{B_{eff}} = \frac{\mathbf{v} \times \mathbf{E}}{c^2} \tag{6.34}$$

where $\mathbf{v}$ is the speed of the electron and c is the speed of light. This magnetic field interacts with the electron spin, leading to the spin-orbit coupling term

$$H_{SO} = \frac{g\mu_B}{2c^2}(\mathbf{v} \times \mathbf{E}) \cdot \sigma \tag{6.35}$$

By defining $\zeta = \frac{g\mu_B E_0}{2c^2}$, then

$$H_{SO} = \zeta(\sigma \times \mathbf{p}) \cdot \hat{z} = \zeta \sum_i^N (p_i^{(y)} \sigma_i^{(x)} - p_i^{(x)} \sigma_i^{(y)}) \tag{6.36}$$

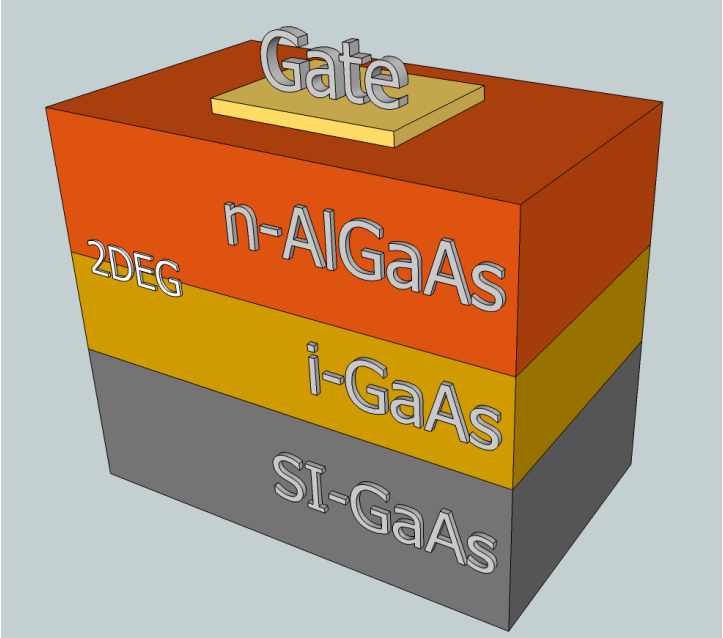

Figure 6.2: Heterostructure corresponding to the band edge diagram in Figure 6.1. The picture is retrieved at http://en.wikipedia.org/wiki/File: Sketch_of_a_heterostructure_-_2DEG_in_GaAs-GaAlAs.png

This is the additional term in the Hamiltonian, called Rashba term. Together with the kinetic and Coulomb terms, the full Hamiltonian of 2D HEG with Rashba is

$$H = H_b + H_{e-b} + H_e \tag{6.37}$$

where

$$H_e = T + V + H_{SO} = \sum_i^N \frac{p^2}{2} + \sum_{i<j} \frac{1}{|\mathbf{r_i} - \mathbf{r_j}|} + \sum_i^N \zeta(p_i^{(y)}\sigma_i^{(x)} - p_i^{(x)}\sigma_i^{(y)}) \tag{6.38}$$

The term H_b and H_{e-b} and the Hartree term associated with H_e are canceled as we showed in the last subsection. The only part of interest is the rest of H_e: kinetic part, exchange part, correlation part and Rashba part. This is the problem we want to solve. The Rashba part contains spin operator explicitly where we want to apply the methodology we discussed above.

Let us first remove the Coulomb part of the Hamiltonian, the rest of all are one-

particle operators only

$$H_e^{'} = \sum_i^N \frac{p^2}{2} + \sum_i^N \zeta(p_i^{(y)}\sigma_i^{(x)} - p_i^{(x)}\sigma_i^{(y)}) \tag{6.39}$$

Note that the Rashba operator commutes with the kinetic operator, so we guess the possible solution has the following form of a planewave and a Dirac spinor

$$\phi_k(r) = \frac{1}{\sqrt{S}} e^{i\mathbf{k}\cdot\mathbf{r}} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \tag{6.40}$$

We substitute the solution into the eigenvalue equation $H_e^{'}\phi(r) = \epsilon\phi(r)$

$$\frac{k^2}{2}\begin{pmatrix} \alpha \\ \beta \end{pmatrix} + \zeta(k_y \begin{pmatrix} \beta \\ \alpha \end{pmatrix} + ik_x \begin{pmatrix} \beta \\ -\alpha \end{pmatrix}) = \epsilon \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \tag{6.41}$$

We solve α, β, ϵ. The one-electron eigenstates are

$$\phi_{k\lambda} = \frac{1}{\sqrt{S}} e^{i\mathbf{k}\cdot\mathbf{r}} \frac{1}{2} \begin{pmatrix} \pm\frac{k_y + ik_x}{|\mathbf{k}|} \\ 1 \end{pmatrix} \tag{6.42}$$

The eigenenergies are

$$\epsilon_{k\lambda} = \frac{k^2}{2} \pm \zeta|\mathbf{k}| \tag{6.43}$$

The plus and minus signs correspond to quasi-spin-up state and quasi-spin-down state, respectively. Figure 6.3 shows the plot of eigenenergy as a function of k.

6.3.3 Calculation Details: Wavefunctions, Periodic Boundary Conditions, Treatment of Coulomb Interactions in Extended Systems

We apply the simplest form of the wavefunction: Slater-Jastrow type wavefunction.

$$\Psi_T(\mathbf{R}, \mathbf{\Omega}) = det(\phi_{k\lambda}(\mathbf{r}, \omega))J(\mathbf{R}) \tag{6.44}$$

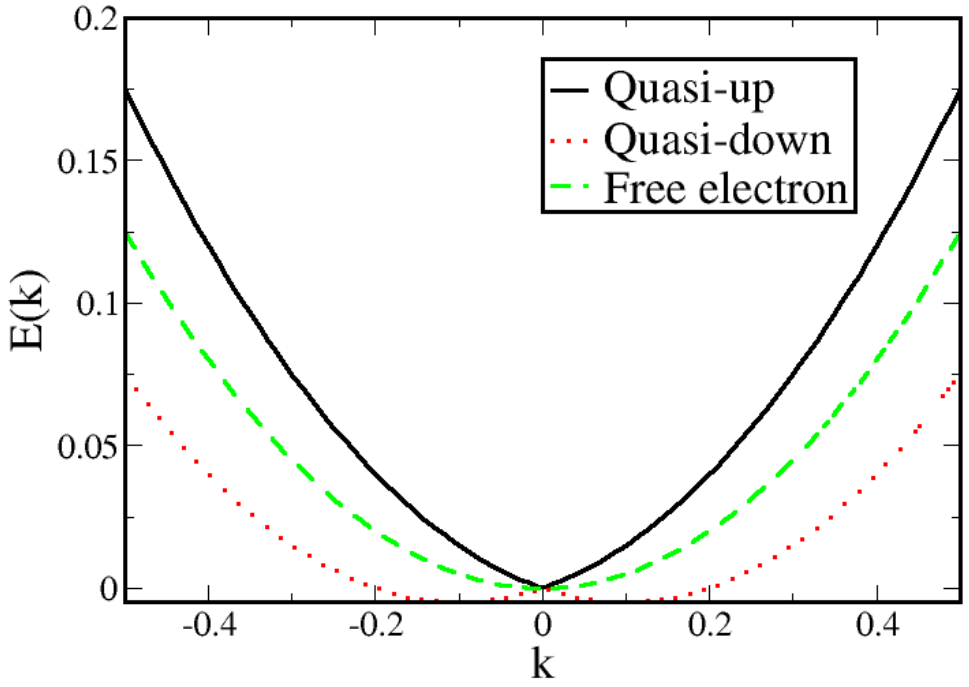

Figure 6.3: Plot of the energy band given by Eq. 6.43 and energy of a free electron.

We build our Slater determinant by taking the one-electron orbitals derived in last subsection(removing the normalization factor, since it does not matter in QMC), and write the orbitals in our new spin representation which we introduced in Section 6.1

$$\phi_{k\lambda} = e^{i\mathbf{k}\cdot\mathbf{r}}(\pm\frac{k_y + ik_x}{|\mathbf{k}|}e^{i\omega} + e^{-i\omega}) \tag{6.45}$$

The Jastrow factor $J(\mathbf{R})$ doesn't contain any spin variables so it is the same thing as in Chapter 3. Our goal is to study the extended quantities of a system with infinite size, however, our current simulation only supports the system with finite size. The approach we use in the simulation is to put electrons in a finite box, called a simulation cell, subjected to an artificially imposed periodic condition. We consider the interaction of the electrons both the ones in the simulation cell and the imaginary images created by the periodic repeat of the simulation cell. This approach generates finite-size errors. The error appears in all the energy components: kinetic, Coulomb potential and Rashba. In an infinite system, the momentum of an electron can take any values from 0 up to the Fermi momentum. But only discrete momentum vectors are allowed to be occupied in the finite systems, which makes the kinetic energy differ from the values in the thermodynamic limit. The difference goes smaller and converges to values in the thermodynamic limit when the size of the systems becomes large enough. To address the error in kinetic energy and avoid the computational cost of simulating systems with a very large size, we use twist average method. Recall the many-body Bloch theorem: the wavefunction of a system which has translational symmetry satisfies

$$\Psi_{T,K}(\mathbf{R}) = e^{i\mathbf{K}\cdot\sum_i \mathbf{r_i}}U_K(\mathbf{R}) \tag{6.46}$$

$$\Psi_{T,K_p}(\mathbf{R}) = e^{i\mathbf{K_p}\cdot\frac{1}{N}\sum_i \mathbf{r_i}}W_K(\mathbf{R}) \tag{6.47}$$

where $\mathbf{K}$ is the crystal wave vector, also called the crystal momentum and L is the lattice vector of the simulation cell. U_K is a periodic function that is invariant under the translation of any single electron by a lattice constant vector: $U_K(\mathbf{r_i} + n\mathbf{L}) = U_K(\mathbf{r_i})$. W_K is a periodic function that is invariant under the simultaneous translation of all the electrons by any primitive lattice vector $\mathbf{R_p}$. $W_K(\mathbf{r_1} + n\mathbf{L}, ..., \mathbf{r_i} + n\mathbf{L}, ..., \mathbf{r_N} + n\mathbf{L}) = W_K(\mathbf{r_1}, ..., \mathbf{r_i}, ..., \mathbf{r_N})$. $\mathbf{K_p}$ is the primitive wave vector distinct from the crystal wave vector K. If we impose periodic condition on the wavefunction: $\Psi_T(\mathbf{r_i} + \mathbf{L}) = \Psi_T(\mathbf{r_i})$, it means

we choose $K = 0$ in the Bloch theorem. When $K = \frac{\pi}{L}$, we have anti-periodic boundary condition. When K takes other values, we have the twisted boundary condition. In the twisted boundary condition, the plane-wave vector of one-electron orbital is

$$k_i = K + \frac{2\pi n_i}{L} = \frac{\pi}{L}(m + 2n_i) \quad i = x, y, z; \;\; n_i = 0, \pm 1, \pm 2; \;\; -1 \leq m \leq 1 \tag{6.48}$$

Twisted-averaged method is to repeatedly sample special K-points in the first Brillouin zone of the simulation-cell reciprocal lattice, which corresponds the case if we occupy the one-particle k vectors starting with different K-points and finally average the energy over all K-points. This repeated sampling is equivalent to doing an integration over the first Brillouin zone. Let $E_{\mathbf{K}} =< \Psi_{T,\mathbf{K}}|H|\Psi_{T,\mathbf{K}} >$, then

$$E = \frac{L^3}{(2\pi)^3} \int_{BZ} d^3K E_{\mathbf{K}} = \sum weight E_{\mathbf{K}} \tag{6.49}$$

The weight is determined by the number of points in the K-point grids which are equivalent to each other by the symmetry of the point group to unique K-point.

Since the Rashba operator is a one-electron operator involving momentum operator, the correct treatment of kinetic energy solves the issue of finite-size error with Rashba operator together.

The potential energy has its problem as well. The long-range property makes it converge very slowly. Even the system is sufficiently large, the potential energy still fluctuates too much[100]. The Coulomb interaction also needs special treatment to converge. In the Hamiltonian, the Coulomb interaction is

$$V = \frac{1}{2} \sum_i \sum_{j \neq i} \frac{1}{|\mathbf{r_i} - \mathbf{r_j}|} \tag{6.50}$$

When it is subjected to periodic boundary condition, the actual form is changed to

$$V = \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N}{}' \frac{1}{|\mathbf{r_i} - \mathbf{r_j} - \mathbf{nL}|} \tag{6.51}$$

The prime here means the term $j = i$ is omitted when $n = 0$, and $\mathbf{L}$ is the lattice constant vector. Therefore considering the Coulomb potential of the ith electron due to the jth one, we also take account of all the image copies of jth electrons in the infinite identical

copies of the simulation cell. There are two ways of treating the potential energy: one way is called Ewald summation[101], the other one is simpler, called model potential[102]. In Ewald summation, the potential energy can be decomposed into two parts:

$$V = \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1,j\neq i}^{N} V_E(|\mathbf{r_i} - \mathbf{r_j} - n\mathbf{L}|) + \frac{1}{2}\sum_{i=1}^{N}\xi \tag{6.52}$$

The first term includes the interaction of the ith electron at $\mathbf{r_i}$ and jth electron at $\mathbf{r_j}$(when $n = 0$), plus the ith electron and all the images of jth electron positions positioned at $\mathbf{r_j} + n\mathbf{L}$. The second term is due to the ith electron and its own images in the infinite copies of the simulation cell. It can be calculated by

$$\xi = \lim_{\mathbf{r}->\mathbf{r_j}} (V_E(\mathbf{r} - \mathbf{r_j} - n\mathbf{L}) - \frac{1}{|\mathbf{r} - \mathbf{r_j} - n\mathbf{L}|}) \tag{6.53}$$

Those two sums V_E and ξ are only conditionally convergent and not easy to evaluate. In order to tackle the problem, we decompose each sum into two parts: short-range part and the long-range part and evaluate the short-range part in real space and the long-range part in reciprocal space. In 3 dimensional case, the formula is

$$V_E = \sum_n \frac{erfc(\alpha|\mathbf{r}^{(ij)} - n\mathbf{L}|)}{|\mathbf{r}^{(ij)} - m\mathbf{L}|} + \sum_G \frac{4\pi}{VG^2} exp(-\frac{G^2}{4\alpha^2}) exp(-i\sum_i \mathbf{G}\cdot\mathbf{r_i}) - \frac{\pi}{V\alpha^2} \tag{6.54}$$

The first term is the sum in real space. $erfc$ is the complementary error function. $\mathbf{r}^{(ij)} = min|\mathbf{r_i} - \mathbf{r_j} - n\mathbf{L}|$, which means when we consider the potential at r_i due to an electron at r_j, we consider all the image of the electron at r_j over the periodic lattice and pick up the one which is closest to r_i. In other words, we use the nearest image. α is chosen in such a way that the sum over m is only needed for the nearest neighbour of the simulation cell, that is, $m = 0, \pm 1$. The second term is the sum in reciprocal space over G, the reciprocal lattice vector. i is the imaginary unit. Finally,

$$\xi = \sum_n \frac{erfc(\alpha|n\mathbf{L}|)}{n|\mathbf{L}|} - \frac{2\alpha}{\sqrt{\pi}} - \frac{\pi}{\alpha^2 V} + \sum_G \frac{4\pi}{VG^2} exp(-\frac{G^2}{4\alpha^2}) \tag{6.55}$$

In 2 dimensional case , the formula is

$$V_E = \sum_n \frac{erfc(\alpha|\mathbf{r}^{(ij)} - n\mathbf{L}|)}{|\mathbf{r}^{(ij)} - n\mathbf{L}|} + \frac{2\pi}{SG} erfc(\frac{G}{2\alpha}) exp(i \sum_i \mathbf{G} \cdot \mathbf{r_i}) - \frac{2\sqrt{\pi}}{\alpha S} \tag{6.56}$$

$$\xi = \sum_n \frac{erfc(\alpha|m\mathbf{L}|)}{|n\mathbf{L}|} + \frac{2\pi}{SG} erfc(\frac{G}{2\alpha}) - \frac{2\alpha}{\sqrt{\pi}} - \frac{2\sqrt{\pi}}{\alpha S} \tag{6.57}$$

where S is the area of the simulation surface.

Model potential is a simpler treatment. The potential that the ith electron feels by the jth electron is written as

$$V_{ij} = \frac{1}{r^{(ij)}} \tag{6.58}$$

where $r^{(ij)} = min|\mathbf{r_i} - \mathbf{r_j} - n\mathbf{L}|$. Then the total potential energy is the sum of the potential the each electron feels then subtract the classical Hartree energy. The total potential energy is

$$V = \sum_{i \neq j} V_{ij} - H \tag{6.59}$$

where in 3 dimensional case $H = \int_V \frac{1}{r} d^3r$, V is the simulation box, while in 2 dimensional case $H = \int_S \frac{1}{r} d^2r$, S is the simulation surface.

6.3.4 Calculation Results

We calculate the energy of 2 dimensional homogeneous electron gas system(2DHEG). The Hamiltonian used is Eq.6.37. Since ζ depends on some constant number and the electric field which can be tuned in experiment, we use ζ as a parameter in the calculation. Coulomb potential is treated as model potential. The wavefunction we use is Slater-Jastrow wavefunction. We take the simplest form of Jastrow, just cusp condition. Figure 6.4 shows the DMC energy per electron of the system with 58 electrons and $r_s = 1$. The results use 4 k-points for twisted-averaging. The energy per electron is a function of quasi-polarization. The quasi-polarization is defined as

$$\xi = \frac{N^+ - N^-}{N} \tag{6.60}$$

N^+ is the number of quasi-spin-up electrons, and N^- is the number of quasi-spin-down electrons. The quasi-spin states are defined in Eq. 6.42. Based on what we discuss the

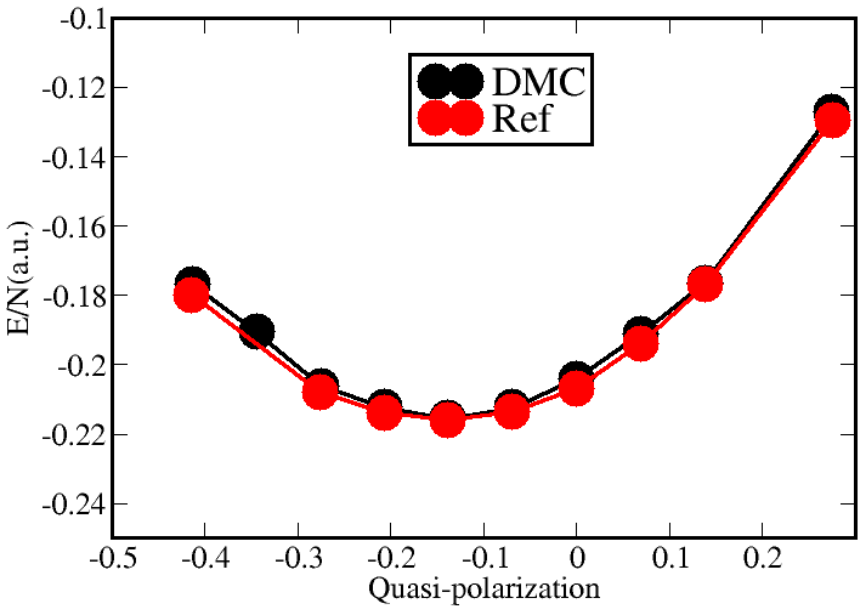

Figure 6.4: DMC energy per electron. The system is 58 electrons with $r_s = 1$, $\zeta = 0.1$. Coulomb potential is treated as model potential. Slater-Jastrow wavefunction is used. Number of configurations: 200. Timestep for space: 0.01. Timestep for spin: 0.0003. Our calculation data is labeled as DMC. Ref data is from [103].

basic theory of jellium, the DMC energy should include kinetic part, exchange part, and the correlation part and Rashba part. We also plot the result of another calculation of the same system in reference [103]. From the Figure 6.4, we see two different methods in treatment of spin converge very well. In Figure 6.5, we calculate the energy per electron for the same system as in Figure 6.4 except $r_s = 5$. We see the fact that for low-density electron gas, the lowest energy happens to the state with more negative quasi-polarization. This can be understood as for a system with given number of electrons, since the Rashba operator is linear function of momentum, the Rashba energy is proportional to $1/r_s$. So when the density is low, the Rashba energy becomes dominant, making the ground state to favor a more negative polarization. The data of our DMC result in Figure 6.4 and Figure 6.5 are shown in Table 6.1 and Table 6.2.

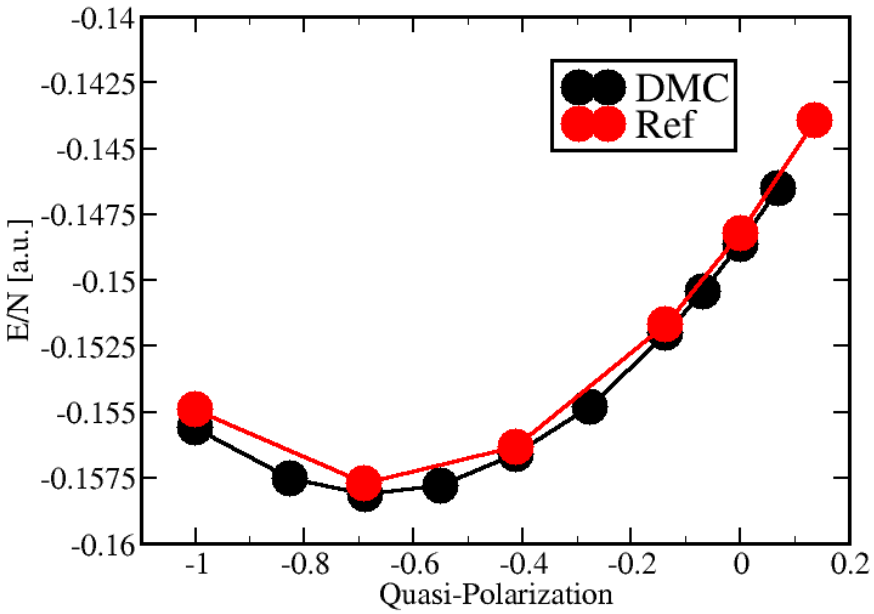

Figure 6.5: DMC energy per electron. The system is 58 electrons with $r_s = 5$, $\zeta = 0.1$. Coulomb potential is treated as model potential. Slater-Jastrow wavefunction is used. Number of configurations: 200. Timestep for space: 0.01. Timestep for spin: 0.0003. Our calculation data is labeled as DMC. Ref data is from [103].

Table 6.1: Data for Figure 6.4 with error bars

Quasi-polarization	N^+	N^-	E/N [a.u.]	E/N[a.u.] [103]
-0.413	17	41	-0.1768(4)	-0.180(2)
-0.344	19	39	-0.1906(4)	n/a
-0.276	21	37	-0.2063(4)	-0.208(3)
-0.207	23	35	-0.2123(6)	-0.214(2)
-0.138	25	33	-0.2155(4)	-0.216(2)
-0.069	27	31	-0.2123(4)	-0.214(2)
0	29	29	-0.2041(3)	-0.207(2)
0.069	31	27	-0.1912(3)	-0.194(2)
0.137	33	25	-0.1764(4)	-0.177(2)
0.275	37	21	-0.1271(4)	-0.129(2)

Table 6.2: Data for Figure 6.5 with error bars

Quasi-polarization	N^+	N^-	E/N [a.u.]	E/N[a.u.] [103]
-1	0	58	-0.1556(1)	-0.1550(3)
-0.827	5	53	-0.1575(1)	n/a
-0.689	9	49	-0.1581(1)	-0.1578(3)
-0.551	13	45	-0.1578(1)	n/a
-0.413	17	41	-0.1566(2)	-0.1564(2)
-0.276	21	37	-0.1548(1)	n/a
-0.137	25	33	-0.1520(1)	-0.1517(2)
-0.069	27	31	-0.1504(1)	n/a
0	29	29	-0.1486(1)	-0.1482(2)
0.069	31	27	-0.1465(1)	n/a

6.3.5 Discussion

The comparisons of the results in Figure 6.4 and Figure 6.5 are very reasonable, although the two methods treat spin sampling differently. The new treatment of spin in QMC introduces another two factors which could affect the quality of the results: Fixed-phase error and spin-timestep bias. The conventional QMC which handles the Hamiltonian without spin operator can use real function and that leads to fixed-node error. In this calculation of Rashba interaction, we artificially introduce a phase in the wavefunction and that causes fixed-phase error. To assess the fixed-phase error, we perform a calculation of 2D HEG whose Hamiltonian does not contain Rashba operator. The system contains 58 electrons and it is non-spin-polarized, which means half the electrons have up-spin($s_z = 1/2$) and half have down-spin($s_z = -1/2$). We use Slater-Jastrow wavefunction, model potential and no twisted averaging. Jastrow factors are as simple as cusp. We show the DMC results(energy per electron) with and without spin sampling, and compare other reference results[104] in Table 6.3. The DMC without spin and reference[104] both use fixed-node approximation, and the results match together. But since DMC with spin uses fixed-phase approximation, the energies per electron appear all higher than fixed-node energy for all r_s values. This fact indicates the presence of fixed-phase error, which appears to be larger than than the fixed-node error. In 2DEG, model potential is usually preferred because it can give smaller finite-size errors than Ewald sum, even if the Ewald

Table 6.3: Comparison of different DMC result: DMC with spin sampling, without spin sampling and reference[104]. Our DMC data is calculated using Slater-Jastrow wavefunction and model potential without twist averaging. The number listed in the table is the energy per electron. The number of configurations is 200. The timestep for space is 0.002, 0.01, 0.1 in the order of increasing r_s

r_s	DMC with spin	DMC without spin	Reference
1	-0.1989(2)	-0.2012(3)	-0.2013(1)
2	-0.2542(1)	-0.2557(2)	-0.255802(4)
5	-0.14853(5)	-0.14907(5)	-0.149134(9)

Table 6.4: Comparison of different DMC result: DMC with spin sampling, without spin sampling. We use Slater-Jastrow wavefunction. The potential is treated as Ewald sum, and no twist averaging is used. The number listed in the table is the energy per electron. The number of configurations is 200. The timestep for space is 0.002, 0.01, 0.1 in the order of increasing r_s

r_s	DMC with spin	DMC without spin
1	-0.1917(2)	-0.1941(3)
2	-0.2447(2)	-0.2445(2)
5	-0.1367(2)	-0.1372(2)

sum is more complex to evaluate[102]. For completeness, though, we also list the data calculated using Ewald sum instead of model potential in Table 6.4.

The timestep error can addressed by choosing a reasonably small timestep for both space and spin. By choosing a small timestep, the DMC results has acceptance value larger than 0.99.

6.3.6 Conclusion

We have introduced a new math representation for spin and incorporated in QMC methods. Then we have calculated the Rashba interaction in the two dimensional electron gas and compared the results with other calculations. The comparisons show a very good agreement overall. The difference of fixed-node and fixed-phased results are compared to assess the quality of the methods. The whole study demonstrates the feasibility of evaluating spin-orbit coupling effect using QMC methods.

REFERENCES

[1] D. M. Ceperley and B.J. Alder. *Phys. Rev. Lett.*, 45:566, 1980.

[2] D. R. Hartree. *Proc. Cambridge Philos. Soc*, 24:89, 1928.

[3] V. Fock. *Z. Phys.*, 61:126, 1930.

[4] J. C. Slater. *Phys. Rev.*, 35:210, 1930.

[5] P. Hohenberg and W. Kohn. *Phys. Rev.*, 136:B864, 1964.

[6] W. Kohn and L. J. Sham. *Phys. Rev.*, 140:A1133, 1965.

[7] R. O. Jones and O. Gunnarsson. *Rev. Mod. Phys*, 61:689, 1989.

[8] D. M. Ceperley. *Phys. Rev. B*, 18:3126, 1978.

[9] J. P. Perdew and Y. Wang. *Phys. Rev. B*, 45:13244, 1992.

[10] J. P. Perdew, K. Burke, and M. Ernzerhof. *Phys. Rev. Lett.*, 77:3865, 1996.

[11] A. B. Becke. *Phys. Rev. A.*, 38:3098, 1988.

[12] C. Lee, W. Yang, and R. G. Parr. *Phys. Rev. B*, 37:785, 1988.

[13] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. *J. Chem. Phys.*, 21:1087, 1953.

[14] J. W. Moskowitz, K. E. Schmidt, M. A. Lee, and M. H. Kalos. *J. Chem. Phys.*, 77:349, 1982.

[15] W. M. C. Foulkes, L. Mitas, R. J. Needs, and G. Rajagopal. *Rev. Mod. Phys.*, 73:33, 2001.

[16] C.-J. Huang, C. J. Umrigar, and M. P. Nightingale. *J. Chem. Phys.*, 107:3007, 1997.

[17] K. Madsen, H. B. Nielsen, and O. Tingleff. *Methods for non-linear least squares problems, 2nd edition.* 2004.

[18] C.J.Umrigar and Claudia Filippi. *Phys. Rev. Lett.*, 94:150102.

[19] Julien Toulouse and C.J.Umrigar. *J. Chem. Phys.*, 126:084102, 2007.

[20] D.M.Ceperley. *J. Stat. Phys*, 43:815.

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

[22] D. R. Hamann, M. Schluter, and C. Chiang. *Phys. Rev. Lett.*, 43:1494, 1979.

[23] P. A. Christiansen, Y. S. Lee, and K. S. Pitzer. *J. Chem. Phys*, 71:4445, 1979.

[24] M. Krauss and W. J. Stevens. *Ann. Rev. Phys. Chem*, 35:357, 1984.

[25] D. Vanderbilt. *Phys. Rev. B*, 41:7892, 1990.

[26] L.Mitas, E.L.Shirley, and D.M.Ceperley. *J. Chem. Phys*, 95:3467.

[27] Lincoln D Carr, David DeMille, Roman V Krems, and Jun Ye. Cold and ultracold molecules: science, technology and applications. *New Journal of Physics*, 11(5):055049, 2009.

[28] E. R. Hudson. Magnetic trapping of calcium monohydride molecules at millikelvin temperatures. *Phys. Rev. A*, 79:032716, 2009.

[29] J. D. Weinstein, R. deCarvalho, T. Guillet, B. Friedrich, and J. M. Doyle. *Nature*, 395:148, 1998.

[30] J. M. Doyle, B. Friedrich, J. Kim, and D. Patterson. *Phys. Rev. A*, 52:R2515, 1995.

[31] J. M. Doyle, B. Friedrich, R. V. Krems, and F. Masnou-Seeuws. *Eur. J. Phys. D.*, 31:149, 2004.

[32] S. A. Rangwala, T. Junglen, T. Rieger, P. W. H. Pinkse, and G. Rempe. *Phys. Rev. A.*, 67:043406, 2003.

[33] H. L. Bethlem, G. Berden, F. M. H. Crompvoets, R. T. Jongma, A. J. A. van Roij, and G. Meijer. *Nature*, 406:491, 2000.

[34] E. S. Shuman, J. F. Barry, D. R. Glenn, and D. DeMille. *Phys. Rev. Lett.*, 103:223001, 2009.

[35] E. S. Shuman, J. F. Barry, and D. DeMille. *Nature*, 467:820, 2010.

[36] M. D. Di Rosa. *Eur. Phys. J. D*, 31:395, 2004.

[37] B. K. Stuhl, B. C. Sawyer, D. Wang, and J. Ye. *Phys. Rev. Lett.*, 101:243002, 2008.

[38] J. M. Sage, S. Sainis, T. Bergeman, and D. DeMille. *Phys. Rev. Lett.*, 94:203001, 2005.

[39] J. M. Hutson and P. Soldan. *Int. Rev. Phys. Chem*, 25:497, 2006.

[40] O. Dulieu, M. Raoult, and E. Tiemann. *J. Phys. B: At. Mol. Opt. Phys*, 39:1, 2006.

[41] O. Dulieu and C. Gabbanini. *Rep. Prog. Phys.*, 72:086401, 2009.

[42] G. Pupillo, A. Micheli, H. P. Büchler, and P. Zoller. *arXiv*, 0805:1896, 2008.

[43] D. DeMille. *Phys. Rev. Lett*, 88:067901, 2002.

[44] A. Andre, D. DeMille, J. M. Doyle, M. D. Lukin, S. E. Maxwell, P. Rabl, R. J. Schoelkopf, and P. Zoller. Hybrid quantum information processing with polar molecules. *International Conference on Atomic Physics (ICAP 2006)*, 20:128, 2006.

[45] E. Kuznetsova, K. Kirby R. Cote, and S. F. Yelin. *Phys. Rev. A.*, 78:012313, 2008.

[46] S. F. Yelin, K. Kirby, and R. Cote. *Phys. Rev. A.*, 74:050301(R), 2006.

[47] A. Andre, D. DeMille, J. M. Doyle, M. D. Lukin, S. E. Maxwell, P. Rabl, R. J. Schoelkopf, and P. Zoller. *Nature Physics*, 2:636, 2006.

[48] P. Rabl, D. DeMille, J. M. Doyle, M. D. Lukin, R. J. Schoelkopf, and P. Zoller. *Phys. Rev. Lett.*, 97:33003, 2006.

[49] M. G. Kozlov and L. N. Labzowsky. *J. Phys. B*, 28:1933, 1995.

[50] J. J. Hudson, B. E. Sauer, M. R. Tarbutt, and E. A. Hinds. *Phys. Rev. Lett.*, 89:023003, 2002.

[51] A. N. Petrov, N. S. Mosyagin, and A. V. Titov. *Phys. Rev. A.*, 79:012505, 2009.

[52] L. V. Skripnikov, A. N. Petrov, A. V. Titov, and N. S. Mosyagin. *Phys. Rev. A.*, 80:060501, 2009.

[53] D. DeMille, S. Sainis, J. Sage, T. Bergeman, S. Kotochigova, and E. Tiesinga. *Phys. Rev. Lett.*, 100:043202, 2008.

[54] V. V. Flambaum and M. G. Kozlov. *Phys. Rev. Lett.*, 99:150801, 2007.

[55] T. Zelevinsky, S. Kotochigova, and J. Ye. *Phys. Rev. Lett.*, 100:043201, 2008.

[56] D. DeMille, S. B. Cahn, D. Murphree, D. A. Rahmlow, and M. G. Kozlov. *Phys. Rev. Lett.*, 100:023003, 2008.

[57] D. Kawall, F. Bay, S. Bickman, Y. Jiang, and D. DeMille. *Phys. Rev. Lett.*, 92:133007, 2004.

[58] E. R. Hudson, H. J. Lewandowski, B. C. Sawyer, and J.Ye. *Phys. Rev. Lett.*, 96:143004, 2006.

[59] A. D. Ludlow, T. Zelevinsky, G. K. Campbell, S. Blatt, M. M. Boyd, M. H. G. de Miranda, M. J. Martin, J. W. Thomsen, S. M. Foreman, Jun Ye1, T. M. Fortier, J. E. Stalnaker2, S. A. Diddams, Y. Le Coq, Z. W. Barber, N. Poli, N. D. Lemke, K. M. Beck, and C. W. Oates. *Science*, 319:1805, 2008.

[60] I. Bloch, J. Dalibard, and W. Zwerger. *Rev. Mod. Phys.*, 80:885, 2008.

[61] A. Micheli, G. K. Brennen, and P. Zoller. *Nature Phys.*, 2:341, 2006.

[62] B. Capogrosso-Sansone, C. Trefzger, M. Lewenstein, P. Zoller, and G. Pupillo. *Phys. Rev. Lett.*, 104:125301, 2010.

[63] K. Goral, L. Santos, and M. Lewenstein. *Phys. Rev. Lett.*, 88:170406, 2002.

[64] D.-W. Wang, M. D. Lukin, and E. Demler. *Phys. Rev. Lett.*, 97:180413, 2006.

[65] A. Pikovski, M. Klawunn, G. V. Shlyapnikov, and L. Santos. *Phys. Rev. Lett.*, 105:215302, 2010.

[66] M.A. Baranov, A. Micheli, S. Ronen, and P. Zoller. *arXiv*, 1012:5589v1, 2010.

[67] R.V. Krems. *Int. Rev. Phys. Chem.*, 24:99, 2005.

[68] S. Ospelkaus, K. K. Ni, D. Wang, M. H. G. de Miranda, B. Neyenhuis, G. Quemener, P. S. Julienne, J. L. Bohn, D. S. Jin, and J. Ye. *Science*, 327:853, 2010.

[69] D. Jin and J. Ye. *Physics Today*, 64:27, 2011.

[70] R. V. Krems. *Phys. Chem. Chem. Phys.*, 10:4079, 2008.

[71] O. Dutta, M. Jaaskelainen, and P. Meystre. *Phys. Rev. A*, 71:051601, 2005.

[72] F. M. Crompvoets, H. L. Bethlem, R. T. Jongma, and G. Meijer. *Nature (London)*, 411:6834, 2001.

[73] A. J. Kerman. *Phys. Rev. Lett.*, 92:033004, 2004.

[74] K. K. Ni, S. Ospelkaus, M. H. G. de Miranda, A. Pe'er, B. Neyenhuis, J. J. Zirbel, S. Kotochigova, P. S. Julienne, D. S. Jin, and J. Ye. *Science*, 322:231, 2008.

[75] D. Wang, J. Qi, M. F. Stone, O. Nikolayeva, B. Hattaway, S. D. Gensemer, H. Wang, W. T. Zemke, P. L. Gould, E. E. Eyler, and W. C. Stwalley. *Eur. Phys. J. D*, 31:165, 2004.

[76] M.W. Mancini, G. D. Telles, A. R. L. Caires, V. S. Bagnato, and L. G. Marcassa. *Phys. Rev. Lett.*, 92:133203, 2004.

[77] S. Ospelkaus, A. Peer, K.-K. Ni, J. J. Zirbel, B. Neyenhuisand S. Kotochigova, P. S. Julienne, J. Ye, and D. S. Jin. *Nature Phys.*, 4:622, 2008.

[78] C. Haimberger, J. Kleinert, M. Bhattacharya, and N. P. Bigelow. *Phys. Rev. A.*, 70:21402, 2004.

[79] J. Deiglmayr, M. Aymar, R. Wester, M. Weidemüller, and O. Dulieu. *J. Chem. Phys.*, 129:064309, 2008.

[80] S. D. Kraft, P. Staanum, J. Lange, L. Vogel, R. Wester, and M. Weidemüller. *J. Phys. B: At. Mol. Opt. Phys.*, 39:S993, 2006.

[81] J. G. Danzl, E. Haller, M. Gustavsson, M. J. Mark, R. Hart, N. Bouloufa, O. Dulieu, H. Ritsch, and H.-C. Nägerl. *Science*, 321:1062, 2008.

[82] S. Inouye, J. Goldwin, M. L. Olsen, C. Ticknor, J. L. Bohn, and D. S. Jin. *Phys. Rev. Lett.*, 93:183201, 2004.

[83] A. Tkatchenko and M. Scheffler. *Phys. Rev. Lett.*, 102:073005, 2009.

[84] M. Dion, H. Rydberg, E. Schroder, D. C. Langreth, and B. I. Lundqvist. *Phys. Rev. Lett.*, 92:246401, 2004.

[85] A. Halkier, W. Klopper, T. Helgaker, and P. J ø rgensen. *J. Chem. Phys.*, 111:4424, 1999.

[86] P. J. Reynolds, D. M. Ceperley, B. J. Alder, W. A. Lester, and Ji. *J. Chem. Phys.*, 77:5593, 1982.

[87] M. Bajdich and L. Mitas. *Acta Physica Slovaca*, 59:81, 2009.

[88] F. Schautz and H.-J. Flad. J. chem. phys. 110:11700, 1999.

[89] Shih-I Lu. J. chem. phys. 114:3898, 2001.

[90] L. Wagner and L. Mitas. J. chem. phys. 126:034105, 2007.

[91] R. N. Barnett, P. J. Reynolds, W. A. Lester, and Jr. J. chem. phys. 96:2141, 1992.

[92] R. N. Barnett, P. J. Reynolds, W. A. Lester, and Jr. *J. Comp. Phys.*, 96:258, 1991.

[93] K. S. Liu, M. H. Kalos, and G. V. Chester. *Phys. Rev. A*, 10:303, 1974.

[94] S. Kotochigova, A. Petrov, M. Linnik, J. Klos, and P. S. Julienne. *J. Chem. Phys.*, 135:164108, 2011.

[95] R. Guerout, M. Aymar, and O. Dulieu. *Phys. Rev. A*, 82:042508, 2010.

[96] Enrique R. Batista, Richard L. Martin, and P. Jeffrey Hay. *J. Chem. Phys.*, 121:11104, 2004.

[97] W.A.De Jong and W.C.Nieuwpoort. *International Journal of Quantum Chemistry*, 58:203, 1996.

[98] K.H.Lau and D.L.Hildenbrand. *J. Chem. Phys*, 93:6124, 1990.

[99] D. L. Hildenbrand and K. H. Lau. *J. Chem. Phys.*, 93:5893, 1990.

[100] N. D. Drummond and R. J. Needs. *Phys. Rev. B*, 78:125106, 2008.

[101] N. D. Drummond, R. J. Needs, A. Sorouri, and W. M. C. Foulkes. *Phys. Rev. B*, 78:125106, 2008.

[102] L. M. Fraser, W. M. C. Foulkes, G. Rajagopal, R. J. Needs, S. D. Kenny, and A. L. Williamson. *Phys. Rev. B*, 53:1814, 1996.

[103] A. Ambrosetti, F. Pederiva, and E. Lipparini. *Phys. Rev. B*, 80:125306, 2009.

[104] Claudio Attaccalite, Saverio Moroni, Paola Gori-Giorgi, and Giovanni B. Bachelet. *Phys. Rev. Lett.*, 88:256601, 2002.