Scuola Internazionale Superiore di Studi Avanzati - Trieste

- ma per seguir virtute e conoscenza -

RVB phase of hydrogen at high pressure: towards the first *ab-initio* Molecular Dynamics by Quantum Monte Carlo

Thesis submitted for the degree of

Philosophiæ Doctor

Candidate:
Claudio Attaccalite

Supervisor:
Prof. S. Sorella

October, 24th 2005

To myself

I want acknowledge:

Sandro Sorella for his patience with me. In these years he taught many things but I spent enough time to understand them (especially his QMC code);

Irina Yegorova: My sweet ljubimaja, she corrected my thesis, and promised to listen to my presentation, poor girl;

Lorenzo Stella: My personal “contarolo”, since he helped me many many times. It’s unbelievable, but he always has an answer to any question. I don’t know if I have to thank him for his sense of humor, for sure, this halved the scientific production in our room;

Michele Casula: Because he was always ready to help me, and because I had a good collaboration with him (I’m not sure he will think the same), but I want to complain one thing about him: He speaks too little, and also his girlfriend agrees with me!

Tania Zykova Timan: Because she is whimsical, wayward and full of glee for the life. She often made me happy in bad moments.

Hamed Seyed Allaei: A very good friend patient and wise. He taught me English, and he is still continuing to do it at this moment. I want to complain about him that he never cut his beard.

Joachim Almergrin: Because he is always happy (I don’t know how) even when he was fired, and because he was a master in knowing new girls and new friends.

Vittoria Colizza: My dear “amichetta”, I was lucky to live with her two times during this PhD, and I hope we will be always friends, I will never forget her strong character.

Sara Bertocco: My second ”mamma”, she always took care of me and helped in difficult moments.

Fabio Affinito: for nice internet chatting during these years.

Davide Ceresoli and Ugo Tartaglino: For the clarity and simplicity of their explanations, and for all the lunches together.

Erio Tosatti: For useful discussions, probably I should have discussed more with him.

The Roman Guys: Ciccio, Nicola, Piero, Adriano, Matteo, Laura, Luca and many others for good weekends spent in Rome.

My family, my sister and my brother: they always were close to me in these years.

Fabbio Giovi, Francesca, Sara, Stefano, Manuela, Evgeny, Oleg (for teaching me many bad russian words), Simona, Federica, Giacomo, Lorenzo, Chiara and all people in Adriatico caffetteria: Sandro, Patrizia, Tiziano and Lucia.

Cyrus Umrigar for correcting English of my acknowledgement.

Contents

Introduction **1**

Thesis outline **3**

Intro. **1**

1 Quantum Monte Carlo and the JAGP wave-function **5**
- 1.1 Variational Monte Carlo 5
  - 1.1.1 Forces with finite variance 6
- 1.2 Functional form of the wave function 8
  - 1.2.1 Pairing determinant 9
  - 1.2.2 One body term 12
  - 1.2.3 Two body Jastrow term 12
  - 1.2.4 Three Body Jastrow term 13

2 Optimization Methods **15**
- 2.1 Stochastic Reconfiguration 16
  - 2.1.1 Setting the SR parameters 19
  - 2.1.2 Stabilization of the SR technique 23
- 2.2 Structural optimization 25
- 2.3 Hessian Optimization 28

3 Results on Molecules **33**
- 3.1 Application of the JAGP to molecules 33
- 3.2 Small diatomic molecules, methane, and water 34
- 3.3 Benzene and its radical cation 38

4 Quantum Monte Carlo on extended systems 47

4.1 Periodic Wave-Function 48

4.1.1 Periodic orbitals 49

4.1.2 The wave-function for high pressure hydrogen 50

4.2 Coulomb Interactions in periodic systems 51

4.2.1 Ewald Sums 51

4.2.2 Forces with finite variance in periodic systems 54

4.3 How to evaluate pressure 55

4.4 Empirical laws of melting 56

5 A new technique for the simulation of electronic systems at finite temperature by means of noisy QMC forces 59

5.1 The Born-Oppenheimer approximation 60

5.2 Dealing with Quantum Monte Carlo noise 61

5.3 Canonical ensemble by Generalized Langevin Dynamics 62

5.3.1 Numerical integration of the Generalized Langevin Equation 64

5.4 Practical implementation of the finite temperature dynamics . . . 66

5.4.1 Setting the parameters in the Langevin dynamics 68

5.4.2 Following the ionic dynamics 69

5.4.3 Reducing the number of parameters 69

6 Preliminary results on high pressure hydrogen 71

6.1 Comparison with previous calculations 71

6.2 Pair Correlation Functions 74

6.3 Another possible phase in liquid hydrogen 77

7 Conclusions 83

7.1 Future developments 85

A Re-sampling methods 87

A.1 Jackknife 87

B Local Energy and its derivatives 89

B.1 Kinetic Energy 89

B.1.1 Derivatives of the Kinetic Energy 89

B.2 Pairing determinant 90
B.3 Three-body 94

C Cusp conditions 97

D Determinant derivatives 99

E Error Analysis due to finite time step in the GLE integration in a simple case 101

References 113

List of Tables

3.1 Total energies in variational (E_{VMC}) and diffusion (E_{DMC}) Monte Carlo calculations; the percentages of correlation energy recovered in VMC ($E_c^{VMC}(\%)$) and DMC ($E_c^{DMC}(\%)$) have been evaluated using the "exact" (E_0) and Hartree–Fock (E_{HF}) energies from the references (1). Here "exact" means the ground state energy of the non relativistic infinite nuclear mass Hamiltonian. The energies are in *Hartree*. 34

3.2 Bond lengths (R) in atomic units; the subscript 0 refers to the "exact" results. For the water molecule R is the distance between O and H and θ is the angle HOH (in deg), for CH_4 R is the distance between C and H and θ is the HCH angle. 35

3.3 Binding energies in eV obtained by variational (Δ_{VMC}) and diffusion (Δ_{DMC}) Monte Carlo calculations; Δ_0 is the "exact" result for the non-relativistic infinite nuclear mass Hamiltonian. Also the percentages ($\Delta_{VMC}(\%)$ and $\Delta_{DMC}(\%)$) of the total binding energies are reported. 36

3.4 Binding energies in eV obtained by variational (Δ_{VMC}) and diffusion (Δ_{DMC}) Monte Carlo calculations with different trial wave functions for benzene. In order to calculate the binding energies yielded by the 2-body Jastrow we used the atomic energies reported in Ref. (2). The percentages ($\Delta_{VMC}(\%)$ and $\Delta_{DMC}(\%)$) of the total binding energies are also reported. 38

3.5 Bond lengths (r) for the two lowest ${}^2B_{2g}$ and ${}^2B_{3g}$ states of the benzene radical cation. The angles α are expressed in degrees, the lengths in a_0. The carbon sites are numerated from 1 to 6. . . . 39

3.6 Total energies for the $^2B_{2g}$ and $^2B_{3g}$ states of the benzene radical cation after the geometry relaxation. A comparison with a BLYP/6-31G* and SVWN/6-31G* all-electron calculation (Ref. (3)) is reported. 44
3.7 Adiabatic ionization potential of the benzene molecule; our estimate is done for the $^2B_{3g}$ relaxed geometries of the benzene radical cation, with an inclusion of the zero point motion correction between the $^2B_{3g}$ state and the $^1A_{1g}$ neutral molecule ground state, calculated in Ref. (4) at the B3LYP/6-31G* level. 45

6.1 Total energies in variational (E_{VMC}) and diffusion (E_{DMC}) Monte Carlo calculations for 16 hydrogen atoms in a BCC lattice at Rs=1.31 and T=0 (i.e. frozen ion positions). The energies are in *Hartree* for atom. 71
6.2 Pressure at different temperatures and densities. We report also the pressure obtained with Gasgun experiment (5), with Silvera-Goldman empirical potential model (6) and CEICM method (7) at Γ point. The pressure are in GPa. 74

List of Figures

2.1 Example of the convergence of the SR method for the variational parameters as a function of the number of stochastic iterations. In the upper(lower) panel the Jastrow (geminal) parameters are shown. For each iteration, a variational Monte Carlo calculation is employed with a bin containing 15000 samples of the energy, yielding at the equilibrium a standard deviation of $\simeq 0.0018H$. For the first 200 iteration $\Delta t = 0.00125H^{-1}$, for the further 200 iterations $\Delta t = 0.0025H^{-1}$, whereas for the remaining ones $\Delta t = 0.005H^{-1}$. 20

2.2 Calculation of the derivative of the energy with respect to the second Z in the $2p$ orbital of the geminal function for the Be atom. The calculation of the force was obtained, at fixed variational parameters, by averaging over 10^7 samples, allowing e.g. a statistical accuracy in the total energy of $0.07mH$. The variational parameters have been obtained by an SR minimization with fixed bin length shown in the x label. The parameter considered has the largest deviation from the Euler conditions. 22

2.3 Plot of the equilibrium distance of the Li_2 molecule as a function of the inverse bin length. The total energy and the binding energy are reported in Tables 3.3 and 3.2 respectively. For all simulations the initial wave-function is optimized at $Li - Li$ distance 6 a.u. . . 27

3.1 Electron density (atomic units) projected on the plane of C_6H_6. The surface plot shows the difference between the resonating valence bond wave function, with the correct $A1g$ symmetry of the molecule, and a non-resonating one, which has the symmetry of the Hartree Fock wave function. 40

3.2 Surface plot of the charge density projected onto the molecular plane. The difference between the non-resonating (indicated as HF) and resonating Kekulé 3-body Jastrow wave function densities is shown. Notice the corresponding change from a dimerized structure to a C_6 rotational invariant density profile. 42

3.3 Plot of the convergence toward the equilibrium geometry for the ${}^2B_{2g}$ acute and the ${}^2B_{3g}$ obtuse benzene cation. Notice that both the simulations start form the ground state neutral benzene geometry and relax with a change both in the $C-C$ bond lengths and in the angles. The symbols are the same of Tab. 3.5. 43

4.1 A simulation box with periodic boundary conditions. 47

5.1 Ionic dynamics of 54 hydrogen atoms using GLQ, with a time step $0.4fs$, starting from a BCC lattice. The trial wave-function contains 2920 variational parameters and we have optimized 300 of them at each step. In the inset the maximum deviation $F_i/\Delta F_i$ of the forces acting on the variational parameters is shown. 67

6.1 Energy per atom of 16 hydrogen atoms at Rs=1.31 calculated on configurations obtained by CEIMC with the method (8). The first 10 configurations are in the atomic liquid phase at 2000k while in the last ten the system is forming clusters at T=500. 72

6.2 Variance per atom of 16 hydrogen atoms at Rs=1.31 calculated on configuration obtained by CEIMC with the method (8). The first 10 configurations are in the atomic liquid phase at 2000k while in the last ten the system is forming clusters at T=500. 73

6.3 Proton-proton correlation function, g(r), at Rs=1.31. The GLQ and CEIMC have used a periodical simulation box with 32 atoms while Hohl et al. with 64 atoms. All the calculations were performed for a single k point (Γ). 75
6.4 Comparison of the proton-proton correlation function, g(r), at Rs=2.1 and T=4350 obtained with different methods CEIMC (8) (7) and GLQ. All the simulations were performed with 32 atoms for a single k point (Γ). 76
6.5 Variational and Diffusion Condensation Energy per atom 79
6.6 Eigenvalues of the λ matrix for 16 hydrogen atoms at Rs=1.31 and 100K as function of the simulation time 80
6.7 Off-Diagonal Long Range Order for 16 hydrogen atoms at Rs=1.31 and 100K, in a box of size $L = 5.3211$, for $\theta = 0$ as function of the distance x 81
6.8 Off-Diagonal Long Range Order for 54 hydrogen atoms at Rs=1.31 at 100K in a box of size $L = 7.9817$, for $\theta = 0$ as function of the distance x 82

Introduction

There is an oral tradition that says shortly after the Shrödinger equation was validated on H atom and H_2 P.A.M. Dirac exclaimed that the chemistry had come to an end. Fortunately some years later, in 1929, he added ”The fundamental laws necessary for the mathematical treatment of a large part of physics and the whole of chemistry are thus completely known, and the difficulty lies only in the fact that application of these laws leads to equations that are too complex to be solved.”.

During the last 70 years, staring from the Thomas-Fermi theory, there were proposed many ways to solve approximately the Shrödinger equation with several electrons. In 1999 W. Kohn and J.Sham received the Nobel prize for the well known Densitiy Functional Theory. They found an innovative way to deal with many electron system using a three-dimensional electronic density instead of a 3N-dimension wave-function. Their approach have opened the possibility to simulate thousands of different systems with an affordable computer effort.

Several other techniques, different from DFT, exist to deal with the time-independent Shrödinger equation with many electrons. The so colled Quantum Monte Carlo (QMC) techniques (9; 10; 11) are some of the most accurate and efficient statistical methods for treating many body quantum systems. In the past they were used to study different systems as quantum liquids (12; 13; 14), atoms (2), molecules (15; 16), solids (17) and lattice models for strongly correlated electron systems (18). Moreover QMC allows to include many important physical informations of the system, such as cusp conditions, symmetries, exact limits, because it deals directly with the many body wave-function. In this thesis we have used these techniques to study molecular systems by introducing a new highly correlated wave-function (19).

Although QMC have led to great progress in understanding the zero temperature physics of strongly correlated electron systems, there is no obvious way to extend it to finite low temperature. Even if there exist Monte Carlo techniques, as Path Integral Monte Carlo, to study quantum system at finite temperature, they are limited to the high temperature regime.

In this thesis following the idea of Car and Parrinello we present a new approach to study many electron systems at low temperature using a classical ionic dynamic combined with a ground state QMC for the electrons. This method is then applied to study high pressure hydrogen.

Hydrogen is the most abundant element in the universe. Despite the simple structure of hydrogen atom, it does not form the simplest solids or liquids. It has a very complex phase diagram that has been widely studied by experimental (20) and theoretical approaches (8; 21).

At low pressure hydrogen crystallizes as an insulating molecular solid. As the pressure increase different molecular phases were encountered (20; 22). At higher pressure electron would no longer remain in localized bound orbitals and would instead delocalize. As predicted by the pioneering work of Wigner and Huntington (1935), at extreme pressure, comparable with the ones present in giant planets, the molecules of solid hydrogen will dissociate to form a mono-atomic metallic solid. Despite the simple interaction, the structure of this solid has been predicted to be a simple hexagonal lattice (23). At higher density some theoretical models has suggested that compressed hydrogen can form unusual two-component quantum fluid, made by electrons and protons, at low and even zero temperature (24; 25). Moreover a recent study (21) predicted a maximum in the melting curve between solid and liquid in high pressure hydrogen, confirming the idea of a possible stable quantum fluid phase at low temperature.

Motivated by the so reach phase diagram we decided to investigate the effects of the correlation in high pressure hydrogen, and to this end, we decided to use a resonating valence bond (RVB) wave-function together with a new technique to simulate finite temperature systems. In the RVB approach the variational trial-function is written as an antisymmetrized combination of bonds. Each bond contains two electrons shared by two orbitals. In fact after the original proposal by Anderson, there is now a large amount of numerical evidence that the simple but general resonating valence bond (RVB) wave function contains just those ingredients missing in uncorrelated theories, so that the main features of electron correlation can be captured by the variational RVB approach. Moreover from the computational point of view the remarkable feature of this approach is that several resonating valence bonds can be dealt simultaneously with a single determinant, at a computational cost growing with the number of electrons similarly to more conventional methods, such as Hartree-Fock or Density Functional Theory.

Thesis outline

The thesis is organized as follows:

- In the first chapter we briefly review QMC methods mainly used in this thesis, the Variational Monte Carlo. Then we introduce the pairing wave-function used to study molecular and solid systems.
- In the second chapter two optimization methods, used in VMC, are presented: Stochastic Reconfiguration(SR) and Stochastic Reconfiguration with Hessian acceleration (SRH). Advantages, limitations, tips and tricks are shown for both of them.
- In the third chapter we show the results obtained by applying the paring wave-function to different molecular systems.
- In the beginning of the fourth chapter we show how to generalize the pairing function to study extended systems. Then we come back for a while to technical aspects of simulation of periodic systems.
- In the fifth chapter a new method to simulate systems at finite temperature is presented. This technique has allowed us to perform ionic dynamics using noisy forces coming from Quantum Monte Carlo.
- In the last chapter we show preliminary results obtained on high-pressure hydrogen, using our new technique. Moreover our results allowed us to guess a new possible exotic phase in high pressure hydrogen driven by the electronic correlation.

Chapter 1

Quantum Monte Carlo and the JAGP wave-function

1.1 Variational Monte Carlo

The VMC is a stochastic method that allows to evaluate expectation values of physical operators on a given wave-function (see Ref. (26)). It is based on a statistical calculation of the integrals that involve the mean values of these operators.

$$\langle A \rangle_{var} = \frac{\int \Psi_T^*(R) \hat{A} \Psi_T(R) dR}{\int \Psi_T^*(R), \Psi_T(R) dR} \tag{1.1}$$

where $R = (r_1, r_2, r_3, ...r_N)$ are the electron coordinates. Monte Carlo integration is necessary because the wave-function contains explicit particle correlations and this leads to non-factoring multi-dimension integrals. Notice that in the case of the Hamiltonian operator, according to the variational principle, the expectation value will be greater than or equal to the exact ground state energy. We can write the integral 1.1 as:

$$\langle A \rangle_{var} = \int P(R) A_L(R) dR$$

where $P(R)$ in a probability density, and $A_L(R)$ is the diagonal element associated to the operator $\hat{A}$.

$$P(R) = \frac{|\Psi_T(R)|^2}{\int \Psi_T^*(R) \Psi_T(R) dR} \tag{1.2}$$

$$A_L(R) = \frac{\hat{A} \Psi_T(R)}{\Psi_T(R)} \tag{1.3}$$

We can sample the probability distribution $P(R)$ using the Metropolis scheme (27) and then evaluate $A_L(R)$ on the obtained configurations. Then using the central limit theorem the integral can be estimate as:

$$\langle A\rangle_{var} = \frac{1}{M}\sum_{m=1}^{M} A_L(R_m)$$

The sampling process continues until the desired statistical error on the expectation value of the operator $\hat{A}$ is reached.

The VMC algorithm is implemented so that only a single electron is moved at each time. In this way only one column or one row of the Pair Determinant is changed at each step. The new determinant can be computed in $\mathcal{O}(N)$ operations, given the inverse of the old pair determinant. This inverse is computed once at the beginning of the simulation and then updated whenever a trial move is accepted. If the trial move is accepted, the inverse matrix is updated in $\mathcal{O}(N^2)$ operations. This trick makes the VMC sampling very efficient. Notice that the direct computation of a determinant takes $\mathcal{O}(N^3)$ operations.

1.1.1 Forces with finite variance

In VMC the expectation value of operators different from the Hamiltonian is usually much less favorable and accurate than the one obtained for the energy. This is due to two kinds of errors: first the statistical one due to the finite sampling in the Monte Carlo integration that behaves as $1/\sqrt{N}$, where N is the number of sampling points and second the systematic error (”bias”) resulting from an approximated wave-function.

In order to understand the behavior of this errors we define the trial wave-function error, $\delta\Psi = \Psi_T - \Psi_0$, where Ψ_0 is the exact wave function. In the case of the energy, applying the variational principle (see for instance ref. (28)), one finds that the systematic error Δ_E goes as $Delta_E \sim \mathcal{O}(\delta\Psi^2)$, where Δ_E can be represented as:

$$\Delta_E = \frac{\langle \Psi_T - \Psi_0 | H - E_0 | \Psi_T - \Psi_0\rangle}{\langle \Psi_T | \Psi_T\rangle}. \tag{1.4}$$

Instead the statistical error is related to the variance of the operator on the trial wave-function Ψ_T. For instance for the energy:

$$\sigma^2(E_L) = \langle (E_L - E_v)^2\rangle_{\Psi_T^2}. \tag{1.5}$$

Using the equality:

$$E_L - E_v = \frac{(H - E_0)(\Psi_T - \Psi_0)}{\Psi_T} - \Delta_E \tag{1.6}$$

it is easy to see that $\sigma^2(E_L) = \mathcal{O}(\delta\Psi^2)$. Thus in the case of the energy both these errors vanish as $\mathcal{O}(\delta\Psi^2)$. For any other operator that not commutes with the Hamiltonian, Ψ_0 is not anymore an eigenstate of $\hat{O}$ and so the systematic error is $\Delta_O = \mathcal{O}(\delta\Psi)$ while the statistical one is $\sigma^2(O) = \mathcal{O}(1)$ (see Ref. (29; 30).

The situation is even worst if we consider atomic or molecular forces. In fact, let us derive the potential energy in the respect to an atomic position:

$$F_A^\nu = -\frac{\partial}{\partial R_i^x} V(r_1, .., r_N; R_1, ..., R_m) = -Z_A \sum_{i \neq A}^{M} \frac{Z_i (R_A^\nu - R_i^\nu)}{R_{Ai}^3} - Z_A \sum_{j}^{N} \frac{(\vec{x}_j - \vec{R}_A^\nu)}{|r_j - R_A|^3}, \tag{1.7}$$

the second term in the right-hand side of this equation is responsible for a infinite variance contribution.

In order to overcome this problem Assaraf and Caffarel (31) proposed an original and ingenious solution. Denoting $\hat{O}$ an arbitrary hermitian operator they showed that is possible to define a new ”renormalized” operator $\tilde{O}$ such that:

$$\langle \tilde{O} \rangle = \langle \hat{O} \rangle \tag{1.8}$$

$$\sigma^2(\tilde{O}) \leq \sigma^2(\hat{O}). \tag{1.9}$$

The new operator $\tilde{O}$ is obtained from the old one by adding to O another operator with zero expectation value and finite variance, namely:

$$\tilde{O} = O + \left[\frac{\tilde{H}\tilde{\Psi}}{\tilde{\Psi}} - \frac{\tilde{H}\Psi_T}{\Psi_T} \right] \frac{\tilde{\Psi}}{\Psi_T}, \tag{1.10}$$

where $\tilde{H}$ is an arbitrary Hermitian operator, and $\tilde{\Psi}$ is an auxiliary square-integrable function. In the case of atomic forces, the simplest and effective choice for $\tilde{H}$ and $\tilde{\Psi}$ is:

$$\tilde{H} = H \tag{1.11}$$

$$\tilde{\Psi} = Q\Psi_T \tag{1.12}$$

with

$$Q_A^\nu = Z_A \sum_{i=1}^{N_{elect}} \frac{(x_i^\nu - R_A^\nu)}{\left|\vec{r}_i - \vec{R}_A\right|}. \tag{1.13}$$

This particular form cancels the pathological part in the bare force 1.7. The renormalized force reads:

$$\tilde{F}_A^\nu = -Z_A \sum_{i \neq A}^{M} \frac{Z_i(R_i^\nu - R_A^\nu)}{R_{Ai}^3} - \frac{\vec{\nabla} Q_A^\nu \cdot \vec{\nabla} \Psi_T}{\Psi_T}. \tag{1.14}$$

Notice that the infinite variance contribution in the bare force 1.7 no longer appears in the latter expression, indeed the new ”renormalized” force 1.14 has now a finite variance. The use of ”renormalized” operators has allowed us to evaluate forces with a finite variance and to perform structural optimization and finite temperature dynamics.

1.2 Functional form of the wave function

In both variational Monte Carlo (VMC) and diffusion Monte Carlo (DMC) (for a review about DMC see Ref. (26)) the trial-function completely determines the quality of the approximation for the physical observables. Because of this, it is extremely important to choose carefully a flexible wave-function that contains as much knowledge as possible of the physics of the system being studied.

In the first part of this thesis we proposed an highly correlated wave function that is able to capture the major part of the correlation energy: The Antisymmetric Geminal Product supplemented by the Jastrow correlation (JAGP). This wave function is an extension of the Antisymmetric Geminal Product AGP, introduced in quantum chemistry by Coleman (32):

$$\Psi_{AGP}(r_1, ..., r_N) = \hat{A}\Pi_{i=1}^{N/2} \Phi(r_{2i}, r_{2i-1}) \tag{1.15}$$

where $\hat{A}$ is the antisymmetrization operator. The AGP wave-function is determined by the geminal, which is usually expanded in a one-particle basis:

$$\Phi(r_i, r_j) = \sum_{1 \leq l,m \leq r} \lambda_{lm} \phi_l(r_i) \phi_m(r_j) \tag{1.16}$$

where r is the size of the orbital basis set. The geminal is then determined by $r(r-1)/2$ coefficients λ. For instance, for the simple hydrogen molecules, using only two orbitals as basic set, the AGP is:

$$\Psi_{H_2} = \lambda_{11}\phi_{1s}^A(r_1)\phi_{1s}^A(r_2) + \lambda_{22}\phi_{1s}^B(r_1)\phi_{1s}^B(r_2) + \lambda_{12}\phi_{1s}^A(r_1)\phi_{1s}^B(r_2) + \lambda_{21}\phi_{1s}^B(r_1)\phi_{1s}^A(r_2) \tag{1.17}$$

where Ψ_{H_2} contains bonding and anti-bonding orbitals and so it is able to reproduce the Heithler-London limit. Notice that AGP wave-function is similar to the Gutzwiller BCS wave function used on lattice system (33).
The full JAGP wave-function is defined by the product of different terms, namely one-body, two-body, three-body Jastrow J_1, J_2, J_3 and an antisymmetric part ($\Psi = J\Psi_{AGP}$). The first term is used to satisfy the nuclear cusp conditions, while the second the electron-electron one. The third one is an explicit contribution to the dynamic electronic correlation, and the latter is able to treat the non-dynamic one arising from near degenerate orbitals through the geminal expansion. Therefore our wave function is highly correlated and it is expected to give accurate results on widely range of systems.

1.2.1 Pairing determinant

As it is well known, a simple Slater determinant provides the exact exchange electron interaction but neglects the electronic correlation, which is by definition the missing energy contribution. In the past different strategies were proposed to go beyond Hartee-Fock theory. In particular a sizable amount of the correlation energy is obtained by applying to a Slater determinant a so-called Jastrow term, that explicitly takes into account the pairwise interaction between electrons.
On the other hand, within the Quantum Chemistry community the Antisymmetric Geminal Product (AGP) is a well known ansatz to improve the HF theory, because it implicitly includes most of the double-excitations of an HF state.

Recently a new trial function was proposed for atoms, that includes both the terms (2). In the first part of this thesis we extend this promising approach to a number of small molecular systems with known experimental properties, that are commonly used for testing new numerical techniques.

The major advantage of this approach is the inclusion of many CI expansion terms with the computational cost of a single determinant. For instance this has allowed us to perform the full structural optimization of benzene without a particularly heavy computational effort on a single processor machine.

For an unpolarized system containing N electrons (the first $N/2$ coordinates are referred to the up spin electrons) the AGP wave function is a $\frac{N}{2} \times \frac{N}{2}$ pairing matrix determinant, which reads:

$$\Psi_{AGP}(\vec{r}_1, ..., \vec{r}_N) = \det\left(\Phi_{AGP}(\vec{r}_i, \vec{r}_{j+N/2})\right). \tag{1.18}$$

Here the geminal function is expanded over an atomic basis:

$$\Phi_{AGP}(\vec{r}^{\uparrow}, \vec{r}^{\downarrow}) = \sum_{l,m,a,b} \lambda^{l,m}_{a,b} \phi_{a,l}(\vec{r}^{\uparrow}) \phi_{b,m}(\vec{r}^{\downarrow}), \tag{1.19}$$

where indices l, m span different orbitals centred on atoms a, b, and i,j are coordinates of spin up and down electrons respectively.

Differently from the previous pairing function formulation (2), appropriate only for simple atoms, here also off-diagonal elements are included in the λ matrix, which must be symmetric in order to define a spin singlet state. Moreover this formulation allows to easily fulfill other symmetries by imposing the appropriate relations among different $\lambda_{l,m}$. For instance in homo-nuclear diatomic molecules, the invariance under reflection in the middle plane perpendicular to the molecular axis yields the following relation:

$$\lambda^{a,b}_{m,n} = (-1)^{p_m + p_n} \lambda^{b,a}_{m,n}, \tag{1.20}$$

where p_m is the parity under reflection of the $m-$th orbital.

An important property of this formalism is the possibility to describe explicitly resonating bonds present in many structures, like benzene. A $\lambda^{a,b}_{m,n}$ different from zero represents a chemical bond formed by the linear combination of the *m*-th and *n*-th orbitals belonging to *a*-th and *b*-th nuclei. It turns out that resonating bonds can be well described through the geminal expansion by switching on the appropriate $\lambda^{a,b}_{m,n}$ coefficients: the relative weight of each bond is related to the amplitude of the corresponding λ.

Also polarized systems can be treated within this framework, by using the spin generalized version of the AGP (GAGP), in which also the unpaired orbitals are expanded as well as the paired ones over the same atomic basis employed in the geminal (34).

Another important property of AGP wave-function is the size consistency: if we smoothly increase the distance between two regions A and B, each containing a given number of atoms, the many-electron wave function Ψ factorizes into the product of space-disjoint terms $\Psi = \Psi_A \bigotimes \Psi_B$ as long as the interaction between the electrons coupling the different regions A and B can be neglected. In this limit the total energy of the wave function approaches the sum of the energies corresponding to the two space-disjoint regions. This property, that is obviously valid for the exact many-electron ground state, is not always fulfilled by a generic variational wave function as for instance configuration-interaction (CI) wave-function. Notice that this property is valid when both the compound and the separated fragments have the minimum possible total spin. This is precisely the relevant case for hydrogen phase diagram studied in this thesis because we have not studied ferromagnetic or partially ferromagnetic phases, that are not believed to be present in the reasonable pressure-temperature range of hydrogen (for a discussion about ferromagnetism in high-pressure hydrogen see Ref. (35)).

Now we want to highlight how it is possible to implement nuclear cusp condition (see Appendix C) for molecular systems with a pairing wave-function. A straightforward calculation shows that the AGP wave function fulfills the cusp conditions around the nucleus a if the following linear system is satisfied:

$$\sum_{j}^{(1s,2s)} \lambda_{a,b}^{j,j'} \hat{\phi}'_{a,j}(\mathbf{r} = \mathbf{R}_a) = -Z_a \sum_{c,j} \lambda_{c,b}^{j,j'} \phi_{c,j}(\mathbf{r} = \mathbf{R}_a), \tag{1.21}$$

for all b and j'; in the LHS the caret denotes the spherical average of the orbital gradient. If we impose that the orbitals satisfy the atomic cusp condition on their atom, this equation reduce to:

$$\sum_{c(\neq a),j} \lambda_{c,b}^{j,j'} \phi_{c,j}(\mathbf{R}_a) = 0, \tag{1.22}$$

and because of the exponential orbital damping, if the nuclei are not close together each term in the previous equations is very small, of the order of $\exp(-|\mathbf{R}_a - \mathbf{R}_c|)$. Therefore in the first part of this thesis, with the aim of making the optimization faster, we have chosen to use $1s$ and $2s$ orbitals satisfying the atomic cusp conditions and to disregard the sum (1.22). In this way, once the energy minimum is reached, also the molecular cusp conditions (1.21) are rather well satisfied. Later in the second part of the thesis we have adopted a different and more efficient strategy to the cusp problem as described in the following section.

1.2.2 One body term

Another important term of our trial-function it is the one-body term. In fact as pointed out in Ref. (36), it is easier to optimize a one-body term explicitly rather than including more orbitals in the determinantal basis set.
Moreover, even if it is possible to satisfy nuclear cusp conditions (see Appendix C) with the pairing determinant, this has to be done iteratively during the optimization process adding constraints to the variational parameters or approximately disregarding the term of the eq.1.22. In order to solve efficiently this problem we included nuclear cusp conditions explicitly in the one-body term, in the same way of ref. (37):

$$J_1(\vec{r}_1, ..., \vec{r}_N) = \exp\left[\sum_{i,a}^{N} (\xi_a(\vec{r}_i) + \Xi_a(\vec{r}_{ia}))\right], \tag{1.23}$$

where $\xi_a(\vec{r}_i)$ orbital is used to satisfy the nuclear cusp conditions on nucleus a:

$$\xi_a(r) = \frac{-Z_a r}{(1 + br)} \tag{1.24}$$

and the $\Xi(\vec{r}_{ia}) = \sum_l \lambda_l \psi_{a,l}(\vec{r}_{ia})$ is a linear combination of atomic orbitals centered on the nucleus a, and that do not effect the nuclear cusp condition. We have used Gaussian and exponential orbitals such to have a smooth behaviour close to the corresponding nuclei, namely as:

$$\psi_{a,i}(\vec{r}) - \psi_{a,i}(\vec{R}_a) \simeq |\vec{r} - \vec{R}_a|^2, \tag{1.25}$$

or with larger power, in order to preserve the nuclear cusp conditions (1.24).
The basis set $\psi_{a,i}(\vec{r})$ is the same used in the so-called three-body term that we are going to describe in the following. The same kind of behavior has been imposed for the orbitals appearing in the determinant. In this way the nuclear cusp conditions are very easily satisfied for a general system containing many atoms, in a simple and efficient way.

1.2.3 Two body Jastrow term

As it is well known, the Jastrow term plays a crucial role in treating many body correlation effects. One of the most important correlation contribution arises from the electron-electron interaction. Therefore it is important to use at least a two-body Jastrow factor in the trial wave function. Moreover this term reduces the probability for two electrons to be close, and so decreases the average value of the repulsive interaction, providing a clear energy gain. The two-body Jastrow function reads:

$$J_2(\vec{r}_1, ..., \vec{r}_N) = \exp\left(\sum_{i<j}^{N} u(r_{ij})\right), \tag{1.26}$$

where $u(r_{ij})$ depends only on the relative distance $r_{ij} = |\vec{r}_i - \vec{r}_j|$ between two electrons and allows to fulfill the cusp conditions for opposite spin electrons as long as $u(r_{ij}) \rightarrow \frac{r_{ij}}{2}$ for small electron-electron distance. The pair correlation function u can be parametrized successfully by few variational parameters.

We have adopted a functional form u proposed by Fahy (38), that we found particularly convenient:

$$u(r) = \frac{r}{2(1+br)}, \tag{1.27}$$

where the variational parameter b has been optimized for each system. In this functional form the cusp condition for anti-parallel spin electrons is satisfied, whereas the one for parallel spins is neglected in order to avoid the spin contamination (for more details about spin contamination see Ref. (39)). This allows to remove the singularities of the local energy due to the collision of two opposite spin electrons, yielding a smaller variance and a more efficient VMC calculation.

1.2.4 Three Body Jastrow term

In order to describe well the correlation between electrons the simple two-body Jastrow factor is not sufficient. Indeed it takes into account only the electron-electron separation and not the individual electronic position $\vec{r}_i$ and $\vec{r}_j$. It is expected that close to atoms the correlation effects deviate significantly from the translational invariant Jastrow. For this reason we introduce a factor, often called three body (electron-electron-nucleus) Jastrow, that explicitly depends on both electronic positions $\vec{r}_i$ and $\vec{r}_j$. The three body Jastrow is chosen to satisfy the following requirements:

- The cusp conditions set up by the two-body Jastrow term and by the one-body term are preserved.

- Similarly to the two-body we do not include any spin dependency in the three-body Jastrow. In the way the wave-function remains a spin singlet.

- Whenever the atomic distances are large it factorizes into a product of independent contributions located near each atom, an important requirement to satisfy the size consistency of the variational wave function.

Analogously to the pairing trial function in Eq. 1.19 we define a three body factor as:

$$\begin{aligned} J_3(\vec{r}_1, ..., \vec{r}_N) &= \exp\left(\sum_{i<j} \Phi_J(\vec{r}_i, \vec{r}_j)\right) \\ \Phi_J(\vec{r}_i, \vec{r}_j) &= \sum_{l,m,a,b} g^{a,b}_{l,m} \psi_{a,l}(\vec{r}_i) \psi_{b,m}(\vec{r}_j), \end{aligned} \tag{1.28}$$

where indices l and m indicate different orbitals located around the atoms a and b respectively. Each Jastrow orbital $\psi_{a,l}(\vec{r})$ is centred on the corresponding atomic position $\vec{R}_a$. We have used Gaussian and exponential orbitals multiplied by appropriate polynomials of the electronic coordinates, related to different spherical harmonics with given angular momentum, as in the usual Slater basis.

The chosen form for the 3-body Jastrow (1.28) has very appealing features: it easily allows to include the symmetries of the system by imposing them on the matrix $g^{a,b}_{l,m}$ exactly as it is possible for the pairing part (e.g. by replacing $\lambda^{a,b}_{m,n}$ with $g^{a,b}_{m,n}$ in Eq. 1.20). It is size consistent, namely the atomic limit can be smoothly recovered with the same trial function when the matrix terms $g^{a,b}_{l,m}$ for $a \neq b$ approach zero in this limit (see Ref. (15)). Notice that a small non zero value of $g^{a,b}_{l,m}$ for $a \neq b$ acting on p-wave orbitals can correctly describe a weak interaction between electrons such as the Van der Waals forces.

Chapter 2

Optimization Methods

QMC calculations crucially depend on the quality of the trial-function, and so it is essential to have an optimized wave-function as close as possible to the ground state.

The problem of function optimization is a very important research topic in numerical simulation. In QMC, in addition to the usual difficulties to find the minimum of multidimensional parametric function, the statistical noise is present in the estimate of the cost function (usually the energy), and its derivatives , required for an efficient optimization.

Different cost functions and different strategies were used to optimize a many-body trial-function. Usually three cost functions were used in QMC optimization energy, variance or a linear combination of them. In this thesis we always used energy optimization. The variance optimization have the advantage to be bounded by below, to be positive defined and its minimum is known, but different authors Ref. (40; 41; 42) recently showed that the energy optimization is more effective than the variance one.

There are different motivations for this: first, usually one is interested in the lowest energy rather than in the lowest variance in both variational and diffusion Monte Carlo; second, variance optimization takes many iterations to optimize determinant parameters and often the optimization can get stuck in multiple local minimum and it suffers of the ”false convergence” problem (41); third energy-minimized wave functions on average yield more accurate values of other expectation values than variance minimized wave functions do (40).

The optimization strategies can be divided into three categories. The first strategy is based on correlated sampling together with deterministic optimization methods (43). Even if this idea yielded very accurate results for the first-row atoms (43), this procedure can have problems if parameters affect the nodes, and moreover density ratio of the current and initial trial-function increases exponentially with the size of the system (44). In the second strategy one use a large bin to evaluate the cost function and its derivatives in such way that the noise can be neglected and deterministic methods can be used (see for instance (45; 46)).

Third approach, the one we used, is based on an iterative technique to handle directly with noise functions. The first example of these methods is the so called Stochastic Gradient Approximation (SGA) Ref. (47), recently used also for structure optimization Ref. (48).

In this thesis we have used two new optimization methods the Stochastic Reconfiguration (SR) method (15; 49) and Stochastic Reconfiguration with Hessian acceleration (SRH) (50).

2.1 Stochastic Reconfiguration

Stochastic Reconfiguration (SR) technique was initial developed to partially solve the sign problem in lattice green function Monte Carlo (51) and then it was used as an optimization method for a generic trial-function (15; 49). An important advantage of this technique is that we use more information about the trial-function than the simple steepest descent allowing a faster optimization of the many-body wave-function.

Given a generic trial-function Ψ_T, not orthogonal to the ground state it is possible to obtain a new one closer to the ground-state by applying the operator $(\Lambda - \hat{H})$ to this wave-function for a sufficient large Λ. The idea of the Stochastic Reconfiguration is to change the parameters of the original trial-function in order to be as close as possible to the projected one.

For this purpose we define:

$$|\Psi_P\rangle = \left(\Lambda - \hat{H}\right) |\Psi_T(\alpha'_k \ldots \alpha_p)\rangle \tag{2.1}$$

$$|\Psi'_T\rangle = \delta\alpha_0 |\Psi_T\rangle + \sum_{k=1}^{p} \delta\alpha_k \frac{\partial}{\partial\alpha_k} |\Psi_T\rangle \tag{2.2}$$

where Ψ_P is the projected one and Ψ'_T is the new trail-function obtained changing variational parameters. We can write the equation eq. 2.2 as:

$$\Psi'_T = \sum_{k=0}^{p} \delta\alpha_k \hat{O}_k \Psi_T \tag{2.3}$$

where

$$\hat{O}_k \Psi_T(x) = \frac{\partial}{\partial \alpha_k} \ln \Psi_T(x) \text{ and } \hat{O}_0 = \hat{I} \tag{2.4}$$

Now we want to choose the new parameters in such a way that Ψ'_T is as close as possible to Ψ_P. Thus we require that a set of mixed average correlation function, corresponding to the two wave-functions 2.2, 2.1, are equal. Here we impose precisely that:

$$\frac{\langle \Psi_T | \hat{O}_k | \Psi'_T \rangle}{\langle \Psi_T | \Psi'_T \rangle} = \frac{\langle \Psi_T | \hat{O}_k | \Psi_P \rangle}{\langle \Psi_T | \Psi_P \rangle} \tag{2.5}$$

for $k = 0, ..., n$. This is equivalent to the equation system:

$$\delta\alpha_0 + \sum_{l=1} \delta\alpha_l \langle \hat{O}^l \rangle = \Lambda - \langle \hat{H} \rangle \tag{2.6}$$

$$\delta\alpha_0 \langle \hat{O}^k \rangle + \sum_{l=1} \delta\alpha_l \langle \hat{O}^k \hat{O}^l \rangle = \Lambda \langle \hat{O}^k \rangle - \langle \hat{O}^k \hat{H} \rangle \text{ for } k \neq 0 \tag{2.7}$$

Because the equation for α_0 is related to the normalization of the trial-function and this parameter doesn't effect any physical observable of the system, we can substitute $\delta\alpha_0$ from the first equation in the others:

$$\sum_{l=1} \delta\alpha_l s_{kl} = \langle \hat{O}^k \rangle \langle \hat{H} \rangle - \langle \hat{O}^k \hat{H} \rangle \tag{2.8}$$

where

$$s_{kl} = \langle (\hat{O}^k - \langle \hat{O}^k \rangle)(\hat{O}^l - \langle \hat{O}^l \rangle) \rangle \tag{2.9}$$

The solution of this equation system defines a direction in the parameters space. If we vary parameters along this direction for a sufficient small step Δt we will decrease the energy.

The matrix $s_{k,l}$ is calculated at each iteration through a standard variational Monte Carlo sampling; the single iteration constitutes a small simulation that will be referred in the following as “bin”. After each bin the wave function parameters are iteratively updated according to

$$\alpha_i^{new} = \alpha_i^{old} + \delta\alpha_k \Delta t \tag{2.10}$$

SR is similar to a standard steepest descent (SD) calculation, where the expectation value of the energy $E(\alpha_k) = \frac{\langle\Psi|H|\Psi\rangle}{\langle\Psi|\Psi\rangle}$ is optimized by iteratively changing the parameters α_i according to the corresponding derivatives of the energy (generalized forces):

$$f_k = -\frac{\partial E}{\partial \alpha_k} = -\frac{\langle\Psi|O_k H + H O_k|\Psi\rangle}{\langle\Psi|\Psi\rangle} + 2\frac{\langle\Psi|O_k|\Psi\rangle\langle\Psi|H|\Psi\rangle}{\langle\Psi|\Psi\rangle^2}, \tag{2.11}$$

namely:

$$\alpha_k \to \alpha_k + \Delta t f_k. \tag{2.12}$$

where Δt is a suitable small time step, which can be taken fixed or determined at each iteration by minimizing the energy expectation value.

Indeed the variation of the total energy ΔE at each step is easily shown to be negative for small enough Δt because, in this limit

$$\Delta E = -\Delta t \sum_i f_i^2 + O(\Delta t^2).$$

Thus the method certainly converges at the minimum when all the forces vanish. In the SR we have

$$\alpha_i^{new} = \alpha_i^{old} + \sum_i \bar{s}^{-1}_{i,k} f_k \Delta t \tag{2.13}$$

Using the analogy with the steepest descent, it is possible to show that convergence to the energy minimum is reached when the value of Δt is sufficiently small and is kept constant for each iteration. Indeed the energy variation for a small change of the parameters is:

$$\Delta E = -\Delta t \sum_{i,j} \bar{s}^{-1}_{i,j} f_i f_j,$$

and it is easily verified that the above term is always negative because the reduced matrix s, as well as s^{-1}, is positive definite, being s an overlap matrix with all positive eigenvalues.

For a stable iterative method, such as the SR or the SD one, a basic ingredient is that at each iteration the new parameters α' are close to the previous α according to a prescribed distance. The fundamental difference between the SR minimization and the standard steepest descent is just related to the definition of this distance. For the SD it is the usual one, that is defined by the Cartesian metric $\Delta_\alpha = \sum_k |\alpha'_k - \alpha_k|^2$, instead the SR works correctly in the physical Hilbert space metric of the wave function Ψ, yielding $\Delta_\alpha = \sum_{i,j} s_{i,j}(\alpha'_i - \alpha_i)(\alpha'_j - \alpha_j)$, namely the square distance between the two wave functions corresponding to the two different sets of variational parameters $\{\alpha'\}$ and $\{\alpha_k\}$. Therefore, from the knowledge of the generalized forces f_k, the most convenient change of the variational parameters minimizes the functional $\Delta E + \bar{\Lambda}\Delta_\alpha$, where ΔE is the linear change in the energy $\Delta E = -\sum_i f_i(\alpha'_i - \alpha_i)$ and $\bar{\Lambda}$ is a Lagrange multiplier that allows a stable minimization with small change Δ_α of the wave function Ψ. Then the final iteration (2.13) is easily obtained.

The advantage of SR compared with SD is obvious because sometimes a small change of the variational parameters corresponds to a large change of the wave function, and the SR takes into account this effect through the Eq. 2.13. In particular the method is useful when a non orthogonal basis set is used, as we have in this thesis. Moreover by using the reduced matrix s it is also possible to remove from the calculation those parameters that imply some redundancy in the variational space, as it is shown in the following sections of this chapter.

2.1.1 Setting the SR parameters

In this thesis we have determined Δt by verifying the stability and the convergence of the SR algorithm for fixed Δt value.

The simulation is stable whenever $1/\Delta t > \Lambda_{\text{cut}}$, where Λ_{cut} is an energy cutoff that is strongly dependent on the chosen wave function and it is generally weakly dependent on the bin length. Whenever the wave function is too much detailed, namely has a lot of variational freedom, especially for the high energy components of the core electrons, the value of Λ_{cut} becomes exceedingly large and too many iterations are required for obtaining a converged variational wave function. In fact a rough estimate of the corresponding number of iterations P is given by $P\Delta t >> 1/G$, where G is the typical energy gap of the system, of the order of few electron Volts in small atoms and molecules. Within the SR method it is therefore extremely important to work with a bin length rather small, so that many iterations can be performed without much effort.

In a Monte Carlo optimization framework the forces f_k are always determined with some statistical noise η_k, and by iterating the procedure several times with a fixed bin length the variational parameters will fluctuate around their mean val-

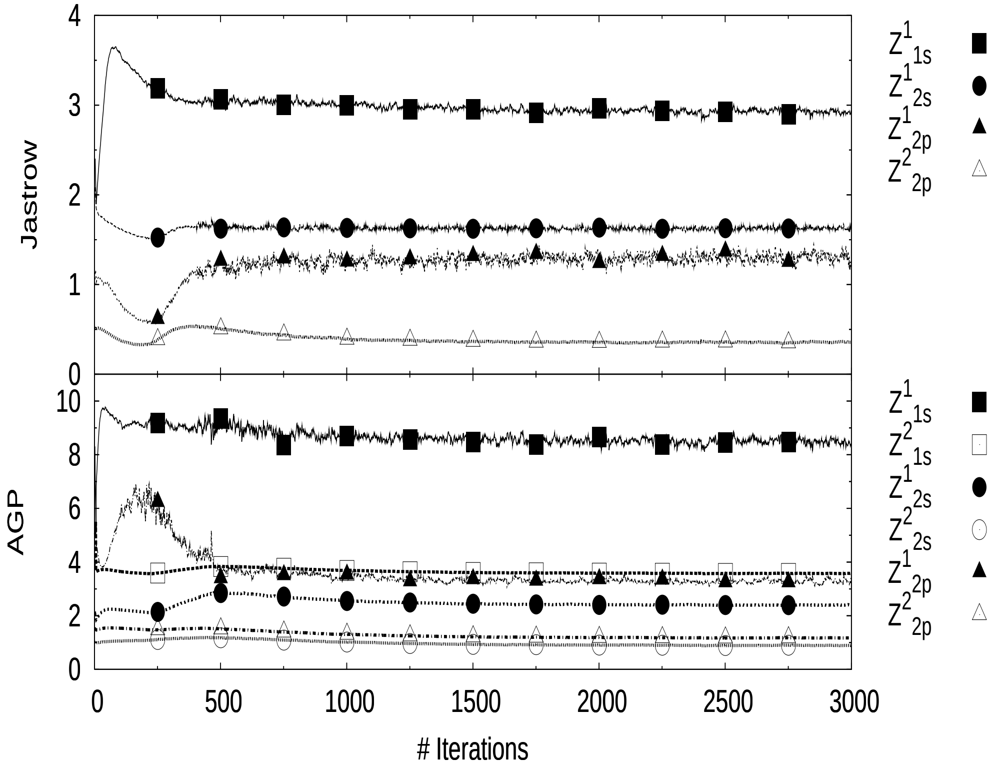

Figure 2.1: Example of the convergence of the SR method for the variational parameters as a function of the number of stochastic iterations. In the upper(lower) panel the Jastrow (geminal) parameters are shown. For each iteration, a variational Monte Carlo calculation is employed with a bin containing 15000 samples of the energy, yielding at the equilibrium a standard deviation of $\simeq 0.0018H$. For the first 200 iteration $\Delta t = 0.00125H^{-1}$, for the further 200 iterations $\Delta t = 0.0025H^{-1}$, whereas for the remaining ones $\Delta t = 0.005H^{-1}$.

ues. These statistical fluctuations are similar to the thermal noise of a standard Langevin equation:

$$\partial_t \alpha_k = f_k + \eta_k, \tag{2.14}$$

where

$$\langle \eta_k(t)\eta_{k'}(t')\rangle = 2T_{noise}\delta(t-t')\delta_{k,k'}. \tag{2.15}$$

The variational parameters α_k, averaged over the Langevin simulation time (as for instance in Fig.2.1), will be close to the true energy minimum, but the corresponding forces $f_k = -\partial_{\alpha_k} E$ will be affected by a bias that scales to zero with the thermal noise T_{noise}. Within a QMC scheme, one needs to estimate T_{noise} by increasing the bin length, as clearly $T_{noise} \propto 1/\mathrm{Bin\ length}$, this noise being directly related to the statistical fluctuations of the forces. Thus there is an optimal value for the bin length, which guarantees a fast convergence and avoid the forces to be biased within the statistical accuracy of the sampling. Moreover in the fluctuation around the minimum also non Gaussian correction will be present, but in analogy to the an-harmonic effects in solids, this error is expected to vanish linearly with the temperature T_{noise}. An example is shown in Fig. 2.1 for the optimization of the Be atom, using a basis two exponentials for each orbital both for the geminal and the three-body Jastrow part. The convergence is reached in about 1000 iteration with $\Delta t = 0.005H^{-1}$. However, in this case it is possible to use a small bin length, yielding a statistical accuracy in the energy much poorer than the final accuracy of about $0.05mH$. This is obtained by averaging the variational parameters in the last 1000 iterations, when they fluctuate around a mean value, allowing a very accurate determination of the energy minimum which satisfies the Euler conditions, namely with $f_k = 0$ for all parameters. Those conditions have been tested by an independent Monte Carlo simulation about 600 times longer than the bin used during the minimization.

As shown in Fig. 2.2 the Euler conditions are fulfilled within statistical accuracy even when the bin used for the minimization is much smaller than the overall simulation. On the other hand if the bin used is too small, as we have already pointed out, the averaging of the parameters is affected by a sizable bias.

Whenever it is possible to use a relatively small bin in the minimization, the apparently large number of iterations required for equilibration does not really matter, because a comparable amount of time has to be spent in the averaging of

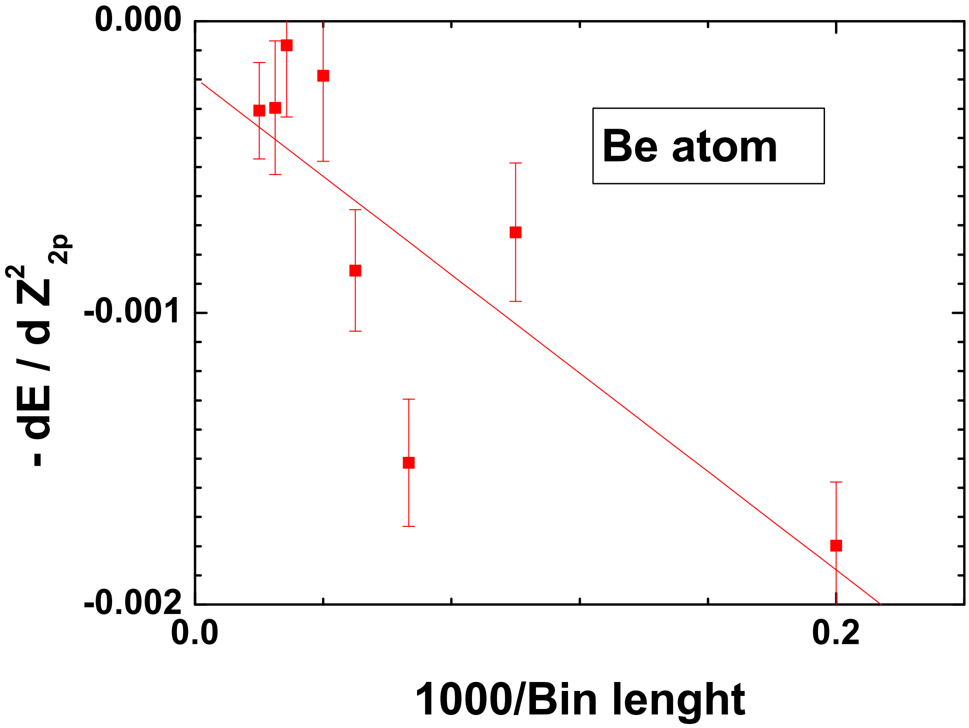

Figure 2.2: Calculation of the derivative of the energy with respect to the second Z in the $2p$ orbital of the geminal function for the Be atom. The calculation of the force was obtained, at fixed variational parameters, by averaging over 10^7 samples, allowing e.g. a statistical accuracy in the total energy of $0.07mH$. The variational parameters have been obtained by an SR minimization with fixed bin length shown in the x label. The parameter considered has the largest deviation from the Euler conditions.

the variational parameters, as shown in Fig. 2.1.

It is easy to convince oneself that for high enough accuracy the number of iterations needed for the equilibration becomes negligible from the computational point of view. In fact, in order to reduce, e.g. by a factor of ten, the accuracy in the variational parameters, a bin ten times larger is required for decreasing the thermal noise T_{noise} by the same factor. Whereas to reduce the statistical errors by the same ratio, it has to be done average on 100 times steps more. This means that the fraction of time spent for equilibration becomes ten times smaller compared with the less accurate simulation.

2.1.2 Stabilization of the SR technique

Whenever the number of variational parameters increases, it often happens that the stochastic $N \times N$ matrix

$$s_{k,k'} = \frac{\langle\Psi|O_k O_{k'}|\Psi\rangle}{\langle\Psi|\Psi\rangle} - \frac{\langle\Psi|O_k|\Psi\rangle}{\langle\Psi|\Psi\rangle}\frac{\langle\Psi|O_{k'}|\Psi\rangle}{\langle\Psi|\Psi\rangle} \tag{2.16}$$

becomes singular, i.e. the condition number, defined as the ratio $\sigma = \lambda_N/\lambda_1$ between its maximum λ_N and minimum eigenvalue λ_1, is too large. In that case the inversion of this matrix generates clear numerical instabilities which are difficult to control especially within a statistical method.

The first successful proposal to control this instability was to remove from the inversion problem (49), required for the minimization, those directions in the variational parameter space corresponding to exceedingly small eigenvalues λ_i. In this thesis we describe a method the is much better. As a first step, we show that the reason of the large condition number σ is due to the existence of ”redundant” variational parameters that do not make changes to the wave function within a prescribed tolerance ϵ.

Indeed in practical calculations, we are interested in the minimization of the wave function within a reasonable accuracy. The tolerance ϵ may represent therefore the distance between the exact normalized variational wave function which minimizes the energy expectation value and the approximate acceptable one, for which we no longer iterate the minimization scheme. For instance $\epsilon = 1/1000$ is by far acceptable for chemical and physical interest.

A stable algorithm is then obtained by simply removing the parameters that do not change the wave function by less than ϵ from the minimization. An efficient scheme to remove the ”redundant parameters” is also given.

Let us consider the N normalized states orthogonal to Ψ, but not mutually orthogonal:

$$|e_i\rangle = \frac{(O_k - \langle O_k\rangle)|\Psi\rangle}{\sqrt{\langle\Psi|(O_k - \langle O_k\rangle)^2|\Psi}}. \tag{2.17}$$

These normalized vectors define N directions in the $N-$dimensional variational parameter manifold, which are independent as long as the determinant S of the corresponding $N \times N$ overlap matrix

$$s_{k,k'} = \langle e_k|e_{k'}\rangle \tag{2.18}$$

$$\langle e_k|e_k\rangle = 1 \tag{2.19}$$

is non zero. The number S is clearly positive and it assumes its maximum value 1 whenever all the directions e_i are mutually orthogonal. On the other hand, let us suppose that there exists an eigenvalue $\bar{\lambda}$ of s smaller than the square of the desired tolerance ϵ^2, then the corresponding eigenvector $|v >= \sum_i a_i|e_i\rangle$ is such that:

$$\langle v|v\rangle = \sum_{i,j} a_i a_j \bar{s}_{i,j} = \bar{\lambda} \tag{2.20}$$

where the latter equation holds due to the normalization condition $\sum_i a_i^2 = 1$. We arrive therefore to the conclusion that it is possible to define a vector v with almost vanishing norm $|v| = \sqrt{\lambda} \le \epsilon$ as a linear combination of e_i, with at least some non zero coefficient. This implies that the N directions e_k are linearly dependent within a tolerance ϵ and one can safely remove at least one parameter from the calculation.

In general whenever there are p vectors v_i that are below the tolerance ϵ the optimal choice to stabilize the minimization procedure is to remove p rows and p columns from the matrix (2.18), in such a way that the corresponding determinant of the $(N-p) \times (N-p)$ overlap matrix is maximum.

From practical purposes it is enough to consider an iterative scheme to find a large minor, but not necessarily the maximum one. This method is based on the inverse of $\bar{s}$. At each step we remove the $i-th$ row and column from $\bar{s}$ for which $\bar{s}^{-1}_{i,i}$ is maximum. We stop to remove rows and columns after p inversions. In this approach we exploit the fact that, by a consequence of the Laplace theorem on determinants, $\bar{s}^{-1}_{k,k}$ is the ratio between the described minor without the $k-th$ row and column and the determinant of the full $\bar{s}$ matrix. Since within a stochastic method it is certainly not possible to work with a machine precision tolerance, setting $\epsilon = 0.001$ guarantees a stable algorithm, without affecting the accuracy of the calculation. The advantage of this scheme, compared with the previous one(18), is that the less relevant parameters can be easily identified after few iterations and do not change further in the process of minimization.

2.2 Structural optimization

In the last few years remarkable progresses have been made to develop Quantum Monte Carlo (QMC) techniques which are able in principle to perform structural optimization of molecules and complex systems (29; 52). Within the Born-Oppheneimer approximation the nuclear positions $\vec{R}_i$ can be considered as further variational parameters included in the set $\{\alpha_i\}$ used for the SR minimization (2.13) of the energy expectation value. For clarity, in order to distinguish the conventional variational parameters from the ionic positions, in this section we indicate with $\{c_i\}$ the former ones, and with $\vec{R}_i$ the latter ones. It is understood that $R_i^\nu = \alpha_k$, where a particular index k of the whole set of parameters $\{\alpha_i\}$ corresponds to a given spatial component ($\nu = 1, 2, 3$) of the $i-$th ion.

We computed the forces $\vec{F}$ acting on each of the M nuclear positions $\{\vec{R}_1, \ldots, \vec{R}_M\}$, being M the total number of nuclei in the system:

$$\begin{aligned} \vec{F}(\vec{R}_a) &= -\vec{\nabla}_{\vec{R}_a} E(\{c_i\}, \vec{R}_a) \qquad (2.21) \\ &= -\frac{\langle\Psi|O_R H + H O_R + \partial_R H|\Psi\rangle}{\langle\Psi|\Psi\rangle} + 2\frac{\langle\Psi|O_R|\Psi\rangle\langle\Psi|H|\Psi\rangle}{\langle\Psi|\Psi\rangle^2}, \qquad (2.22) \end{aligned}$$

where operator O_R are defined as logarithmic derivatives respect to nuclear position of the trial-function in analogy to the operator O_k 2.4. This generalized forces were than used to perform structural optimization using the the iteration (2.13). In the first part of this thesis we have used a finite difference operator $\frac{\vec{\Delta}}{\Delta R_a}$ for the evaluation of the force acting on a given nuclear position a:

$$\vec{F}(\vec{R}_a) = -\frac{\vec{\Delta}}{\Delta R_a} E = -\frac{E(\vec{R}_a + \vec{\Delta R}_a) - E(\vec{R}_a - \vec{\Delta R}_a)}{2\Delta R} + O(\Delta R^2) \qquad (2.23)$$

where $\Delta\vec{R}_a$ is a 3 dimensional vector. Its length ΔR is chosen to be 0.01 atomic units, a value that is small enough for negligible finite difference errors.

In order to evaluate the energy differences in Eq. 2.23 with a finite variance we have used the Space-Warp coordinate transformation (46; 53). This transformation was also used in the evaluation of the wave-function derivatives respect to nuclear positions O_R. Even if Space-Warp transformation is a very efficient technique to reduce the variance of the forces, it is very time consuming and so for larger systems we preferred to use Zero Variance forces (29), as it was described in the chapter 1.

The O_R operators are used also in the definition of the reduced matrix $\bar{s}$ for those elements depending on the variation with respect to a nuclear coordinate. In this way it is possible to optimize both the wave function and the ionic positions at the same time, in close analogy with the Car-Parrinello(54) method applied to the minimization problem. Also Tanaka (48) tried to perform Car-Parrinello like simulations via QMC, within the less efficient steepest descent framework.

An important source of systematic errors is the dependence of the variational parameters c_i on the ionic configuration $\vec{R}$, because for the final equilibrium geometry all the forces f_i corresponding to c_i have to be zero, in order to guarantee that the true minimum of the potential energy surface (PES) is reached (55; 56). As shown clearly in the previous subsection, within a QMC approach it is possible to control this condition by increasing systematically the bin length, when the thermal bias T_{noise} vanishes. In Fig. 2.3 we report the equilibrium distance of the Li molecule as a function of the inverse bin length, so that an accurate evaluation of such an important quantity is possible even when the number of variational parameters is rather large, by extrapolating the value to an infinite bin length. However, as it is seen in the picture, though the inclusion of the 3s orbital in the atomic AGP basis substantially improves the equilibrium distance and the total energy by $\simeq 1mH$, this larger basis makes our simulation less efficient, as the time step Δt has to be reduced by a factor three.

We have not attempted to extend the geometry optimization to the more accurate DMC, since there are technical difficulties (57), and it is computationally much more demanding.

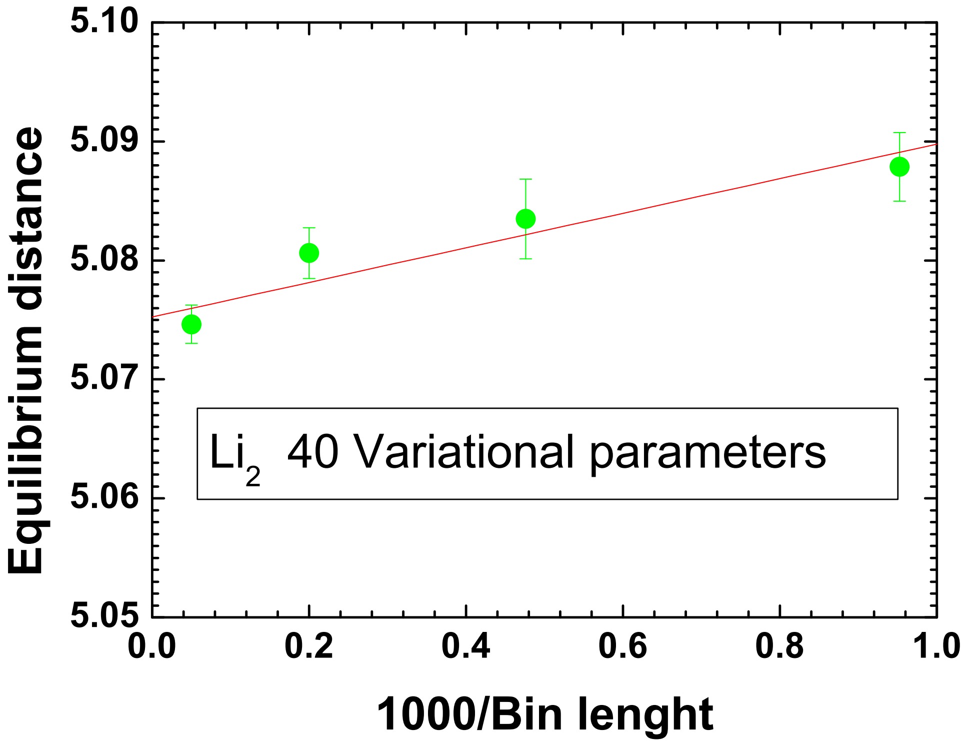

Figure 2.3: Plot of the equilibrium distance of the Li_2 molecule as a function of the inverse bin length. The total energy and the binding energy are reported in Tables 3.3 and 3.2 respectively. For all simulations the initial wave-function is optimized at $Li - Li$ distance 6 a.u.

2.3 Hessian Optimization

The SR method generally performs very well, whenever there is only one energy scale in the variational wave function. However if there are several energy scales in the problem, some of the variational parameters, e.g. the ones defining the low energy valence orbitals, converge very slowly with respect to the others, and the number of iterations required for the equilibration becomes exceedingly large. Moreover the time step Δt necessary for a stable convergence depends on the high energy orbitals, whose dynamics cannot be accelerated beyond a certain threshold. Futhermore the SR method is based on a first order dynamics, and as will be illustrated in the following (see section 5.4.2), it is not adapted to perform parameters optimization during a ion Langevin Dynamics. In this thesis to overcome these difficulties we have used a very efficient optimization method the Stochastic Reconfiguration with Hessian acceleration (SRH) (50).

The central point of the SRH is to use not only directions given by the generalized forces 2.11, to lower the energy expectation value, but also the information coming from the Hessian matrix to accelerate the convergence. The idea to use the Hessian matrix is not new, already Lin, Zhang and Rappe (45) proposed to use analytic derivatives to optimize the energy, but their implementation was inefficient and unstable.

Now we will review the SRH method and we will explain the reason of its efficiency.

Given an Hamiltonian H and a trial-function $\psi_\alpha(x) = \langle x|\psi_\alpha\rangle$ depending on a set of parameters $\alpha = \alpha_1, \alpha_2, ...\alpha_n$, we want to optimize the energy expectation value of the energy on this wave-function:

$$E_\alpha = \frac{\langle\psi_\alpha|H|\psi_\alpha\rangle}{\langle\psi_\alpha|\psi_\alpha\rangle} \tag{2.24}$$

respect to the parameters set.

To simplify the notation henceforth the symbol $<>$ indicates the quantum expectation value over ψ_α, so that $E_\alpha = \langle H\rangle$. In order the optimize the energy and to find the new parameters $\alpha' = \alpha + \gamma$, we expand the trial-function up to second order in γ:

$$|\psi_{\alpha+\gamma}\rangle = \left\{1 + \left[\sum_{k}^{p} \gamma_k (O^k - \langle O^k \rangle) + \frac{\beta}{2} \sum_{k,k'=1}^{p} \gamma_k \gamma_{k'} (O^k - \langle O^k \rangle)(O^{k'} - \langle O^{k'} \rangle)\right]\right\} |\psi_\alpha\rangle \tag{2.25}$$

with $\beta = 1$, where O^k is the operator with associated diagonal elements (15):

$$O_k(x) = \frac{\partial_{\alpha_k} \psi_\alpha(x)}{\psi_\alpha(x)} \tag{2.26}$$

Here the constant β will be used in order to highlight the various terms in the energy expansions. Using the fact that:

$$\begin{aligned}
\langle \psi_\alpha | \psi_\alpha \rangle &= 1 \\
\langle \psi_{\alpha+\gamma} | \psi_{\alpha+\gamma} \rangle &= 1 + (1+\beta) \sum_{k,k'=1}^{p} \gamma_k \gamma_{k'} \langle (O^k - \langle O^k \rangle)(O^{k'} - \langle O^{k'} \rangle) \rangle + O(\gamma^3)
\end{aligned}$$

we can expand up second order the energy given by the new wave-function $\psi_{\alpha+\gamma}(x)$ and obtain:

$$\begin{aligned}
E_{\alpha+\gamma} &= \frac{\langle \psi_{\alpha+\gamma} | H | \psi_{\alpha+\gamma} \rangle}{\langle \psi_{\alpha+\gamma} | \psi_{\alpha+\gamma} \rangle} \\
&= E_\alpha + 2 \sum \gamma_k \langle (H - E_\alpha)\, O^k \rangle \\
&+ (1+\beta) \sum_{k,k'=1}^{p} \gamma_k \gamma_{k'} \langle (H - E_\alpha)\, (O^k - \langle O^k \rangle)(O^{k'} - \langle O^{k'} \rangle) \rangle \\
&+ \frac{1}{2} \langle \left[O^k - \langle O^k \rangle, \left[H - E_\alpha, O^k - \langle O^k \rangle\right]\right] \rangle
\end{aligned}$$

We can define:

$$\begin{aligned}
S_h^{k,k'} &= \langle \left[O^k - \langle O^k \rangle, \left[H - E_\alpha, O^k - \langle O^k \rangle\right]\right] \rangle && (2.27) \\
G^{k,k'} &= 2 \langle (H - E_\alpha)\, (O^k - \langle O^k \rangle)(O^{k'} - \langle O^{k'} \rangle) \rangle && (2.28) \\
f_k &= -2 \langle (H - E_\alpha)\, O^k \rangle && (2.29)
\end{aligned}$$

and so the expansion of the energy reads:

$$\Delta E = -\sum_k \gamma_k f_k + \frac{1}{2} \sum_{k,k'} (1+\beta)\, [S_h + (1+\beta) G]^{k,k'} \tag{2.30}$$

The wave-function parameters can then iteratively changed to stay at the minimum of the above quadratic form whenever it is positive defined, and in such case the minimum energy is obtained for:

$$\vec{\gamma} = B^{-1}\vec{f} \tag{2.31}$$

where

$$B = S_h + (1+\beta)G \tag{2.32}$$

It can happen that the quadratic form is not positive definite and so the energy expansion 2.30 is not bounded from below, this can due to different reasons: non quadratic corrections; statistical fluctuations of the Hessian matrix expectation value; or because we are far from the minimum. In this cases the correction due to the equation 2.31 may lead to an higher energy than E_α. To overcame this problem the matrix B is changed in such way to be always positive definite $B' = B + \mu S$, where S is the Stochastic Reconfiguration matrix 2.9. The use of the S matrix guarantees that we are moving in the direction of a well defined minimum when the change of the wave-function is small, moreover in the limit of large μ we recover the Stochastic Reconfiguration optimization method. To be sure that the change of the wave-function is small we use a control parameter to impose a constraint to the variation of the wave-function ΔWF by means of the inequalities

$$|\Delta WF|^2 \leq r^2 \tag{2.33}$$

where, using 2.9 and 2.25, $|\Delta WF|^2 = \langle \phi_\alpha | \phi_{\alpha+\gamma} \rangle = \sum_{k,k'} \gamma_k \gamma_{k'} S^{k,k'}$. This constraint always allows to work with a positive definitive matrix B', and for small r the energy is certainly lower than E_α. We want to emphasize that the condition $\mu \geq 0$ is non zero both when 2.32 is not positive defined and when $|\Delta WF|$ corresponding to eq. 2.31 exceeds r. This is equivalent to impose a Lagrange multiplier to the energy minimization, namely $\Delta E + \mu |\Delta WF|^2$, with the condition $|\Delta WF| = r$.

There is another important ingredients for an efficient implementation of the Hessian technique to QMC. In fact, as pointed out in Ref.(50; 58) is extremely important to evaluate the quantities appearing in the Hessian 2.32 in the form of correlation function $< AB > - < B >< A >$. This because the fluctuation of operators in this form are usually smaller than the one of $< AB >$ especially if A and B are correlated. Therefore using the fact that the expectation value of the derivative of a local value $O_L = \hat{O}\Psi/\Psi$ of an Hermitian operator $\hat{O}$ respect to any real parameter c in a real wave function Ψ is always zero (see for instance (45)), we can rearrange the Hessian terms in more symmetric way in form of correlation function:

$$\begin{aligned} f_k &= \langle E_L(x)O^k(x)\rangle - \langle E_L(x)\rangle\langle O^k(x)\rangle \\ S_h^{k,k'} &= \langle \partial_{\alpha_k} E_L(x)O^{k'}(x)\rangle - \langle \partial_{\alpha_k} E_L(x)\rangle\langle O^{k'}(x)\rangle \\ &+ \langle \partial_{\alpha_{k'}} E_L(x)O^{k}(x)\rangle - \langle \partial_{\alpha_{k'}} E_L(x)\rangle\langle O^{k}(x)\rangle \\ S^{k,k'} &= \langle O^{k'}(x)O^k(x)\rangle - \langle O^{k'}(x)\rangle\langle O^k(x)\rangle \\ G &= \langle \delta E_L(x)\delta O^{k'}(x)\delta O^k(x)\rangle \end{aligned}$$

Because the G matrix 2.28 is zero for the exact ground state and therefore is expected to be very small for a good variational wave-function, it is possible, following the suggestion of Ref. (50), to chose $\beta = -1$, so that $B = S_h + \mu S$. As shown by Ref. (50) this choice can even lead to faster convergence than the full Hessian matrix.

The matrix G is the only part that is not in the form of a correlation function, for this reason is important that B does not depend on it, in such way to reduce the fluctuation of the Hessian matrix, and this can naively explain the suggestion of Ref. (50) to chose $\beta = -1$.

As for the SR method the parameters are iteratively updated using the equation:

$$\vec{\gamma} = [S_h + \mu S]^{-1} \vec{f} \tag{2.34}$$

where the forces $\vec{f}$ and the matrix B are evaluated using VMC.

Chapter 3

Results on Molecules

3.1 Application of the JAGP to molecules

In the first part of this thesis we study correlation and atomization energies, accompanied with the determination of the ground state optimal structure for a restricted ensemble of molecules. For each of them we performed a full all-electron SR geometry optimization, starting from the experimental molecular structure. After the energy minimization, we carried out all-electron VMC and DMC simulations at the optimal geometry within the so-called ”fixed node approximation”. The basis that we used was composed by exponential and Gaussian orbitals for both the three-body and the pairing determinant, in this way both the antisymmetric and the bosonic part are well described. However, both in the AGP and in the Jastrow part we never used a large basis set, in order to keep the wave function as simple as possible. The accuracy of our wave function can be obviously improved by an extension of the one particle basis set. Nevertheless, for most of the molecules studied with a simple JAGP wave function, a DMC calculation is able to reach the chemical accuracy in the binding energies and the SR optimization yields very precise geometries already at the VMC level.

In the first part of this section some results will be presented for a small set of widely studied molecules and belonging to the G1 database. In the second subsection we will treat the benzene and its radical cation $C_6H_6^+$, by taking into account its distortion due to the Jahn-Teller effect, that is well reproduced by our SR geometry optimization.

Table 3.1: Total energies in variational (E_{VMC}) and diffusion (E_{DMC}) Monte Carlo calculations; the percentages of correlation energy recovered in VMC ($E_c^{VMC}(\%)$) and DMC ($E_c^{DMC}(\%)$) have been evaluated using the "exact" (E_0) and Hartree–Fock (E_{HF}) energies from the references (1). Here "exact" means the ground state energy of the non relativistic infinite nuclear mass Hamiltonian. The energies are in *Hartree*.

	E_0	E_{HF}	E_{VMC}	$E_c^{VMC}(\%)$	E_{DMC}	$E_c^{DMC}(\%)$
Li	-7.47806 (59)	-7.432727 (59)	-7.47721(11)	98.12(24)	-7.47791(12)	99.67(27)
Li_2	-14.9954 (16)	-14.87152 (16)	-14.99002(12)	95.7(1)	-14.99472(17)	99.45(14)
Be	-14.66736 (59)	-14.573023 (59)	-14.66328(19)	95.67(20)	-14.66705(12)	99.67(13)
Be_2	-29.33854(5) (16)	-29.13242 (16)	-29.3179(5)	89.99(24)	-29.33341(25)	97.51(12)
O	-75.0673 (59)	-74.809398 (59)	-75.0237(5)	83.09(19)	-75.0522(3)	94.14(11)
H_2O	-76.438(3) (60)	-76.068(1) (60)	-76.3803(4)	84.40(10)	-76.4175(4)	94.46(10)
O_2	-150.3268 (16)	-149.6659 (16)	-150.1992(5)	80.69(7)	-150.272(2)	91.7(3)
C	-37.8450 (59)	-37.688619 (59)	-37.81303(17)	79.55(11)	-37.8350(6)	93.6(4)
C_2	-75.923(5) (16)	-75.40620 (16)	-75.8293(5)	81.87(10)	-75.8810(5)	91.87(10)
CH_4	-40.515 (61)	-40.219 (61)	-40.4627(3)	82.33(10)	-40.5041(8)	96.3(3)
C_6H_6	-232.247(4) (62)	-230.82(2) (63)	-231.8084(15)	69.25(10)	-232.156(3)	93.60(21)

3.2 Small diatomic molecules, methane, and water

Except from Be_2 and C_2, all the molecules presented here belong to the standard G1 reference set; all their properties are well known and well reproduced by standard quantum chemistry methods, therefore they constitute a good case for testing new approaches and new wave functions.

The Li dimer is one of the easiest molecules to be studied after the H_2, which is exact for any Diffusion Monte Carlo (FN DMC) calculation with a trial wave function that preserves the node-less structure. Li_2 is less trivial due to the presence of core electrons that are only partially involved in the chemical bond and to the $2s - 2p$ near degeneracy for the valence electrons. Therefore many authors have done benchmark calculation on this molecule to check the accuracy of the method or to determine the variance of the inter-nuclear force calculated within a QMC framework. In this thesis we start from Li_2 to move toward a structural analysis of more complex compounds, thus showing that our QMC approach is able to handle relevant chemical problems.

With our approach more than 99% of the Li_2 correlation energy is recovered by a DMC simulation (Table 3.1), and the atomization energy is exact within few thousands of eV ($0.02\ kcal\ mol^{-1}$) (Table 3.3). Similar accuracy have been previously reached within a DMC approach(16), only by using a multi-reference CI like wave function, that before our work, was the usual way to improve the electronic nodal structure. As stressed before, the JAGP wave function includes many resonating configurations through the geminal expansion, beyond the $1s\ 2s$ HF ground state. The bond length has been calculated at the variational level through the fully optimized JAGP wave function: the resulting equilibrium geometry turns out to be highly accurate (Table 3.2), with a discrepancy of only $0.001 a_0$ from the exact result.

Table 3.2: Bond lengths (R) in atomic units; the subscript 0 refers to the "exact" results. For the water molecule R is the distance between O and H and θ is the angle HOH (in deg), for CH_4 R is the distance between C and H and θ is the HCH angle.

	R_0	R	θ_0	θ
Li_2	5.051	5.0516(2)		
O_2	2.282	2.3425(18)		
C_2	2.348	2.366(2)		
H_2O	1.809	1.8071(23)	104.52	104.74(17)
CH_4	2.041	2.049(1)	109.47	109.55(6)
	R_0^{CC}	R^{CC}	R_0^{CH}	R^{CH}
C_6H_6	2.640	2.662(4)	2.028	1.992(2)

The good bond length, we obtained, is partially due to the energy optimization that is often more effective than the variance minimization, as shown by different authors (40; 41; 42), and partially due to the quality of the trial-function.

Indeed within our scheme we obtain good results without exploiting the computationally much more demanding DMC, thus highlighting the importance of the SR minimization described in Subsection 2.2.

Table 3.3: Binding energies in eV obtained by variational (Δ_{VMC}) and diffusion (Δ_{DMC}) Monte Carlo calculations; Δ_0 is the "exact" result for the non-relativistic infinite nuclear mass Hamiltonian. Also the percentages ($\Delta_{VMC}(\%)$ and $\Delta_{DMC}(\%)$) of the total binding energies are reported.

	Δ_0	Δ_{VMC}	$\Delta_{VMC}(\%)$	Δ_{DMC}	$\Delta_{DMC}(\%)$
Li_2	-1.069	-0.967(3)	90.4(3)	-1.058(5)	99.0(5)
O_2	-5.230	-4.13(4)	78.9(8)	-4.56(5)	87.1(9)
H_2O	-10.087	-9.704(24)	96.2(1.0)	-9.940(19)	98.5(9)
C_2	-6.340	-5.530(13)	87.22(20)	-5.74(3)	90.6(5)
CH_4	-18.232	-17.678(9)	96.96(5)	-18.21(4)	99.86(22)
C_6H_6	-59.25	-52.53(4)	88.67(7)	-58.41(8)	98.60(13)

Let us now consider larger molecules. Both C_2 and O_2 are poorly described by a single Slater determinant, since the presence of the non-dynamic correlation is strong. Instead with a single geminal JAGP wave function, including implicitly many Slater-determinants(15), it is possible to obtain a quite good description of their molecular properties. In both the cases, the variational energies recover more than 80% of the correlation energy, the DMC ones yield more than 90%, as shown in Tab. 3.1. These results are of the same level of accuracy as those obtained by Filippi *et al*(16) with a multi-reference wave function by using the same Slater basis for the antisymmetric part and a different Jastrow factor. From the Table 3.3 of the atomization energies, it is apparent that DMC considerably improves the binding energy with respect to the VMC values, although for these two molecules it is quite far from the chemical accuracy ($\simeq$ 0.1 eV): for C_2 the error is 0.60(3) eV, for O_2 is 0.67(5) eV. Indeed, it is well known that the electronic structure of the atoms is described better than the corresponding molecules if the basis set remains the same, and the nodal error is not compensated by the energy difference between the separated atoms and the molecule. In a benchmark DMC calculation with pseudo-potentials (64), Grossman found an error of 0.27 eV in the atomization energy for O_2, by using a single determinant wave function. Probably, pseudo-potentials allow the error between the pseudo-atoms and the pseudo-molecule to compensate better, thus yielding more accurate energy differences. As a final remark on the O_2 and C_2 molecules, our bond lengths are in between the LDA and GGA precision, and still worse than the best CCSD calculations, but our results may be considerably improved by a larger atomic basis set.

Methane and water are very well described by the JAGP wave function. Also for these molecules we recover more than 80% of correlation energy at the VMC level, while DMC yields more than 90%, with the same level of accuracy reached in previous Monte Carlo studies (61; 65; 66; 67). Here the binding energy is almost exact, since in this case the nodal energy error arises essentially from only one atom (carbon or oxygen) and therefore it is exactly compensated when the atomization energy is calculated. Also the bond lengths are highly accurate, with an error lower then 0.005 a_0.

For Be_2 we applied a large Gaussian and exponential basis set for the determinant and the Jastrow factor and we recovered, at the experimental equilibrium geometry, the 90% of the total correlation energy in the VMC, while DMC gives 97.5% of correlation, i.e. a total energy of -29.33341(25) H. Although this value is better than the one obtained by Filippi *et al* (16) (-29.3301(2) H) with a smaller basis ($3s$ atomic orbitals not included), it is not enough to bind the molecule, because the binding energy remains still positive (0.0069(37) H). Instead, once the molecular geometry has been relaxed, the SR optimization finds a bond distance of $13.5(5)$ a_0 at the VMC level; therefore the employed basis allows the molecule to have a Van der Waals like minimum, quite far from the experimental value. In order to have a reasonable description of the bond length and the atomization energy, one needs to include at least a $3s2p$ basis in the antisymmetric part, as pointed out in Ref. (68). Indeed an atomization energy compatible with the experimental result (0.11(1) eV) has been obtained within the extended geminal model (69) by using a much larger basis set (9s,7p,4d,2f,1g) (70). This suggests that a complete basis set calculation with JAGP may describe also this molecule. However our SR method can not cope with a very large basis in a feasible computational time. Therefore we believe that at present the accuracy needed to describe correctly Be_2 is out of the possibilities of the approach.

Table 3.4: Binding energies in eV obtained by variational (Δ_{VMC}) and diffusion (Δ_{DMC}) Monte Carlo calculations with different trial wave functions for benzene. In order to calculate the binding energies yielded by the 2-body Jastrow we used the atomic energies reported in Ref. (2). The percentages ($\Delta_{VMC}(\%)$ and $\Delta_{DMC}(\%)$) of the total binding energies are also reported.

	Δ_{VMC}	$\Delta_{VMC}(\%)$	Δ_{DMC}	$\Delta_{DMC}(\%)$
Kekule + 2body	-30.57(5)	51.60(8)	-	-
resonating Kekule + 2body	-32.78(5)	55.33(8)	-	-
resonating Dewar Kekule + 2body	-34.75(5)	58.66(8)	-56.84(11)	95.95(18)
Kekule + 3body	-49.20(4)	83.05(7)	-55.54(10)	93.75(17)
resonating Kekule + 3body	-51.33(4)	86.65(7)	-57.25(9)	96.64(15)
resonating Dewar Kekule + 3body	-52.53(4)	88.67(7)	-58.41(8)	98.60(13)
full resonating + 3body	-52.65(4)	88.869(7)	-58.30(8)	98.40(13)

3.3 Benzene and its radical cation

We studied the ${}^1A_{1g}$ ground state of the benzene molecule by using a very simple one particle basis set: for the AGP, a 2s1p DZ set centered on the carbon atoms and a 1s SZ on the hydrogen, instead for the 3-body Jastrow, a 1s1p DZ-GTO set centered only on the carbon sites. C_6H_6 is a peculiar molecule, since its highly symmetric ground state, which belongs to the D_{6h} point group, is a resonance among different many-body states, each of them characterized by three double bonds between carbon atoms. This resonance is responsible for the stability of the structure and therefore for its aromatic properties. We started from a non resonating 2-body Jastrow wave function, which dimerizes the ring and breaks the full rotational symmetry, leading to the Kekulé configuration. As we expected, the inclusion of the resonance between the two possible Kekulé states lowers the VMC energy by more than 2 eV. The wave function is further improved by adding another type of resonance, that includes also the Dewar contributions connecting third nearest neighbor carbons.

As reported in Tab. 3.4, the gain with respect to the simplest Kekulé wave

Table 3.5: Bond lengths (r) for the two lowest $^2B_{2g}$ and $^2B_{3g}$ states of the benzene radical cation. The angles α are expressed in degrees, the lengths in a_0. The carbon sites are numerated from 1 to 6.

	$^2B_{2g}$ acute	$^2B_{3g}$ obtuse	Computational method
$r(C_1 - C_2)$	2.616	2.694	B3LYP/cc-pVTZ (4)
	2.649	2.725	BLYP/6-31G* (3)
	2.659(1)	2.733(4)	SR-VMC [1]
$r(C_2 - C_3)$	2.735	2.579	B3LYP/cc-pVTZ (4)
	2.766	2.615	BLYP/6-31G* (3)
	2.764(2)	2.628(4)	SR-VMC [2]
$\alpha(C_6C_1C_2)$	118.4	121.6	B3LYP/cc-pVTZ (4)
	118.5	121.5	BLYP/6-31G* (3)
	118.95(6)	121.29(17)	SR-VMC [1]

function amounts to 4.2 eV, but the main improvement arises from the further inclusion of the three-body Jastrow factor, which allows to recover the 89% of the total atomization energy at the VMC level. The main effect of the three body term is to keep the total charge around the carbon sites to approximately six electrons, thus penalizing the double occupation of the p_z orbitals. The same important correlation ingredient is present in the well known Gutzwiller wave function already used for polyacetylene (71; 72). Within this scheme we have systematically included in the 3-body Jastrow part the same type of terms present in the AGP one, namely both $g^{a,b}$ and $\lambda^{a,b}$ are non zero for the same pairs of atoms. As expected, the terms connecting next nearest neighbour carbon sites are much less important than the remaining ones because the VMC energy does not significantly improve (see the full resonating + 3-body wave function in Tab. 3.4). A more clear behaviour is found by carrying out DMC simulations: the interplay between the resonance among different structures and the Gutzwiller-like correlation refines more and more the nodal surface topology, thus lowering the DMC energy by significant amounts.

Therefore it is crucial to insert into the variational wave function all these in-

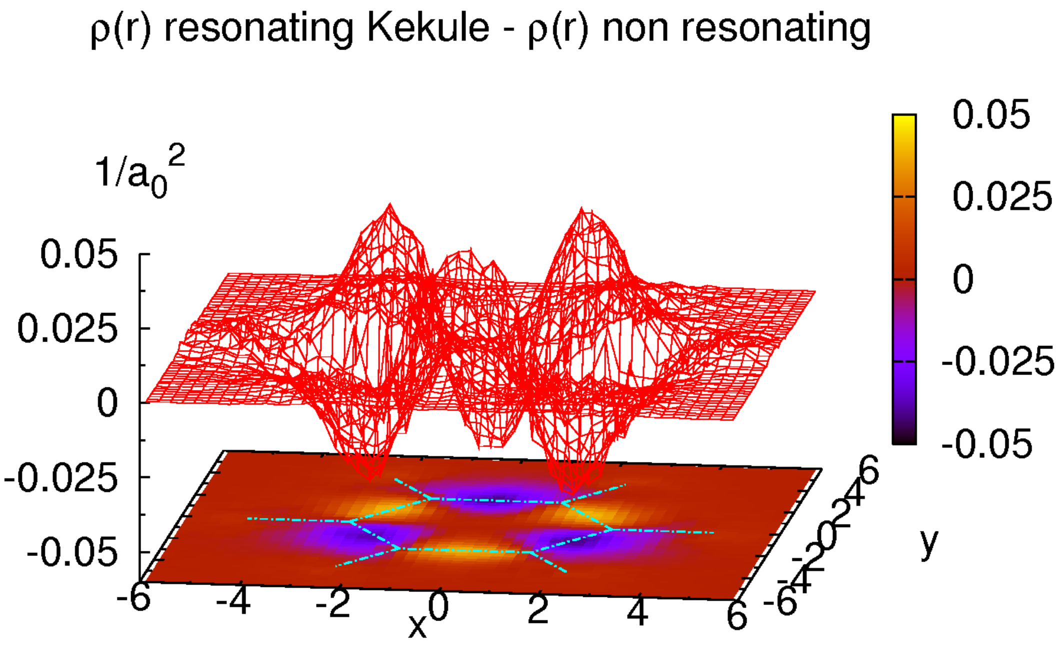

Figure 3.1: Electron density (atomic units) projected on the plane of C_6H_6. The surface plot shows the difference between the resonating valence bond wave function, with the correct $A1g$ symmetry of the molecule, and a non-resonating one, which has the symmetry of the Hartree Fock wave function.

gredients in order to have an adequate description of the molecule. For instance, in Fig. 3.2 we report the density surface difference between the non-resonating 3-body Jastrow wave function, which breaks the C_6 rotational invariance, and the resonating Kekulé structure, which preserves the correct A_{1g} symmetry: the change in the electronic structure is significant. The best result for the binding energy is obtained with the Kekulé Dewar resonating 3 body wave function, which recovers the $98,6\%$ of the total atomization energy with an absolute error of 0.84(8) eV. As Pauling (73) first pointed out, benzene is a genuine RVB system, indeed it is well described by the JAGP wave function. Moreover Pauling gave an estimate for the resonance energy of 1.605 eV from thermochemical experiments in qualitative agreement with our results. A final remark about the error in the total atomization energy: the latest frozen core CCSD(T) calculations (62; 74) are able to reach a precision of 0.1 eV, but only after the complete basis set extrapolation and the inclusion of the core valence effects to go beyond the psudopotential approximation. Without the latter corrections, the error is quite large and even in the CCSD approach it is 0.65 eV (74). In our case, such an error arises from the fixed node approximation, whose nodal error is not compensated by the difference between the atomic and the molecular energies, as already noticed in the previous subsection.

The radical cation $C_6H_6^+$ of the benzene molecule has been the subject of intense theoretical studies(3; 4), aimed to focus on the Jahn-Teller distorted ground state structure. Indeed the ionized $^2E_{1g}$ state, which is degenerate, breaks the symmetry and experiences a relaxation from the D_{6h} point group to two different states, $^2B_{2g}$ and $^2B_{3g}$, that belong to the lower D_{2h} point group. In practice, the former is the elongated acute deformation of the benzene hexagon, the latter is its compressed obtuse distortion. We applied the SR structural optimization, starting from the $^2E_{1g}$ state, and the minimization correctly yielded a deformation toward the acute structure for the $^2B_{2g}$ state and the obtuse for the $^2B_{3g}$ one; the first part of the evolution of the distances and the angles during those simulations is shown in Fig.3.3. After this equilibration, average over 200 further iterations yields bond distances and angles with the same accuracy as the all-electron BLYP/6-31G* calculations reported in Ref. (3) (see Tab. 3.5).

As it appears from Tab. 3.6 not only the structure but also the DMC total energy is in perfect agreement with the BLYP/6-31G*, and much better than

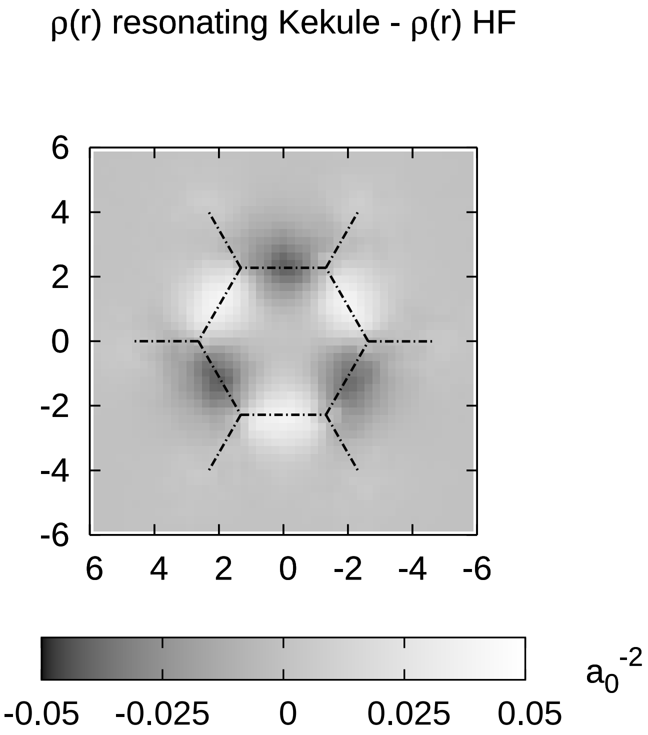

Figure 3.2: Surface plot of the charge density projected onto the molecular plane. The difference between the non-resonating (indicated as HF) and resonating Kekulé 3-body Jastrow wave function densities is shown. Notice the corresponding change from a dimerized structure to a C_6 rotational invariant density profile.

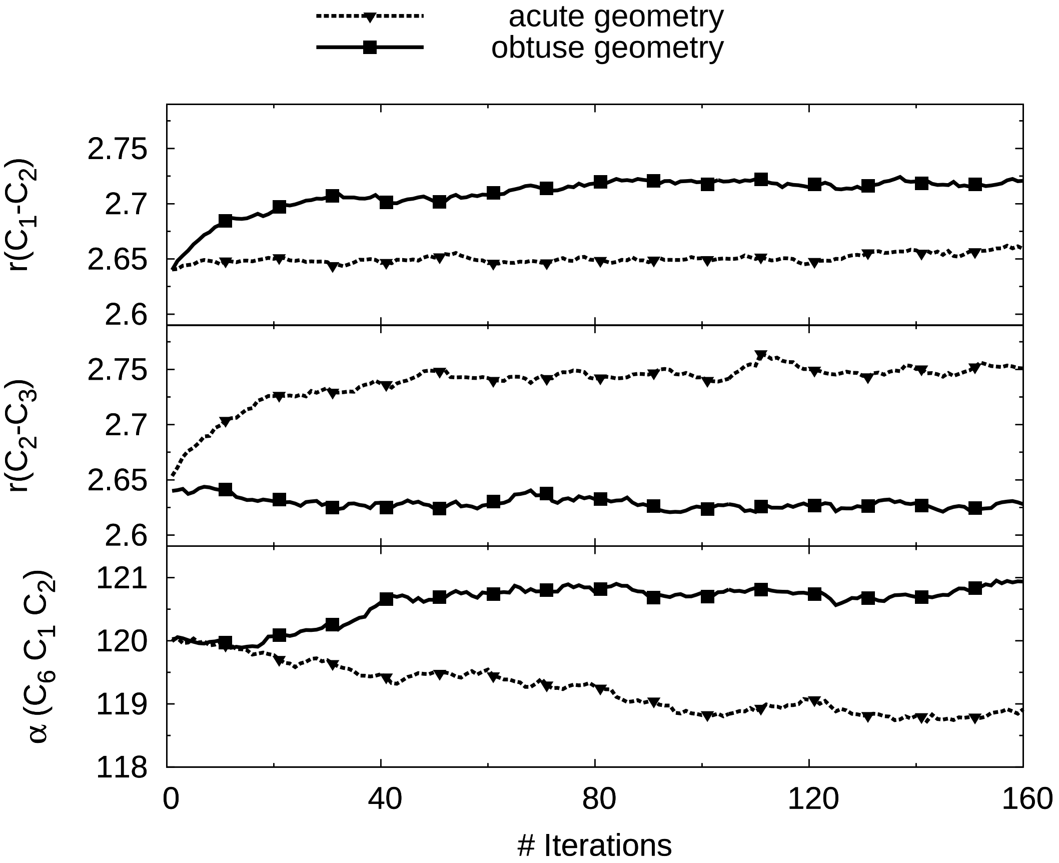

Figure 3.3: Plot of the convergence toward the equilibrium geometry for the ${}^2B_{2g}$ acute and the ${}^2B_{3g}$ obtuse benzene cation. Notice that both the simulations start form the ground state neutral benzene geometry and relax with a change both in the $C-C$ bond lengths and in the angles. The symbols are the same of Tab. 3.5.

Table 3.6: Total energies for the $^2B_{2g}$ and $^2B_{3g}$ states of the benzene radical cation after the geometry relaxation. A comparison with a BLYP/6-31G* and SVWN/6-31G* all-electron calculation (Ref. (3)) is reported.

	VMC	DMC	BLYP/6-31G*	SVWN/6-31G*
$^2B_{2g}$	-231.4834(15)	-231.816(3)	-231.815495	-230.547931
$^2B_{3g}$	-231.4826(14)	-231.812(3)	-231.815538	-230.547751

SVWN/6-31G* that does not contain semi empirical functionals, for which the comparison with our calculation is more appropriate, being fully ab-initio.

The difference of the VMC and DMC energies between the two distorted cations are the same within the error bars; indeed, the determination of which structure is the real cation ground state is a challenging problem, since the experimental results give a difference of only few meV in favor of the obtuse state and also the most refined quantum chemistry methods are not in agreement among themselves (3). A more affordable problem is the determination of the adiabatic ionization potential (AIP), calculated for the $^2B_{3g}$ state, following the experimental hint. Recently, very precise CCSD(T) calculations have been performed in order to establish a benchmark theoretical study for the ionization threshold of benzene (4); the results are reported in Tab. 3.7. After the correction of the zero point energy due to the different structure of the cation with respect to the neutral molecule and taken from a B3LYP/cc-pVTZ calculation reported in Ref. (4), the agreement among our DMC result, the benchmark calculation and the experimental value is impressive. Notice that in this case there should be a perfect cancellation of nodal errors in order to obtain such an accurate value; however, we believe that it is not a fortuitous result, because in this case the underlying nodal structure does not change much by adding or removing a single electron.

Therefore we expect that this property holds for all the affinity and ionization energy calculations with a particularly accurate variational wave function as the one we have proposed here. Nevertheless DMC is needed to reach the chemical accuracy, since the VMC result is slightly off from the experimental one just by few tenths of eV. The AIP and the geometry determination for the $C_6H_6^+$ are encouraging to pursue this approach, with the aim to describe even much more

Table 3.7: Adiabatic ionization potential of the benzene molecule; our estimate is done for the $^2B_{3g}$ relaxed geometries of the benzene radical cation, with an inclusion of the zero point motion correction between the $^2B_{3g}$ state and the $^1A_{1g}$ neutral molecule ground state, calculated in Ref. (4) at the B3LYP/6-31G* level.

	VMC	DMC	CCSD(T)/cc-pV∞Z (4)	experiment (75)
AIP	8.86(6)	9.36(8)	9.29(4)	
ΔZPE_{ad}	-0.074	-0.074	-0.074	
best estimate	8.79(6)	9.29(8)	9.22(4)	9.2437(8)

interesting and challenging chemical systems.

Chapter 4

Quantum Monte Carlo on extended systems

A naive and certainly very inefficient way to study extended system is to simulate clusters of atoms and to investigate the property of the cluster as the number of atoms increase. In this limit the collective behaviour should asymptotically approach to the bulk solid one. However the number of atoms that can be simulated by QMC is so small that the properties of the cluster will be dominated by the surface effects.

An alternative and efficient way to approximate the bulk properties of an infinite system is the use of periodic boundary conditions (PBC) on a finite box. These boundary conditions mean that the simulation cell is wrapped onto itself and, as an electron moves out of one side of the super-cell it immediately moves back through the opposite side (see figure 4.1). The advantage of using such boundary

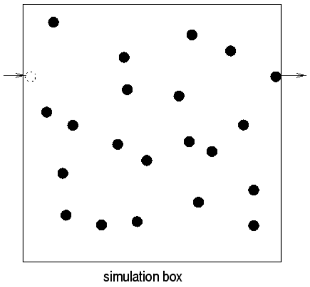

Figure 4.1: A simulation box with periodic boundary conditions.

conditions is that there are no longer ”surface electrons” and hence no surface effects. However even with PBC size effects are still present. This is due to the lack of long wavelength fluctuations in the charge density. For a simulation box of linear dimension L, the periodicity will remove any correlation length greater than L.

In this thesis we used a cubic simulation cell with volume L^3 with PBC, and the size effects are partially taken into account by increasing the size of the super-cell. The general hydrogen Hamiltonian with periodic boundary condition is written as:

$$\begin{aligned}\hat{H} &= \sum_{i=1}^{N} -\frac{1}{2}\nabla_i^2 + \frac{1}{2}\sum_{\vec{r}_i \neq \vec{r}_j + \vec{R}_s, \vec{R}_s} \frac{1}{|\vec{r}_i - \vec{r}_j - \vec{R}_s|} - \sum_{\vec{R}_s}\sum_{i}^{N}\sum_{j}^{N_{ions}} \frac{1}{|\vec{r}_i - \vec{R}_j - \vec{R}_s|} \\ &+ \frac{1}{2}\sum_{\vec{R}_i \neq \vec{R}_j + \vec{R}_s, \vec{R}_s} \frac{1}{|\vec{R}_i - \vec{R}_j - \vec{R}_s|} \end{aligned} \tag{4.1}$$

where $\vec{R}_s$ are the vectors of the periodic lattice associated with the simulation box, $\vec{r}_i$ are electron coordinates and $\vec{R}_i$ are the proton coordinates, and N is the number of electrons in the simulation cell. Infinite mass of the protons is assumed so that the kinetic term contains only the electronic contribution. Notice that the Hamiltonian 4.1 is invariant under the translation of any electron coordinate by a vector in R_s. Moreover if the one body potential is generated by a ionic lattice, the Hamiltonian 4.1 has to be invariant also with respect to a translation given by a vector of the ionic lattice. Notice that only for neutral systems the sum of the one and the two body potential 4.1 is well defined and convergent.

4.1 Periodic Wave-Function

As far as the electron part is concerned, by applying the Bloch’s theorem one finds that the eigen-functions for the Hamiltonian 4.1 can be written as:

$$\Psi(\vec{r}_1, \vec{r}_2, ...) = e^{-i\vec{k}\sum_i \vec{r}_i}\Phi(\vec{r}_1, \vec{r}_2, ...), \tag{4.2}$$

where $\Phi(\vec{r}_1, \vec{r}_2, ...)$ is a function invariant for translation of any electron coordinates by a vector $\vec{R}_s$, and $\vec{k}$ is a vector in the first Brillouin Zone. Although better choices are possible, as Baldereschi’s points, or using Twisted Boundary Conditions (76), in this preliminary work on solids we adopted the simplest choice

$\vec{k} = 0$. The correct thermodynamic limit, within the Bohr-Oppenaimer approximation, can be obviously reached for $L \rightarrow \infty$ at fixed density $\rho = N/L^3$. In our simulations the distances are evaluated from the closest image of a given particle. One has to choose carefully an appropriate wave-function for a periodic system. In fact, as the minimal inter-particle distance switches from one image to another, there could be a discontinuity in the derivatives of the wave function. If this happen, the VMC energy can become lower than the true ground state one. This is due to the fact that the discontinuity leads to δ functions that produce a finite positive contribution to the kinetic energy that however will be missed by the VMC sampling procedure, because it occurs in an irrelevant surface of the configuration space. In the past this problem has been solved by making use of different approaches: either by summing over all possible images by the Ewald sums or by requiring that the trial-function and its derivatives vanish at the surface of the sphere inscribed within the Wigner-Seitz cell (77). Instead in the present thesis we introduced a simple and more efficient approach by using periodic orbitals with the correct behavior at $L/2$ without resorting the expensive evaluation of the Ewald sums.

4.1.1 Periodic orbitals

In order to obtain periodic orbitals, starting from non periodic ones it is sufficient to replace the Cartesian coordinates x_i with a simple periodic function $x'_i(x)$ that take into account the appropriate periodicity of the box. In this thesis we used:

$$x'_i = \frac{L}{\pi} \sin\left(\frac{\pi x_i}{L}\right) \tag{4.3}$$

and the new distance is defined as

$$r' = \frac{L}{\pi} \sqrt{\sum_{i=1}^{3} \sin^2\left(\frac{\pi x_i}{L}\right)} \tag{4.4}$$

In doing so, we have only to compute gradients and Laplacian with the chain rule:

$$\begin{aligned} \frac{\Phi(r')}{\partial x_i} &= \frac{\partial \Phi(r')}{\partial r'} \frac{\partial r'}{\partial x'_i} \frac{\partial x'_i}{\partial x_i} \\ \frac{\partial^2 \Phi(r')}{\partial x_i^2} &= \frac{\partial^2 \Phi(r')}{\partial r'^2} \left(\frac{\partial r'}{\partial x'_i} \frac{\partial x'_i}{\partial x_i} \right)^2 + \frac{\partial \Phi(r)}{\partial r'} \left[\frac{\partial^2 r'}{\partial x_i'^2} \left(\frac{\partial x'_i}{\partial x_i} \right)^2 + \frac{\partial r'}{x'_i} \frac{\partial^2 x'_i}{\partial x_i^2} \right] \end{aligned}$$

where

$$\begin{aligned}\frac{\partial x'_i}{\partial x_i} &= \cos\left(\frac{\pi\, x_i}{L}\right)\\ \frac{\partial^2 x'_i}{\partial x_i^2} &= -\frac{\pi}{L}\sin\left(\frac{\pi\, x_i}{L}\right)\end{aligned}$$

This transformation has been applied to all orbitals appearing in the wave-function and also to the one-body term and the two-body Jastrow.

We remark here that also the normalization constant of a given orbital has to be changed in a periodic system. Namely its integral over the simulation cell has to be equal to one.

$$\int_{-L/2}^{L/2}\int_{-L/2}^{L/2}\int_{-L/2}^{L/2}\phi^2(r_{iA})\,dx\,dy\,dz = 1 \tag{4.5}$$

For instance a normal Gaussian in three-dimension:

$$\Phi(r) = \left(\frac{2k}{\pi}\right)^{3/4} e^{-kr^2} \tag{4.6}$$

becomes after the substitution 4.3:

$$\Phi'(r') = \left(Le^{-\frac{kL^2}{\pi^2}} I_0\left[\frac{kL^2}{\pi^2}\right]\right)^{-3/2} e^{-kr'^2} \tag{4.7}$$

where I_0 is the modified Bessel function of the first kind and L is the size of the simulation box and r' is the periodic distance 4.4.

4.1.2 The wave-function for high pressure hydrogen

In order to study the high pressure hydrogen we used the periodic generalization of the JAGP wave-function defined in the chapter 1. In the two-body 1.2.3 terms and one-body 1.2.2 terms the distances electron-electron and electron-ion are replaced with the periodic distance:

$$\left|\vec{\xi}_i - \vec{\xi}_j\right| = \frac{L}{\pi}\sqrt{\sin^2\left[\frac{\pi}{L}(\xi_i^1 - \xi_j^1)\right] + \sin^2\left[\frac{\pi}{L}(\xi_i^2 - \xi_j^2)\right] + \sin^2\left[\frac{\pi}{L}(\xi_i^3 - \xi_j^3)\right]} \tag{4.8}$$

where ξ indicates electron and proton coordinates r_i,R_i. Both in the pairing determinant 1.2.1 and in the three-body Jastrow 1.2.4 we used one orbital per atom as basis set. In the first one we used periodic Gaussian orbitals 4.6 while in the second $2s$ orbitals defined as:

$$\phi_{2s}(r) = r^2 e^{-zr}$$

where the distance between electrons is defined as 4.8. We found that this basis set was sufficient to describe accurately the systems studied.

4.2 Coulomb Interactions in periodic systems

In the evaluation of the potential energy in a periodic system the interaction with all possible images has to be considered. This fact could make very inefficient the simulation of periodic systems. The Coulomb interaction ion-ion, ion-electron and electron-electron can be generally written as:

$$U = \frac{1}{2} \sum_{\vec{\xi}_i \neq \vec{\xi}_j + \vec{R}_s, \vec{R}_s} \frac{q_i q_j}{\left| \vec{\xi}_i - \vec{\xi}_j + \vec{R}_s \right|}, \tag{4.9}$$

where ξ_i indicates electron coordinates $\vec{r}$ corresponding to $q_i = -e$ and proton coordinates R_i corresponding to $q_i = +e$ and$\vec{R}_s$ are the vectors of the periodic lattice associated with the simulation box. Notice that this summation converges only for neutral systems $\sum q_i = 0$. For short range interaction it is possible to consider only the closest images, that represents an efficient and accurate way to calculate the potential energy. For long range interaction the equation eq. 4.9 cannot be used in a numerical simulation because the sum is very slowly convergent, so other approaches are necessary.

It is not possible to use a truncated Coulomb potential. In fact, large inaccuracies are introduced by neglecting the long-range part (see Ref. (78)).

In the following we present the well known Ewald method that allows to evaluate in an efficient way the potential energy in periodic systems.

4.2.1 Ewald Sums

In 1921 Ewald (79) proposed an efficient way to recast the summation 4.9 in two rapidly converging series. Here in order to derive in a systematic and controlled way the final result we consider a Yukawa potential $v(r) = e^{-\epsilon|r|}/|r|$ and take the limit $|\epsilon| \rightarrow 0$ only in the final expression. Following the Ewald's idea we split the potential in two parts:

$$v(|r|) = v_{long}(|r|) + v_{short}(|r|) \tag{4.10}$$

where

$$\begin{aligned} v_{short}(|r|) &= v(|r|)erfc(\sqrt{\alpha}|r|) \quad (4.11)\\ v_{long}(|r|) &= v(|r|) - v_{short}(|r|) = erf(\sqrt{\alpha}|r|)v(|r|), \quad (4.12) \end{aligned}$$

erf is the error function and $erfc$ the complementary one. Notice that the long range part has several important properties:

$$\begin{aligned} \lim_{r\rightarrow 0} v_{long}(r) &= 2\sqrt{\alpha/\pi} \quad (4.13)\\ v_{long}(k) &= 4\pi/k^2 e^{-k^2/(4\alpha)} \text{ for } \epsilon \rightarrow 0 \quad (4.14)\\ v_{long}(k=0) &= 4\pi/\epsilon^2 \ \text{ finite only for } \ \epsilon > 0 \quad (4.15) \end{aligned}$$

On the other hand the short range potential decays very fast in real space and the sum converges very quickly. Since U in Eq.(4.9) depends linearly on the potential, we can easily decompose two contributions: a short-range and a long range one. Then the latter can be more easily evaluated in Fourier space:

$$\begin{aligned} U &= U_{short-range} + U_{long-range}\\ U_{short-range} &= \frac{1}{2}\sum_{\xi_i \neq \xi_j + R_s, R_s} q_i q_j v_{short}(|\xi_i - \xi_j + R_s|)\\ U_{long-range} &= \frac{1}{2}\sum_{\xi_i,\xi_j,R_s} q_i q_j v_{long}(|\xi_i - \xi_j + R_s|) - \frac{1}{2}\sum_i v_{long}(r \rightarrow 0) q_i^2\\ &= \frac{1}{2V}\sum_{\vec{k}\neq 0}\sum_{i,j} q_i q_j e^{i\vec{k}(\xi_i - \xi_j)} v_{long}(k) - \sqrt{\frac{\alpha}{\pi}}\sum_i q_i^2 \end{aligned}$$

where V is the volume of the unit cell and the sum over the momenta are on the discrete $\vec{k}$ values allowed by the periodicity $\vec{k}\cdot\vec{R}_s = 2\pi n$. In the latter expression we have used Eq.(4.15) and the fact that the charge neutrality $\sum_i q_i = 0$ implies that the $\vec{k} = 0$ term can be omitted in the sum for any $\epsilon > 0$. In this way the limit $\epsilon \rightarrow 0$ can be found consistently also for long range potentials by replacing expression (4.14) in the corresponding Fourier transform for v_{long}. For a non neutral system instead the Ewald sum is divergent as expected.

For Coulomb interaction the potential energy becomes:

$$\begin{aligned} U &= \frac{1}{2}\sum_{i,j}^{N}\sum_{R_s}\frac{q_iq_j}{\vec{r}_{ij}+\vec{R}_s}erfc\left(\sqrt{\alpha}(\vec{r}_{ij}+\vec{R}_s)\right) \\ &+ \frac{1}{2}\sum_{\vec{k}\neq 0}\sum_{i,j}^{N}\frac{4\pi q_iq_j}{V|\vec{k}|^2}e^{i\vec{k}(\vec{r}_i-\vec{r}_j)}e^{-|\vec{k}|^2/4\alpha}-\sqrt{\frac{\alpha}{\pi}}\sum_{i=1}^{N}q_i^2. \end{aligned} \tag{4.16}$$

In the potential energy 4.16 the parameters α determines the convergence speed in the real and Fourier space series. For a given choice of α we have chosen a real-space cutoff distance r_c and a k_c cutoff in the Fourier space. The cutoff k_c determines the total number of Fourier components, $(4\pi/3)n_c^3$, where n_c is a positive integer. This parameter has been choosen in such a way that the error on the Ewald summation is much smaller than the Quantum Monte Carlo statistical one.

A careful choice of the parameter α can minimize the error in the summation (see Ref. (80)). In our simulation we have chosen $\alpha = L/5$, where L is the size of the simulation box. With this cutoff it is sufficient to sum the short range part in the eq. 4.16 only on the first image of each particle. Notice that during each VMC or DMC simulation the ionic coordinates are fixed throughout the calculation. Therefore the contribution of the ion-ion Coulomb interaction in the short-range part can be evaluated only at the beginning of the simulation. As an electron i is moved during a VMC calculation the sum of the short range part of the eq. 4.16 is easily updated subtracting the old contribution electron-electron and electron-ion due to the electron i, and adding the new one.

The sum in Fourier space can be written as:

$$U_k = \sum_{\vec{k}\neq 0}\frac{4\pi}{V|\vec{k}|^2}e^{-|\vec{k}|^2/4\alpha}\frac{1}{2}\left[\left(\sum_{i}^{N}\sin(\vec{k}\vec{r}_i)\right)^2+\left(\sum_{i}^{N}\cos(\vec{k}\vec{r}_i)\right)^2\right], \tag{4.17}$$

then for each $\vec{k}$ vector all sin and cos are stored in such a way that when an electron moves, the sum can be easily updated without calculating all the elements from scratch.

It is easy to understand that the Ewald summation scales as $O(N^2)$. In fact the updating of the eq. 4.16 costs N times the number of Fourier's components. Then the number of Fourier component goes as $(1/\alpha)^3$ where α is proportional L and for a given density scales as N. The Ewald sums are faster than the QMC sweep and so, even if nowadays other faster techniques exist, as for instance particle-mesh-based one, it was not necessary to adopt other more complicated methods in our calculation.

4.2.2 Forces with finite variance in periodic systems

The method present in section 1.1.1 can be easily generalized to periodic systems. It is sufficient use an auxiliary periodic function $\tilde{\Psi}$ with the same behavior of 1.12 close to nuclei. We have used the following form:

$$\begin{aligned} \tilde{\Psi}_{PBC} &= Q_{PBC}\Psi_T \\ Q^{\nu}_{PBC} &= Z_A \sum_{i=1}^{N_{elect}} \frac{\frac{L}{2\pi}\sin\left(\frac{2\pi}{L}(x_\nu - R^{\nu}_A)\right)}{r'_{iA}} \end{aligned}$$

where r'_{iA} is the periodic distance between the nucleus A and the electron i:

$$r'_{iA} = \frac{L}{\pi}\sqrt{\sin^2\left(\frac{\pi}{L}(x^1_i - R^1_A)\right) + \sin^2\left(\frac{\pi}{L}(x^2_i - R^2_A)\right) + \sin^2\left(\frac{\pi}{L}(x^3_i - R^3_A)\right)}. \tag{4.18}$$

Notice that Ψ_{PBC} is a periodic function because we are using a periodic trial wave-function Ψ_T (see section 4.1.2). This auxiliary function Ψ_{PBC} removes the divergence in the bare force and it is consistent with the periodicity of the system. At variance of the case without periodic boundary conditions $\nabla^2 Q_{PBC}$ does not cancel exactly the term coming from the derivative of the ion-electron potential. Therefore we have to included both the Laplacian of the Q_{PBC} and the derivatives of the ion-electron potential in the calculation of the forces. More precisely the expression we used for the force is:

$$\tilde{F}^{\nu} = F^{\nu}_{bare} - \frac{1}{2}\nabla^2 Q^{\nu}_{PBC} - \frac{\vec{\nabla} Q^{\nu}_{PBC}\vec{\nabla}\Psi_T}{\Psi_T} \tag{4.19}$$

where:

$$\begin{aligned}
\nabla_i^{\mu} Q_{PBC}^{\nu} &= Z_A \frac{2\pi^2}{L^2} \frac{\sin(\frac{2\pi x_i^{\nu}}{L}) \sin(\frac{\pi x_i^{\mu}}{L}) \cos(\frac{\pi x_i^{\mu}}{L})}{r_i'^3} + \delta_{\mu,\nu} Z_A \frac{2\pi}{L} \frac{\cos(\frac{2\pi x_i^{\nu}}{L})}{r'} \\
\nabla_i^2 Q_{PBC}^{\nu} &= Z_A \sum_{\mu=1,3} -\frac{\pi^2}{L^2} \frac{\sin(\frac{2\pi x^{\nu}}{L})}{2} \left\{ 3 \left[\sin\left(\frac{\pi x_i^{\mu}}{L}\right) \cos\left(\frac{\pi x_i^{\mu}}{L}\right) \right]^2 \frac{1}{r'^2} \right. \\
&- \left. \cos^2\left(\frac{\pi x_i^{\mu}}{L}\right) + \sin^2\left(\frac{\pi x_i^{\mu}}{L}\right) \right\} \frac{1}{r'^3} \\
&- 2Z_A \frac{\pi^2}{L^2} \left[-\sin\left(\frac{2\pi x_i^{\nu}}{L}\right) - \sin\left(\frac{\pi x_i^{\nu}}{L}\right) \cos\left(\frac{\pi x_i^{\nu}}{L}\right) \cos\left(\frac{2\pi x_i^{\nu}}{L}\right) \frac{1}{r'^2} \right] \frac{1}{r'}
\end{aligned}$$

4.3 How to evaluate pressure

Following the seminal paper of O.H. Nielsen an R.M. Martin Ref. (81), the pressure can be expressed as the negative trace of the stress tensor $T_{\alpha\beta}$:

$$T_{\alpha\beta} = \frac{\partial}{\partial \epsilon_{\alpha\beta}} \frac{\langle \Psi_\epsilon | H | \Psi_\epsilon \rangle}{\langle \Psi_\epsilon | \Psi_\epsilon \rangle} \tag{4.20}$$

where the Ψ_ϵ is the wave-function ”stretched” by the transformation on each particle $r_{i\alpha} \rightarrow r_{i\alpha} + \sum_\beta \epsilon_{\alpha\beta} r_{i\beta}$ where $\epsilon_{\alpha\beta}$ is a symmetric strain tensor. The pressure is then defined through the negative trace of the stress tensor:

$$3PV = -Tr[T_{\alpha\beta}] \tag{4.21}$$

For an isotropic system, as the one we studied, the pressure can be easily written as:

$$3PV = \frac{\partial \langle E_l \rangle}{\partial V} = \frac{\partial \langle E_l \rangle}{\partial L} 3L^2. \tag{4.22}$$

where $V = L^3$ and L is the edge of our cubic box. To evaluate this derivative it is convenient to write the energy using rescaled distances that are invariant for stretching of the simulation box, namely $r' = r/L$. After this transformation the expectation value of the energy can be written as:

$$\langle E_l \rangle = \frac{\langle E'_{kin} \rangle}{L^2} + \frac{\langle E'_{pot} \rangle}{L}, \tag{4.23}$$

where E'_{kin} and E'_{pot} are the kinetic and the potential energy in the new coordinates. Thus the pressure will be:

$$P = \frac{1}{3V} \left[2\frac{\langle E'_{kin} \rangle}{L^2} + \frac{\langle E'_{pot} \rangle}{L} \right] + \frac{2}{3V} \left[\langle O_L H \rangle - \langle O_L \rangle \langle H \rangle \right], \tag{4.24}$$

where O_L is the logarithmic derivative of the wave-function respect to the simulation box size L:

$$O_L = \frac{\partial_L \Psi_T}{\Psi_T} \tag{4.25}$$

Notice that the box size L appears in the one-body terms and in the two-body terms due to the cusp condistions, see Appendix C. The first part of the expression 4.24 is the usual Virial Theorem and the second one is given by the contribution due to the Pulay stress tensor. In fact although we used rescaled coordinates, it is not possible to cancel the dependencies of the wave-function from the simulation box size. This is due to the nuclear and electronic cusp conditions that depend explicitly on L (see Appendix C).

In the expression 4.24 only the electronic kinetic energy is considered. When we perform a dynamic on the ionic part, we have to add the pressure due to the momentum flux carried by the ions. This part can be easily calculate with the standard kinetic theory (see Appendix B of Ref. (82)) of gas, and it reads:

$$\langle P_{ionic} \rangle = \frac{2N}{3V} \langle E_k^{ionic} \rangle, \tag{4.26}$$

the total pressure will be the sum of the ionic 4.26 and electronic part 4.24.

4.4 Empirical laws of melting

Although in general the melting and freezing transition is non-universal, there are some useful phenomenological criteria which are usually based on the properties of only one of the two coexisting phases. The advantages of these empirical rules is that they permit an estimation of the solid-liquid coexistence line, without carrying out any free energy calculation.

The empirical rule of bulk melting is the so called Lindeman criterion according to which a crystal melts when the amplitude of thermal vibrations (r.m.s.) exceeds a given thresholds of the order of the lattice space (83). For many materials this Lindemann ratio is about of $\simeq 0.15$ of the lattice space.

Another interpretation of the Lindemann criterion is that an infinite solid will become mechanically unstable at a sufficiently high temperature. Although the ideal mechanical instability temperature of a solid is different, and of course somewhat higher than the true melting temperature T_m (where the free energy crossing of solid and liquid phases takes place),nonetheless it can be heuristically taken as a qualitative indicator of the tendency of the solid to melt.

Chapter 5

A new technique for the simulation of electronic systems at finite temperature by means of noisy QMC forces

The most common application of computer simulations is to predict the properties of materials. Since the first works, by Metropolis et al. and Fermi et al. (27; 84), Molecular Dynamic (MD) techniques turned out to be a powerful tool to reproduce the properties of materials in different conditions and also to predict them.
The combination of these techniques with the density functional theory (DFT) has become a widely accepted and powerful *ab-initio* method: the Car-Parrinello Molecular Dynamics (CPMD) (54) that has allowed to study a broad range of chemical, physical and biological systems. The CPMD approach offers a balance of accuracy and computational efficiency that is well suited for both static and dynamic calculations of numerous properties of systems with hundreds and even thousands of atoms. Although in principle DFT is an exact theory for the electron correlation, it relies on an unknown exchange and correlation functional that must be approximated. The widely used Local Density Approximation (LDA) is difficult to improve systematically. Therefore, in some cases (see for instance Ref. (85)), one requires a more accurate computational approach, such as the quantum Monte Carlo (QMC) approach to solve the Schrödinger equation very accurately.

In this thesis, we present a new method that treats the electrons within the many-body QMC and perform Molecular Dynamic ”on the fly” on the ions. This method provides improved dynamical trajectories and significantly more accurate total energies.

In the past two different approaches were proposed to couple Quantum Monte Carlo with ionic Molecular Dynamic. The first, called Coupled Electronic-Ionic Monte Carlo (CEIMC) (8), is based on a generalized Metropolis algorithm that takes into account the statistical noise present in the QMC evaluation of the Bohr-Oppenheimer surface energy. In the second approach, called Continuous Diffusion Monte Carlo (CDMC) (64), the Molecular Dynamics trajectories are generated with some empirical models or by CPMD-DFT, and then the CDMC technique is used to efficiently evaluate energy along the trajectories. Both methods present some drawbacks. In the second method even if all the properties are evaluated using the Diffusion Monte Carlo, the trajectories are generated using empirical models without the accuracy given by the QMC for the structural properties, as radial distribution, bonding lengths and so on. Instead, in the first one the QMC energies are used to perform the Monte Carlo sampling leading to accurate static properties. In order to have a reasonable acceptance rate within this scheme simulations have to be carryed out with a statistical error on the energy of the order of $K_b T$ Furthermore, in order to have a fixed acceptance rate the amplitude of the ionic move has to be decreased with the size of the system.

The method we present here, allows to solve two major drawbacks of the previous two techniques. Following the idea of Car and Parrinello (54) we will show that it is possible to perform a feasible *ab-initio* Molecular Dynamics and structural optimization in the framework of the Quantum Monte Carlo by using noisy ionic forces, and with a method that do not contain any rejection scheme, at the expense of a time discretization error, that is present in any type of MD scheme.

5.1 The Born-Oppenheimer approximation

The idea of treating ionic dynamics classically, while electrons are in the ground-state of the Hamiltonian 4.1, is based on two very reasonable approximations: the Born-Oppenheimer Approximation(BO) and the Adiabatic one.

In a system of interacting electrons and nuclei there will be usually small momentum transfer between the two types of particles due to their very different masses.

On the time-scale of nuclear motion, one can therefore consider the electrons to relax in the ground-state given by the Hamiltonian with the nuclei at fixed locations. This separation of the electronic and nuclear degrees of freedom is known as the Born-Oppenheimer approximation. Moreover since the energy scale associated with the electronic excitations is usually much larger than to the one related to the ionic motion, one can safely consider the electron in their own ground-state. Although this approximation is not always fulfillid even in the worst case, for the simulation of the lightest atom, hydrogen, Galli et al. (21), using the Car-Parrinello molecular dynamics with DFT, showed that the electronic band-gap is about $2eV$ and that the first order correction due to the quantistic effects on ions is about $2meV$ for pressure up to $200GPa$.

Although there are techniques, as Path Integral Monte Carlo, to treat finite temperature quantum systems, they become extremely inefficient for low temperature regime, therefore we have preferred to simply neglect quantum effects due to the finite protons mass.

5.2 Dealing with Quantum Monte Carlo noise

Recently different method were proposed to evaluate forces by Quantum Monte Carlo with a finite and small variance (29),(52),(86) (see also section 1.1.1 for a discussion about zero-variance principle).

It is well known that noisy forces can be used in different way for obtaining, following a first order stochastic differential equation, the Canonical distribution. For instance it is possible to use the Langevin dynamic defined by:

$$\begin{aligned} \dot{x}_i &= \beta_{ij}\left(-\frac{\partial V}{\partial x_j} + \eta_j\right) & (5.1) \\ \langle \eta_i(t) \rangle &= 0 & (5.2) \\ \langle \eta_i(t)\eta_j(t') \rangle &= \alpha_{ij}(x)\delta(t-t'), & (5.3) \end{aligned}$$

where η is a random noise with variance $\alpha_{ij}(x)$ and zero mean. It is easy to show, using the Fokker-Plank equation associated to this equation, that in order to obtain the usual Boltzmann distribution the matrix β has to be chosen as$\beta = \alpha^{-1}K_bT$. The problem to obtain the desired canonical distribution may be therefore solved in this way. In QMC one can calculate the covariance matrix $\alpha_{ij}(x) = \langle f_i f_j \rangle - \langle f_i \rangle \langle f_j \rangle$ and then invert this matrix to obtain β and continue the dynamics. This method is unfortunately very unstable, because the matrix α^{-1} can be ill-defined because of statistical fluctuations. Moreover it is not possible to estimate the error on the temperature simulated.

Here we present a new method that uses these QMC forces to perform a Molecular Dynamics at finite temperature. In the past the major problem of using QMC to perform *ab-initio* Molecular Dynamic was the presence of the statistical noise, but now we will show that this noise can be efficiently used as thermal bath. We called this method Generalized Langevin Dynamics by Quantum Monte Carlo (GLQ).

In our simulation there exists a correlated noise associated to the forces. We rely on the central limit theorem implying the noise in all component of the forces evaluated by QMC is Gaussian with a given covariance matrix. We used the Jackknife re-sampling method (see Appendix A) to estimate the covariance matrix.

The idea of the GLQ is to use this noise to produce a given finite temperature using a Generalized Langevin Equation. The use of the Generalized Langevin Equation (GLE) as thermostat is not new. In the past some authors have used this approach to simulate different systems. This method was applied for the fist time by Schneider and Stoll (87), to study distortive-phase transitions. Later the GLE was used to simulate different systems and also to stabilize the usual Molecular Dynamic method (88).

5.3 Canonical ensemble by Generalized Langevin Dynamics

In order to simulate the canonical ensemble we use a Langevin dynamics and we assume that our system is coupled with a thermal bath due to Quantum Monte Carlo statistical noise plus an additional friction term:

$$\begin{cases} \dot{v}_i(t) = -\gamma_{ij}(x) v_j(t) + \frac{f_i(x(t))}{m_i} + \Gamma_i(t) \\ \dot{x}_i(t) = v_i(i) \end{cases} \tag{5.4}$$

with

$$\begin{aligned} \langle \Gamma_i(t)\Gamma_j(t')\rangle &= \alpha_{ij}(x)\delta(t-t') && (5.5)\\ \langle \Gamma_i(t)\rangle &= 0. && (5.6) \end{aligned}$$

where Γ is a generic Gaussian noise that implicitly contains the one associated to the statistical evaluation of the forces by QMC. Notice that in this case an explicit dependence on x of the noise has been taken in account. This is a realistic case that has not been considered so far not even in (89). In the following we determinate a form for the friction matrix $\gamma_{ij}(x)$ that allows to converge the usual Boltzmann distribution at a given temperature.

To this purpose we write down the corresponding Fokker-Planck equation. Following (90) we have to evaluate the drift and the diffusion coefficient:

$$\begin{aligned} D_i(x,t) &= \lim_{\tau\to 0}\frac{1}{\tau}\langle x_i(t+\tau)-x_i(t)\rangle && (5.7)\\ D_{ij}(x,t) &= \frac{1}{2}\lim_{\tau\to 0}\frac{1}{\tau}\langle [x_i(t+\tau)-x_i(t)][x_j(t+\tau)-x_j(t)]\rangle && (5.8) \end{aligned}$$

A straightforward calculation shows that:

$$\begin{aligned} D_{x_i}(t) &= v_i(t) && (5.9)\\ D_{v_i}(t) &= -\gamma_{ij}(x)v_j(t)+\frac{f_i(t)}{m_i} && (5.10)\\ D_{x_i,x_j} &= D_{x_i,v_j}=0 && (5.11)\\ D_{v_i,v_j} &= \frac{\alpha_{ij}(x)}{2} && (5.12) \end{aligned}$$

And so the corresponding Fokker-Plank equation will be:

$$\frac{\partial W(x,v,t)}{\partial t} = \sum_i \left\{ \frac{\partial}{\partial x_i}v_i + \frac{\partial}{\partial v_i}\left[-\gamma_{ij}(x)v_j+\frac{f_i}{m_i}\right] + \frac{\partial}{\partial v_i}\left[\frac{\alpha_{ij}(x)}{2}\frac{\partial}{\partial v_j}\right]\right\} W(x,v,t) \tag{5.13}$$

Then the friction matrix $\hat{\gamma}$ is chosen in a way that the stationary solution of the Fokker-Planck equation is the canonical distribution:

$$p_{eq}(v_1,...,v_n,x_1,...,x_n) \simeq e^{-\beta H}. \tag{5.14}$$

More precisely by substituting the Boltzmann distribution

$$W_{eq}(x,v) = e^{\frac{-\sum_i \frac{m_i v_i^2}{2} - V(x)}{KT}} \tag{5.15}$$

in the equation 5.13 we obtain:

$$\gamma_{ij}(x) = \frac{\alpha_{ij}}{2}(x)\beta m_j \tag{5.16}$$

So for a given noise on the forces α_{ij} and the desired temperature we can set the friction tensor using eq. 5.16 in order to obtain the Boltzmann distribution.
Notice that the external random noise scale as $\sqrt{\tau}$ (see also Ref. (90)) whereas the QMC noise scales as τ. Therefore in the limit $\tau \to 0$ if we do not add any external noise and we set $\hat{\gamma}$ according to eq. 5.16, the system will converge to the Newton dynamics at zero temperature. Nevertheless adding an appropriate external noise it is possible equilibrate the system to the canonical ensemble at the desired temperature.
A peculiar feature of this approach is that in the limit of small τ the statistical error on the forces becomes irrelevant because $\tau \leq \sqrt{\tau}$.
The stability of this approach compared with the first order Langevin Dynamic is now evident, in fact there is no need to calculate the inverse of the covariance matrix. Moreover in the second order Langevin Dynamics the temperature can be estimated at posteriori by equality:

$$\frac{3}{2}K_B T = \frac{1}{2}M\langle V^2\rangle, \tag{5.17}$$

compensating the error in the integration of the GLE.

5.3.1 Numerical integration of the Generalized Langevin Equation

In the literature there are different algorithms to integrate numerically the Generalized Langevin Equation. The most common ones are the BBK (91), vGB82 (92) and the Impulse Integrator (LI) (93). All of them, in the limit of Newtonian dynamics, as $\gamma \to 0$, reduce to the well known Verlet method:

$$x^{n+1} = 2x^n - x^{n-1} + \tau^2\frac{F(x^n)}{M}. \tag{5.18}$$

Although these methods offer good numerical accuracy, other criteria have to be considered for the choice of the algorithm. In our approach the friction is related to the quantum Monte Carlo noise, and so, in order to simulate low temperate phases, it can happen that we are forced to use a large friction matrix. Due to this, an integration algorithm that allows to work accurately with a large friction is required. In the limit of large friction, or equivalently $m_i \rightarrow 0$ the eq. 5.4 reduces to the simple Brownian dynamics. In this case the BBK scheme becomes unconditionally unstable (94), and this fact automatically excludes this algorithm. Instead the other two schemes reduce respectively to the second-order explicit Adams formula and to the Euler-Maruyama method (see Ref. (93)). In this thesis we have decided to use the Impulse Integrator proposed by Skell and Izaguirre (93) that achieves a good accuracy (see Appendix E) and it is simple to implement. Moreover in order to integrate the Langevin dynamics 5.4 we have generalized this algorithm to a non-diagonal friction tensor. So we rewrite the equation 5.4 in the simpler form:

$$\dot{v} + Av = g \tag{5.19}$$

where A is the friction tensor and $g = \frac{f}{m} + \Gamma$. The forces F are evaluated with QMC, Γ contains therefore the noise of the QMC evaluation of f and a possible additional external generated noise. The friction tensor is assumed to be determined without noise but can explicitly depend on the atomic positions. [1]. Then we factorize the matrix $A = L\Lambda L^T$, where Λ is a diagonal matrix and L contains the corresponding eigenvectors. Substituting this factorization into eq. 5.19:

$$L^T\dot{v} + \Lambda L^T v = L^T g \tag{5.20}$$

Defining the vectors $w = L^T v$ we obtain:

$$\dot{w} + \Lambda w = L^T g. \tag{5.21}$$

Now the equation is in the usual diagonal form with the new forces $s = L^T g$. We assume that the friction matrix is slowly varying compared to the forces g and so the usual integral scheme LI(93) is used for the variable w. Finally the transformation $v = Lw$ is applied to come back to the original variables. The final integration formula is:

$$r^{n+1} = L(I + e^{-\Lambda\tau})L^T r^n - Le^{-\Lambda\tau}L^T r^{n-1} + L\tau\Lambda^{-1}(I - e^{-\Lambda\tau})L^T g \tag{5.22}$$

---

[1] Strictly speaking in the following we relate γ to the covariance matrix, that then is evaluated statistically. We neglect here this statistical noise.

Then we can write the final formula to obtain velocities:

$$v^n = L\frac{\Lambda e^{-\Lambda\tau}}{I - e^{-\Lambda\tau}}L^T(r^n - r^{n-1}) + L\tau\left(I - \frac{e^{-\Lambda\tau} - I + \Lambda\tau}{\Lambda\tau(I - e^{-\Lambda\tau})}\right)L^T g \qquad (5.23)$$

5.4 Practical implementation of the finite temperature dynamics

In the same spirit of the Car-Parrinello dynamics, at each ionic move, eq. 5.22, the parameters are optimized using eq. 2.31. This has allowed us to relax the wave-function to the energy minimum during the ionic dynamics . We call this technique Generalized Langevin Dynamics using Quantum Monte Carlo noisy forces(GLQ).

In order to simulate finite temperature systems using the GQL dynamics some control parameters has to be fixed for an efficient scheme. Let us imagine that we want to simulate a system at given temperature T and density r_S. First of all we have to choose the friction tensor. Two choices are possible: the first one is to work with the friction as small as possible, compatibly with the QMC noise, in order not to affect dynamic properties; the second choice is to choose the friction in such a way to achieve the maximum convergence speed. In this thesis we did not investigate dynamical properties and so we opted for the second choice. Second, we have to determine the two parameters r and Δ see eq. 2.33 and 5.22. We have chosen r as big as possible to have a stable optimization. Instead Δ has been chosen enough small to allow the Hessian optimization to follow the ionic dynamics. This can be easily checked controlling if the forces 2.11 are zero within a given accuracy. For example for a system of 16 hydrogen atoms we have used a time step Δ of the order of $0.3fs$ for a temperature around $100K$. For higher temperature in order to maintain the same precision between ionic dynamics and optimization of parameters the time step Δ is roughly rescaled as $\sim 1/\sqrt{T}$ in such a way to maintain the same mean ionic step $\langle|\Delta R|\rangle \sim \sqrt{\langle v^2\rangle}\Delta$. Moreover to have a stable minimization not all parameters have to be changed at each ionic step but only the most relevant, see the forthcoming section 5.4.3.

Thanks to GLQ technique we were able to simulate reasonably large systems, by using highly correlated wave-functions with many parameters (see figure 5.1) and with essentially a single processor machine.

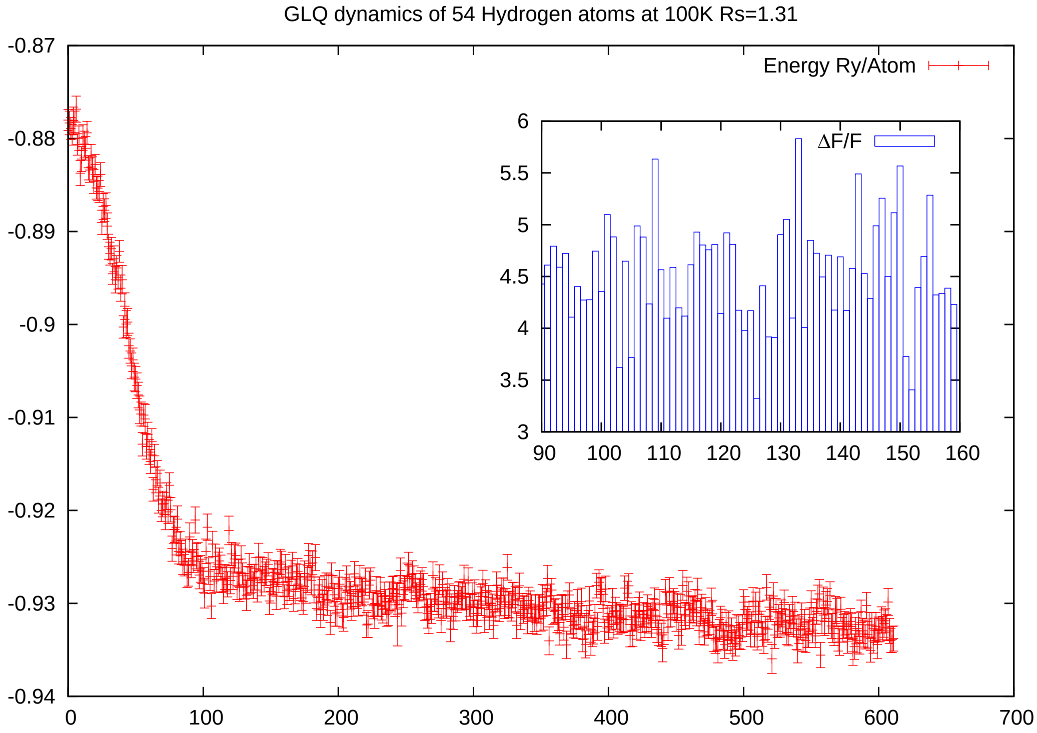

Figure 5.1: Ionic dynamics of 54 hydrogen atoms using GLQ, with a time step $0.4fs$, starting from a BCC lattice. The trial wave-function contains 2920 variational parameters and we have optimized 300 of them at each step. In the inset the maximum deviation $F_i/\Delta F_i$ of the forces acting on the variational parameters is shown.

5.4.1 Setting the parameters in the Langevin dynamics

Parameters of the Eq. 5.22 can be tuned in different ways according to what we want to simulate. As shown by Ref (87; 88) the Generalized Langevin Dynamic allows to study dynamical properties with the condition that the friction term γ is small compared to the typical frequencies of the system, in fact too large γ overdamps the low-frequency vibrational modes (see Ref. (95)). Therefore in order to have the desired small Γ the only possible solution is to increase the VMC bin length in such a way that the VMC noise decrease to the desired value.
On the other hand if only static properties are required it is possible to tune GLE parameters in a more efficient way. In fact, although by our method it is possible to work only with QMC noise, it can be convenient sometimes to add external noise to the forces for different reasons.
Let us imagine to simulate a system at high temperature. In that case there are two possibilities: either the bin length used to evaluate the forces has to be decreased in order to increase the QMC noise, or the friction has to be rescaled according to Eq. 5.16 that implies a corresponding reduction of the dissipation in the system. But both these solutions present some problems. In the first case one cannot decrease the bin length below a certain threshold otherwise the hypothesis of Gaussian noise is no longer fulfilled and the optimization algorithm becomes unstable; in the second case the use of a very small friction matrix leads to a long convergence time that is related to the smallest eigenvalues of the γ matrix $\tau \simeq \frac{1}{\gamma_{min}}$.
For these reasons we found that is more convenient to add external noise to the system, in such a way to produce the desired temperature with a given bin length and to maintain the friction matrix not too small. To this purpose we have added a Gaussian noise with diagonal correlation matrix:

$$\Gamma'_i(t) = \Gamma_i(t) + \eta(t) \tag{5.24}$$

$$\langle \eta_i(t)\eta_j(t')\rangle = \beta_{ij}\delta(t-t') \tag{5.25}$$

$$\langle \Gamma_i(t)\eta_j(t')\rangle = 0 \tag{5.26}$$

$$\langle \Gamma'_i(t)\Gamma'_j(t')\rangle = (\alpha_{ij} + \beta_{ij})\,\delta(t-t') \tag{5.27}$$

This procedure allows us to achieve the maximum efficiency when it is possible to set the γ as close as possible to the critical dumping of the system (see Tassone, Car and Mauri Ref. (96)).

5.4.2 Following the ionic dynamics

A reasonable approximation for the physical forces acting on the ions are the ones calculated when the electronic system is at the energy minimum with the chosen variational ansatz. Therefore in order to generate the correct dynamical trajectories for the ions it is extremely important to relax the trial function to the minimum energy at each ionic time step. As shown by different authors first-order optimization methods, as stochastic gradient or stochastic reconfiguration, defined by:

$$\dot{\psi} = S^{-1}\vec{f}, \tag{5.28}$$

where S is a positive definite matrix, and $\vec{f}$ is the vector of the generalized forces eq. 2.11 fails to follow a second order dynamics (97). This causes a systematic error in the ionic forces, because the wave-function is not at the energy minimum.

To overcome this problem many techniques were used, such as Car-Parrinello dynamics ((54)), or conjugate gradients ((97)). In our work we used a new technique that is suited very much for energy optimization by means of Quantum Monte Carlo, the SRH method. The major advantage of this method is that it uses the information coming from the Hessian. Thanks to the Hessian matrix this method is able not only to follow the direction of the minimum but also to estimate the distance from the minimum for each parameter in such a way to converge in few steps. Moreover, if we start from an optimized trial function and we move the ions not too fast, we can maintain the system in a regime in which the quadratic approximation is always valid so that in principle the Hessian optimization converges always in one step.

5.4.3 Reducing the number of parameters

After each ionic move it is important to optimize the trial function to be as close as possible to the variational ground state. The point is that there are some parts of the wave function that vary a bit between different ionic configurations, as for example the two-body Jastrow factor or the core orbitals for large ions. So it is not important to move all parameters at each step, but it is fundamental to recognize which ones have to be optimized because they are far from the minimum. Moreover since we are performing an optimization in the presence of noise it is possible to know the exact minimum only within a given statistical error due to the finite sampling. So we fix a tolerance ξ and move only such parameters whose generalized forces satisfy:

$$\frac{|F_i|}{\Delta F_i} \geq \xi \tag{5.29}$$

we have chosen $\xi \simeq 4$ that amounts to change on average the 20% of the variational parameters. This procedure has allowed us to follow the ion dynamic with a stable and fast optimization on the variational parameters. Moreover to stabilize the optimization procedure other two cautions were used:

1. It is very difficult to parametrize an highly correlated wave-function with many parameters. In fact it happens often that in this case the stochastic matrix becomes singular because there are too many parameters redundant to describe the wave-function and some of them have to be eliminated for a stable optimization scheme, see section 2.1.2.

2. It can happen that the Hessian Matrix is not positive definite, this can be due to different reasons: a statistical fluctuation due the finite sampling or because the trial-function is far from the minimum, in this case the correction, proposed in Ref. (50) and described at the end of section 2.3, is used. In our simulations we have used 0.2 as a threshold value for the r parameters that rules the stability of the SRH optimization (see section 2.3). This has achieved a fast and stable convergence of the wave-function.

Chapter 6

Preliminary results on high pressure hydrogen

6.1 Comparison with previous calculations

In order to check the quality of our trial-function we compared the energy and the variance on different configurations with the ones obtained using other functional form for the trial wave-function (7; 37). In the table 6.1 the energy and the variance for a BCC lattice with 16 hydrogen is reported. As one can see our wave function gives a very good energy. The variance is not so low and this is probably due to the lack of back-flow correlation or to the energy optimization.

Moreover we compare the energies and variances obtained on different con-

Table 6.1: Total energies in variational (E_{VMC}) and diffusion (E_{DMC}) Monte Carlo calculations for 16 hydrogen atoms in a BCC lattice at Rs=1.31 and T=0 (i.e. frozen ion positions). The energies are in *Hartree* for atom.

WF	E_{VMC}	σ^2_{VMC}	E_{DMC}
SJ	-0.4742(2)	0.0764(2)	-0.4857(1)
$SJ3B$	-0.4857(2)	0.0274(2)	-0.4900(1)
LDA	-0.4870(10)		-0.4890(5)
$JAGP$	-0.4871(5)	0.0700(1)	-0.49019(5)
$JAGP-reduced$	-0.4846(2)	0.067(1)	-0.4880(1)

figurations generated with the CEICM method (8). In the figure 6.1 the first ten configurations are obtained at $T = 2000K$. Then the system is cooled to $500K$ and it starts to clusterize. If we compare the wave-function of Holzmann et al. (8) with the JAGP wave-function we can see that the latter one gives an accurate description of both the liquid and the cluster phase. This is due to the resonating nature of the JAGP trial-function that allows to describe the liquid phase through resonating bonds among different atoms.

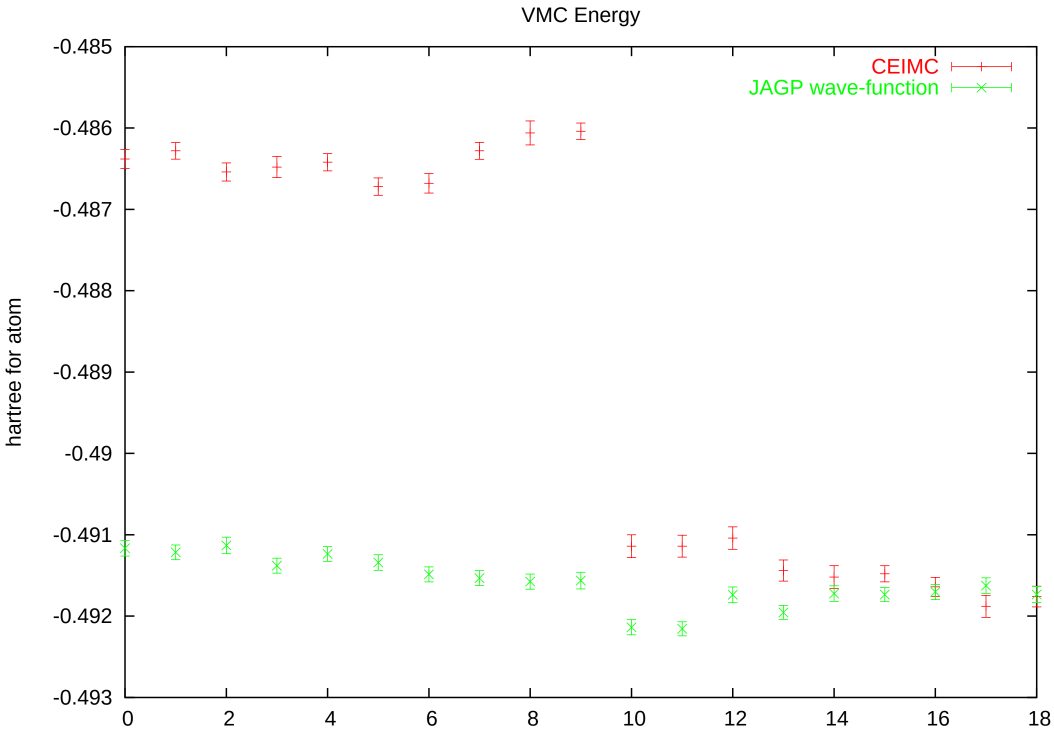

Figure 6.1: Energy per atom of 16 hydrogen atoms at Rs=1.31 calculated on configurations obtained by CEIMC with the method (8). The first 10 configurations are in the atomic liquid phase at 2000k while in the last ten the system is forming clusters at T=500.

Then we compared the pressure obtained by using the GLQ technique, obtained after equilibration at given temperature and density, with both Gas-Gun

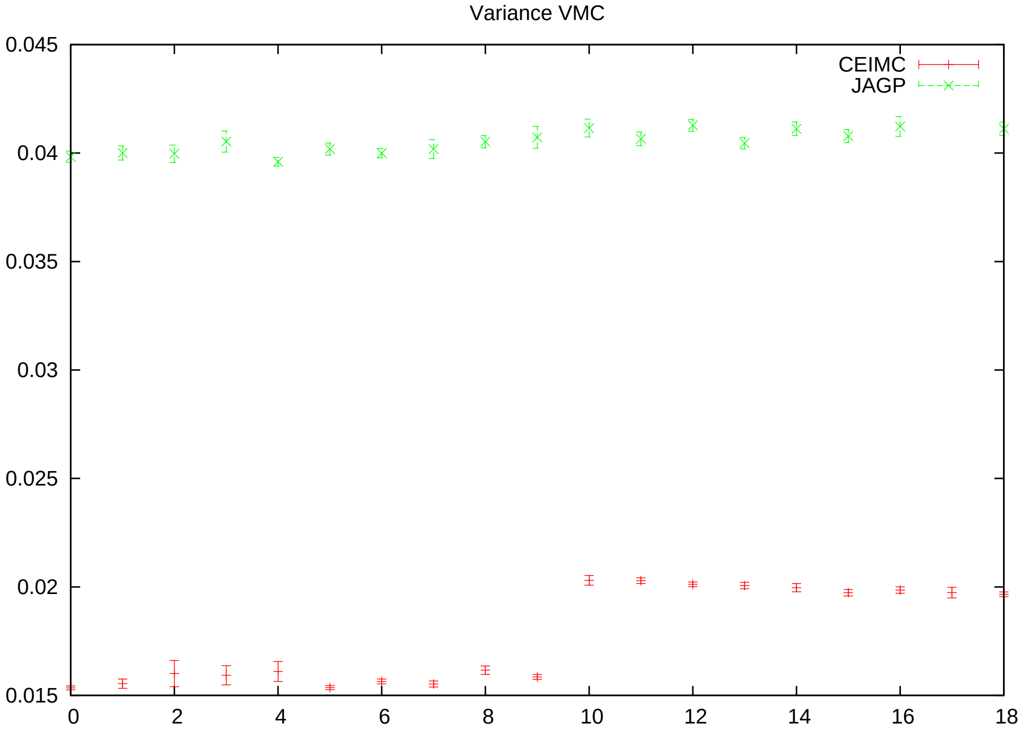

Figure 6.2: Variance per atom of 16 hydrogen atoms at Rs=1.31 calculated on configuration obtained by CEIMC with the method (8). The first 10 configurations are in the atomic liquid phase at 2000k while in the last ten the system is forming clusters at T=500.

Table 6.2: Pressure at different temperatures and densities. We report also the pressure obtained with Gasgun experiment (5), with Silvera-Goldman empirical potential model (6) and CEICM method (7) at Γ point. The pressure are in GPa.

r_s	T	Gasgun	S-G	CEICM-VMC	CEICM-DMC	GLE-VMC
2.202	2820	0.120	0.116	0.105(6)	0.10(5)	0.144(8)
2.1	4530	0.234	0.234	0.226(4)	0.225(3)	0.246(9)
1.8	3000	-	0.528	-	0.433(4)	0.410(8)

experimental results [1] and numerical simulations done with CEICM method (7). In the table 6.2 we report the pressures calculated for different densities and temperatures. The pressure was obtained as explained in section 4.3.
As it is shown the combination of GLQ with the JAGP wave-function provides a good agreement with the experimental values already with this small size. [2].

6.2 Pair Correlation Functions

In this section we report the proton-proton pair correlation functions for different densities and temperatures obtained by using the GLQ technique. The pair correlation function is defined (see Allen Tildesley (98)) as:

$$g(r) = \frac{V}{N^2}\langle\sum_{j\neq i}\delta(r - r_{ij})\rangle \tag{6.1}$$

The pair distribution function is a useful property because it provides insights for the liquid or solid structure. We compared the obtained proton-proton distribution functions with the ones reported in Ref. (7; 99) for different densities .

Hohl et al. (1993) have performed DFT-LDA simulations at rs = 1.78 and T=3000K, the resulting proton-proton distribution functions are compared in 6.3. The lack of accuracy of Local Density Approximation (LDA), used by Hohl et.

---

[1] In the Gas-Gun experiments extraordinarily high pressures are created by the gas gun, occurring during explosions. The high pressures of a shock wave make materials denser and heat them to thousands of degrees.

[2] Notice that our calculations are done at Γ point. Moreover we used 32 atoms and this does not fulfill the closed-shell condition increasing further the size effects.

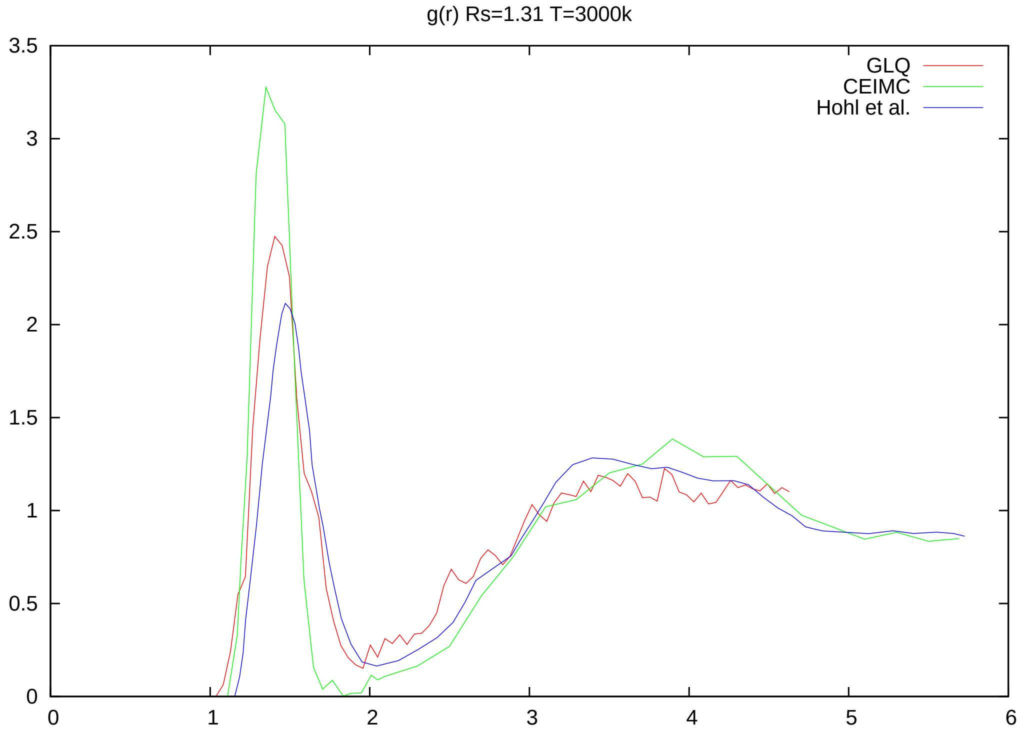

Figure 6.3: Proton-proton correlation function, g(r), at Rs=1.31. The GLQ and CEIMC have used a periodical simulation box with 32 atoms while Hohl et al. with 64 atoms. All the calculations were performed for a single k point (Γ).

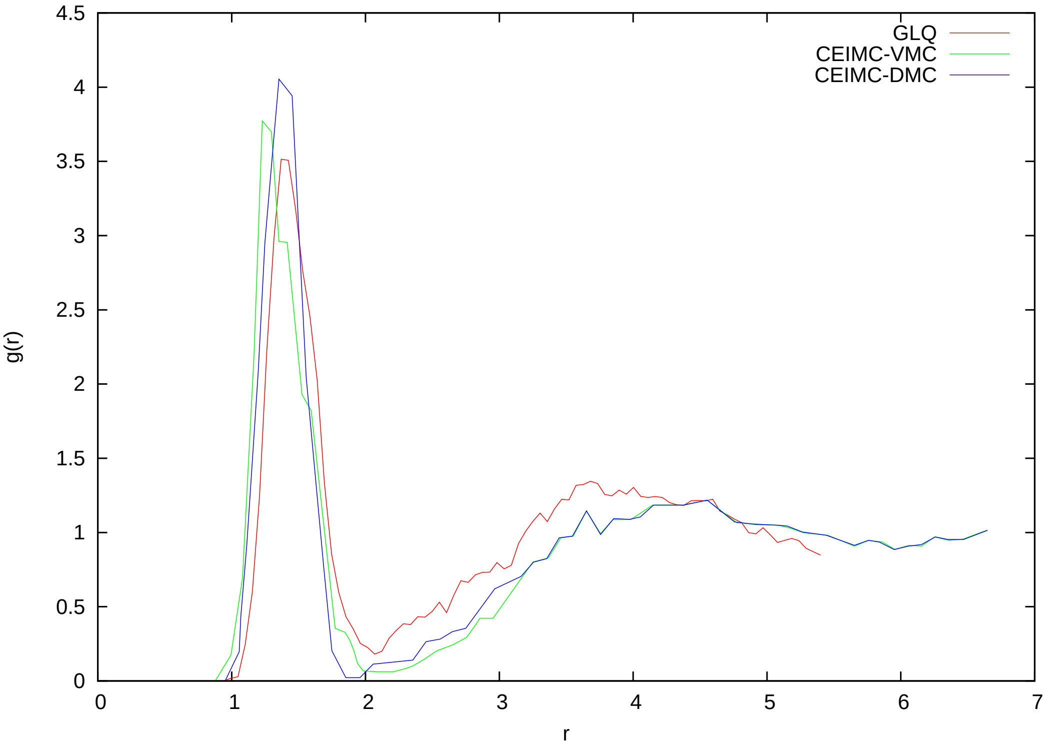

Figure 6.4: Comparison of the proton-proton correlation function, g(r), at Rs=2.1 and T=4350 obtained with different methods CEIMC (8) (7) and GLQ. All the simulations were performed with 32 atoms for a single k point (Γ).

al, and the small size, we simulated, can account for the differences in the location of the peak. The discrepancy between CEIMC and GLQ is not clear at present. It can be explained by the nature of the wave-function used in CEICM that does not describe well the molecular disassociation (7). Nevertheless at density $Rs = 1.31$ the inter-atomic distances are comparable with the typical bonding length of the hydrogen molecule and we have found a tendency to filament formation in agreement with the results of Hohl and Ceperly (99).

6.3 Another possible phase in liquid hydrogen

The Resonant Valence Bond state has been shown to describe accurately frustrated spin systems. It appears that in many systems superconductivity is realized when the system in the normal state is quite close to metal-insulator transition even in presence of strong disorder, see Ref. (100). The large mass difference between ions and electrons leads, in a good approximation, to two distinct dynamics. This allows us to consider always the electrons in the ground state of a disordered ionic system at finite temperature, and from this point of view it is ”similar” to an average on disorder. Following this hypotheses we tried to study the possibility of stable superconducting phases driven by correlation close to the metal-insulator transition and solid-liquid transition. Moreover as in high-Tc superconductor, we expect that due to the small size of the Cooper pair a superconductive state can be more stable close to a disordered phase (101).

In order to detect superconductivity we have calculated the condensation energy on different configurations in the liquid phase, with VMC and DMC. The condensation energy, in a variational scheme, is defined as the energy difference obtained between the best Slater determinant, the normal state, and the AGP. In order to estimate the condensation energy we have reduced the rank of the pairing matrix to $N/2$. In such limit the pairing determinant is equivalent to a Slater one, as shown in (2). We optimized the Slater wave-function on given configurations and we found a non-zero condensation energy (see figure 6.5), namely a gain in energy given by allowing pairing within a variational ansatz. At present we do not even know whether this gain in energy is macroscopic or is just a finite size effect even a the VMC level. The corresponding energy gain, obtained in this way, certainly overestimate a possible true condensation energy due to a real superconducting ground state for high pressure hydrogen. In fact the variational approach is certainly biased towards superconducting phases even though recently in the 2D Hubbard model there are other evidences of superconducting phase from non variational approaches (102).

As shown in the famous paper of Yang (103) a peculiar property of superconductors is the Off-Diagonal Long-Range Order (ODLRO) exhibited by the reduced density matrices in the coordinate space representation (103). For Fermions ODLRO corresponds to the appearance of an eigenvalue which scales with the number of particles in the two-body density matrix (103):

$$\begin{aligned} \rho_2(x_1,x_2;x_1',x_2') &= \langle\Psi_N|a^+(x_1)a^+(x_2)a(x_1')a(x_2')|\Psi_N\rangle & (6.2)\\ Tr\rho_2 &= \frac{1}{2}\int \rho_2(x_1,x_2;x_1,x_2)dx_1dx_2 = \frac{N(N-1)}{2}, & (6.3)\end{aligned}$$

where N is the number of particles. In a translational invariant system, ODLRO implies for the two-body density matrix the following asymptotic behavior:

$$\rho_2(x_1',x_2';x_1,x_2) = \alpha f^*(\vec{x}_1' - \vec{x}_2')f(\vec{x}_1 - \vec{x}_2) \tag{6.4}$$

$$\text{for } |x_1 - x_2|\,, |x_1' - x_2'| \le \xi \text{ and } |x_1 - x_1'| \to \infty \tag{6.5}$$

where α/N is the pair condensate fraction and ξ is the size of the pair defined by the pairing function f. The function f is zero for large separation $|x_1 - x_2|$ and is $\simeq 1/V^{1/2}$ for microscopic separation for x_1 and x_2. In order to estimate the condensate fraction, following De Palo et al. (104), we resort the two-body density matrix to the projected density matrix:

$$h(x,\theta,\phi) = \frac{1}{N}\int dx_1dx_2\rho_2(x_1'+x,x_2'+x;x_1,x_2) \tag{6.6}$$

which tends to α in the large x limit. The presence of non positive eigenvalues in the λ matrix, see figure 6.6, led us to investigate the presence of non s-wave superconductivity. Therefore we also evaluated $h(x,\theta,\phi)$ as a function of rotations angle θ,ϕ of the electron pair in order to investigate the possibility of different symmetries.

$$\vec{x}_1' - \vec{x}_2' = U_{\theta,\phi}(\vec{x}_1 - \vec{x}_2) \text{ and } \frac{\vec{x}_1' + \vec{x}_2'}{2} = \frac{\vec{x}_1 + \vec{x}_2}{2} + \vec{x} \tag{6.7}$$

A simple estimator of $h(x,\theta,\phi)$ is given by

$$h(x,\theta,\phi) = \frac{1}{M_c} \sum_{i<j\{r_{ij}<\xi\}} \frac{\Psi_t(r_1, r_2, ..., r'_i, ..., r'_j)}{\Psi(r_1, ..., r_n)} \tag{6.8}$$

where r'_i, r'_j is an electron pair translated of x and rotated of θ, ϕ, and M_c the number of pairs translated. A cutoff ξ is introduced to speed up the calculation excluding contributions coming by pairs of far electrons that do not contribute to the ODLRO because the pairing is short range. We have verified that the cutoff used does not effect the final result. In practice, for each pair we generate a few translations x uniformly distributed in the simulation box and, for better statistics, we also average over all pairs with the condition $r_{ij} \leq \xi$.

We argue the possibility of a non s-wave symmetry in the ODLRO, see figure 6.7 and 6.8. Unfortunately the size of the studied system is too small to give a conclusive answer.

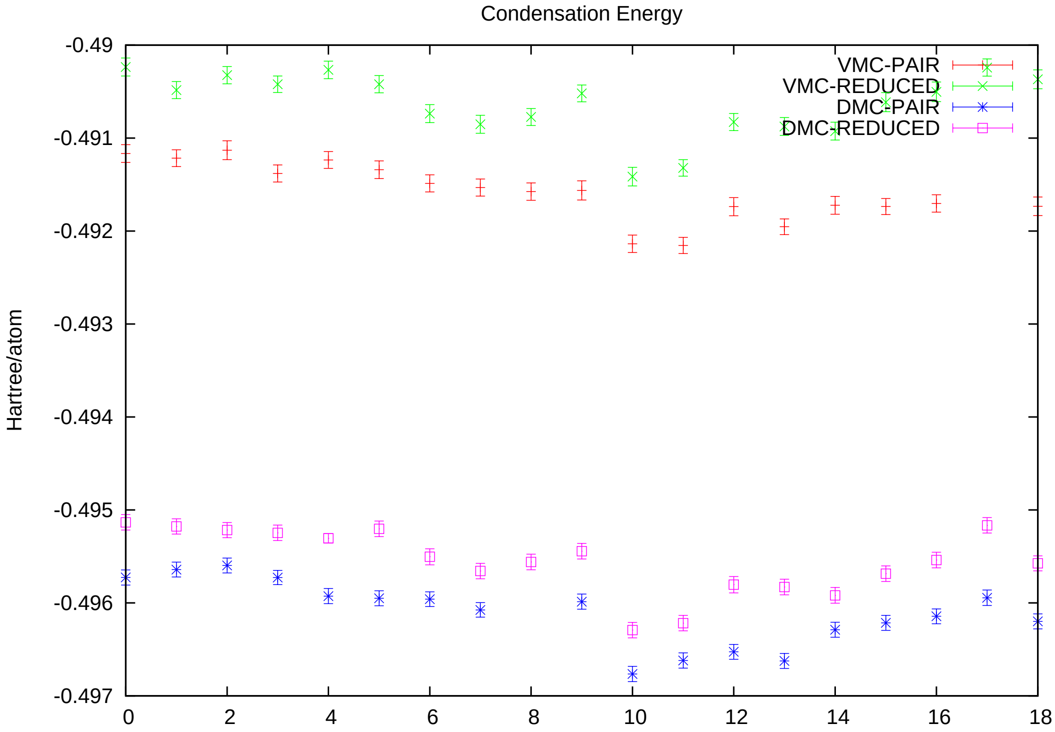

Figure 6.5: Variational and Diffusion Condensation Energy per atom

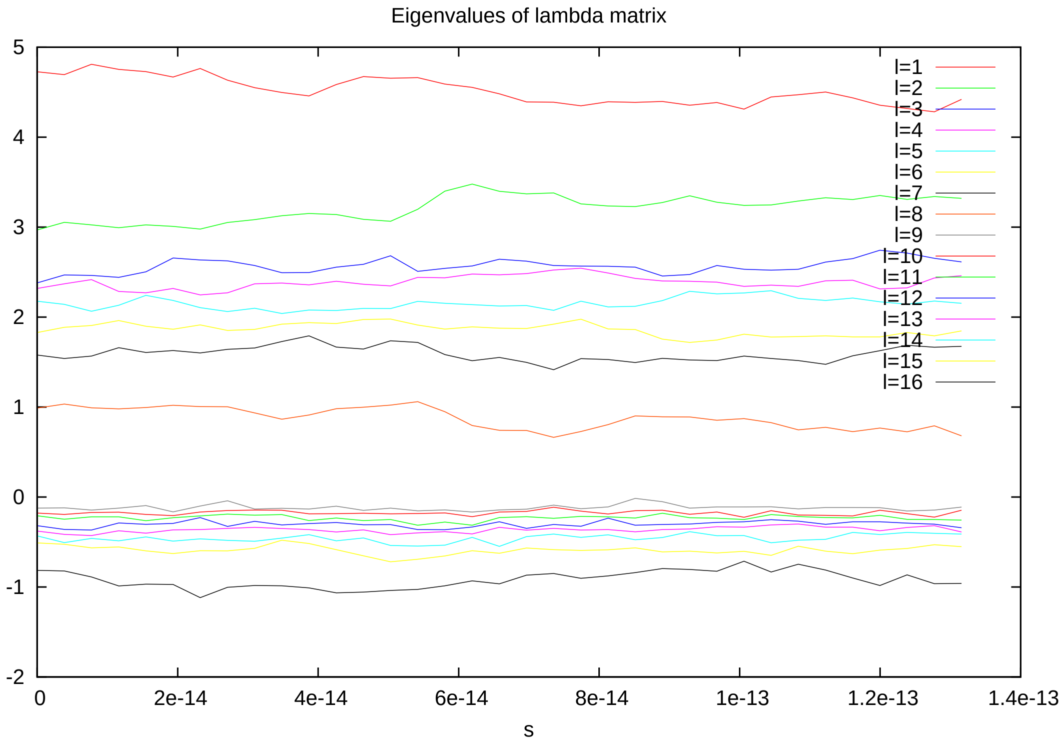

Figure 6.6: Eigenvalues of the λ matrix for 16 hydrogen atoms at Rs=1.31 and 100K as function of the simulation time

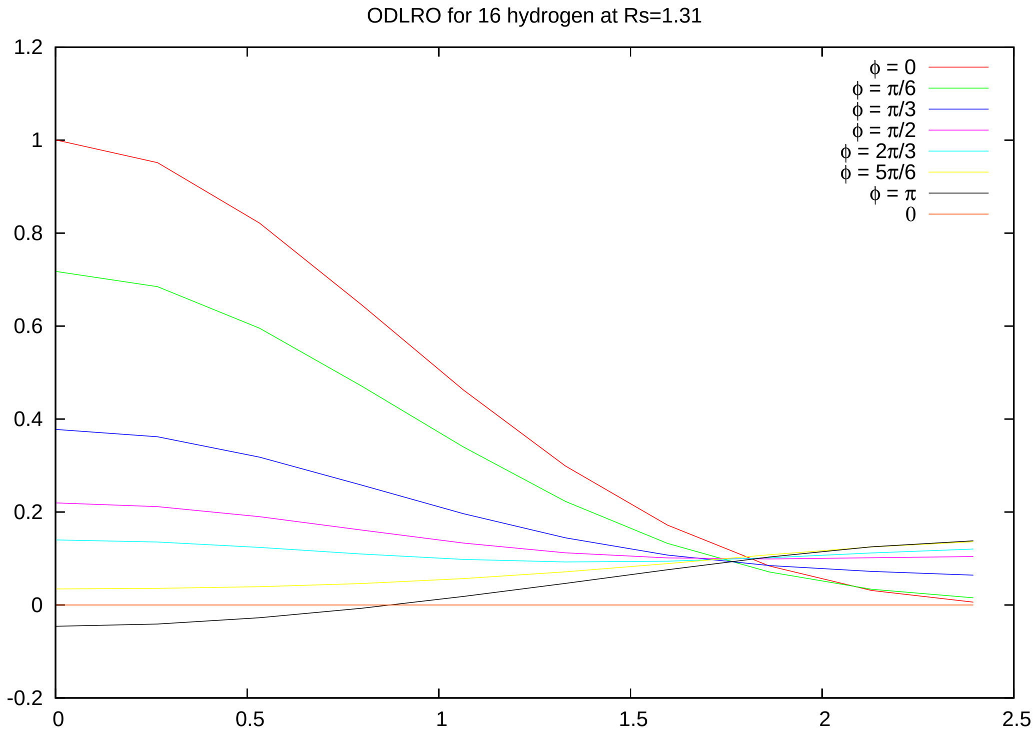

Figure 6.7: Off-Diagonal Long Range Order for 16 hydrogen atoms at Rs=1.31 and 100K, in a box of size $L = 5.3211$, for $\theta = 0$ as function of the distance x

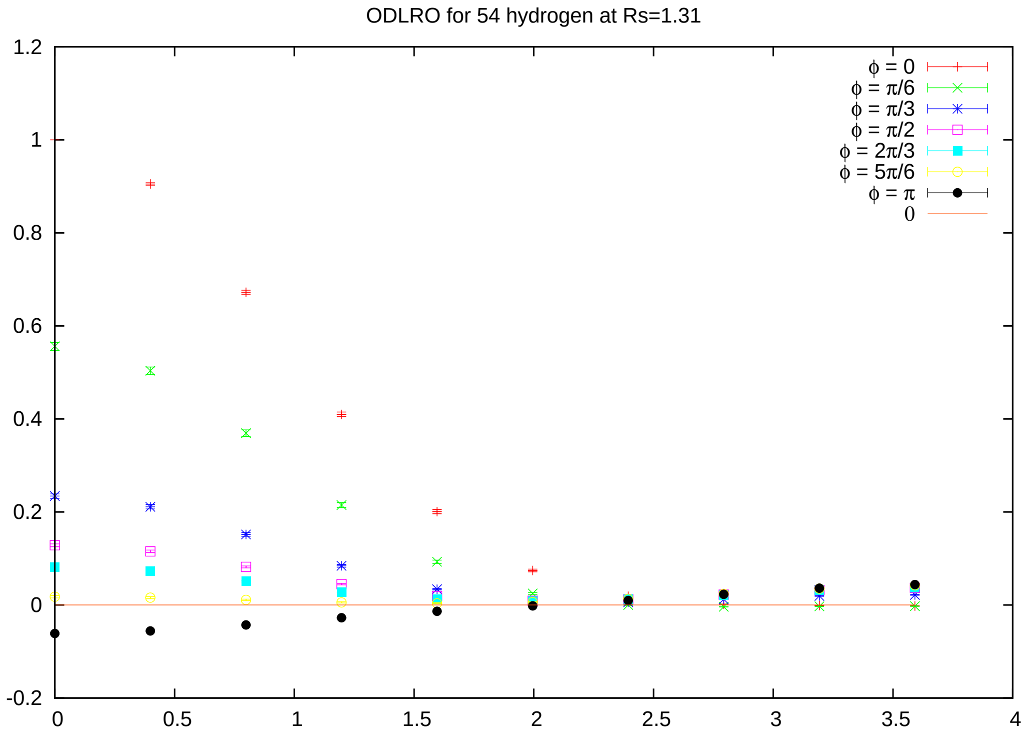

Figure 6.8: Off-Diagonal Long Range Order for 54 hydrogen atoms at Rs=1.31 at 100K in a box of size $L = 7.9817$, for $\theta = 0$ as function of the distance x

Chapter 7

Conclusions

In this first part of this thesis, we have proposed a new kind a trial-function the JAGP for QMC. We have tested this wave-function on simple molecular systems where accurate results were obtained. Within this formulation it was possible to recover a large amount of the correlation energy at the variational level with a computationally very efficient and feasible method. Indeed within the JAGP ansatz, it is sufficient to sample a single determinant whose leading dimension scales only with the number of electrons. Moreover the interplay between the Jastrow and the geminal part has been shown to be very effective in all cases studied and particularly in the non trivial case of the benzene molecule. Only when both the Jastrow and the AGP terms are accurately optimized together, the AGP nodal structure of the wave function is considerably improved. In fact the Jastrow factor is an important ingredient because: it takes into account the local conservation of the charge around each molecule; it allows a fast convergence in the basis set for the determinant because the electron-electron and the electron-nucleus cusp conditions are satisfied. Nevertheless, in some cases, as for instance Be_2, the used basis set was not sufficiently large. Anyway all results presented here can be systematically improved with larger basis set. Moreover we showed that, by using the Stochastic Reconfiguration optimization, it is possible to perform geometry optimization as well, and obtain very accurate geometries for the molecules studied.

In the second part of the thesis we applied the JAGP wave-function to study high pressure hydrogen. The JAGP wave-function is a crucial ingredient to study correlation effects. In fact, as it is known from lattice models with electronic repulsion, it is not possible to obtain a superconducting ground state at the mean-field Hartree-Fock level. Instead as soon as a correlated Jastrow term is applied to the BCS wave function (equivalent to the AGP wave function in momentum space(33)), the stabilization of a d-wave superconducting order parameter is possible. Furthermore the presence of the Jastrow factor can qualitatively change the wave function especially at one electron per site filling, by converting a BCS superconductor to a Mott insulator with a finite charge gap(105). When the charge is locally conserved the phase of the BCS-AGP wave-function cannot have a definite value and phase coherence is correctly forbidden by the Jastrow factor.

In the second part of the thesis we studied the hydrogen close to the transition between the molecular solid to the atomic one, where it is expected a metal-insulator transition due to the closure of the band gap. We introduced a new technique to perform a Car-Parrinello like dynamics on ions by Quantum Monte Carlo noisy forces. This technique opens the possibility to use QMC to study finite temperature system with a reasonable computational effort. The combination of GLQ technique and JAGP wave-function has allowed us to study the electronic pairing structure during the nuclear motion. We have observed a non trivial behaviour on the eigenvalues of the λ matrix, see figure 6.6. This has led us to study the Off-Diagonal Long Range Order in this system. The study of the ODLRO evidences a non conventional superconductivity. Because of the classical nuclei, this superconductivity can be due only to correlation effects as in lattice models used to describe High Tc superconductor (see for instance (106)). Moreover our results showed that the dominant channel for superconductivity may be not be s-wave. Unfortunately the small size of the systems studied does not allow a conclusive answer. In fact as for lattice models, a finite size scaling is very difficult to perform (107). However motivated by recent results obtained on lattice models using renormalization group (102) we are planing to study larger systems to clarify our results. This implies the solution of some technical problems and the reduction of the size effects as discussed in the following.

The new advances in this thesis can be summarized in three points: a new highly correlated wave-function; a new technique to study finite temperature system with QMC; and the possibility, combining the two previous point,s to study exotic phases due to electronic correlation effects.

7.1 Future developments

- **Reduction of the number of parameters**

  In the present implementation the JAGP wave-function is parametrized with an exceedingly large number of parameters. In fact both in the three-body and in the pair determinant the number of coefficients λ increases as $N(N-1)/2$ where N is the number of orbitals in the basis set. So even for a system with only 54 hydrogen atoms we have to optimize about 3000 parameters. Although it is still possible (using the strategies showed in Chapter 6, to optimize this wave-function following the ion dynamic with a reasonable accuracy the computational cost) the amount of memory needed to work with so large matrices make impossible to extend this approach to system larger than 54 protons.

  At this stage we are testing different strategies to overcome this problem: the most promising is to use a parametric form for the λ coefficients that, to a first approximation, can be chosen to be dependent only from the atomic distances.

- **Improving the wave-function optimization.**

  Because QMC calculation are very computer demanding, one has to accelerate the ionic dynamics as much as possible. But this is not possible up to a certain threshold otherwise the optimization procedure is not able anymore to follow the Bohr-Oppenheimer ionic dynamics. In order to overcome this difficulty it is possible to change the optimization procedure to converge not to the minimum of the current ionic configuration, but to be as close as possible to the one of the following ionic configuration. This approach can be partially realized with the information we have from the Hessian matrix, and should allow the use of much larger time steps in the GLE.

- **Size effects and TABC.**

  The computational cost of the Quantum Monte Carlo integration does not allow to study very large system. In order to make the QMC competitive one has to reduce as much as possible the size effects. The size effects derive from the kinetic and the potential energy. We are planing to apply the Twisted Average Boundary Conditions (76) to our system using a Twist sampling in such a way to integrate dynamically the boundary conditions during the simulation. Moreover in these years different strategies were proposed to reduce the finite size effect due to the long-range potential, and we are planing to use some of these strategies (108).

Appendix A

Re-sampling methods

There is a simple motivation to use re-sampling methods. In fact let us consider a set of independent and identically distributed data samples n of an unknown probability distribution F:

$$X_1, X_2, ..., X_n \sim F \tag{A.1}$$

We can compute the sample average $\bar{x} = \sum_{i=1}^{n} x_i/n$, and then we can estimate the accuracy of $\bar{x}$ using the standard deviation:

$$\hat{\sigma} = \sqrt{\frac{1}{n(n-1)} \sum_{i=1}^{n} (x_i - \bar{x})^2} \tag{A.2}$$

The trouble with this formula is that it does not, in any obvious way, extend to estimators other than $\bar{x}$. For this reason a generalized version of A.2 is introduced such that it reduces to the usual standard deviation when the chosen estimator is the average.

A.1 Jackknife

Now we briefly describe how it is possible to obtain the standard deviation of a generic estimator using the Jackknife method. For simplicity we consider the average estimator. Let us consider the variables:

$$\bar{x}_{(i)} = \frac{n\bar{x} - x_i}{n-1} = \frac{1}{n-1} \sum_{j \neq i} x_j, \tag{A.3}$$

where $\bar{x}$ is the sample average. $\bar{x}_i$ is the sample average of the data set deleting the *i*th point. Then we can define the average of $\bar{x}_i$:

$$\bar{x}_{(.)} = \sum_{i=1}^{n} x_{(i)}/n. \tag{A.4}$$

The jackknife estimate of standard deviation is then defined as:

$$\hat{\sigma}_{JACK} = \sqrt{\frac{n-1}{n}\sum_{i=1}^{n}(\bar{x}_{(i)} - \bar{x}_{(.)})^2} \tag{A.5}$$

The advantage of this formula is that it can be used for any estimator, and it reduces to the usual standard deviation for the mean value estimator.

In this thesis we always used the Jackknife re-sampling method. Here we want to show that the connection between the Jackknife and another very used re-sampling method the Bootstrap. Consider a generic estimator $\theta(F)$ evaluated on set of data $x_1, x_2, ..., x_n$ of the unknown distribution F. Let us take a *re-sampling vector*

$$P^* = (P_1^*, P_2^*, ..., P_n^*) \tag{A.6}$$

such that

$$P_i^* \geq 0$$
$$\sum_{i=1}^{n} P_i^* = 1$$

in other words, a probability vector. We can re-weight our data sample with the vector P^* and then evaluate the estimator $\hat{\theta}$ on the re-sampled data:

$$\hat{\theta}^* = \hat{\theta}(P^*) \tag{A.7}$$

The difference between Bootstrap and Jackknife is in the choice of this re-sampling probability vector. In the Bootstrap we use:

$$P^0 = \left(\frac{1}{n}, \frac{1}{n}, ..., \frac{1}{n}\right) \tag{A.8}$$

while in the Jackknife

$$P_{(i)} = \left(\frac{1}{n-1}, \frac{1}{n-1}, ..., 0, \frac{1}{n-1}, ..., \frac{1}{n-1}\right). \tag{A.9}$$

The estimate of the standard deviation is then given by eq. A.2, for a good discussion about Jackknife, Bootstrap and other re-sampling methods see Ref. (109).

Appendix B

Local Energy and its derivatives

B.1 Kinetic Energy

To evaluate the kinetic energy we rewrite the the kinetic operator as:

$$-\frac{1}{2}\frac{\nabla_i^2\Psi}{\Psi} = -\frac{\nabla_i^2 \ln\Psi}{2} - \frac{\left(\vec{\nabla}_i \ln\Psi\right)^2}{2} \tag{B.1}$$

Because our trial-function is made as product of different terms:

$$\Psi = e^J e^T P \tag{B.2}$$

we can rewrite the kinetic energy through gradients and laplacian of the logarithm of each term, namely:

$$\begin{aligned} \ln\Psi &= J(r_{ij}) + T(r_i, r_j, r_{ij}) + \ln P \\ \vec{\nabla}\ln\Psi &= \vec{\nabla}J(r_{ij}) + \vec{\nabla}T(r_i, r_j, r_{ij}) + \frac{\vec{\nabla}P}{P} \\ \nabla^2\ln\Psi &= \nabla^2 J(r_{ij}) + \nabla^2 T(r_i, r_j, r_{ij}) + \frac{\nabla^2 P}{P} - \left(\frac{\vec{\nabla}P}{P}\right)^2 \end{aligned} \tag{B.3}$$

B.1.1 Derivatives of the Kinetic Energy

We want to calculate the derivatives of the Kinetic Energy respect to a variational parameter of the wave-function:

$$\frac{\partial}{\partial a}\frac{\nabla_i^2\Psi}{\Psi} = \frac{\partial}{\partial a}\nabla_i^2 \ln\Psi + \frac{\partial}{\partial a}\left(\vec{\nabla}_i \ln\Psi\right)^2 \tag{B.4}$$

using B.3 we have that:

$$\begin{aligned}\frac{\partial}{\partial a}\nabla_i^2 \ln\Psi &= \frac{\partial}{\partial a}\left(\nabla_i^2 \ln e^J + \nabla_i^2 \ln e^T + \nabla_i^2 \ln P\right) \\ &= \frac{\partial}{\partial a}\nabla_i^2 J + \frac{\partial}{\partial a}\nabla_i^2 T + \frac{\partial}{\partial a}\nabla_i^2 \ln P\end{aligned} \tag{B.5}$$

and

$$\frac{\partial}{\partial a}\left(\vec{\nabla}_i \ln\Psi\right)^2 = 2\sum_{l=1}^{3}\left[\frac{\partial}{\partial a}\left(\nabla_i^l J + \nabla_i^l T + \nabla_i^l \ln P\right)\right]\left(\nabla_i^l J + \nabla_i^l T + \nabla_i^l \ln P\right) \tag{B.6}$$

For the pairing determinant the terms we have to evaluate will be:

$$\begin{aligned}\frac{\partial}{\partial a}\nabla_i^l \ln P &= \frac{\partial_a \nabla_i^l P}{P} - \frac{\partial_a P}{P}\frac{\nabla_i^l P}{P} \\ \frac{\partial}{\partial a}\nabla_i^2 \ln P &= \frac{\partial_a \nabla_i^2 P}{P} - \frac{\partial_a P}{P}\frac{\nabla_i^2 P}{P} + 2\left(\frac{\vec{\nabla}_i P}{P}\right)^2\frac{\partial_a P}{P} - 2\sum_{\mu=1}^{3}\frac{\partial_a \nabla_i^\mu P}{P}\frac{\nabla_i^\mu P}{P}\end{aligned}$$

So to evaluate the gradient of the local energy we need only to know these vectors:

$$\frac{\partial_a P}{P}, \frac{\partial_a \vec{\nabla}_j P}{P}, \frac{\partial_a \nabla_j^2 P}{P} \tag{B.7}$$

B.2 Pairing determinant

Let us define the matrix A_{ij} as:

$$A_{ij} = \Phi(r_i, r_j) = \sum_{l,m}\lambda_{l,m}\phi_l(r_i)\phi_m(r_j) \tag{B.8}$$

where i are coordinates of spin up electrons and j of spin down electrons. For polarized system is possible extend the definition of the geminal wave-function. This generalization was first proposed by Coleman (110). In practise if $N^\uparrow > N^\downarrow$ we can define a $N^\uparrow x N^\uparrow$ matrix A_{ij} in the following way:

$$\begin{aligned}A_{ij} &= \Phi(r_i^\uparrow, r_j^\downarrow) \text{ for } j = 1, N^\downarrow && \text{(B.9)} \\ &= \bar{\phi}_j(r_i^\uparrow) \text{ for } j = N^\downarrow + 1, N^\uparrow && \text{(B.10)}\end{aligned}$$

When we move an electron to ratio between the old and the new determinant will be given for a spin down electron:

$$\begin{aligned}
|A(r_i, r'_k)| &= \sum_i \Phi(r_i, r'_k) \, |A(r_i, r_k)| \, A^{-1}_{ik} \\
\frac{|A(r_i, r'_k)|}{|A(r_i, r_k)|} &= \sum_i \Phi(r_i, r'_k) A^{-1}_{ik} \\
\frac{|A(r_i, r'_k)|}{|A(r_i, r_k)|} &= \sum_i \sum_{l,m} \lambda_{l,m} \phi_l(r_i) \phi_m(r'_k) A^{-1}_{ik} \qquad \text{(B.11)}
\end{aligned}$$

and for spin up:

$$\begin{aligned}
|A(r'_i, r_k)| &= \sum_k \Phi(r'_i, r_k) \, |A(r'_i, r_k)| \, A^{-1}_{ik} \\
\frac{|A(r'_i, r_k)|}{|A(r'_i, r_k)|} &= \sum_k \Phi(r'_i, r_k) A^{-1}_{ik}
\end{aligned}$$

For updating the inverse matrix A^{-1}, used here, after a move we follow the simple formula used for the Slater determinant(38) with indices that depend from the spin of the electron.

Gradients and Laplacian

Using the formula D.3 and the fact the only a column or a row may depends from a given electronic coordinate we obtain:

$$\begin{aligned}
\nabla^2_i \ln |A| &= \sum_k \nabla^2_i \Phi(r_i, r_k) A^{-1}_{ik} \\
\vec{\nabla}_i \ln |A| &= \sum_k \vec{\nabla}_i \Phi(r_i, r_k) A^{-1}_{ik} \qquad \text{(B.12)}
\end{aligned}$$

where:

$$\nabla^2_i \Phi(r_i, r_k) = \sum_{l,m} \lambda_{l,m} \nabla^2_i \phi_l(r_i) \phi_m(r_j) \qquad \text{(B.13)}$$

$$\vec{\nabla}_i \Phi(r_i, r_k) = \sum_{l,m} \lambda_{l,m} \vec{\nabla}_i \phi_l(r_i) \phi_m(r_j) \qquad \text{(B.14)}$$

for spin down we have to exchange i with k in all these equations.

The logarithmic derivatives

In pairing trial-function the variation of a parameter involves all terms of the matrix and so using D.3 the logarithmic derivatives will be:

$$\frac{\partial \ln det(A)}{\partial \beta} = \sum_{ij} \frac{\partial \ln det(A)}{\partial a_{ij}} \frac{\partial a_{ij}}{\partial \beta} = \sum_{ij} \frac{\partial \Phi(r_i, r_j)}{\partial \beta} A^{-1}_{ij}$$

If only the k-th orbital depends by β we obtain:

$$\frac{\partial \Phi(r_i, r_j)}{\partial \beta} = \frac{\partial \phi_k(r_i)}{\partial \beta} \sum_m \lambda_{km} \phi_m(r_j) + \frac{\partial \phi_k(r_j)}{\partial \beta} \sum_m \lambda_{mk} \phi_m(r_i) \tag{B.15}$$

If β is one of the λ parameter the derivative will be:

$$\frac{\partial \Phi(r_i, r_j)}{\partial \lambda_{ab}} = \phi_a(r_i)\phi_b(r_j) + \phi_b(r_i)\phi_a(r_j) \tag{B.16}$$

because λ matrix is symmetric.

Second derivatives

In the general case in which all elements of the matrix depend from the parameters β, γ the derivative is:

$$\frac{1}{|A|} \frac{\partial^2 |A|}{\partial \beta \partial \gamma} = \frac{1}{|A|} \sum_{n,k,j,m} \frac{\partial |A|}{\partial a_{nk} \partial a_{jm}} \frac{\partial a_{nk}}{\partial \beta} \frac{\partial a_{jm}}{\partial \gamma} + \frac{1}{|A|} \sum_{nk} \frac{\partial |A|}{\partial a_{nk}} \frac{\partial^2 a_{nk}}{\partial \beta \partial \gamma} \tag{B.17}$$

now using equations D.3 and D.5

$$\begin{aligned} \frac{1}{|A|} \sum_{n,k,j,m} \frac{\partial |A|}{\partial a_{nk} \partial a_{jm}} \frac{\partial a_{jm}}{\partial \beta} \frac{\partial a_{nk}}{\partial \gamma} &= \sum_{n,k,j,m} \left(A^{-1}_{kn} A^{-1}_{mj} - A^{-1}_{km} A^{-1}_{nj} \right) \frac{\partial a_{nk}}{\partial \beta} \frac{\partial a_{jm}}{\partial \gamma} \\ \frac{1}{|A|} \sum_{nk} \frac{\partial |A|}{\partial a_{nk}} \frac{\partial^2 a_{nk}}{\partial \beta \partial \gamma} &= \sum_{nk} A^{-1}_{nk} \frac{\partial^2 a_{nk}}{\partial \beta \partial \gamma} \end{aligned} \tag{B.18}$$

where the derivatives $\partial a_{nk}/\partial \gamma$, $\partial a_{nk}/\partial \beta$ are given by B.15, B.16, B.14. The second derivatives can be evaluate from B.8.
If β and γ are two λ_{lm} using B.16 the second derivative is obviously zero.

If β and γ are both orbital parameters, and for example the $k-th$ orbital depends from β and $l-th$ orbital from γ, using B.15 we have:

$$\begin{aligned} \frac{\partial^2\Phi(r_i,r_j)}{\partial\beta\partial\gamma} &= \lambda_{kl}\frac{\partial\phi_k(r_i)}{\partial\beta}\frac{\partial\phi_l(r_j)}{\partial\gamma}+\lambda_{lk}\frac{\partial\phi_l(r_i)}{\partial\beta}\frac{\partial\phi_k(r_j)}{\partial\gamma} && \text{(B.19)} \\ &= \lambda_{kl}\left(\frac{\partial\phi_k(r_i)}{\partial\beta}\frac{\partial\phi_l(r_j)}{\partial\gamma}+\frac{\partial\phi_l(r_i)}{\partial\beta}\frac{\partial\phi_k(r_j)}{\partial\gamma}\right) && \text{(B.20)} \\ & && \text{(B.21)} \end{aligned}$$

because $\lambda_{kl}=\lambda_{lk}$.
If β and γ are parameters of the same orbital we have:

$$\frac{\partial^2\Phi(r_i,r_j)}{\partial\beta^2}=\frac{\partial^2\phi_k(r_i)}{\partial\beta\partial\gamma}\sum_m\lambda_{km}\phi_m(r_j)+\frac{\partial^2\phi_k(r_j)}{\partial\beta\partial\gamma}\sum_m\lambda_{mk}\phi_m(r_i) \qquad \text{(B.22)}$$

If γ is one of the λ parameter, using the fact the λ matrix is symmetric we obtain:

$$\frac{\partial^2\Phi(r_i,r_j)}{\partial\beta\partial\lambda_{ab}}=\frac{\partial\phi_a(r_i)}{\partial\beta}\phi_b(r_j)+\phi_b(r_i)\frac{\partial\phi_a(r_j)}{\partial\beta} \qquad \text{(B.23)}$$

Derivatives of the Local Energy

To evaluate the derivatives of the local energy we need to calculate the following terms:

$$\frac{\partial_a\vec{\nabla}_i\left|A\right|}{\left|A\right|},\frac{\partial_a\nabla_i^2\left|A\right|}{\left|A\right|} \qquad \text{(B.24)}$$

we obtain:

$$\begin{aligned} \frac{\partial_a\vec{\nabla}_i\left|A\right|}{\left|A\right|} &= \frac{1}{\left|A\right|}\sum_{n,l,m}\frac{\partial\left|A\right|}{\partial a_{ni}\partial a_{lm}}\vec{\nabla}_i\Phi(r_n,r_i)\partial_a\Phi(r_l,r_m)+\sum_n\frac{1}{\left|A\right|}\frac{\partial\left|A\right|}{\partial a_{in}}\partial_a\vec{\nabla}_i\Phi(r_i,r_n) \\ &= \sum_{n,l,m}\left(A_{in}^{-1}A_{ml}^{-1}-A_{il}^{-1}A_{mn}^{-1}\right)\vec{\nabla}_i\Phi(r_n,r_i)\partial_a\Phi(r_l,r_m)+\sum_n A_{ni}^{-}1\partial_a\vec{\nabla}_i\Phi(r_i,r_n) \end{aligned}$$

and a similar formula for the derivative of the laplacian. If only the orbital k depends by a we have:

$$\partial_a\vec{\nabla}_i\Phi(r_i,r_j)=\partial_a\vec{\nabla}_i\phi_k(r_i)\sum_m\lambda_{km}\phi_m(r_j)+\partial_a\phi_k(r_j)\sum_m\lambda_{mk}\vec{\nabla}_i\phi_m(r_i)$$

B.3 Three-body

In the same spirit of the pairing determinant we built a three-body factor as:

$$U = \exp\left(\sum_{i,j}^{N_{elec}} u(\vec{r}_i,\vec{r}_j)\right) \tag{B.25}$$

$$u(\vec{r}_i,\vec{r}_j) = \sum_{m,n}^{N_{orb}} \lambda_{mn}\phi_m(r_i)\phi_n(r_j) \tag{B.26}$$

When you move an electron r_k the ratio between the two three-body factor is given by:

$$\frac{U(r'_k)}{U(r_k)} = \exp\left(\sum_{m,n}\left[\sum_{(j\neq k)}\lambda_{mn}\phi_m(r_j)\left(\phi_n(r'_k)-\phi_n(r_k)\right)+\phi_m(r'_k)\phi_n(r'_k)-\phi_m(r_k)\phi_n(r_k)\right]\right) \tag{B.27}$$

and if you accept the move to update the value of three-body you have only to update N_{orb} orbitals.

Gradients and Laplacian

The gradients and laplacian of the logarithm of three-body term are given by:

$$\vec{\nabla}_k \ln U = \sum_j\sum_{m,n}\lambda_{mn}\vec{\nabla}_k\phi_m(r_k)\phi_n(r_j) \tag{B.28}$$

$$\nabla^2_k \ln U = \sum_j\sum_{m,n}\lambda_{mn}\nabla^2_k\phi_m(r_k)\phi_n(r_j) \tag{B.29}$$

Three-body Derivatives

The derivative respect to a parameter α_m of an orbital m is given by:

$$\frac{\partial \ln U}{\partial\alpha_m} = \sum_{i,j}\sum_l\left(\lambda_{lm}\phi_l(r_i)\frac{\partial\phi_m(r_j)}{\partial\alpha_m}+\lambda_{ml}\frac{\partial\phi_m(r_i)}{\partial\alpha_m}\phi_l(r_j)\right) \tag{B.30}$$

and the derivative respect to λ_{ab} is

$$\frac{\partial \ln U}{\partial\lambda_{ab}} = \sum_{i,j}\phi_a(r_i)\phi_b(r_j)+\phi_b(r_i)\phi_a(r_j) \tag{B.31}$$

because λ matrix is symmetric.

Derivatives of the local energy

It is very simple to evaluate the terms appearing in B.5 and B.6 in fact using B.29 and B.28 and considering only the m orbital dependent by the parameter a we obtain:

$$\begin{aligned}
\partial_a \vec{\nabla}_k \ln U &= \sum_j \sum_n \lambda_{mn} \partial_a \vec{\nabla}_k \phi_m(r_k) \phi_n(r_j) + \sum_j \sum_n \lambda_{nm} \vec{\nabla}_k \phi_n(r_k) \partial_a \phi_m(r_j) \\
\partial_a \nabla_k^2 \ln U &= \sum_j \sum_n \lambda_{mn} \partial_a \nabla_k^2 \phi_m(r_k) \phi_n(r_j) + \sum_j \sum_n \lambda_{nm} \nabla_k^2 \phi_n(r_k) \partial_a \phi_m(r_j)
\end{aligned}$$

Appendix C

Cusp conditions

When two Coulomb particles get close, the potential has $1/r$ singularity. We want modify the wave function in such a way to cancel this singularity. Let us consider the case of an electron close to a nucleus, the Schrödinger equation reduces to:

$$\left[-\frac{1}{2L^2}\nabla_e^2 - \frac{Ze^2}{rL}\right]\psi = E\psi \tag{C.1}$$

where Z is the nuclear charge, notice that we used rescaled distances (see Eq. 4.23). Writing the first term in spherical coordinates, we get

$$-\frac{1}{2}\frac{\psi''}{L^2} - \frac{1}{rL}\left(Ze^2\psi + \frac{\psi'}{L}\right) = E\psi \tag{C.2}$$

To cancel the singularity at small r the term multiplying by $1/r$ must vanish. So we have

$$\frac{1}{\psi}\psi' = -ZLe^2 \tag{C.3}$$

If $\psi = e^{-cr}$ we must have $c = ZLe^2$. For the case of two electrons, when they are close each other the Schrödinger equation, using relative coordinates $r_{12} = r_1 - r_2$, reduces to

$$\left[-\frac{\nabla_{12}^2}{L^2} + \frac{e^2}{Lr_{12}}\right]\psi = E\psi \tag{C.4}$$

Electrons with unlike spins have an extra factor of $1/2$ in the cusp condition compared with the electron-nucleus case. So we have $c = -e^2L/2$. In the antisymmetric case, the electrons will be in a relative p state, reducing the cusp condition by $1/2$, so $c = -e^2L/4$. Since the antisymmetry requirement keeps them apart anyway having the correct cusp for like spin electrons leads to a very little in the energy or the variance(see Ref. (39).

Appendix D

Determinant derivatives

Consider a matrix A, we want find a simple way to express derivatives of the its determinant respect to the matrix elements a_{ij}. The determinant can be expanded in the elements a_{ij}:

$$\det A = |A| = \sum_j a_{ij} (-1)^{i+j} C_{ji} \tag{D.1}$$

where C_{ij} is the minor of the matrix A respect to the element a_{ij} and therefore does not depend explicitly by the elements of the raw i^{th}. So the derivative with respect to a_{kj} will be:

$$\frac{\partial |A|}{\partial a_{kj}} = (-1)^{k+j} C_{kl} \tag{D.2}$$

and for the logarithmic derivative we have:

$$\frac{\partial \ln |A|}{\partial a_{kj}} = (-1)^{k+j} C_{kl} \frac{1}{|A|} = A^{-1}_{kj} \tag{D.3}$$

We want to find a simple relation to evaluate second derivatives of the determinant. We write the relation

$$\sum_j a_{ij} A^{-1}_{jk} = \sum_j a_{ij} \frac{1}{|A|} \frac{\partial |A|}{\partial a_{kj}} = \delta_{ik} \tag{D.4}$$

if we derive this equation for a_{ln} we obtain:

$$\begin{aligned} \frac{\partial}{\partial a_{ln}} \left(\sum_j a_{kj} A^{-1}_{ji} \right) &= \sum_j a_{kj} \left(\frac{-1}{|A|^2} \frac{\partial |A|}{\partial a_{ij}} \frac{\partial |A|}{\partial a_{ln}} + \frac{1}{|A|} \frac{\partial^2 |A|}{\partial a_{ij} \partial a_{ln}} \right) + \delta_{lk} \delta_{nj} \frac{1}{A} \frac{\partial |A|}{\partial a_{ji}} \\ &= \sum_j a_{kj} \left(\frac{-1}{|A|^2} \frac{\partial |A|}{\partial a_{ij}} \frac{\partial |A|}{\partial a_{ln}} + \frac{1}{|A|} \frac{\partial^2 |A|}{\partial a_{ij} \partial a_{ln}} + \frac{1}{|A|^2} \frac{\partial |A|}{\partial a_{lj}} \frac{\partial |A|}{\partial a_{ni}} \right) \end{aligned}$$

where we substitute the δ_{lk} with the eq. D.4. Because of this equation is zero for all a_k this means that the expression in parentheses is zero, and this yields to:

$$\frac{1}{|A|}\frac{\partial^2 |A|}{\partial a_{kn}\partial a_{jm}} = A^{-1}_{nk}A^{-1}_{mj} - A^{-1}_{mk}A^{-1}_{nj} \tag{D.5}$$

Appendix E

Error Analysis due to finite time step in the GLE integration in a simple case

When we discretize the equation 5.13 we introduce an error due to the finite integration time step τ. Following the idea of Ref. (111) we can evaluate this error analytically in the case of a simple harmonic oscillator. Consider the equation:

$$x^{n+1} - (1 + e^{-\lambda\tau})x^n + e^{-\lambda\tau}x^{n-1} = \frac{\tau}{\lambda}(1 - e^{-\lambda\tau})(\omega^2 x^n + r^n). \tag{E.1}$$

We are interested in a statistically stationary process and so we proceed to evaluate mean average energies and correlation functions as function of τ. To do so we multiply the equation E.1 for x^n and x^{n-1} , respectively, and then take the average. We obtain the following pair of equations:

$$\begin{aligned} \langle x^{n+1}x^n\rangle + \left[\frac{\tau\omega^2}{\lambda}(e^{-\lambda\tau} - 1) - (1 + e^{-\lambda\tau})\right]\langle x^n x^n\rangle + e^{-\lambda\tau}\langle x^{n-1}x^n\rangle &= 0 \\ \langle x^{n+1}x^{n-1}\rangle - (1 + e^{-\lambda\tau})\langle x^n x^{n-1}\rangle + \left[\frac{\tau\omega^2}{\lambda}(e^{-\lambda\tau} - 1) + e^{-\lambda\tau}\right]\langle x^{n-1}x^{n-1}\rangle &= 0 \end{aligned}$$

Becuase we are interested in the equilibrium distribution, we can assume that $\langle x^n x^{n-1}\rangle = \langle x^{n+1}x^n\rangle$ and $\langle x^{n+1}x^{n+1}\rangle = \langle x^n x^n\rangle = \langle x^{n-1}x^{n-1}\rangle$. Thus we have three unknown quantities $\langle x^{n+1}x^n\rangle$,$\langle x^n x^n\rangle$ and $\langle x^{n+1}x^{n-1}\rangle$. To get a third relation among this quantities we square the Eq. E.1 to obtain the relation:

$$\begin{aligned} 2 &\quad \left[\frac{\tau\omega^2}{\lambda}(e^{-\lambda\tau}-1)-(1+e^{-\lambda\tau})\right]\left(1+e^{-\lambda\tau}\right)\langle x^{n+1}x^n\rangle \\ + &\quad \left[1+e^{-2\lambda\tau}+\left(\frac{\tau\omega^2}{\lambda}(1-e^{-\lambda\tau})+(1+e^{-\lambda\tau})\right)^2\right]\langle x^n x^n\rangle \\ + &\quad 2e^{-\lambda\tau}\langle x^{n+1}x^{n-1}\rangle = \frac{\tau^2}{\lambda^2}(1-e^{-\lambda\tau})^2\langle r^n r^n\rangle \end{aligned}$$

After solving this equation system we can evaluate the potential and the kinetic energy:

$$\langle E_{pot}\rangle = \frac{1}{2}\omega^2\langle x^n x^n\rangle \tag{E.2}$$

$$\langle E_{kin}\rangle = \frac{1}{2}\langle V^n V^n\rangle = \frac{1}{2}\frac{\langle (x^n - x^{n-1})^2\rangle}{\tau^2} \tag{E.3}$$

$$= \frac{1}{\tau^2}\left[\langle x^n x^n\rangle - \langle x^n x^{n-1}\rangle\right] \tag{E.4}$$

It is easy to show that in the limit of small τ the potential and the kinetic energies converge to:

$$\langle E_{kin}\rangle = \frac{1}{2}k_bT\left(1+\frac{1}{4}\omega^2\tau^2+O(\tau^4)\right) \tag{E.5}$$

$$\langle E_{pot}\rangle = \frac{1}{2}k_bT\left(1+O(\tau^2)+O(\tau^3)\right) \tag{E.6}$$

This show that at least in this simple model the impulse integrator leads to a quadratic error in τ in both kinetic and poterntial energy.

Bibliography

[1] S. J. Chakravorty S. R. Gwaltney E. R. Davidson F. A. Parpia and C. F. Fischer. Ground state correlation energies for atomic ions with 3 to 18 electrons. *Phys. Rev. A*, 47:3649, 1993. (document), 3.1

[2] M. Casula S. Sorella. Geminal wave functions with jastrow correlation: A first application to atoms. *J. Chem. Phys.*, 119(13):6500, 2003. (document), 1.2.1, 1.2.1, 3.4, 6.3

[3] K. Muller-Dethlefs J. B. Peel. Calculations on the jahn teller configurations of the benzene cation. *J. Chem. Phys.*, 111:10550, 1999. (document), 3.5, 3.3, 3.6, 3.3

[4] M. S. Deleuze L. Cleas E. S. Kryachko and J.P. Francois. Benchmark theoretical study of the ionization threshold of benzene and oligoacenes. *J. Chem. Phys.*, 119:3106, 2003. (document), 3.5, 3.3, 3.3, 3.7

[5] N. C. Holmes M. Ross and W. J. Nellis. Temperature measurements and dissociation of shock-compressed liquid deuterium and hydrogen. *Phys. Rev. B*, 52:15835, 1995. (document), 6.2

[6] I.F. Silvera and V.V. Goldman. The isotropic intermolecular potential for h_2 and d_2 in the solid and gas phase. *J. Chem. Phys.*, 69:1, 1978. (document), 6.2

[7] M.D. Dewing. *Monte Carlo Methods: Application to Hydrogen Gas and Hard Spheres*. PhD thesis, Michigan Technological University, 1993. (document), 6.1, 6.2, 6.1, 6.2, 6.4, 6.2

[8] D. M. Ceperley C. Pierleoni and M. Holzmann. Coupled electron ionmontecarlo calculations ofatomic hydrogen. *Phys. Rev. Lett.*, 93(14):146402, 2004. (document), 5, 6.1, 6.1, 6.2, 6.4

[9] W. M. C. Foulkes L. Mitas R. J. Needs and G. Rajagopal. Quantum monte carlo simulations of solids. *Rev. Mod. Phys.*, 73:33, 2001. (document)

[10] M. P. Nightingale and C. J. Umrigar. *Quantum Monte Carlo Methods in Physics and Chemistry*, volume 525. Nato Science Series C:Mathematical and Physical Sciences, 1999. (document)

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

[12] C. Attaccalite S. Moroni P. Gori-Giorgi and G.B. Bachelet. Correlation energy and spin polarization in the 2d electron gas. *Phys. Rev. Lett.*, 88:256601, 2002. (document)

[13] D. M. Ceperley and B.J. Alder. Ground state of the electron gas by a stochatic method. *Phys. Rev. Lett.*, 45:556, 1980. (document)

[14] D.M. Ceperley. Path integrals in the theory of condensed helium. *Rev. Mod. Phys.*, 67:279, 1995. (document)

[15] M. Casula C. Attaccalite and S. Sorella. Correlated geminal wave function for molecules: An efficient resonating valence bond approach. *Journal. Chem. Phys.*, 121:7110, 2004. (document), 1.2.4, 2, 2.1, 2.3, 3.2

[16] C. Filippi and C. J. Umrigar. Multiconfiguration wave functions for quantum monte carlo calculations of first row diatomic molecules. *J. Chem. Phys.*, 105:123, 1996. (document), 3.1, 3.2, 3.2

[17] W. M. C. Foulkes L. Mitas R. J. Needs and G. Rajagopal. Quantum monte carlo simulations of solids. *Rev. Mod. Phys.*, 73:33, 2001. (document)

[18] M. Calandra and S. Sorella. Numerical study of the two dimensional heisemberg model by green function monte carlo at fixed number of walkers. *Phys. Rev. B*, 1998. (document), 2.1.2

[19] Casula M. Yunoki S. Attaccalite C. Sorella S. Resonating valence bond wave function: from lattice models to realistic systems. *Phys. Rev. Lett.*, 169:386, 2005. (document)

[20] H. Mao and R. J. Hemley. Ultrahigh-pressure transitions in solid hydrogen. *Rev. Mod. Phys.*, 66:671, 1994. (document)

[21] E. Schwegler T.Ogitzu S. A. Bonev and G. Galli. A quantum fluid of metallic hydrogen suggested by first principles calculations. *Nature*, 431:669, 2004. (document), 5.1

[22] J. Kohanoff S. Scandolo S. de Gironcoli and E. Tosatti. Dipole-quadrupole interactions and the nature of phase iii of compressed hydrogen. *Phys. Rev. Lett.*, 83(20):4097, 1999. (document)

[23] V. Natoli R. M. Martin and D.M. Ceperley. Crystal structure of atomic hydrogen. *Phys. Rev. Let.*, 70:1952, 1993. (document)

[24] N. W. Ashcroft. The hydrogen liquids. *J. Phys. Condens. Matter*, 12:129, 2000. (document)

[25] N. W. Ashcroft E. Babaev, A. SudbÃ¸. Observability of a projected new state of matter: a metallic superfluid. 2005. (document)

[26] L. Mitas R. J. Needs W. M. C. Foulkes and G. Rajagopal. Quantum monte carlo simulations of solids. *Rev. Mod. Phys.*, 73:33, 2001. 1.1, 1.2

[27] N. Metropolis A.W. Rosenbluth M.N. Rosenbluth A.N. Teller and E. Teller. Equation of state calculations by fast computing machines. *J. Chem. Phys.*, 21:1078, 1953. 1.1, 5

[28] J.J. Sakurai. *Modern Quantum Mechanics*. Addison-Wesley Publishing Company Inc., 1994. 1.1.1

[29] R. Assaraf and M. Caffarel. Zero-variance principle for monte carlo algorithms. *Phys. Rev. Lett.*, 83:4682, 1999. 1.1.1, 2.2, 2.2, 5.2

[30] R. Assaraf and M. Caffarel. Computing forces with quantum monte carlo. *J. Chem. Phys.*, 113:4028, 2000. 1.1.1

[31] R. Assaraf and M. Caffarel. Zero-variance zero-bias principle for observables in quantum monte carlo: Application to forces. *J. Chem. Phys.*, 119:10536, 2003. 1.1.1

[32] A. J. Coleman. Structure of fermion density matrices. *Rev. Mod. Phys.*, 35:668, 1963. 1.2

[33] B. Weiner and O. Goscinski. Calculation of optimal antisymmetrized geminal power (projected bcs) functions and their associated excitation spectrum. *Phys. Rev. A*, 22:2374, 1980. 1.2, 7

[34] A. J. Coleman. Spin contamination in quantum monte carlo wave functions. *J. Math. Phys.*, 13:214, 1972. 1.2.1

[35] Steven G. Louie X. W. Wang Jing ZhuB and S. Fahy. Magnetic structure and equation of state of bcc solid hydrogen:a variational quantum monte carlo study. *Phys. Rev. Lett.*, 65(19):2414, 1990. 1.2.1

[36] W. Kohn E. Krotscheck and G. X. Qian. Theory of inhomogeneous quantum systems. iv. variational calculations of metal surfaces. *Phys. Rev. B*, 32:5693, 1985. 1.2.2

[37] D. M. Ceperley C. Pierleoni Holzmann M. and K. Esler. Backflow correlations for the electron gas and metallic hydrogen. *Phys. Rev. E*, 68:046707, 2003. 1.2.2, 6.1

[38] S.G. Louise S. Fahy, X.W. Wang. Variational quantum monte carlo nonlocal pseudopotential approach to solids: Formulation and application do diamond, graphite, and silicon. *Phys. Rev. B*, 42:3503, 1990. 1.2.3, B.2

[39] C. Filippi C. J. Huang and C. J. Umrigar. Spin contamination in quantum monte carlo wave functions. *Journal. Chem. Phys.*, 108(21):8838, 1998. 1.2.3, C

[40] M. Snajdr. S.M. Rothstein. Are properties derived from variance-optimized wave functions generally more accurate? monte carlo study of non-energy-related properties of h2, he, and lih. *J. Chem. Phys.*, 112:4935, 2000. 2, 3.2

[41] D. Bressanini *et al.* Robust wave function optimization procedures in quantum monte carlo methods. *J. Chem. Phys.*, 116:5345, 2002. 2, 3.2

[42] F.J. Galvez *et al.* *Mol. Phys.*, 17:627, 2001. 2, 3.2

[43] anf J. W. Wilkins C.J. Umrigar, K. G. Wilson. Optimized trial wave functions for quantum monte carlo calculations. *Phys. Rev. Lett.*, 60:1719, 1988. 2

[44] P. R. C. Kent R. J. Needs and G. Rajagopal. Monte carlo energy and variance-minimization techniques for optimizing many-body wave functions. *Phys. Rev. B*, 59:12344, 1999. 2

[45] Xi Lin Hongkai Zhang and Andrew M. Rappe. Optimization of quantum monte carlo wave functions using analytical energy derivatives. *J. Chem. Phys.*, 112(6):2650, 2000. 2, 2.3, 2.3

[46] F. Schautz and C. Filippi. Optimized jastrow-slater wave functions for ground and excited states: Application to the lowest states of ethene. *J. Chem. Phys.*, 120:10931, 2004. 2, 2.2

[47] A. Harju, B. Barbiellini, S. Siljamaki, R. M. Nieminen, and G. Ortiz. Stochastic gradient approximation: An efficient method to optimize many body wave functions. *Phys. Rev. Lett.*, 79:1173, 1997. 2

[48] S. Tanaka. Structural optimization in variational quantum monte carlo. *J. Chem. Phys.*, 100:7416, 1994. 2, 2.2

[49] S. Sorella. Stochastic reconfiguration. *Phys. Rev. Lett.*, 64:024512, 2001. 2, 2.1, 2.1.2

[50] S. Sorella. Wave function optimization in Variational Monte Carlo. *Phys. Rev. B*, 71:241103(R), 2005. 2, 2.3, 2.3, 5.4.3

[51] L Sorella, S Capriotti. Green function monte carlo with stochastic reconfiguration: An effective remedy for the sign problem. *Phys. Rev. B*, 61:2599, 2000. 2.1

[52] C. Filippi and C. J. Umrigar. Correlated sampling in quantum monte carlo: A route to forces. *Phys. Rev. B Rapid Communications*, 61:16291, 2000. 2.2, 5.2

[53] M.P. Nightingale C. J. Umrigar and K.J Runge. A diffusion monte carlo algorithm with very small time-step errors. *J. Chem. Phys.*, 99:2865, 1998. 2.2

[54] R. Car and M. Parrinello. Unified approach for molecular dynamics and density-functional theory. *Phys. Rev. Lett.*, 55:2417, 1985. 2.2, 5, 5.4.2

[55] M. Mella M. Casalegno and A. M. Rappe. Computing accurate forces in quantum monte carlo using pulay's corrections and energy minimization. *J. Chem. Phys.*, 118:7193, 2003. 2.2

[56] M. Mella M. Won Lee and A. M. Rappe. Electronic quantum monte carlo calculations of atomic forces, vibrations and anharmonicities. 2005. 2.2

[57] S. Baroni S. De Palo, S. Moroni. Derivatives of the fixed-node energy. 2001. 2.2

[58] C. J. Umrigar and C. Filippi. Energy and variance optimization of many-body wave functions. *Phys. Rev. Lett.*, 94:150201, 2005. 2.3

[59] S. J. Chakravorty S. R. Gwaltney E. R. Davidson F. A. Parpia and C. F. Fischer. Ground-state correlation energies for atomic ions with 3 to 18 electrons. *Phys. Rev. A*, 47:3649, 1993. 3.1

[60] D. Feller C. M. Boyle and E. R. Davidson. One-electron properties of several small molecules using near hartree fock limit basis sets. *J. Chem. Phys.*, 86:3424, 1987. 3.1

[61] H. Huang and Z. Cao. A novel method for optimizing quantum monte carlo wave functions. *J. Chem. Phys.*, 104:200, 1996. 3.1, 3.2

[62] Srinivasan Parthiban and J.M.L. Martin. Fully ab initio atomization energy of benzene via weizmann-2 theory. *J. Chem. Phys.*, 115:2051, 2001. 3.1, 3.3

[63] W. C. Ermler and C. W. Kern. Properties of the benzene molecule near the hartree-fock limit. *J. Chem. Phys.*, 58:3458, 1973. 3.1

[64] J.C. Grossman and L. Mitas. Efficient quantum monte carlo energies for molecular dynamics simulation. *Phys. Rev. Lett.*, 94:56403, 2005. 3.2, 5

[65] D.R.Garmer and J.B. Anderson. Quantum chemistry by random walk: Methane. *J. Chem. Phys.*, 86:4025, 1987. 3.2

[66] ShihI Lu. Accurate atomization energies and dipole moments from ornstein uhlenbeck diffusion quantum monte carlo calculations for small first row polyatomic molecules. *J. Chem. Phys.*, 118:9528, 2003. 3.2

[67] A. Luchow and D. Feller J. B. Anderson. Improved estimates of the total correlation energy in the ground state of the water molecule. *J. Chem. Phys.*, 106:7706, 1997. 3.2

[68] J. M. L. Martin. *Chem. Phys. Lett.*, 303:399, 1999. 3.2

[69] I. Røeggen. Derivation of an extended geminal model. *J. Chem. Phys.*, 79:5520, 1983. 3.2

[70] I. Røeggen and J. Almlof. Interatomic potential for the x(1)sigma(+)(g) state of be-2. *Int. J. Quantum. Chem.*, 60:453, 1996. 3.2

[71] P. Horsch. Correlation effects on bond alternation in polyacetylene. *Phys. Rev. B*, 24:7351, 1981. 3.3

[72] D. Baeriswyl and K. Maki. Electron correlations in polyacetylene. *Phys. Rev. B*, 31:6633, 1985. 3.3

[73] L. Pauling. *The nature of the chemical bond*, page 204. Third edition, Cornell University Press, Ithaca, New York. 3.3

[74] D. Feller and D. A. Dixon. A monte carlo study of titrating polyelectrolytes. *J. Chem. Phys.*, 104:3048, 2000. 3.3

[75] P. J. Linstrom and W. G. Mallard, editors. *NIST Chemistry Webbook, NIST Sandard Reference Database*. National Institute of Standards and Technology, Gaithersburg, 2001. 3.7

[76] F.H. Zong C. Lin and D.M. Ceperley. Twist-averaged boundary conditions in continuum quantum monte carlo algorithms. *Phys. Rev. E*, 64:16702, 2001. 4.1, 7.1

[77] G. Ortiz and P. Ballone. Correlation energy, structure factor, radial distribution function, and moment distribution of the spin-polarized uniform electron gas. *Phys. Rev. B*, 50(3):1391, 1994. 4.1

[78] P.J. Steinbach and B.R. Brooks. New sperichal cutoff methods for long range forces in macromoleculer simulation. *J. Comp. Chem.*, 15:667, 1994. 4.2

[79] P.P. Ewald. Die berechnung optischer und elektrostatischer gitterpotentiale. *Ann. Phys.*, 64:253, 1921. 4.2.1

[80] D. Frenkel and B. Smit. *Understanding Molecular Simulation: From Algorithms to Applications*. Academic Press: A division of Harcourt, Inc., 1996. 4.2.1

[81] O.H. Martin and R.M. Martin. Quantum mechanical thoery of stress and force. *Phys. Rev. B*, 32:3780, 1985. 4.3

[82] J.M. Haile. *Molecular Dynamics Simulation: Elementary Methods*. A Wiley Interscience Publication, 1992. 4.3

[83] F. A. Lindemann. *Phys. Z*, 11:609, 1910. 4.4

[84] J.G. Pasta E. Fermi and S.M. Ulam. Studies of non-linear problems. *LASL Report LA-1940*, 1955. 5

[85] I. Stich L. Mitas, J. Grossman and J. Tobik. Silicon clusters of intermediate size: Energetics, dynamics and thermal effects. *Phys. Rev. Lett.*, 84:1479, 2000. 5

[86] S. Chiesa and D. M. Ceperley S. Zhang. Accurate, efficient, and simple forces computed with quantum monte carlo methods. *Phys. Rev. Lett.*, 94:036404, 2005. 5.2

[87] T. Schneider and E. Stoll. Molecular-dynamics study of a three-dimensional one-component model for distortive phase transitions. *Phys. Rev. B*, 17(3):1302, 1978. 5.2, 5.4.1

[88] J. M. Wozniak J. A. Izagiurre, D. P. Caterello and R. D. Skeel. Langevin stabilization of molecular dynamics. *J. Chem. Phys.*, 114(5):2090, 2001. 5.2, 5.4.1

[89] F. R. Krajewski and M. Parrinello. Linear scaling electronic structure calculations and accurate sampling with noisy forces. 2005. 5.3

[90] Hannes Risken. *The Fokker-Planck Equation*. Springer, 1984. 5.3, 5.3

[91] C. L. Brooks A. Brunger and M. Karplus. Stochastic boundary conditions for molecular dynamics simulations of st2 water. *Chem. Phys. Lett.*, 105:495, 1984. 5.3.1

[92] W. F. van Gunsteren and H. J. C. Berendsen. *Molec. Phys.*, 45:637, 1982. 5.3.1

[93] J. A. Izaguirre R.D. Skeel. An impulse integrator for langevin dynamics. *Mol. Phys.*, 100(24):3885, 2002. 5.3.1, 5.3.1, 5.3.1

[94] G. Dahlquist and A. Bjorck. *Numerical Methods*. Englewood Cliffs, NJ: Prentice-Hall, 1974. 5.3.1

[95] C. L. Brooks III M. Berkowitz and S.A. Adelman. Generalized langevin theory for many-body problems in chemical dynamics: Modelling of solid and liquid state response functions. *J. Chem. Phys.*, 72:3889, 1980. 5.4.1

[96] F. Tassone F.Mauri and R. Car. Acceleration schemes for ab initio molecular-dynamics simulations and electronic-structure calculations. *Phys. Rev. B*, 50:15, 1994. 5.4.1

[97] M.C. Payne M.P. Teter D. C. Allan T.A. Arias J.D. Joannopoulos. Iterative minimization techniques for ab initio total-energy calculations: molecular dynamics and conjugate gradients. *Rev. Mod. Phys.*, 64(4):1045, 1992. 5.4.2

[98] M.P. Allen and D.J. Tildesley. *Computer Simulation of Liquids*. Oxford Univeristy Press, 1991. 6.2

[99] D. Hohl V. Natoli D. M. Ceperlay and R.M. Martin. Molecular dissosation in dense hydrogen. *Phys. Rev. Lett.*, 71:541, 1993. 6.2, 6.2

[100] M. V. Sadovskii. Superconductivity and localization. *Rev. Mod. Phys.*, 0:0, 1999. 6.3

[101] Sadovskii MV. Superconductivity and localization. *PHYSICS REPORTS-REVIEW SECTION OF PHYSICS LETTERS*, 282:226, 1997. 6.3

[102] Daniel Rohe and Walter Metzner. Pseudogap at hot spots in the two-dimensional hubbard model at weak coupling. *Phys. Rev. B*, 71:115116, 2005. 6.3, 7

[103] C.N. Yang. Concept of off-diagonal long-range order and quantum phase of liquid he and of superconductors. *Rev. Mod. Phys.*, 34:694, 1962. 6.3

[104] S. De Palo F. Rapisarda and Gaetano Senatore. Excitonic condensation in a symmetric electron-hole bilayer. *Phys. Rev. Lett.*, 88:206401, 2002. 6.3

[105] P Capello M Becca F Fabrizio M Sorella S Tosatti E. Variational description of mott insulators. *Phys. Rev. Lett.*, 94(2):026406, 2005. 7

[106] S. Sorella G. Martins F. Becca C. Gazza L. Capriotti A. Parola and E. Dagotto. Superconductivity in the two-dimensional t-j model. *Phys. Rev. Lett*, 88:117002, 2002. 7

[107] T.A. Maier M. Jarrell T.C. Schulthess P. R. C. Kent and J.B. White. A systematic study of superconductivity in the 2d hubbard model. 2005. 7

[108] P. R. C. Kent R. Q. Hood A. J. Williamson R. J. Needs W. M. C. Foulkes and G. Rajagopal. Finite-size errors in quantum many-body simulations of extended systems. *Phys. Rev. B*, 59:1917, 1999. 7.1

[109] Bradley Efron. *The Jackknife, the Bootstrap and Other Resampling Plans*. Society for Industrial and Applied Mathematics, 1982. A.1

[110] A. J. Coleman. *J. Math. Phys.*, 13:214, 1972. B.2

[111] B. Mishra and T. Schlick. The notion of error in langevin dynamics. i. linear analisys. *J. Chem. Phys.*, 105:299, 1996. E