Scuola Internazionale Superiore
di Studi Avanzati

Ab initio molecular dynamics of water by quantum Monte Carlo

Author: Ye Luo
Supervisor: Prof. Sandro Sorella

A thesis submitted for the degree of
Doctor of Philosophy

October 2014

Acknowledgments

I would like to acknowledge first my supervisor Prof. Sandro Sorella. During the years in SISSA, he guided me in exploring the world of QMC with great patience and enthusiasm. He's not only a master in Physics but also an expert in high performance computing. He never exhausts new ideas and is always advancing the research at the light speed. For four years, I had an amazing journey with him.

I would like to appreciate Dr. Andrea Zen, Prof. Leonardo Guidoni and my classmate Guglielmo Mazzola. I really learned a lot from the collaboration on various projects we did.

I would like to thank Michele Casula and W.M.C. Foulkes who read and correct my thesis and also give many suggestive comments.

I would like to express the gratitude to my parents who give me unlimited support even though they are very far from me. Without their unconditional love, I can't pass through the hardest time. I feel very sorry for them because I went home so few times in the previous years. Therefore, I would like to dedicate this thesis to them.

I remember the pleasure with my friends in Trieste who are from all over the world. Through them, I have access to so many kinds of food and culture. They help me when I meet difficulties and they wipe out my loneliness by sharing the joys and tears of my life. Especially, I would like to thank all my Italian friends because they often have to bear my poor spoken Italian.

I would like to thank all my colleagues. They are always very kind to spare their precious time to share their wealth of knowledge with me in order to resolve my doubts.

Last but not the least, I would like to thank SISSA and Trieste. SISSA is by far the best place I have ever worked or studied. Besides its academic prestige, it's attractive because of its wonderful location and facilities. The beautiful garden and the whole landscape view of Trieste gulf alleviate a lot my stress during the working hours. Trieste city is charming and the local people (in general, Italian people) are warmhearted. I had so many unforgettable memories of Trieste. For this reason, every time I see the lights on the Trieste coast when my train runs across the historic bridge in Barcola, I'm filled with happiness since my beloved city is approaching and my home is there.

Contents

Chapter 1

Introduction

"The Blue Planet" is how we describe our beloved homeland planet — Earth. When we see its pictures taken from the satellites, the planet is wrapped by a beautiful blue color under the sunshine with a black background of the dark universe. The blue color comes from the oceans which cover 71% of the surface of Earth and all the oceans contain 97.1% of the total amount of the water on Earth in a liquid form. The rest part (less than 3%) is found in groundwater, glaciers or ice caps and other large water bodies like lakes and rivers. Water circulation is one of the most important circulations on Earth besides the carbon circulation. Water evaporates from the sea and forms clouds in the sky. Some of those clouds stay still and some others are carried to the lands by the wind and later turn into rainfalls. Rain drops accumulate and join rivers or groundwater and finally return back to the sea. During this circulation, the flowing rivers cut the ground and valleys form. Meanwhile, the mud and sand from the erosion of river banks are carried by the river to the lower course and plains form. In this way the landscape of Earth was reshaped gradually in the past billions of years. This circulation also causes the weather changes. Having clouds, rains, snows and fogs totally depends on the amount of water in the air and sky. Moreover, the abundance of water strongly affects the climate of certain areas. For instance, rain forests around the amazon river have plenty of rain while the Sahara deserts have almost zero precipitation.

The other contribution of the water to Earth, even more important than that to the geographic evolution, is that it brings life to Earth. Even though there are many models of abiogenesis, the earliest life is discovered in the oceans without any doubts. From bacteria, later algae to marine plants and fishes, the forms of lives evolve their sizes and complexity during the billions of years. On the land, insects, plants and animals started to appear a little bit later than in the oceans. They no longer live in an aquatic environment but they need acquiring water for sustaining their lives since a very large portion of the body weights consists of water. For instance, the human body contains from 55% to 78% water. If one stops drinking water for a week, he will die soon because of organ failures due to the dehydration. Many

biological processes happen with the participation of water. As one of the most important processes in cyanobacteria and plants, the photosynthesis

$$2n\mathrm{CO_2} + 2n\mathrm{H_2O} \rightarrow 2n\mathrm{CH_2O} + 2n\mathrm{O_2}$$

which generates the oxygen in the atmosphere needs water as one of the raw materials. In another case, when leaves fall off the trees and get dry, their color turns from green into yellow. According to our current knowledge about those lives we are familiar with, the existence of liquid water is a key prerequisite for any kind of lives. Therefore, whenever we send robot rovers to other planets, looking for the signs of water is always a very important subject in order to discover alien lives. When the Cassini-Huygens arrived Enceladus, a satellite of Saturn, we were so encouraged by this piece of inspiring news that water vapor was discovered there. Since water plays such an important role in lives on Earth, it deserves a very deep understanding.

Our journey on understanding water started very early. In 1783, Cavendish published a paper on the production of water [1] by burning inflammable air (hydrogen) in dephlogisticated air (oxygen). Later physicists found that atoms could be divided into electrons and nuclei. As the positive charges of nuclei increase, all the elements can be sorted in the periodic table. We know that oxygen is the richest element on Earth and hydrogen is the lightest element and also the first element appeared at the original stage of the Universe according to the Big Bang theory. Not only in chemistry but also in physics, water has left so many footprints because of its easy accessibility. In the early days, the SI unit kilogram was chosen to be as heavy as $1\,\mathrm{dm}^3$ of water. In the Celsius measurement of temperature, zero and a hundred Celsius are the frozen and boiling points of water at the atmosphere. Besides, the lifting power, buoyancy, of liquids was first explained by Archimedes with an experiment in a bathtub filled with water.

All our knowledge about water in physics and chemistry came from experiments till 1925 when quantum mechanics was developed to describe how the quantum states of a physical system evolve with time. It laid the foundation of studying materials by a theoretical method. By solving the Schrödinger equation [2], in principle, all the electronic properties of materials can be studied theoretically. However, in practice, solving such many body problem is very difficult. Thanks to fast evolving computers, it becomes possible to simulate the liquid water theoretically and compare the results with experiments.

It's been almost seven decades since the first general purpose electronic computer ENIAC was born in 1946. Ten years later, the history of supercomputer started with the Atlas at the University of Manchester and several computers designed by Seymour Cray to exploit innovative parallelism to achieve superior peak performance. The period of 1970s and 1980s was the Cray era when Cray-1/2 supercomputer achieved gigaflops performance. The main boost of supercomputer came in the early 1990s with the emerging computer clusters which consist of computers connected by network. Before

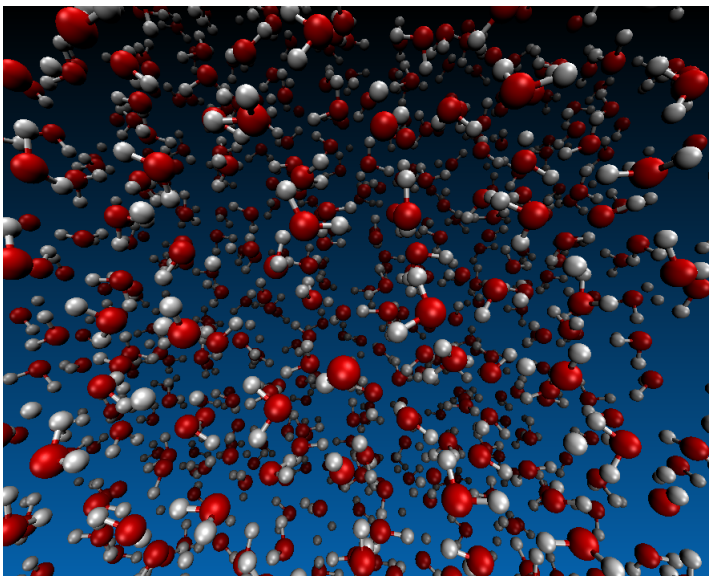

Figure 1.1: A snapshot of the simulation of liquid water.

that, supercomputer had excluded clusters and relied on shared memory and vector processors with internal parallelism but the cost of those components are extremely expensive. Instead, computer cluster has distributed memory and its computing power can be dramatically increased by means of integrating more computers which are much less expensive. Nowadays on the TOP500 organization's semiannual list of the 500 fastest supercomputers, almost all the supercomputers including all the top 10 have cluster architecture. The peak performance has reached $\sim$ 34 petaflops by Tianhe-2 in China which deploys 32,000 processors with 48,000 accelerators. In two or three years, new exaflops supercomputer will appear.

When equipped with supercomputer, we have several methods with various accuracy for study materials at different scales. Usually, the more accurate the method is, the more its computation costs. In figure 1.2, the system size of a typical simulation now and the qualitative accuracy of the method used by this simulation show a negative correlation if the computing power is fixed. Simulations of systems with millions of atoms can be easily performed by empirical force field methods but are not affordable for other ab initio approaches. The former ones have been largely applied to study interesting properties in physics, chemistry and biology, including the thermal conductivity of materials, viscoelasticity of polymers and the folding process of proteins. However, such classical simulations treat atoms as effective potentials without any explicit quantum description of electrons and nuclei and thus is inadequate to describe the electronic properties of real systems. Density functional theory (DFT), one of the most popular ab initio methods, usually gives more accurate results compared with empirical force fields and is capable of studying the electronic structure of a material but it's limited by the quality of the density

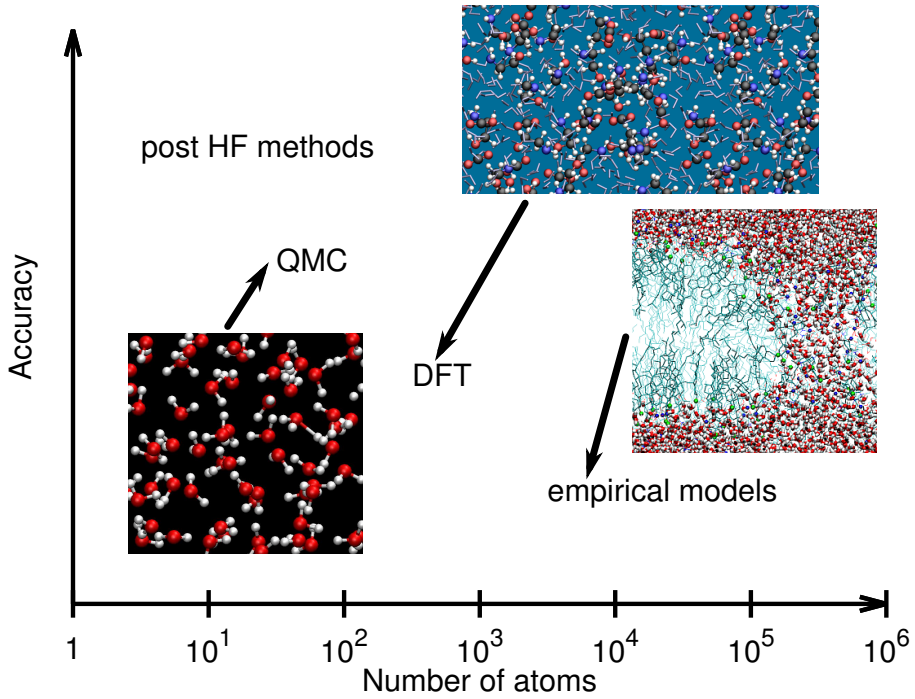

Figure 1.2: Accuracy and system size capability of various simulation methods.

functional. By further climbing up to the next steps, quantum Monte Carlo (QMC) methods which provide better accuracy because of less approximation compared with DFT while the most accurate and expensive methods are the configuration interaction (CI) and coupled cluster (CC) often called as post-Hartree-Fock (post-HF) methods. Usually a plain wave DFT algorithm, QMC and post-HF methods have typical scaling as n^3, $n^{3\sim4}$ and $n^{5\sim}$ respectively.

As the computing power increases, we can choose to simulate systems of larger sizes with tools we already have or start refining our old knowledge by studying the same system with approaches at a higher level of accuracy. This thesis takes the later course and we have simulated the liquid water for the first time with a molecular dynamics approach driven by the forces evaluated by quantum Monte Carlo.

Before starting the simulation, we have solved the following issues

a) Determine a good wavefunction ansatz, which provides a perfect balance between accuracy and computational cost.

b) Devise efficient optimization methods, since we need to optimize at least ten thousand parameters and the optimization is intensively employed in order to fulfill the Born-Oppenheimer approximation when ions are moving during the molecular dynamics.

c) Develop a robust molecular dynamics scheme even if the forces are eval-

uated in presence of statistical errors.

The chapter 2, we deal with the challenge b). It focuses on the variational Monte Carlo (VMC) and the wavefunction optimization methods based on VMC. The performance of different methods are displayed through the optimization of the Jastrow factor in our test case Beryllium dimer and the efficiency is improving surprisingly during the evolution of these methods.

In chapter 3, we focus on the challenge a). It describes the wavefunction ansatz used by our simulation. In this thesis, we introduce the atomic hybrid orbitals which significantly increase the compactness of our wavefunction without hurting accuracy. This chapter also explain how to optimize the determinant in a way that the number of variational parameters scales only linearly with the system size. This further helps the efficiency of the wavefunction optimization.

In chapters 4 and 5, the issue c) is explained in detail. In chapter 4, a second order Langevin dynamics (SLD) scheme is devised particularly for QMC and this thesis improves this scheme by developing a better integration method. Here, we also highlight the remarkable power of the force covariance matrix which can be defined only in QMC and is capable of accelerating the slow modes of a dynamics. In chapter 5, this SLD for QMC is validated through intensive benchmarking on the calculation of the vibrational frequencies of water and other small molecules. It is shown that many systematic biases in our MD scheme and QMC evaluation can be controlled so that we are confident to push forward this ab initio molecular dynamics for applications on large systems.

Finally in chapter 6, we perform the simulation of liquid water with all the preparation done in the previous chapters. The results are encouraging since we've closed the discrepancy of the peak positions of RDFs between experiments and ab initio simulations. The power of QMC is also demonstrated by the fact that the shapes of our RDFs are much less structured than previous DFT-based ab initio simulations even if the two water molecule interaction is dealt with the same level of accuracy as the DFT/BLYP calculation. In this chapter, we have also studied the features of hydrogen bonds in our simulation of liquid water. All our results indicate that it is important to consider the quantum nature of the ions for a faithful description of liquid water. This will be left for future studies, possible in principle even within the QMC approach.

Chapter 2

Quantum Monte Carlo

2.1 Introduction

The nature of atoms, molecules, clusters, solids and liquids in condensed matter physics is determined by the fundamental particles, electrons and nuclei whose understanding is based upon theoretical methods of quantum mechanics and statistical mechanics. The Hamiltonian of such a system consisting of M ions and N electrons can be expressed in the following form by neglecting the relativistic effect

$$\begin{aligned}\hat{H} = &-\sum_{a=1}^{M}\frac{\hbar^2}{2M_a}\nabla^2_{\boldsymbol{R}_a} + \frac{1}{2}\sum_{a\neq b}^{M}\frac{Z_a Z_b e^2}{|\boldsymbol{R}_a - \boldsymbol{R}_b|} \\ &-\sum_{i=1}^{N}\frac{\hbar^2}{2m_e}\nabla^2_{\boldsymbol{r}_i} - \sum_{a=1}^{M}\sum_{i=1}^{N}\frac{Z_a e^2}{|\boldsymbol{R}_a - \boldsymbol{r}_i|} + \frac{1}{2}\sum_{i\neq j}^{N}\frac{e^2}{|\boldsymbol{r}_i - \boldsymbol{r}_j|} \qquad (2.1)\\ = &\hat{T}_N + \hat{V}_{NN} + \hat{T}_e + \hat{V}_{Ne} + \hat{V}_{ee} \qquad (2.2)\end{aligned}$$

where $\boldsymbol{R}_a$, Z_a and M_a are the coordinates, proton number and mass of the ion a while $\boldsymbol{r}_i$ and m_e are the coordinates of the electron i and the electron mass. e is the one electron charge. In the first line of eq. (2.1), the two terms are the ionic kinetic energy operator $\hat{T}_N$ and potential operator $\hat{V}_{NN}$ of all ion-ion pairs. In the second line of eq. (2.1), the three terms are the electronic kinetic energy operator $\hat{T}_e$ and potential operators $\hat{V}_{Ne}$ and $\hat{V}_{ee}$ of all electron-ion and electron-electron pairs. With this Hamiltonian, all properties of the system can be obtained in principle by solving the time-independent Schrödinger equation

$$\hat{H}\Psi_{\text{full}}(\boldsymbol{R}, \boldsymbol{r}) = E\Psi_{\text{full}}(\boldsymbol{R}, \boldsymbol{r})\,. \qquad (2.3)$$

However, solving this equation in the full quantum mechanic framework, in practice, is very difficult since there are too many degrees of freedom. A practical way to overcome this difficulty is to invoke the so called Born-Oppenheimer approximation by recognizing the big difference in the time scales between

nuclear and electronic motion. The full wavefunction in eq. (2.3) can be factorized into the product of a nuclear part $\chi(\boldsymbol{R})$ and an electronic part $\Psi(\boldsymbol{r};\boldsymbol{R})$ which depends parametrically on the nuclear positions $\boldsymbol{R}$

$$\Psi_{\text{full}}(\boldsymbol{R},\boldsymbol{r}) = \chi(\boldsymbol{R})\Psi(\boldsymbol{r};\boldsymbol{R})\,. \tag{2.4}$$

Since the nuclei are much heavier than the electrons by at least three orders of magnitude, electrons can follow immediately the motion of the ions. The nuclear kinetic term could be approximated as

$$\hat{T}_N\Psi_{\text{full}}(\boldsymbol{R},\boldsymbol{r}) \approx \Psi(\boldsymbol{r};\boldsymbol{R})\hat{T}_N\chi(\boldsymbol{R})\,. \tag{2.5}$$

Thus, eq. (2.3) can be disentangled into two coupled equations

$$(\hat{T}_e + \hat{V}_{Ne} + \hat{V}_{ee})\Psi(\boldsymbol{r};\boldsymbol{R}) = E_e\Psi(\boldsymbol{r};\boldsymbol{R}) \tag{2.6}$$

$$(\hat{T}_N + \hat{V}_{NN} + E_e(\boldsymbol{R}))\chi(\boldsymbol{R}) = E\chi(\boldsymbol{R})\,. \tag{2.7}$$

The first equation is the eigenvalue equation for the electronic Hamiltonian $\hat{H}_e = \hat{T}_e + \hat{V}_{Ne} + \hat{V}_{ee}$ which yields a set of eigenfunctions $\Psi(\boldsymbol{r};\boldsymbol{R})$ and eigenvalues E_e depending parametrically on $\boldsymbol{R}$. Together with the repulsive Coulomb interaction between the nuclei it provides an effective potential in which the nuclei move

$$V^n = E_e^n(\boldsymbol{R}) + V_{NN}(\boldsymbol{R}) \tag{2.8}$$

The effective potential $V^0(\boldsymbol{R})$ associated with the electronic ground state is called the Born-Oppenheimer surface or *potential energy surface* (PES). On the PES, the nuclear eigenvalue problem can be solved, which yields a set of rotational and vibrational levels in the nuclear motion.

Since there is no known analytic solution to the many body system described by $\hat{H}_e$, a numerical solution is then needed. The first attempt was done by Fock [3] in 1930 with the so called Hartree-Fock method. If there's no spin-orbit interaction, the determinant wavefunction Ψ_{SD} can be written as a Slater determinant

$$\Phi_{\text{SD}} = \frac{1}{(N!)^{1/2}}\begin{vmatrix} \phi_1(\boldsymbol{r}_1,\sigma_1) & \phi_1(\boldsymbol{r}_2,\sigma_2) & \phi_1(\boldsymbol{r}_3,\sigma_3) & \cdots \\ \phi_2(\boldsymbol{r}_1,\sigma_1) & \phi_2(\boldsymbol{r}_2,\sigma_2) & \phi_2(\boldsymbol{r}_3,\sigma_3) & \cdots \\ \phi_3(\boldsymbol{r}_1,\sigma_1) & \phi_3(\boldsymbol{r}_2,\sigma_2) & \phi_3(\boldsymbol{r}_3,\sigma_3) & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{vmatrix}, \tag{2.9}$$

where ϕ_i are single particle "spin orbitals". The Hartree-Fock approach turns therefore to a minimization of the functional

$$E_{\text{HF}} = \min \frac{\langle\Phi_{\text{SD}}|\hat{H}|\Phi_{\text{SD}}\rangle}{\langle\Phi_{\text{SD}}|\Phi_{\text{SD}}\rangle} \tag{2.10}$$

with an iterative algorithm. Since the Slater determinant takes into account all the particle exchange, the difference between exact non-relativistic ground

state energy E_{exact} and Hartree-Fock E_{HF} in the limit of a complete basis set is the residual energy — the so-called electron *correlation energy* a term coined by Löwdin [4]

$$E_{\text{corr}} = E_{\text{exact}} - E_{\text{HF}} \,. \tag{2.11}$$

In QMC calculations, the correlation energy is recovered by correlating the Slater determinant with a Jastrow factor discussed in chapter 3.

The term "Quantum Monte Carlo (QMC)" stands for a number of stochastic algorithms which solve quantum many-body problems on various systems ranging from model Hamiltonian to molecules and solids [5, 6]. The most distinctive and common feature of these algorithms is to solve Schrödinger equation and evaluate its properties in a stochastic way.

The first QMC, variational Monte Carlo (VMC) [7], was introduced half a century ago when the history of computer was only two decades old. In the following decades, more sophisticated and costly projection methods emerged such as Green function Monte Carlo (GFMC) [8], diffusion Monte Carlo (DMC) [9], reptation Monte Carlo (RMC) [10] and lattice regularized diffusion Monte Carlo (LRDMC) [11, 12]. Apart from these methods suitable for properties at zero temperature, path integral Monte Carlo (PIMC) [6] was developed for calculating thermodynamic properties. Unlike all those QMC variants mentioned above, which sample the wavefunction in the configuration space, full configuration interaction quantum Monte Carlo (FCIQMC) [13] and auxiliary-field quantum Monte Carlo (AFQMC) [14, 15] do the sampling in the determinant space.

In the last decade, QMC methods have been rapidly gaining interest among physics, chemistry and even biology for electronic structure calculations because of their favorable scaling with the system size, accuracy comparable to high level quantum chemistry methods and readiness for the implementation on modern parallel computers. The computational cost of QMC algorithms scales as $N_{\text{el}}^3 \sim N_{\text{el}}^4$ with the number N_{el} of electrons while post-Hartree-Fock quantum chemistry methods like configuration interaction (CI) [16] and couple cluster (CC) [17] scale at least as N_{el}^5. Within the same accuracy, QMC is capable of studying much larger systems far beyond only small compounds. Obviously QMC methods are computationally more demanding than mean field ones, such as density functional theory (DFT) and Hartree-Fock (HF), but the accuracy of QMC, comparable with post-HF methods, is much better than them since it's devised as a many-body technique. We've seen much QMC success in solving problems where the mean field approximation fails. Nowadays parallel supercomputers increase their computational power mainly by integrating more compute units and executing calculations on them simultaneously. QMC algorithms are the ideal applications for these machines because they can be very easily implemented in parallel thanks to its way of generating samples.

This chapter is organized as follows. Section 2.2 introduces the variational Monte Carlo method. Section 2.3 describes the Metropolis algorithm and

shows how VMC is implemented. Then the scheme of wavefunction optimization is presented in section 2.4.

2.2 Variational Monte Carlo

Variational Monte Carlo (VMC) is the cheapest and first developed QMC algorithm. In 1964, W.L. McMillan [7] introduced it from classical liquid theory to calculate the ground state of liquid ^{4}He. A decade later, D. Ceperley et al. [18] generalized it to fermions. VMC is based on the 'Variational principle'. With a given *trial wavefunction* Ψ_T, the expectation value of the Hamiltonian is the variational energy E_V which provides an upper bound to the exact ground state energy E_gs

$$E_\mathrm{V} = \frac{\langle \Psi_\mathrm{T}|\hat{H}|\Psi_\mathrm{T}\rangle}{\langle \Psi_\mathrm{T}|\Psi_\mathrm{T}\rangle} \geq E_\mathrm{gs}\,. \tag{2.12}$$

A meaningful Ψ_T should have the proper symmetry and must be normalizable.

If the configuration space $|x\rangle$ is chosen as the basis set where x denotes a many-electron configuration where all the positions of the electrons and their spins are known, then by inserting the basis set completeness $\int_x \mathrm{d}x\ |x\rangle\langle x| = 1$, eq. (2.12) can be recast as

$$E_\mathrm{V} = \frac{\int_x \mathrm{d}x\ \langle \Psi_\mathrm{T}|x\rangle\langle x|\hat{H}|\Psi_\mathrm{T}\rangle}{\int_x \mathrm{d}x\ \langle \Psi_\mathrm{T}|x\rangle\langle x|\Psi_\mathrm{T}\rangle} = \frac{\int_x \mathrm{d}x\ e_\mathrm{L}(x)\psi_\mathrm{T}^2(x)}{\int_x \mathrm{d}x\ \psi_\mathrm{T}^2(x)}\,, \tag{2.13}$$

where $\psi_\mathrm{T}(x) = \langle x|\Psi_\mathrm{T}\rangle$ and $e_\mathrm{L}(x)$ is the so-called *local energy*

$$e_\mathrm{L}(x) = \frac{\langle x|\hat{H}|\Psi_\mathrm{T}\rangle}{\langle x|\Psi_\mathrm{T}\rangle} = \frac{\hat{H}\psi_\mathrm{T}(x)}{\psi_\mathrm{T}(x)}\,. \tag{2.14}$$

$\mathrm{d}x$ is a shorthand notation for the $3N$-dimensional integral over all the electron coordinates. The $3N$-dimensional integral of eq. (2.13) can be evaluated with Monte Carlo integration. We rewrite it as

$$E_\mathrm{V} = \int \mathrm{d}x\ \pi(x)e_\mathrm{L}(x) = \langle e_\mathrm{L}\rangle, \qquad \pi(x) = \frac{\psi_\mathrm{T}^2(x)}{\int_x \mathrm{d}x\ \psi_\mathrm{T}^2(x)}\,, \tag{2.15}$$

where $\pi(x)$ is a probability density. As we will show in the next section, a stochastic process (Markov chain $\{x_n\}$) is capable of generating configurations according to the desired probability $\pi(x)$. Then the expectation value of the energy is the mean of $e_\mathrm{L}(x)$ during the process

$$E_\mathrm{V} = \frac{1}{N_\mathrm{MC}} \sum_n^{N_\mathrm{MC}} e_\mathrm{L}(x_n)\,. \tag{2.16}$$

Suppose all the samples $\{x\}$ are independent, according to the *central limit theorem* (CLT), the average value $\bar{y}$ of a random variable y approaches Gaussian distribution when the number N_{MC} of samples goes to infinity. Its standard deviation $\sigma(\bar{y})$ (also called error bar) falls off like $\frac{\sigma(y)}{\sqrt{N_{\mathrm{MC}}}}$ as N_{MC} increases. In our case,

$$\sigma(E_{\mathrm{V}}) = \frac{\sigma(e_{\mathrm{L}})}{\sqrt{N_{\mathrm{MC}}}}\,. \tag{2.17}$$

Thus the statistical uncertainty in the estimation of the variational energy can be suppressed to arbitrarily small values as long as N_{MC} is large enough. Therefore, the Monte Carlo integration can be pushed to any desirable accuracy in principle.

Suppose that Ψ_{T} coincides with an eigenstate Ψ_{eg} of $\hat{H}$, namely $\hat{H}|\Psi_{\mathrm{eg}}\rangle = E_{\mathrm{eg}}|\Psi_{\mathrm{eg}}\rangle$ where E_{eg} is the corresponding energy eigenvalue, the local energy

$$e_{\mathrm{L}}(x) = \frac{\hat{H}\psi_{\mathrm{eg}}(x)}{\psi_{\mathrm{eg}}(x)} = E_{\mathrm{eg}} \tag{2.18}$$

keeps constant regardless of configuration $|x\rangle$, which clearly implies that its variance is exactly zero and its average value is E_{eg}. It should be noted that this *zero variance property* holds for all eigenstates of $\hat{H}$, not necessarily the ground state. It provides a rigorous criterion for examining how far Ψ_{T} is from Ψ_{eg}. Let us see the quantum average of the Hamiltonian squared on a variational state Ψ_{T}

$$\begin{aligned}\langle\hat{H}^2\rangle &= \frac{\langle\Psi_{\mathrm{T}}|\hat{H}^2|\Psi_{\mathrm{T}}\rangle}{\langle\Psi_{\mathrm{T}}|\Psi_{\mathrm{T}}\rangle} = \frac{\int_x \mathrm{d}x\ \langle\Psi_{\mathrm{T}}|x|\hat{H}\rangle\langle x|\hat{H}|\Psi_{\mathrm{T}}\rangle}{\int_x \mathrm{d}x\ \langle\Psi_{\mathrm{T}}|x\rangle\langle x|\Psi_{\mathrm{T}}\rangle} \\ &= \frac{\int_x \mathrm{d}x\ e_{\mathrm{L}}^2(x)\psi_{\mathrm{T}}^2(x)}{\int_x \mathrm{d}x\ \psi_{\mathrm{T}}^2(x)} = \int \mathrm{d}x\ \pi(x)e_{\mathrm{L}}^2(x) = \langle e_{\mathrm{L}}^2\rangle\end{aligned} \tag{2.19}$$

which is exactly equal to the average of local energy squared. Thus, the quantum variance of the Hamiltonian

$$\begin{aligned}\sigma^2(\hat{H}) &= \langle(\hat{H} - E_{\mathrm{V}})^2\rangle = \langle\hat{H}^2\rangle - E_{\mathrm{V}}^2 \\ &= \langle e_{\mathrm{L}}^2\rangle - \langle e_{\mathrm{L}}\rangle^2 = \sigma^2(e_{\mathrm{L}})\end{aligned} \tag{2.20}$$

is indeed exactly equal to the variance of the local energy. This relation further fortifies our criterion that the smaller the variance, the closer the variational state is to the exact eigenstate.

This variational Monte Carlo method is very general and can be extended not only to discretized systems by replacing all multi-dimensional integrals with summations, but also to general Hermitian operators $\hat{O}$ (e.g. $\hat{S}^2$) by replacing $e_{\mathrm{L}}(x)$ with the corresponding local estimator $O_{\mathrm{L}}(x)$ defined as

$$O_{\mathrm{L}}(x) = \frac{\langle x|\hat{O}|\Psi_{\mathrm{T}}\rangle}{\langle x|\Psi_{\mathrm{T}}\rangle} \tag{2.21}$$

2.3 Metropolis algorithm

To compute the integral (2.15) with a Monte Carlo integration, we need to generate a Markov chain with the given probability distribution $\pi(x)$. The master equation for the iteration n is

$$\rho_{n+1}(x_{n+1}) = \int \mathrm{d}x \; T(x_{n+1}|x_n)\rho_n(x_n) \tag{2.22}$$

where $T(x_{n+1}|x_n)$ is a conditional probability which describes the transition probability from $x_n \to x_{n+1}$. In order to have a stationary distribution $\pi(x)$, it should follow that

$$\rho_{n+1}(x) = \rho_n(x) \to \pi(x) \qquad \text{for large } n. \tag{2.23}$$

For this purpose, it is sufficient to satisfy the *detailed balance* condition

$$T(x_{n+1}|x_n)\pi(x_n) = T(x_n|x_{n+1})\pi(x_{n+1})\,. \tag{2.24}$$

If a Markov chain satisfies detailed balance and is ergodic, then the equilibrium distribution $\pi(x)$ will be always reached, for large enough n, independently of the initial distribution.

In practice, the implementation of generating such Markov chain was first introduced by Metropolis et al. [19]. In their scheme, the transition probability $T(x_{n+1}|x_n)$ can be chosen with great freedom. To ensure the detail balance, the new configuration x_{n+1} generated from x_n with a transition probability $T(x_{n+1}|x_n)$ is only accepted with a probability

$$A(x_{n+1}|x_n) = \min\left\{1, \frac{\pi(x_{n+1})T(x_n|x_{n+1})}{\pi(x_n)T(x_{n+1}|x_n)}\right\}. \tag{2.25}$$

With this scheme, after large enough iterations, the equilibrium distribution $\pi(x)$ is reached.

The typical flow chart of variational Monte Carlo is described as algorithm 1.

It should be emphasized that this algorithm only cares about the ratios of 1) probability distribution and 2) transitional probability between the proposed and current configurations. For the first ratio, e.g. VMC, we can simply calculate the ratio as $\psi_{\mathrm{T}}^2(x'_{n+1})/\psi_{\mathrm{T}}^2(x_n)$ without computing the normalization. For the second ratio, we can choose scheme like smart Monte Carlo [20] to have asymmetric transition probability or choose simple schemes to obtain symmetric transition probability $T(x_{n+1}|x_n) = T(x_n|x_{n+1})$ such that the ratio is constantly 1.

In our implementation of the generation of new electronic configurations, we move electrons one by one and determine if the proposed move should be accepted or rejected. Updating the wavefunction and computing the acceptance ratio after proposing a single electron move is implemented in a very

Initial setup;
for *iteration i less than required statistics* **do**
  Propose a move from $x_i \to x'$ with probability $T(x'|x_i)$;
  Evaluate the new probability $\pi(x')$ and $A(x'|x_i)$;
  Generate a random number $\xi \in [0, 1)$;
  if $\xi < A(x'|x_n)$ **then**
    accept: $x_{i+1} = x'$;
  else
    reject: $x_{i+1} = x_i$;
  end
  Calculate the local energy $e_{\mathrm{L}}(x_{i+1})$;
end
Compute variational energy E_{V} as (2.16);

Algorithm 1: Variational Monte Carlo.

efficient way. Instead, if we propose a global move and it is finally rejected, a lot of computation is wasted. For this reason, we have a more efficient sampling by moving electrons individually. In our scheme, the proposed move of an electron is determined inversely proportional to its distance from the nearest ion. If the electron is close to an ion, it moves with a small step but if it is far away from ions, it correspondingly moves in a large step. Since the moves are asymmetric, the ratio of the transition probability is adjusted accordingly.

2.4 Wavefunction optimization

QMC calculations depend not only on the quality of the wavefunction ansatz, which will be discussed in the following chapter 3, but also on how well the wavefunction is optimized, in order to be as close as possible to the exact ground state.

Optimization is a very popular topic in numerical algorithms. All optimization methods are divided into global optimization and local optimization ones. Methods like simulated annealing, evolutionary algorithms and swarm algorithms are global optimization methods which stochastically explore a very wide range of the landscape. Finding the extrema with these methods is guaranteed only after an infinite number of steps. Usually such optimization methods are applied on NP-hard problems which have non-polynomial scaling with systems size. On the other hand, the local optimization methods focus on searching the local extrema in a deterministic way. They are capable of reaching the extrema within a given time. Commonly used local optimization method are steepest decent, quasi-Newton and conjugate gradient methods. Here we only focus on optimizing the wavefunction with local optimization

methods adapted to QMC.

In QMC, local optimization methods become more difficult since the statistical noise is present in the estimation of the cost function (variance, energy or their linear combination) and its first and second derivatives. There are two types of wavefunction optimization widely used in QMC — the variance and the energy minimization. The former was introduced by C. Umrigar [21] in 1988. This method minimizes the variance of the local energy by varying the wavefunction parameters with fixed Monte Carlo samples. According to zero variance principle, the wavefunction closer to the ground state has a smaller variance. In principle, variance minimization can reach the ground state but, in practice, since the variance can hardly approaches zero due to the limitation of the wavefunction ansatz, we can never check if the optimized wavefunction is the closest possible one to the ground state or just a local minimum in the hyper surface of the variance. Indeed, different authors [22, 23] found that variance minimization failed to give the best possible wavefunction but energy minimization could. In this thesis, we use only energy minimization methods which bring the possibility of structural optimizations and also molecular dynamics. Our target is to devise a scheme that is robust and converging with a minimal number of iterations in the presence of QMC noise. In the following subsections, we will introduce the stochastic reconfiguration (SR) method and SR with Hessian accelerator, conjugate gradient and sign/noise filter. On our test case Be2 (the physics of this molecule is discussed in the last part of section 3.5), their performances (see figure 2.1) are well displayed and their features are discussed.

2.4.1 Stochastic reconfiguration method

Stochastic reconfiguration (SR) was introduced as an optimization method for a generic trial wavefunction in [24, 25]. This technique takes advantage of the knowledge of the wavefunction and allows a much more efficient optimization compared with other conventional methods like steepest decent and Newton methods.

The SR method is derived as follows. In projection methods, as long as the trial wavefunction is not orthogonal to the ground state wavefunction,[1] we can obtain a new state closer to the ground state by applying the projection operator $(\Lambda - \hat{H})$ to the current trial wavefunction for Λ large enough. If we could evolve the wavefunction as closer as possible to the new wavefunction after the projection, energy minimization could be achieved. For this purpose,

[1] In practice, this is always true because the machine precision will deteriorate the orthogonality after several iterations.

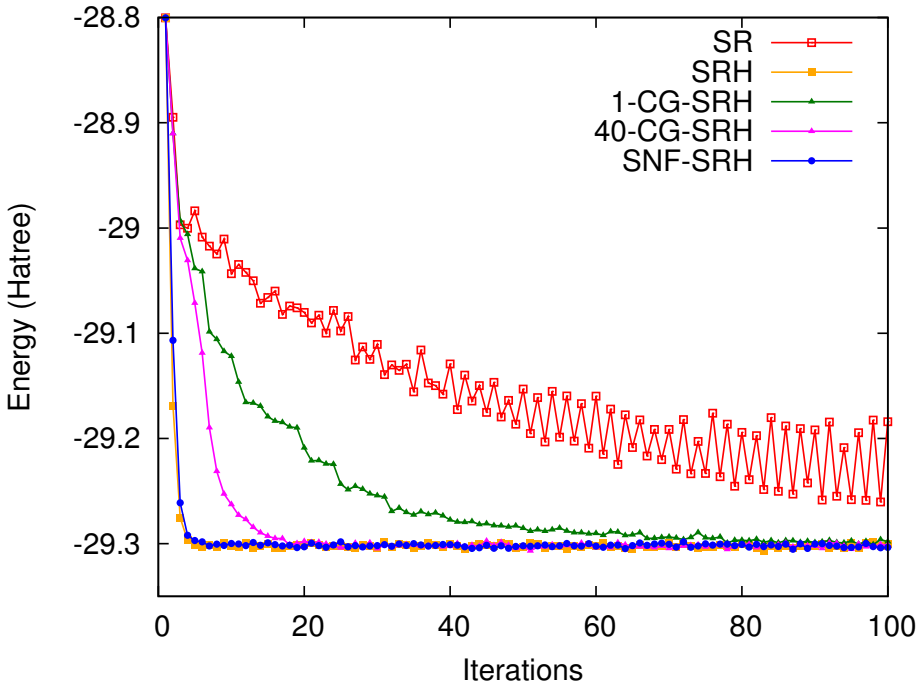

Figure 2.1: The performance test of the optimization of the Jastrow factor with the stochastic reconfiguration method and its variants. The test case is Be_2 and the initial wavefunction is a Slater determinant with the Jastrow factor equal to 1. We choose 0.1 as the trial step in SR and, for stability reasons, we reduce the trial step automatically generated by SRH, SRH-CG and SRH-SNF by a factor of 0.7. In the SRH-CG-n, we use n directions while in the SRH-SNF, we use 40 significant parameters among a total of 94 ones.

we define

$$|\Psi_{\mathrm{T}}^{P}\rangle = (\Lambda - \hat{H})|\Psi_{\mathrm{T}}^{(\alpha_1,\ldots,\alpha_p)}\rangle \tag{2.26}$$

$$|\Psi_{\mathrm{T}}^{\prime(\alpha_1',\ldots,\alpha_p')}\rangle = |\Psi_{\mathrm{T}}^{(\alpha_1,\ldots,\alpha_p)}\rangle + \sum_{i=1}^{p} \delta\alpha_i \frac{\partial}{\partial\alpha_i}|\Psi_{\mathrm{T}}^{(\alpha_1,\ldots,\alpha_p)}\rangle \tag{2.27}$$

where Ψ_{T}^{P} is the new wavefunction after projection and Ψ_{T}' is the optimized wavefunction and $\alpha_i (i = 1,\ldots,p)$ are the parameters of the wavefunction. Eq. (2.27) can be rewritten as

$$|\Psi_{\mathrm{T}}^{\prime(\alpha_1',\ldots,\alpha_p')}\rangle = \sum_{i=0}^{p} \delta\alpha_i \hat{O}^i |\Psi_{\mathrm{T}}^{(\alpha_1,\ldots,\alpha_p)}\rangle \tag{2.28}$$

where

$$O^i(x) = \frac{\partial}{\partial\alpha_i} \ln \Psi_{\mathrm{T}}^{(\alpha_1,\ldots,\alpha_p)}(x) \quad (i > 0) \quad \text{and} \quad O^0 \equiv 1, \delta\alpha_0 = 1\,. \tag{2.29}$$

In order to make $|\Psi'_{\mathrm{T}}\rangle$ as closer as possible to $|\Psi^P_{\mathrm{T}}\rangle$ in the subspace spanned by $\hat{O}^i|\Psi_{\mathrm{T}}\rangle$, we impose the following SR constraints

$$\langle\Psi_{\mathrm{T}}|\hat{O}^k|\Psi'_{\mathrm{T}}\rangle = \langle\Psi_{\mathrm{T}}|\hat{O}^k|\Psi^P_{\mathrm{T}}\rangle \qquad \text{for} \quad k = 0, \ldots, p \tag{2.30}$$

which are equivalent to the following equations

$$\delta\alpha_0 + \sum_{i=1}^{p}\langle\hat{O}^i\rangle = \Lambda - \langle\hat{H}\rangle \tag{2.31}$$

$$\delta\alpha_0\langle\hat{O}^k\rangle + \sum_{i=1}^{p}\langle\hat{O}^k\hat{O}^i\rangle = \Lambda\langle\hat{O}^k\rangle - \langle\hat{O}^k\hat{H}\rangle \quad \text{for} \quad k = 1, \ldots, p\,. \tag{2.32}$$

Since $\delta\alpha_0$ is only a normalization of the wavefunction, by replacing $\delta\alpha_0$ in eqs. (2.32) from eq. (2.31), we obtain a set of linear equations

$$\sum_{i=1}^{p}\delta\alpha_i S_{i,k} = \frac{1}{2}f_k \tag{2.33}$$

where $S_{i,k} = \langle(\hat{O}^i - \langle\hat{O}^i\rangle)(\hat{O}^k - \langle\hat{O}^k\rangle)\rangle$, a covariance matrix, and $f_k = 2(\langle\hat{O}^k\rangle\langle\hat{H}\rangle - \langle\hat{O}^k\hat{H}\rangle)$. By inverting the matrix $\boldsymbol{S}$, the direction along which the parameters should be updated during the optimization reads

$$\delta\boldsymbol{\alpha} = \frac{1}{2}\boldsymbol{S}^{-1}\boldsymbol{f}\,. \tag{2.34}$$

The vector $\boldsymbol{f}$ actually is the generalized force with respect to the parameters $\boldsymbol{\alpha}$

$$f_k \equiv -\frac{\partial E}{\partial\alpha_k} = -\frac{\langle\Psi_{\mathrm{T}}|\hat{O}^k\hat{H} + \hat{H}\hat{O}^k + \frac{\partial\hat{H}}{\partial\alpha_k}|\Psi_{\mathrm{T}}\rangle}{\langle\Psi_{\mathrm{T}}|\Psi_{\mathrm{T}}\rangle} + 2\frac{\langle\Psi_{\mathrm{T}}|\hat{O}^k|\Psi_{\mathrm{T}}\rangle\langle\Psi_{\mathrm{T}}|\hat{H}|\Psi_{\mathrm{T}}\rangle}{\langle\Psi_{\mathrm{T}}|\Psi_{\mathrm{T}}\rangle^2} \tag{2.35}$$

$$= 2(\langle\hat{O}^k\rangle\langle\hat{H}\rangle - \langle\hat{O}^k\hat{H}\rangle) \tag{2.36}$$

since $\hat{H}$ is independent of the parameters $\boldsymbol{\alpha}$. Within each optimization step, the matrix $\boldsymbol{S}$ and vector $\boldsymbol{f}$ are evaluated by standard VMC. At the end of the step, all the parameters are updated as

$$\boldsymbol{\alpha}' = \boldsymbol{\alpha} + \Delta\delta\boldsymbol{\alpha} = \boldsymbol{\alpha} + \frac{1}{2}\Delta\boldsymbol{s}^{-1}\boldsymbol{f}\,. \tag{2.37}$$

Given a sufficient small Δ at each step, this iterative method gains an energy difference

$$-\boldsymbol{f}^{\mathbf{T}}(\boldsymbol{\alpha}' - \boldsymbol{\alpha}) = -\frac{1}{2}\Delta\boldsymbol{f}^{\mathbf{T}}\boldsymbol{S}^{-1}\boldsymbol{f} < 0 \tag{2.38}$$

where the matrix $\boldsymbol{S}$ is positive definite. Compared to a conventional steepest decent algorithm where the parameters are updated as

$$\boldsymbol{\alpha}' = \boldsymbol{\alpha} + \frac{1}{2}\Delta\boldsymbol{f}\,, \tag{2.39}$$

SR takes into account the inverse matrix $\boldsymbol{S}^{-1}$ and yields a more efficient optimization. However, since $\boldsymbol{S}$ is computed in presence of statistical error, it may be an ill-conditioned matrix with very small eigenvalues, which affect the accuracy of its inversion $\boldsymbol{S}^{-1}$. This issue can be solved by regularizing the elements of $\boldsymbol{S}$ on the diagonal, as suggested by ref. [26]

$$S^*_{i,i} = (1+\epsilon)S_{i,i} \tag{2.40}$$

where ϵ is the regularization parameter. If ϵ is small, more accurate estimation of $\boldsymbol{S}$ is required. On the other hand, if ϵ is too large the optimization slows down. A typical value of ϵ we usually chose ranges 10^{-2} to 10^{-5}.

In figure 2.1, the performance of SR shows that the energy goes down very fast in the beginning but then starts to fluctuate due to a too large Δ. In order to continue a efficient optimization, Δ needs to be reduced when the minimum is approaching. This is very inconvenient for users but it can be easily avoided by the Hessian accelerator discussed in the next subsection.

2.4.2 Stochastic reconfiguration with Hessian accelerator

In order to reduce the number of iterations required to achieve convergence to the minimum energy, we have also used the SR with Hessian accelerator method described in ref. [27], which utilizes the information of the second derivatives of the energy with an efficient sampling of the Hessian matrix. This method is quite similar to the one introduced in ref. [28] that has been recently improved in ref. [29].

The details of this method will not be described here but we give an idea how this method works. The energy as a function of the variational parameters is expressed by employing the Taylor expansion up to second order as

$$E = E_0 - \sum_{i=1}^{p} \delta\alpha_i f_i + \frac{1}{2}\sum_{i,j=1}^{p} \delta\alpha_i H_{i,j} \delta\alpha_j \,. \tag{2.41}$$

By imposing $\frac{\partial E}{\partial \alpha_i} = 0$ for all $i = 1, \ldots, p$, the minimum of E can be easily reached by changing all the parameters $\boldsymbol{\alpha}$ by $\delta\boldsymbol{\alpha}$

$$\delta\boldsymbol{\alpha} = \boldsymbol{H}^{-1}\boldsymbol{f} \,. \tag{2.42}$$

With this method, we no longer need to set the amplitude of each step Δ since it has already been determined by the force and Hessian. As a result, in figure 2.1, the curve of SRH never has the fluctuation like SR curve and converges fastest among all methods.

2.4.3 Stochastic reconfiguration with conjugate gradients

When the number of parameters is quite large, the direct application of the above techniques based on the inversion[2] of a large matrix with its leading dimension equal to the number of variational parameters becomes very expensive. It is convenient to implement the scheme — SRH with conjugate gradients — introduced in ref. [26]. This method requires a matrix inversion only in a subspace constructed by the directions of the parameter updates in the most recent n steps.

In the first iteration, the parameters are changed in the direction of $\boldsymbol{g}_1 = \boldsymbol{S}_1^{-1} \cdot \boldsymbol{f}_1$ with $\boldsymbol{S}$ and $\boldsymbol{f}$ given by eqs. (2.40) and (2.36). At this step, the curvature of energy for a small change in the variational parameters $\delta\boldsymbol{\alpha} = \gamma_1 \boldsymbol{g}_1$ can be computed efficiently in the way described by the SRH method. In this way, by assuming the quadratic behavior of the energy, always valid close to the minimum, it is possible to determine the optimal amplitude γ_1 of the parameter change which minimize the energy expectation value. In the following n-th ($n > 1$) iteration, the variation of the parameters is given by

$$\delta\boldsymbol{\alpha} = \sum_{i=1}^{n} \gamma_i \boldsymbol{g}_i \tag{2.43}$$

where $\boldsymbol{g}_i$ are defined analogously as

$$\boldsymbol{g}_i = \boldsymbol{S}_i^{-1} \cdot \boldsymbol{f}_i \tag{2.44}$$

and those $\boldsymbol{g}_i (i < n)$ are saved during the previous iterations. Then at the current iteration, we evaluate the subspace Hessian matrix $\boldsymbol{H}_{\text{sub}}$ corresponding to the second derivatives of the energy with respect to all $\boldsymbol{g}_i (i = 1, \ldots, n)$ directions. With this information, we can easily obtain the optimal set of γ_i which minimizes the energy.

We remark that this subspace technique allows us to change all the variational parameters without any restriction though the number of directions are limited by the number of iterations. In fact, only the most recent directions are important since they are sampled close to each other in a relevant region. In practice, we only keep the latest several directions with the largest signal noise ratio. In our test case (see figure 2.1), the efficiency of this method improves as the number of directions increases. However, the convergence is much slower than SRH due to the fact that in the beginning, for the i-th iteration, the parameters are allowed to change only in these limited i directions.

2.4.4 Stochastic reconfiguration with signal noise filter

It has been shown in the previous subsection that the way to select the dominant directions and evaluate the Hessian matrix only in a relevant subspace

[2]Matrix inversion has a complexity between $O(n^{2.373})$ and $O(n^3)$.

can save computational cost but requires more iterations to converge. In this thesis we show another variant of the SRH method which also constructs a Hessian in the subspace by applying the signal noise filter (SNF) but it has an efficiency as good as the SRH and is also more robust than SRH.

Given a cutoff $\eta > 0$, we collect the parameters with their corresponding generalized forces f_k owning signal noise ratios (SNR)

$$\frac{|f_k|}{\sigma(f_k)} > \eta \,. \tag{2.45}$$

We call these parameters 'significant' ones and construct the Hessian matrix by computing the second derivatives of energy with respect to these parameters only. In this way efficient optimization can be performed within this subspace. However, usually at the beginning of the optimization, almost all of the parameters are significant because they are all far from the values corresponding to the minimum. Therefore we can also impose a restriction on the number of significant parameters to limit the leading dimension of the Hessian matrix subspace.

In the SRH-CG method, the directions computed by $\boldsymbol{S}^{-1}\boldsymbol{f}$ are not actually the optimal ones since they might be large because of a large noise while the SNF directly excludes such possibilities. Figure 2.1 displays that SRH-SNF requires less than 10 iterations to reach the convergence. It is by far the best optimization given its efficiency, robustness and also low computational cost.

Chapter 3

Jastrow correlated antisymmetrized geminal power wavefunction

3.1 Introduction

The wavefunction ansatz is of fundamental importance for the efficiency, reliability and accuracy of both variational and diffusion Monte Carlo calculations. The repeated evaluation of the wavefunction with also its gradient and Laplacian is one of the most consuming part of the computation. The physical observables measured by VMC are directly affected by the properties of the trial wavefunction. Therefore, the optimal choice of such wavefunction should maximize the quality in the description of the physics for the systems of interest. As a projection technique, DMC is less affected by the quality of the wavefunction ansatz, but it is still not free from it. A good wavefunction guides the random sampling in the phase space in an efficient way with the importance sampling technique [9]. In the fermionic systems, the accuracy of the fixed node approximation depends very much on the quality of the nodal surface (defined by $\Psi = 0$) of the guiding wavefunction. Only when the guiding wavefunction coincides with the exact wavefunction, the FNA becomes exact. For the above reasons, we should seek wavefunctions that are both accurate enough and cheap to evaluate.

A typical wavefunction choice consists of an antisymmetric determinant part $\psi_A(\boldsymbol{r})$ and a symmetric part $J(\boldsymbol{r})$ called Jastrow factor. The whole wavefunction is written as their product

$$\Psi_{\mathrm{QMC}}(\boldsymbol{r}) = \Psi_A(\boldsymbol{r}) J(\boldsymbol{r})\,. \tag{3.1}$$

The determinant part can be chosen simply as a product of spin-up and spin-down single Slater determinants [5] or more complex forms like antisymmetrized geminal power (AGP) [30, 25], Pfaffian [31] or linear combination of

multiple determinants. The Jastrow factor mainly takes into account the electron correlation which is missing in the determinant part of the wavefunction and it recovers most part of the correlation energy. The interplay between determinant and Jastrow provides good wavefunctions with favorable accuracy and feasible computational cost.

This chapter is organized as following: sections 3.2–3.4 describe AGP, hybrid orbitals and molecular orbitals used in the determinant part of our wavefunction ansatz. Section 3.5 focuses on the parametrization and physical correspondence of the Jastrow factor employed in our calculation.

3.2 Antisymmetrized geminal power wavefunction

H_2 is a very simple example to show how a chemical bond forms. When two hydrogen atoms approach to a distance where their valence electrons interact with each other, these two electrons form a singlet pairing state. The spin part $\frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle)$ satisfies the antisymmetry and its orbital part is a symmetric function of two valence electron coordinates $f(\boldsymbol{r}, \boldsymbol{r}') = f(\boldsymbol{r}', \boldsymbol{r})$, called *pairing function*. In fact, this wavefunction describes the singlet ground state exactly at all inter-atomic distances, a well known case where the restricted Hartree-Fock solution $f_{\mathrm{HF}}(\boldsymbol{r}, \boldsymbol{r}') = \phi(\boldsymbol{r})\phi(\boldsymbol{r}')$ fails at large distance.

Since the 50's, many theoretical studies explaining superconductivity have highlighted the role of pairing electrons in material science. The BCS wavefunction is an ansatz in which the correlation is introduced through the product of pairing functions. This ansatz had already been exploited in quantum chemistry in the pioneering work of Hurley et al. [32] to treat correlation effects in molecular properties. This particle-conserving version of the BCS ansatz is called antisymmetrized geminal power (AGP).

The AGP wavefunction is constructed with the pairing function f in real space. For an unpolarized system with N electrons in a singlet state, the AGP wavefunction is written as

$$\boldsymbol{\Psi}_{\mathrm{AGP}}(\boldsymbol{r}_1^{\uparrow}, \ldots, \boldsymbol{r}_{N/2}^{\uparrow}, \boldsymbol{r}_1^{\downarrow}, \ldots, \boldsymbol{r}_{N/2}^{\downarrow}) = \mathcal{A}\left[F(\boldsymbol{r}_1^{\uparrow}, \boldsymbol{r}_1^{\downarrow})F(\boldsymbol{r}_2^{\uparrow}, \boldsymbol{r}_2^{\downarrow}) \cdots F(\boldsymbol{r}_{N/2}^{\uparrow}, \boldsymbol{r}_{N/2}^{\downarrow})\right], \tag{3.2}$$

where the operator $\mathcal{A}$ antisymmetrizes the product of geminals in the square bracket and the geminal is a singlet

$$F(\boldsymbol{r}_i^{\uparrow}, \boldsymbol{r}_j^{\downarrow}) = f(\boldsymbol{r}_i^{\uparrow}, \boldsymbol{r}_j^{\downarrow})\frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle)\,. \tag{3.3}$$

Therefore, one can easily write the spatial part of $\boldsymbol{\Psi}_{\mathrm{AGP}}$ in a compact form

$$\Psi_{\mathrm{AGP}}(\boldsymbol{r}_1^{\uparrow}, \ldots, \boldsymbol{r}_{N/2}^{\uparrow}, \boldsymbol{r}_1^{\downarrow}, \ldots, \boldsymbol{r}_{N/2}^{\downarrow}) = \det \boldsymbol{A} \tag{3.4}$$

where $\boldsymbol{A}$ is a $\frac{N}{2} \times \frac{N}{2}$ square matrix defined as

$$\boldsymbol{A}_{ij} = f(\boldsymbol{r}_i^{\uparrow}, \boldsymbol{r}_j^{\downarrow}) \,. \tag{3.5}$$

As commonly done in previous studies, the pairing function is expanded in a basis set of single particle orbitals as

$$f(\boldsymbol{r}_i^{\uparrow}, \boldsymbol{r}_j^{\downarrow}) = \sum_{a,b}^{M} \sum_{p}^{L_a} \sum_{q}^{L_b} \lambda_{\{a,p\},\{b,q\}} \phi_p^a(\boldsymbol{r}_i^{\uparrow}) \phi_q^b(\boldsymbol{r}_j^{\downarrow}) \tag{3.6}$$

where indices $p/q = 1, \ldots, L_{a/b}$ label different atomic orbitals ϕ residing at the position $\boldsymbol{R}_{a/b}$ of nuclei a/b respectively and M is the total number of nuclei. The matrix $\boldsymbol{\lambda}$ has a size $L \times L$, where $L = \sum_a^M L_a$. Its non-zero element $\lambda_{\{a,p\},\{b,q\}}$ represents the strength of a valence bond formed by an orbital p on the atom a and an orbital q on the atom b. Therefore, $\boldsymbol{\lambda}$ describes the resonance of all the valence bonds contained in the system. If there's no breaking of spin symmetry, $\boldsymbol{\lambda}$ is symmetric and this implies that the number of independent parameters is $L(L+1)/2$. Moreover, if the system satisfies other symmetries, the corresponding constraints can be applied to the elements of $\boldsymbol{\lambda}$ for reducing the number of independent parameters.

3.3 Hybrid orbitals

The orbital hybridization is one of the key concepts in valence bond theory. In 1932, Linus Pauling first developed this theory to explain the structure of simple molecules such as methane (CH_4) using atomic orbitals. Pauling pointed out that the reason why CH_4 has four equal strength C-H bonds separated by the same tetrahedral angle is that the one s and three p orbitals of the carbon atom form four equivalent combinations or *hybrid orbitals*. The concept was developed for such simple molecule, but its way of describing the wavefunction is general and it has been widely applied.

In order to obtain reliable results by QMC, a good wavefunction should be not only accurate but also compact with a limited number of variational parameters because all these parameters will be optimized within VMC. One way to achieve a compact wavefunction is to use atomic hybrid orbitals. The idea is very straightforward, namely that the contribution to a valence bond from many atomic orbitals which sit on the same atom can be represented as a single atomic hybrid orbital. Since the number of bonds formed by an atom and its neighbors is limited, only a few hybrid orbitals for each atom are needed for constructing the wavefunction. Hybrid orbitals can be easily extended in terms of a very large basis set at much lower cost, in principle up to the so called complete basis set limit. Therefore, atomic hybrid orbitals are very suitable for constructing compact and accurate wavefunction for QMC.

An atomic hybrid orbital $\bar{\psi}^a$ centered on the nucleus a is expanded as a linear combination of single particle orbitals centered on a

$$\bar{\psi}_{\{a,p\}}(\boldsymbol{r}) = \sum_{l}^{L_a} c_l^p \phi_l^a(\boldsymbol{r})\,. \tag{3.7}$$

We can use several such collective orbitals instead of many individual orbitals to describe valence bonds. By replacing the atomic orbitals in eq. (3.6) with hybrid orbitals, the AGP wavefunction can be constructed with pairing functions expanded as

$$f(\boldsymbol{r}_i^\uparrow, \boldsymbol{r}_j^\downarrow) = \sum_{a,b}^{M} \sum_{p}^{L_a^{\mathrm{hyb}}} \sum_{q}^{L_b^{\mathrm{hyb}}} \lambda_{\{a,p\},\{b,q\}} \bar{\psi}_{\{a,p\}}(\boldsymbol{r}_i^\uparrow) \bar{\psi}_{\{b,q\}}(\boldsymbol{r}_j^\downarrow)\,, \tag{3.8}$$

which contains a much smaller number of parameters to be optimized. The size of $\boldsymbol{\lambda}$ reduces to $L \times L$, where $L = \sum_a^M L_a^{\mathrm{hyb}}$ since the number L_a^{hyb} of hybrid orbitals belonging to the atom a is much smaller than the number L_a of atomic orbitals.

To see how hybrid orbitals help in reducing the number of optimizable parameters, let us compare the number of independent parameters in the AGP wavefunction of liquid water constructed with and without hybrid orbitals. Suppose the average number of atomic orbitals per atom is L_{avg}, the size of $\boldsymbol{\lambda}$ is $ML_{\mathrm{avg}} \times ML_{\mathrm{avg}}$ and the number of independent parameters is $ML_{\mathrm{avg}}(ML_{\mathrm{avg}}+1)/2 \approx (ML_{\mathrm{avg}})^2/2$. Since the liquid has very low symmetry, the only symmetry imposed on the wavefunction is that $\boldsymbol{\lambda}$ is symmetric. When hybrid orbitals are used to construct the pairing function, we suppose the average number of hybrid orbitals per atom is $L_{\mathrm{avg}}^{\mathrm{hyb}}$, the size of $\boldsymbol{\lambda}$ becomes $ML_{\mathrm{avg}}^{\mathrm{hyb}} \times ML_{\mathrm{avg}}^{\mathrm{hyb}}$ and $\boldsymbol{\lambda}$ contains $\approx (ML_{\mathrm{avg}}^{\mathrm{hyb}})^2/2$ independent elements. Even though the hybrid orbitals require $ML_{\mathrm{avg}}^{\mathrm{hyb}} L_{\mathrm{avg}}$ extra parameters, this is only a minor amount compared with those contained by $\boldsymbol{\lambda}$ when M is large. Thus, the total number of parameters is reduced by a factor of $(L_{\mathrm{avg}}^{\mathrm{hyb}}/L_{\mathrm{avg}})^2$.

However, the scaling of the number of parameters with the system size M remains as M^2. Is there a way to further improve this scaling? The answer is yes. The locality approximation (LA) can achieve a linear scaling. The idea is very similar to the 'neighbor list' algorithm used by molecular dynamics that the interaction between two particles is considered only when one is within a range R_{MAX} (distance cutoff) of the other. Analogously, if two atomic or hybrid orbitals reside on two nuclei close enough, the corresponding element in $\boldsymbol{\lambda}$, namely their interaction term, is non-zero and optimized. Otherwise, that element is fixed to zero forever. Suppose each nucleus has at maximum M_{NB} neighbor nuclei within a safe distance cutoff, the number of parameters to be optimized in $\boldsymbol{\lambda}$ is only $\approx MM_{\mathrm{NB}}(L_{\mathrm{avg}}^{\mathrm{hyb}})^2/2$, proportional to M.

In summary, with the help of atomic hybrid orbitals and LA, we are able to obtain a wavefunction ansatz which has a perfect balance between accuracy

and computational cost. We finally remark that expressing the pair function f with atomic hybrid orbitals is not a prerequisite for applying LA. Indeed, LA can be directly applied to $\boldsymbol{\lambda}$ in eq. (3.6) with a basis of atomic orbitals.

3.4 Molecular orbitals

By employing very simple linear algebra, the AGP wavefunction can be viewed as an extension, introduced by M. Marchi et al. [33], of HF wavefunction which is based on *molecular orbitals* (MO). Molecular orbitals are usually constructed by combining 1) atomic orbitals

$$\tilde{\psi}(\boldsymbol{r}) = \sum_{a}^{M} \sum_{l}^{L_a} c_{a,l} \phi_l^a(\boldsymbol{r}) \,, \tag{3.9}$$

or 2) hybrid orbitals from each atom of the molecule,

$$\tilde{\psi}(\boldsymbol{r}) = \sum_{a}^{M} \sum_{p}^{L_a^{\mathrm{hyb}}} c_{a,p} \bar{\psi}_{\{a,p\}}(\boldsymbol{r}) \,, \tag{3.10}$$

or 3) other molecular orbitals from groups of atoms. Because MOs are capable of describing a state extending over all atoms, they have the advantage over atomic (hybrid) orbitals for giving a clear Fermi surface by occupying the number of MOs corresponding to the number of electrons.

In order to rewrite the pairing function into a diagonal form with MO $\tilde{\psi}$ as

$$f(\boldsymbol{r}_i^{\uparrow}, \boldsymbol{r}_j^{\downarrow}) = \sum_{k}^{L} \tilde{\lambda}_k \tilde{\psi}_k(\boldsymbol{r}_i^{\uparrow}) \tilde{\psi}_k(\boldsymbol{r}_j^{\downarrow}) \,, \tag{3.11}$$

we need to solve a generalized eigenvalue problem considering that the atomic orbitals or hybrid orbitals are not necessarily orthogonal to each other. The eigenvalue λ_k represents the weight of the k^{th} MO. If f is expanded with atomic or hybrid orbitals, MOs are constructed with the corresponding type of orbitals given in eqs. (3.9) and (3.10). After the transformation, the pairing function turns into a one particle density matrix. If the summation is truncated at $N/2$, namely $N/2$ MOs, the AGP wavefunction coincides with HF wavefunction. $\tilde{\psi}_k(\boldsymbol{r}_i^{\uparrow})$ and $\tilde{\psi}_k(\boldsymbol{r}_j^{\downarrow})$ are the elements of two matrices $\boldsymbol{D}^{\uparrow}$ and $\boldsymbol{D}^{\downarrow}$ for spin-up and spin-down electrons respectively. Indeed in this case, the AGP wavefunction can be factorized in the conventional form of a spin-up determinant times a spin-down determinant

$$\Psi_{\mathrm{AGP}} = \det \boldsymbol{A} = \det(\boldsymbol{D}^{\uparrow,\mathrm{T}} \boldsymbol{\lambda} \boldsymbol{D}^{\downarrow}) = (\det \boldsymbol{\lambda}) |D^{\uparrow}\rangle |D^{\downarrow}\rangle \tag{3.12}$$

where the determinant of the diagonal matrix $\boldsymbol{\lambda}$ with $\tilde{\lambda}_k$ on its diagonal contributes to the wavefunction only as an irrelevant prefactor. If the MOs for

spin-up and spin-down electrons are identical, our AGP wavefunction represents a restricted Hartree Fock (RHF) solution. On the other hand, if the electrons with opposite spins occupy different sets of MOs, the AGP represents an unrestricted Hartree Fock (UHF) solution. As a matter of fact, many QMC packages implement $|D^{\uparrow}\rangle|D^{\downarrow}\rangle$ as the antisymmetric determinant part of their wavefunction ansatz.

Before discussing the scheme for the optimization of MOs, let's see why MO is not an optimal representation for optimization. Again, we measure the number of independent parameters which should be optimized in our wavefunction. Given the same assumption of L_{avg} and $L_{\text{avg}}^{\text{hyb}}$ that we made in the previous section, each MO requires ML_{avg} parameters with atomic orbitals or $ML_{\text{avg}}^{\text{hyb}}$ with hybrid orbitals. The number of MOs is also proportional to the number of electrons as well as M. Therefore, the total number of free parameters scales as M^2 and we are not allowed to use LA because MOs are global instead of being localized around the nuclei.

As discussed in subsection 3.3, the quadratic scaling of optimizable parameters in the AGP representation turns into a much more favorable linear scaling when LA is employed. For this reason, the VMC optimization in the AGP representation performs better. During the optimization, the full pair function is recast as a combination of AGP and MOs

$$f(\boldsymbol{r}_i^{\uparrow}, \boldsymbol{r}_j^{\downarrow}) = \sum_{k,l}^{L} \lambda_{k,l} \psi_k(\boldsymbol{r}_i^{\uparrow}) \psi_l(\boldsymbol{r}_j^{\downarrow}) \tag{3.13}$$

$$= \sum_{k,l}^{L} \delta\lambda_{k,l} \psi_k(\boldsymbol{r}_i^{\uparrow}) \psi_l(\boldsymbol{r}_j^{\downarrow}) + \sum_{k}^{U} \tilde{\lambda}_k \tilde{\psi}_k(\boldsymbol{r}_i^{\uparrow}) \tilde{\psi}_k(\boldsymbol{r}_j^{\downarrow}) \,, \tag{3.14}$$

where ψ_k can be atomic or hybrid orbitals and the orbital index notation has been simplified by merging atomic index and atomic (hybrid) orbital index and the number U of MOs is the chosen rank of full AGP $\boldsymbol{\lambda}$ matrix and ranges from $N/2$ (Slater determinant) to L (full-rank AGP). The initial MOs can be imported from HF or DFT calculations and then optimized by VMC. For each VMC optimization step, only the parameters in the AGP part are updated from 0 to $\delta\boldsymbol{\lambda}$ in such a way that the rank of $\boldsymbol{\lambda}$ is constrained to be fixed to U within linear order in $\delta\boldsymbol{\lambda}$, while the MO part is frozen. After that, the summation of those two parts is diagonalized again and the new set of MOs is filled. After the diagonalization, in the case of HF like wavefunction shown in figure 3.1, the weights of the MOs below the Fermi surface are redistributed from all values equal to 1 to a wide range in order to gain variational energy while the above ones are not occupied by electrons and thus have weights almost zero.[1] If the determinant part of the wavefunction is desired to be HF like, the elements of $\delta\boldsymbol{\lambda}$ are all reset to zero and all $\tilde{\lambda}_k$ are set 1 before the next QMC optimization step starts. With this approach, the MOs are optimized in a very stable and efficient way.

[1] Not exact zero due to the statistical noise.

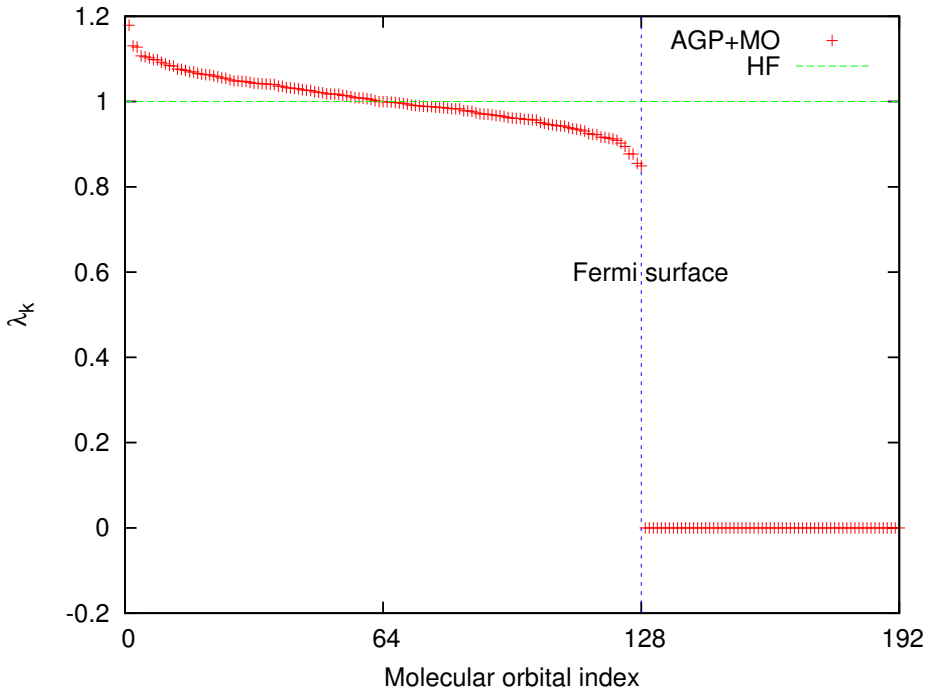

Figure 3.1: $\tilde{\lambda}_k$ associated with the k^{th} MO. This wavefunction for the simulation of 32 water molecules contains 128 MOs.

3.5 Jastrow factor

In QMC calculations the multi-dimensional integral for the estimation of observables (usually energy) is computed statistically. This allows us to use wavefunctions constructed by any types of parametrization. In principle, we could use an extremely complicated wavefunction as close as possible to the ground state but, in practice, the evaluation of such a wavefunction will be computationally prohibitive. As a feasible choice, we have to take a reasonably accurate wavefunction which implies the presence of a sizable statistical noise in the evaluation of energy and its derivatives. According to the zero variance principle, a better wavefunction provides a smaller variance of the energy. Thus, in order to reach a certain accuracy, the amount of statistics can be reduced and save some computation. However, a better wavefunction often costs more on its own evaluation due to a much more complicated parametrization. So the optimal choice for a good wavefunction should reach a balance between its computational cost and accuracy. Fortunately, we know that although the AGP or SD wavefunctions built by Hartree-Fock (HF) or density functional theory (DFT) calculations are usually not accurate enough, they are fairly good as a starting point for building more accurate ones. As is well known, the missing key component of these wavefunction is the electron correlation. Therefore, one of the largest gains in accuracy comes from the introduction of the so-called Jastrow factor which recovers most of the

dynamical electron correlation.

The Jastrow factor is a product of several components expressed as

$$J(\boldsymbol{r}_1^{\sigma_1}, \boldsymbol{r}_2^{\sigma_2}, \ldots, \boldsymbol{r}_N^{\sigma_N}) = J_1 J_2 J_{34} \tag{3.15}$$

where J_1, J_2 and J_{34} are the one-body, two-body and three/four-body Jastrow. With one-body and two-body Jastrow, the wavefunction satisfies nuclear (electron-nucleus) and electron (electron-electron) cusp conditions [34], respectively. Cusp conditions are constraints imposed on the derivatives of the wavefunction. For particle-particle coalescence, it can be shown that

$$\left.\frac{\mathrm{d}\Psi(r)}{\mathrm{d}r}\right|_{r=0} = \xi \Psi(r)\big|_{r=0} \tag{3.16}$$

where Ψ is the many-body wavefunction and r is the particle-particle separation. ξ is $-Z$ for the cusp condition of an electron and a nucleus of charge Z, and $1/4$ and $1/2$ for parallel and anti-parallel spin electrons, respectively. When the wavefunction fulfills the cusp conditions, its local energy doesn't diverge at the coalescent positions, which is always a favorable condition for QMC. The three/four-body Jastrow further improves the description of the correlation between electron-ion pairs including dipole-dipole interaction, namely the van der Waals force. All components of the Jastrow factor are discussed one by one in this chapter.

We've mentioned in section 3.2 that the HF fails for H_2 molecular when two hydrogen atoms stay at large distance. When the Slater determinant obtained from HF is correlated with a Jastrow factor, the dissociation of these two Hydrogen atoms goes to the right limit, as shown in ref. [35].

It should be noted that in this section $\phi_{\{a,q\}}(\boldsymbol{r}_j)$ refers to the atomic Jastrow orbitals and the electron coordinates are labeled as $\boldsymbol{r}_i^{\sigma_i}, i = 1, \ldots, N$ where σ_i is the spin of the electron i.

3.5.1 One-body and two-body Jastrow

The one-body Jastrow is designed to satisfy the nuclear (electron-ion) cusp condition. Even though it's also possible to impose the cusp condition directly on the determinant part of the wavefunction by introducing more involved atomic orbitals, this increases the computational cost and also brings inconvenience during the optimization because certain constraints must be preserved on the orbital coefficients. Instead, by introducing a set of free parameters in the Jastrow, the nuclear cusp condition can be easily satisfied and these parameters can be optimized without constraints by efficient algorithms.

The form of the one-body Jastrow is chosen as

$$J_1(\boldsymbol{r}_1^{\sigma_1}, \boldsymbol{r}_2^{\sigma_2}, \ldots, \boldsymbol{r}_N^{\sigma_N}) = \exp\left[\sum_i^N \sum_a^M \left((2Z_a)^{\frac{3}{4}} u\left((2Z_a)^{\frac{1}{4}} |\boldsymbol{r}_{ia}|\right) + \sum_p^{L_a} \phi_{\{a,p\}}(r_{ia}) \right)\right] \tag{3.17}$$

$$r_{ia} = |\boldsymbol{r}_i - \boldsymbol{R}_a| \tag{3.18}$$

where the function u which is used to satisfy the nuclear cusp conditions is written in the same way of ref. [36] as

$$u(r) = \frac{1}{2b}(1 - \mathrm{e}^{-br}) \tag{3.19}$$

and $\phi_{\{a,p\}}$ is the same set of Jastrow atomic orbitals as those used by three/four-body Jastrow which will be discussed later. Notice that in the one-body Jastrow we have chosen an exponentially converging function because the one body term is supposed to correct the charge density near the nuclei and negative electronic charge is obviously exponentially localized around the positive nuclear charge. For the same reason, we have chosen them as STO or GTO such that they have smooth behavior around nuclei when $\boldsymbol{r}_{ia} \to 0$ and do not affect the mentioned cusp conditions.

The most widely used Jastrow term is the two-body Jastrow which takes into account the many-body effect of interacting electrons and recovers a large portion of the correlation energy. The form of two-body Jastrow reads

$$J_2(\boldsymbol{r}_1^{\sigma_1}, \boldsymbol{r}_2^{\sigma_2}, \ldots, \boldsymbol{r}_N^{\sigma_N}) = \exp\left(\sum_{i<j}^{N} u(\boldsymbol{r}_i^{\sigma_i}, \boldsymbol{r}_j^{\sigma_j})\right), \tag{3.20}$$

where the pair correlation function u is chosen as the Fahy form [37]

$$u(\boldsymbol{r}_i^{\sigma_i}, \boldsymbol{r}_j^{\sigma_j}) = \frac{a_{\sigma_i\sigma_j} r_{ij}}{1 + b_{\sigma_i\sigma_j} r_{ij}}, \tag{3.21}$$

$$r_{ij} = |\boldsymbol{r}_i - \boldsymbol{r}_j|. \tag{3.22}$$

In order to fulfill both electron cusp conditions, a should be $1/4$ for parallel spins and $1/2$ for anti-parallel spins. However, such choice introduces a spin contamination, namely the wavefunction ansatz no longer describes an eigenstate of the spin square operator $\hat{S}^2$ (its efficient evaluation is described in appendix B) as pointed out in ref. [38]. Actually the exchange interaction of electrons has been considered in the determinant part of our wavefunction ansatz, and therefore the probability of coalescence of two electrons with parallel spins is very low due to the Pauli exclusion principle. Therefore, we'd rather avoid spin contamination and fulfill only electron cusp condition for anti-parallel spins by choosing $a = 1/2$ for both cases. As for the other parameters $b_{\sigma_i\sigma_j}$, we fix all of them equal to b which becomes the only optimizable parameter in our two-body Jastrow. The final function u is simplified as $u(\boldsymbol{r}_i^{\sigma_i}, \boldsymbol{r}_j^{\sigma_j}) = \frac{r_{ij}}{2(1+br_{ij})}$.

3.5.2 Three/four-body Jastrow

The three/four- body Jastrow is an inhomogeneous two electron interaction term which has the following form

$$J_{34}(\boldsymbol{r}_1^{\sigma_1}, \boldsymbol{r}_2^{\sigma_2}, \ldots, \boldsymbol{r}_N^{\sigma_N}) = \exp\left(\sum_{i,j}^{N} g(\boldsymbol{r}_i^{\sigma_i}, \boldsymbol{r}_j^{\sigma_j})\right) \tag{3.23}$$

$$\begin{aligned} g(\boldsymbol{r}_i^{\sigma_i}, \boldsymbol{r}_j^{\sigma_j}) = &\sum_{a\neq b}^{M}\sum_{p}^{L_a}\sum_{q}^{L_b} \lambda^{\text{jas}}_{\{a,p\},\{b,q\}} \phi_{\{a,p\}}(r_{ia})\phi_{\{b,q\}}(r_{jb}) \\ &+ \sum_{a}^{M}\sum_{p,q}^{L_a} \lambda^{\text{jas}}_{\{a,p\},\{a,q\}} \phi_{\{a,p\}}(r_{ia})\phi_{\{a,q\}}(r_{ja}) \end{aligned} \tag{3.24}$$

where $\phi_{\{a,p\}}, p = 1, \ldots, L_a$ are Slater or Gaussian types of atomic Jastrow orbitals centered on the nucleus a. Those orbitals are exactly the same ones that appear also in the one-body Jastrow (3.17). The pair correlation function g consists of the four body part for the electron-nucleus-nucleus-electron interaction if $a \neq b$, namely the first term in (3.24), and the three body part for the electron-nucleus-electron interaction if $a = b$, namely the second term in (3.24). The elements of the matrix $\boldsymbol{\lambda}^{\text{jas}}$ connect two electron-nucleus pairs. It should be noticed that the form of the pair function g (see (3.6)) is very similar to the pair function f in the determinant. The symmetry of $\boldsymbol{\lambda}$ can be applied to $\boldsymbol{\lambda}^{\text{jas}}$ analogously.

This choice of three/four-body Jastrow has the following features:

- The presence of these terms doesn't affect the nuclear and electron cusp conditions already fulfilled by the one-body and the two-body Jastrow factors.
- If $\boldsymbol{\lambda}^{\text{jas}}$ is spin independent, the wavefunction remains intact without spin contamination. However, it's also possible to include a spin-sensitive part in the pair correlation function by modifying $\boldsymbol{\lambda}^{\text{jas}}$ as

  $$\boldsymbol{\lambda}^{\text{jas}} = \boldsymbol{\lambda}^{\text{jas}}_{\text{charge}} + \boldsymbol{\lambda}^{\text{jas}}_{\text{spin}} \sigma_i \sigma_j \,, \tag{3.25}$$

  in order to gain variational energy.
- When atoms fall apart, it factorizes into a product of individual contributions located on each atom and, therefore, meets the requirement of the size consistency.

Now let's see an interesting example, the Beryllium dimer. Beryllium atom has two electrons in both 1s and 2s orbitals, namely completely closed shells. So the weak interaction between two atoms leads to a very shallow binding. The main contribution of the binding energy is given by the so called van der Waals interaction. For this reason, the Beryllium dimer is a very good

method	equilibrium distance [Å]	binding energy [mHa]
VMC	2.54	4.05(4)
DMC [40]	2.46	2.82(3)
DMRG [41]	2.45	4.30(4)
i-FCIQMC [41]	2.45	4.21(4)
Exp. [42]	2.44	4.24
Exp. [43]	2.44	4.26

Table 3.1: The equilibrium distance and binding energy of Be_2.

test case for benchmarking both methods for electronic structure calculations and experiments. In the theoretical calculation, many methods like HF and DFT fail to provide a good binding and the correct dissociation limit. DFT significantly overbinds Be_2 by a factor 3 and the total energy of the molecule is higher than two individual atoms at large inter-atomic distance. As for experiments, the correct measurement of the Be_2 binding was only achieved in 2009, very recently. The old experiments gave very large equilibrium distance. So we challenge this interesting system with our VMC calculations.

In the full CI calculation [39], one billion of determinants was used with a (9s,2p,1d) basis set. We use a much cheap but still accurate wavefunction to recover the full binding energy of two Be atoms. It only requires one determinant built from (5s,2p,1d) atomic orbitals for each atom and 4 MOs for 8 electrons plus a full Jastrow factor with its three/four body terms expanded in the (4s,4p,2d) basis set. Figure 3.2 clearly shows that two Be atoms dissociate into the correct limit — twice the energy of Be atom — at large distance. More interestingly, if the four-body part of the Jastrow is removed from the wavefunction, the curve shows similar defects as DFT calculation. It means the four-body Jastrow plays a key role in describing van der Waals interactions for weakly bonded systems.

Our VMC calculation of Be_2 gives a molecular binding at 2.54(4) Bohr with a binding energy 4.05(4) mHa which are very close to the values obtained by DMRG and *i*-FCIQMC and also the experiments listed in table 3.1.

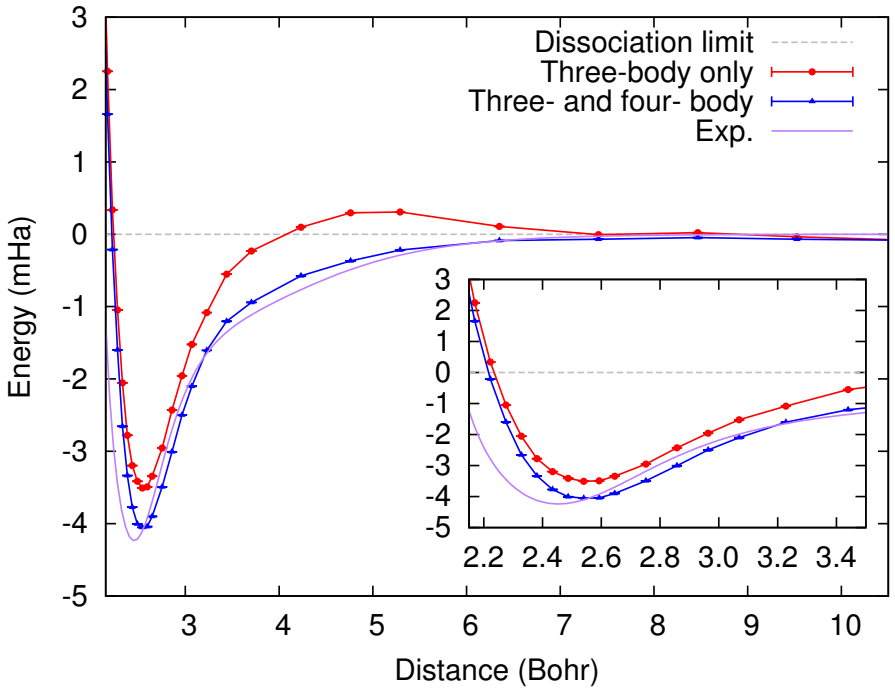

Figure 3.2: The binding curve of the Be_2 molecule calculated under a wave-function ansatz with and without four-body Jastrow factor.

Chapter 4

Molecular dynamics with quantum Monte Carlo

4.1 Introduction

In the last decade, much progress has been made for the simulation of electronic systems by quantum Monte Carlo (QMC), namely by a fully ab-initio approach aimed to solve in a stochastic way the Schrödinger equation, with an appropriate and consistent description of the electron correlation. Only a few years ago, a very general and robust method of optimization was introduced by C. Umrigar et al. [29], that has made it possible to determine by QMC a variational wavefunction containing up to several thousands of parameters [44, 26]. This progress is particularly remarkable, as the variational Monte Carlo (VMC) method was introduced in the early 60's [7] and, until a few years ago, only a few tens of parameters were optimized within the VMC approach. Another recent and important development in QMC was the solution [45, 46, 47] of the infinite variance problem occurring in the straightforward calculation of nuclear forces in the simplest variational Monte Carlo scheme [45, 47]. Moreover, thanks to the algorithmic differentiation [48, 49], the cost of computing all the force components in a system containing several atoms, can be afforded with a computational time at most a factor four larger than the one corresponding to the energy. This progress has led to several works, where structural optimization and highly accurate evaluations of the equilibrium configurations as well as related properties were possible even for quite large systems containing several atoms [50, 51, 52, 53, 54, 55].

Despite this remarkable progress, we notice that ab-initio molecular dynamics (MD) simulation based on quantum Monte Carlo remains so far at a very early stage, as only a few simulations on liquid hydrogen [56, 47, 35, 57] are known. Instead, within the DFT community, MD simulations in the Born-Oppenheimer (BO) approximation, are quite well established, due to almost three decades of achievements from the pioneering work of Roberto Car and Michele Parrinello [58]. Indeed, DFT-based MD simulations are

routinely used to study several properties of condensed matter systems at ambient conditions up to extremely high pressures and temperatures [59, 60, 61, 62, 63, 64, 65, 66, 67], and represents nowadays a quite reliable tool to predict new materials, sometimes more effective or at least much cheaper than experiments.

The application of quantum Monte Carlo for ab-initio simulation of materials or large chemical compounds remains difficult not only because of the heavy computational cost, but also, in our opinion, due to the theoretical difficulties in applying the Newton's equations of motion when the forces are given with a statistical uncertainty. For instance, the basic law of energy conservation cannot be met at all, when the forces are not exactly given at each step. In this context, it is worth mentioning that Ceperley and Dewing have introduced the penalty method [68] that does not rely on any dynamics, and therefore is not affected by this problem. In their method the canonical distribution is directly sampled without using forces, while the statistical uncertainty in the knowledge of the energy is compensated by rejecting the proposed moves more frequently than in the standard Metropolis algorithm. Unfortunately this method is very expensive, especially in the low temperature regime, because of too many rejected moves, and so far applications have been limited to hydrogen with up to 54 protons in this regime [69, 70, 71].

Generally speaking it is clear that, when the computational cost for the calculation of the nuclear forces is comparable to that of the energy, MD should be more efficient, because with the same cost all the atoms are moved in a statistically relevant region of the phase space, without any rejection. For instance in DFT, where the forces are obtained almost for free by applying the Hellmann-Feynman theorem, MD is a common practice to sample the canonical distribution, and, to our knowledge, only hybrid methods based on Monte Carlo and MD [72] can be competitive.

In this chapter, the method of computing forces within QMC and the improvements on its technical implementation will be discussed in section 4.2. In the following section 4.3, we will show how to reduce the infinite variance of the Pulay forces. In the section 4.4, the new type of molecular dynamics scheme which is driven by forces evaluated by QMC will be discussed. In the last section 4.5, the features of the covariance matrix are discussed and the advantage of its use is displayed.

Our ab-initio MD simulations are performed via variational quantum Monte Carlo (VMC) by employing the `TurboRVB` QMC package [73].

4.2 Force evaluation with quantum Monte Carlo

Unlike in DFT, obtaining forces with QMC is by far more difficult. By assuming the Born-Oppenheimer approximation, the nuclei are treated as classical

particles and the $3M$-dimensional force acting on all the M atoms is defined as

$$\boldsymbol{F} \equiv -\nabla_{\boldsymbol{R}} E_{\mathrm{V}}[\Psi]\,, \tag{4.1}$$

where $\nabla_{\boldsymbol{R}}$ is the gradient relative to the Cartesian coordinates $\boldsymbol{R}$ of all the nuclei and $E_{\mathrm{V}}[\Psi]$ is the variational energy associated to the electronic wavefunction Ψ corresponding to $\boldsymbol{R}$. As a functional of Ψ, $E_{\mathrm{V}}[\Psi]$ depends on $\boldsymbol{R}$ through the Hamiltonian $\hat{H}$ and the wavefunction Ψ which has an implicit dependence in the parameter set $\boldsymbol{\alpha}$ that has to be optimized variationally by minimizing the variational energy for the given $\boldsymbol{R}$, and an explicit dependence if ψ is defined in the localized basis set, as our case. Therefore, the local energy e_{L} which appears in the evaluation of $E_{\mathrm{V}}[\Psi]$ also depends on $\boldsymbol{R}$ through both the Hamiltonian and the wavefunction.

Eq. (4.1) can be factorized in the following analytic expression

$$\boldsymbol{F} = \boldsymbol{F}^{\mathrm{HF}} + \boldsymbol{F}^{\mathrm{Pulay}} + F^{\boldsymbol{\alpha}} \tag{4.2}$$

$$\boldsymbol{F}^{\mathrm{HF}} = -\frac{\langle\Psi|\frac{\partial}{\partial\boldsymbol{R}}\hat{H}|\Psi\rangle}{\langle\Psi|\Psi\rangle} \tag{4.3}$$

$$\boldsymbol{F}^{\mathrm{Pulay}} = -2\frac{\langle\Psi|\hat{O}_{\boldsymbol{R}}\hat{H}|\Psi\rangle - \langle\Psi|\hat{O}_{\boldsymbol{R}}|\Psi\rangle E_{\mathrm{V}}}{\langle\Psi|\Psi\rangle} \tag{4.4}$$

$$\boldsymbol{F}^{\boldsymbol{\alpha}} = -\frac{\partial E_{\mathrm{V}}}{\partial\boldsymbol{\alpha}}\frac{\partial\boldsymbol{\alpha}}{\partial\boldsymbol{R}} \tag{4.5}$$

where the three components $\boldsymbol{F}^{\mathrm{HF}}$, $\boldsymbol{F}^{\mathrm{Pulay}}$ and $\boldsymbol{F}^{\boldsymbol{\alpha}}$ are given respectively by the explicit dependence on $\boldsymbol{R}$ of the Hamiltonian and the explicit and implicit dependence of the wavefunction.

In principle, the term $\boldsymbol{F}^{\boldsymbol{\alpha}}$ is the most complicated because the derivatives $\frac{\partial\boldsymbol{\alpha}}{\partial\boldsymbol{R}}$ are very difficult to evaluate. Fortunately, when the parameter set $\boldsymbol{\alpha}$ is optimized to reach the minimum of E_{V}, $\frac{\partial E_{\mathrm{V}}}{\partial\boldsymbol{\alpha}} = 0$ for each component and $\boldsymbol{F}^{\boldsymbol{\alpha}}$ is exactly zero. Therefore, we can safely ignore $\boldsymbol{F}^{\boldsymbol{\alpha}}$ in our calculations. The other two terms $\boldsymbol{F}^{\mathrm{HF}}$, $\boldsymbol{F}^{\mathrm{Pulay}}$ are referred as the Hellmann-Feynman term and the Pulay term. The Hellmann-Feynman term resembles the force computed by applying the Hellmann-Feynman theorem. Actually, in VMC calculations, the Hellmann-Feynman theorem is not applicable because the wavefunction is neither an eigenstate of $\hat{H}$ nor normalized. Only when the wavefunction approaches an eigenstate of $\hat{H}$, the Pulay term becomes zero and the Hellmann-Feynman term converges to the exact Hellmann-Feynman force.

We can also consider the Cartesian coordinates $\boldsymbol{R}$ as optimizable parameters and the forces as energy derivatives with respect to them. By joining them with the parameters of the wavefunction we can achieve the structural and wavefunction optimization at the same time.

In practice, the forces can be computed with the finite difference method. The force component $p = 3(a-1)+i$ acting on atom a in the direction of

$i \in \{1, 2, 3\}$ which represents x, y or z, is computed as

$$F_p = -\frac{E_{\mathrm{V}}(\boldsymbol{R} + \vec{i}_p\Delta) - E_{\mathrm{V}}(\boldsymbol{R})}{\Delta} + \mathcal{O}(\Delta) \tag{4.6}$$

where $\vec{i}_p$ is the p^{th} unit vector in the $3M$-dimensional coordinate space and Δ is the displacement of the atom considered. Δ is chosen small enough such that the finite difference error is negligible. However, the straightforward application of any finite difference method is very inefficient. When the energy values of two configurations are obtained by independent measurements, the error of the energy difference remains similar to the error of the energy due to the propagation of errors. Thus, the error in the energy derivatives diverges as $\frac{1}{\Delta}$ as $\Delta \to 0$. This issue has been solved by the introduction of two technical improvements.

The first one is the correlated sampling (CS) [74] which allows the computation of the energy derivatives with errors much smaller than those obtained in the straightforward way. By employing the same Markov chain for two sets of nuclear coordinates R[1] within a very small distance (Δ in our way of computing forces), the statistical error in the energy difference goes to zero as $\Delta \to 0$ and the corresponding error in the energy derivative remains finite.

The other improvement is the space warp coordinate transformation (SWCT) [75]. We will not explain this method in detail and the interested readers can refer to ref. [74]. The CS plus SWCT implementation was first introduced for structural optimization [25]. In this thesis we just remark that the net force felt by an isolated molecule should be exactly zero. In practice, only after the SWCT the estimator of the net force gives zero variance, namely the translation invariance is fulfilled.

In the limit $\Delta \to 0$, the finite difference forces computed by the correlated sampling converge to the analytic forces. However, the analytic differentiation of forces also has intrinsic infinite variance. In the Hellmann-Feynman term, the derivative $\frac{\partial}{\partial \boldsymbol{R}}\hat{H}$ diverges when the electrons are very close to the nuclei and also when they are close to the nodal surface defined by $\Psi_{\mathrm{T}} = 0$. Meanwhile, the Pulay term diverges when a configuration approaches the nodal surface. To solve these issues, large improvements have been done using the reweighting methods for the stochastic sampling [45, 46, 47, 50]. Ref. [45] solves the infinite variance problem of $\boldsymbol{F}^{\mathrm{HF}}$ contribution due to very close electron-ion distance while ref. [47] and ref. [50] solve the infinite variance problem at the nodal surface of both $\boldsymbol{F}^{\mathrm{HF}}$ and $\boldsymbol{F}^{\mathrm{Pulay}}$ contributions for periodic boundary systems and open boundary systems. In the next section, the solution to the infinite variance on the $\boldsymbol{F}^{\mathrm{Pulay}}$ term for both boundary conditions will be discussed.

A further step for efficient and accurate QMC forces has been recently introduced by S. Sorella and L. Capriotti [48]. In the finite difference way, $3M$

[1]We can also measure the energy difference of two consecutive sets of wavefunction parameters by correlated sampling during the wavefunction optimization.

energy evaluations are required to compute all the forces by displacing each ionic coordinate individually. They introduced the algorithmic differentiation (AD) which is capable of computing all the components of the ionic forces in a computational time only four times larger than the time of a VMC energy calculation.

By employing all the techniques above, we are able to realize the calculation of forces at an affordable computational cost with increasing system size.

4.3 Solution of the infinite variance problem for the forces

As mentioned in the introduction section of this chapter, the reweighting method was introduced to solve the analytic infinite variance problem on the evaluation of forces. In VMC, a generic operator $\hat{O}$ is evaluated statistically by averaging samples distributed according to a probability $\pi(\boldsymbol{x}) = \Psi_{\mathrm{T}}^2(\boldsymbol{x})$ as

$$\langle \hat{O} \rangle = \langle O_{\mathrm{L}} \rangle_{\pi(\boldsymbol{x})} \,. \tag{4.7}$$

Instead, we can use another arbitrary distribution $W(\boldsymbol{x})$ to accumulate samples and all the samples are reweighted by a factor $\frac{\pi(\boldsymbol{x})}{W(\boldsymbol{x})}$ in order to recover the correct evaluation of $\hat{O}$ as

$$\langle \hat{O} \rangle = \left\langle O_{\mathrm{L}} \frac{\pi(\boldsymbol{x})}{W(\boldsymbol{x})} \right\rangle_{W(\boldsymbol{x})} \Bigg/ \left\langle \frac{\pi(\boldsymbol{x})}{W(\boldsymbol{x})} \right\rangle_{W(\boldsymbol{x})} . \tag{4.8}$$

This method is called reweighting method and how good $W(\boldsymbol{x})$ is chosen determines the efficiency of the sampling. C. Attaccalite et al. [47] employed this reweighting method to solve the infinite variance issue in the proximity of the nodal surface by using a different probability distribution $P(\boldsymbol{x}) \propto W(\boldsymbol{x}) = \Psi_{\mathrm{G}}(\boldsymbol{x})^2$, defined in terms of a guiding function $\Psi_{\mathrm{G}}(\boldsymbol{x})$, rather than the standard sampling $\pi(\boldsymbol{x}) = \Psi_{\mathrm{T}}(\boldsymbol{x})^2$.

The guiding function $\Psi_{\mathrm{G}}(\boldsymbol{x})$ is defined in terms of the wavefunction $\Psi_{\mathrm{T}}(\boldsymbol{x})$ as follows:

$$\Psi_{\mathrm{G}}(\boldsymbol{x}) = \frac{R^{\epsilon}(\boldsymbol{x})}{R(\boldsymbol{x})} \Psi_{\mathrm{T}}(\boldsymbol{x}) \tag{4.9}$$

where $R(\boldsymbol{x})$ is a 'measure' of the distance δ from the nodal surface and is assumed to work for $\delta \ll 1$, where $\Psi_{\mathrm{T}}(\boldsymbol{x}) \propto R(\boldsymbol{x})$. The $R^{\epsilon}(\boldsymbol{x})$ is the function that regularizes Ψ_{T} in the vicinity of the nodal surface, namely for $\delta \propto R(\boldsymbol{x}) < \epsilon$, and it is defined as:

$$R^{\epsilon}(\boldsymbol{x}) = \begin{cases} R(\boldsymbol{x}) & \text{if } R(\boldsymbol{x}) \geq \epsilon \,, \\ \epsilon [R(\boldsymbol{x})/\epsilon]^{R(\boldsymbol{x})/\epsilon} & \text{if } R(\boldsymbol{x}) < \epsilon \,, \end{cases} \tag{4.10}$$

where the nontrivial regularization for $R(\boldsymbol{x}) < \epsilon$ is introduced in order to satisfy the continuity of the first derivative of $\Psi_{\mathrm{G}}(\boldsymbol{x})$. The guiding function $\Psi_{\mathrm{G}}(\boldsymbol{x})$ defined in this way and its corresponding probability density function $P(\boldsymbol{x}) \propto \Psi_{\mathrm{G}}(\boldsymbol{x})^2$ define a *reweighting factor*

$$\left(\frac{\Psi_{\mathrm{T}}(\boldsymbol{x})}{\Psi_{\mathrm{G}}(\boldsymbol{x})}\right)^2 = \left(\frac{R(\boldsymbol{x})}{R^{\epsilon}(\boldsymbol{x})}\right)^2 = \min\left[1, \left(\frac{R(\boldsymbol{x})}{\epsilon}\right)^{2\left(1-\frac{R(\boldsymbol{x})}{\epsilon}\right)}\right] \tag{4.11}$$

that in the proximity of the nodal surface, i.e. $R(\boldsymbol{x}) \to 0$, is $\propto \delta^2$, whereas the probability density function $P(\boldsymbol{x}) \propto \epsilon^2 \left(\frac{\Psi_{\mathrm{T}}(\boldsymbol{x})}{R(\boldsymbol{x})}\right)^2 \propto \epsilon^2$ remains constant but finite. This $P(\boldsymbol{x})$ slightly enhances the sampling in the vicinity of the nodal surface where $\pi(\boldsymbol{x})$ vanishes. So far, our reweighting method removes the singularities up to δ^{-2} and provides finite variance.

The regularization scheme which [47] proposed to evaluate $R(\boldsymbol{x})$ is based on the matrix $\boldsymbol{A}$ (defined as (3.5)) that appears in the determinant (anti-symmetric) part of the QMC wavefunction. As soon as the configuration of electrons approaches the nodal surface, $\det(\boldsymbol{A}) \to 0$ and the elements of the $\boldsymbol{A}^{-1}$ become extremely large. The reason why $\det(\boldsymbol{A})$ is not chosen as $R(\boldsymbol{x})$ is that $\det(\boldsymbol{A})$ scales exponentially with the dimension N of the matrix $\boldsymbol{A}$ (associated with the system size) because $\det(\boldsymbol{A}) = \Pi_{i=1}^{N}\lambda_i$ where λ_i is the i^{th} eigenvalue of the matrix $\boldsymbol{A}$. Instead, the summation of all the elements of the $\boldsymbol{A}^{-1}$

$$\sum_{i,j}^{N} {A^{-1}}_{ij} \approx \sum_{i=1}^{N} \lambda_i^{-1} \tag{4.12}$$

has a polynomial scaling. According to this feature, the 'measure' function $R(\boldsymbol{x})$ for the regularization is chosen to be controlled by ${A^{-1}}_{ij}$ in the following way:

$$R(\boldsymbol{x}) = \left(\sum_{i,j} \left|{A^{-1}}_{ij}\right|^2\right)^{-1/2}. \tag{4.13}$$

However, this scheme (4.13) does not take into account the case of open systems like isolated atoms and molecules (the singularity type 4 in ref. [76]). As an electron i samples a region very far from the center of mass $\boldsymbol{R}_0$ of the nuclei, namely $r_{i0} = \|\boldsymbol{r}_i - \boldsymbol{R}_0\| \gg 1$, the decay of the many-body wavefunction is dominated by the determinantal part as the Jastrow correlation is identically one in this limit. A simple inspection shows that $\det(\boldsymbol{A})$ behaves as $\propto \exp(-\tilde{z}_{\min} r_{i0})$ [$\propto \exp(-\tilde{z}_{\min} r_{i0}^2)$], where $\tilde{z}_{\min}$ is the minimum exponent in the Slater [Gaussian] basis. The $R(\boldsymbol{x})$ choice in the old regularization scheme defined by eq. (4.13) vanishes clearly in the same way. In order to show this property, it is enough to apply the Rouché-Capelli theorem stating that the inverse matrix elements A_{ij}^{-1} can be expressed with the ratio of the adjugate

matrix element $\mathrm{adj}(\boldsymbol{A})_{ij}$ and the determinant itself, namely:

$$A^{-1}{}_{ij} = \frac{\mathrm{adj}(\boldsymbol{A})_{ij}}{\det(\boldsymbol{A})}.$$

Now we immediately arrive to the bad conclusion that the probability distribution $P(\boldsymbol{x})$ is ill defined as it converges to a constant in the limit when $r_{i0} \gg 1$, because, $R(\boldsymbol{x}) \to 0$ in the same way as $\Psi_{\mathrm{T}}(\boldsymbol{x}) \to 0$ (as discussed above), and the resulting distribution $P(\boldsymbol{x})$ is not normalizable. In practice this means that the random walk for long enough simulation will be unstable, and all electrons are pushed to very large distance from the atoms, providing unpredictable and certainly biased results.

In order to overcome this clear instability we replace the $\boldsymbol{A}$ in (4.13) with $\boldsymbol{A}'$. The new matrix $\boldsymbol{A}'$ is defined by changing its asymptotic behavior for large r_{i0}:

$$A'_{ij} = A_{ij} \exp(z r_{i0} + z r_{j0}) \,, \tag{4.14}$$

where z can be any positive value. In fact the new regularization will act in the same way close to the nodes of Ψ_{T}, whereas when $r_{i0} \gg 1$, $\det(\boldsymbol{A}')$ decays as $\exp[-\tilde{z}_{\min} r_{i0}^2 + z r_{i0}]$ for a Gaussian basis, and for a Slater basis, if $\tilde{z}_{\min} > z$, it decays as $\exp[-(\tilde{z}_{\min} - z) r_{i0}]$, and diverges otherwise. Therefore $P(\boldsymbol{x})$, by using this new definition of $R(\boldsymbol{x})$, will decay as $\exp(-2 z r_{i0})$ in the former cases, or as Ψ_{T}^2 itself in the latter case, yielding *in any case* a perfectly defined and normalizable distribution.

In practice, if z is too small, $\boldsymbol{A}'$ behaves too much like $\boldsymbol{A}$ and the instability remains. On the other hand if z is too large, the probability distribution $P(x)$, as we have seen, remains too close to the original one $\simeq \Psi_{\mathrm{T}}^2$ for electron-ion distances $\gg 1/z$, and therefore in this region the singularities in the nodal surfaces remain, and the regularization is not effective also in this case. Therefore, with this simple trick, and a reasonable value of $z \simeq 1/\xi$, where ξ is the linear dimension of the important region of non vanishing charge density, this numerical instability, present in open systems, is readily removed, and the singularities around the nodal surfaces are perfectly controlled, because the proposed regularization works exactly as the previous one [47] adopted for PBC. Indeed, if electrons are close to this nodal surface $\det(\boldsymbol{A}) = 0$ and r_{i0} are all finite, the following equality

$$\det(\boldsymbol{A}') = \det(\boldsymbol{A}) \prod_i^N \exp(z r_{i0}) \tag{4.15}$$

implies that the new regularization works as well as the previous one, being the factor $\prod_i^N \exp(z r_{i0})$ just an irrelevant term.

With this regularization scheme based on a new choice of $\boldsymbol{A}'$ for determining the reweighting factor, the infinite variance problem on the ionic forces for open boundary systems is also solved in a simple and efficient way. Other schemes are also possible like the one introduced by J.R. Trail [76] which is

very inefficient from the computational aspect because it involves the calculation of the local energy at each Metropolis step. In the following calculations, the regularization choice of ref. [47] is used for the simulation of liquid water and the new regularization is employed for the calculation of variational frequencies of small molecules.

4.4 Molecular dynamics with noisy forces

A second order Langevin dynamics (SLD) is used in the sampling of the ionic configurations within ground state Born-Oppenheimer approach. Ionic forces are computed with finite and small variance by algorithmic differentiation [48], which allows feasible simulations of a large number of atoms. Moreover the statistical noise, corresponding to the forces, is used to drive the dynamics at finite temperature by means of an appropriate generalized Langevin dynamics [47]. A similar approach has been proposed in ref. [77] and ref. [78] where a SLD algorithm has been devised also at the DFT level. In this thesis we adopt a different numerical integration scheme for the SLD which allows us to use large time steps, even in presence of large friction matrices. For reason of clarity and completeness, we present in this section the method introduced in the original paper of C. Attaccalite and S. Sorella [47], with more details in the derivations, whereas the more advanced techniques, that can be straightforwardly derived following the same analysis, are described in the Appendix A.2.

Let us consider solving the set of differential equations of the SLD:

$$\begin{aligned} \dot{\boldsymbol{v}} &= -\boldsymbol{\gamma}(\boldsymbol{R}) \cdot \boldsymbol{v} + \boldsymbol{f}(\boldsymbol{R}) + \boldsymbol{\eta}(t) && (4.16) \\ \dot{\boldsymbol{R}} &= \boldsymbol{v} && (4.17) \\ \langle \boldsymbol{\eta}(t) \rangle &= 0 \\ \langle \eta_i(t)\eta_j(t') \rangle &= \alpha_{ij}(\boldsymbol{R})\, \delta(t-t') && (4.18) \end{aligned}$$

where $\boldsymbol{R}, \boldsymbol{v}, \boldsymbol{f}, \boldsymbol{\eta}$ are the $3N$-dimensional vectors made by the positions, the velocities, the deterministic and the stochastic forces of the N nuclei, respectively, and the indices i, j run over all the $3N$ nuclear coordinates. The symbol $\langle \cdots \rangle$ indicates the average over the ensemble of possible realizations, and it is used to define properties of the stochastic force $\boldsymbol{\eta}$, which are determined by the fluctuation-dissipation theorem, namely its instantaneous correlation $\boldsymbol{\alpha}$ is given by:

$$\boldsymbol{\alpha}(\boldsymbol{R}) = 2T\, \boldsymbol{\gamma}(\boldsymbol{R}) \tag{4.19}$$

where $T \equiv 1/\beta$ is the temperature[2] and both $\boldsymbol{\gamma}(\boldsymbol{R})$ and $\boldsymbol{\alpha}(\boldsymbol{R})$ are $3N$-dimensional square matrices, implicitly depending on the atomic positions.

Notice that in the above equations we have assumed that all the masses of the particles are set to unit values in atomic Rydberg units, namely twice the

[2] In this section the Boltzmann constant k is conventionally set to one for simplicity.

electronic mass $2m_e$ is one in our conventions. In the following we will always use unit masses, because, in order to sample the canonical distribution the actual values of the masses are immaterial. In order to match the usual atomic units, for instance in the hydrogen case already studied in Refs. [47, 35], the time units have to be scaled by the square root of the ratio between the proton mass and twice the electron mass ($\sqrt{m_p/2m_e} \sim 30.3$). In a polyatomic molecule — like water — the inverse mass of each different atom multiples the force components $\boldsymbol{f}(\boldsymbol{R})$ in the commonly adopted Langevin equations. However, also in this case, it is possible to reduce back to the case studied, by a further appropriate scaling of the length of each particle (distinguishable in classical dynamics). Thus our formulation is quite general up to an appropriate scaling of time and lengths,[3] and therefore can be also used to study the physical Newtonian dynamics — e.g. necessary to compute the diffusion constant in liquid water — with $\boldsymbol{\gamma} \to 0$ and physical masses.

In eq. (4.19) one of the two matrices is arbitrary, and we can choose:

$$\boldsymbol{\alpha}(\boldsymbol{R}) = \alpha_0 \mathbf{I} + \Delta_0 \boldsymbol{\alpha}^{\mathrm{QMC}}(\boldsymbol{R}) \tag{4.20}$$

$$\boldsymbol{\gamma}(\boldsymbol{R}) = \frac{\boldsymbol{\alpha}(\boldsymbol{R})}{2T} \tag{4.21}$$

where $\mathbf{I}$ is the identity matrix, α_0 and Δ_0 are two constants that should be suitably defined in order to minimize the autocorrelation time and therefore the efficiency of the sampling, and the $3N$-dimensional matrix $\boldsymbol{\alpha}^{\mathrm{QMC}}(\boldsymbol{R})$ is the variance-covariance matrix of the nuclear forces $\boldsymbol{f}(\boldsymbol{R})$ evaluated by QMC at the nuclear configuration $\boldsymbol{R}$, and it is defined as:

$$\alpha_{ij}^{\mathrm{QMC}}(\boldsymbol{R}) = \langle (f_i(\boldsymbol{R}) - \langle f_i(\boldsymbol{R}) \rangle)\,(f_j(\boldsymbol{R}) - \langle f_j(\boldsymbol{R}) \rangle) \rangle \tag{4.22}$$

where $\langle \cdots \rangle$ refers to the average over the QMC sampling. In practice, $\boldsymbol{\alpha}^{\mathrm{QMC}}$ is computed as $\boldsymbol{C}_{\boldsymbol{s}}(\boldsymbol{\delta})$ in eq. (5.21) (see Section 5.2.3 for more details).

We now assume only that in the time interval

$$t_n - \tau/2 < t < t_n + \tau/2,$$

n indexing the time steps $t_n = n \times \tau$, the positions $\boldsymbol{R}$ are changing very little and, within a good approximation, we can neglect the $\boldsymbol{R}$ dependence in the RHS of eq. (4.16), and indicate $R(t_n) = R_n$. The second equation (4.17) can be integrated easily once the value of a velocity is known at a given time:

$$\boldsymbol{R}_{n+1} - \boldsymbol{R}_n \simeq \tau \boldsymbol{v}(t) \tag{4.23}$$

[3] In an electronic system containing atoms with different masses, we can scale each length x_i corresponding to a mass M_i by $x_i = \sqrt{\frac{1}{M_i}} x_i'$. After a little algebra one obtains the Langevin equation with unit masses given in eq. (4.17) and a scaled friction matrix $\gamma'_{i,j} = \gamma_{i,j}/\sqrt{M_i M_j}$ with the fluctuation dissipation theorem written in the same form as in eq. (4.19).

where $t_n \leq t \leq t_{n+1}$. A better way to integrate the equation is given in the Appendix A.2. For the time being we assume the above simple form, and for a better accuracy it is useful to consider that the velocities $\boldsymbol{v}_n$ are computed at half-integer times $t_n - \tau/2$:

$$\boldsymbol{v}_n \equiv \boldsymbol{v}(t_n - \tau/2) \tag{4.24}$$

and the quantities that are functions of $\boldsymbol{R}$ in eq. (4.16) are calculated in $\boldsymbol{R}_n$:

$$\boldsymbol{f}_n \equiv \boldsymbol{f}(\boldsymbol{R}_n) \tag{4.25}$$
$$\boldsymbol{\gamma}_n \equiv \boldsymbol{\gamma}(\boldsymbol{R}_n). \tag{4.26}$$

In this small time integration interval once the values of $f(R) = \boldsymbol{f}_n$ and $\gamma(R) = \boldsymbol{\gamma}_n$ are assumed constant, the solution to eq. (4.16) is given in a closed form in eq. (A.7) of the Appendix, with the initial time $\bar{t}$ and the final one t arbitrary:

$$\begin{aligned} \boldsymbol{v}(t) = & \exp\left[\boldsymbol{\gamma}_n(\bar{t} - t)\right] \boldsymbol{v}_n \\ & + \int_{\bar{t}}^{t} \exp\left[\boldsymbol{\gamma}_n(t' - t)\right] \left[\boldsymbol{f}_n + \boldsymbol{\eta}(t')\right] dt' \end{aligned} \tag{4.27}$$

In this way, after substituting the initial and final time with $t_n \mp \tau/2$ a Markov chain of the following form is obtained:

$$\boldsymbol{v}_{n+1} = e^{-\boldsymbol{\gamma}_n \tau} \boldsymbol{v}_n + \boldsymbol{\Gamma}_n \cdot (\boldsymbol{f}_n + \tilde{\boldsymbol{\eta}}) \tag{4.28}$$
$$\boldsymbol{R}_{n+1} = \boldsymbol{R}_n + \tau v_{n+1} \tag{4.29}$$

namely we have singled out the "noisy" corrections to the force components in eq. (4.28) $(\boldsymbol{f}_n + \tilde{\boldsymbol{\eta}})$ by defining the following quantities:

$$\boldsymbol{\Gamma}_n = \boldsymbol{\gamma}_n^{-1}(\mathbf{I} - e^{-\boldsymbol{\gamma}_n \tau}) \tag{4.30}$$

$$\tilde{\boldsymbol{\eta}} = \frac{\boldsymbol{\gamma}_n}{2\sinh(\boldsymbol{\gamma}_n \tau/2)} \int_{t_n - \tau/2}^{t_n + \tau/2} e^{\boldsymbol{\gamma}_n (t - t_n)} \boldsymbol{\eta}(t)\, dt \tag{4.31}$$

By using that $[\boldsymbol{\alpha}, \boldsymbol{\gamma}] = 0$ and a little algebra, the correlator defining the discrete (time integrated) noise can be computed by substitution of eq. (4.31) in eq. (4.18) and is given by the following $3N \times 3N$ matrix:

$$\langle \tilde{\eta}_i \tilde{\eta}_j \rangle = \bar{\boldsymbol{\alpha}} = T \boldsymbol{\gamma}_n^2 \coth(\boldsymbol{\gamma}_n \tau/2) \tag{4.32}$$

The simulation temperature T appearing in the above expression is an input parameter of the dynamics. If the discretization of the SLD is accurate enough this temperature should be related to the mean square velocities measured during the dynamics ($\langle v_i^2 \rangle = T/2$ for each Cartesian component). In the

forthcoming sections we refer to this quantity as the "effective temperature" T_{mes}, as $T_{\mathrm{mes}} - T$ can be used to judge the quality of the approximations in discretizing the SLD.

As discussed also in ref. [47] (see also Section 5.2.3), all the QMC force evaluations $\boldsymbol{f}$ are affected by an intrinsic stochastic noise, that usually determines an effective temperature higher than the target one. This problem can be avoided, by means of the *noise correction* introduced in ref. [47]. Indeed we can follow the correct dynamics by adding to the QMC noise of the force the external noise $\tilde{\boldsymbol{\eta}}_{\mathrm{ext}}$ so that the total noise $\tilde{\boldsymbol{\eta}}$ satisfies the correct expressions in eq. (4.32). In this way, we have to subtract the $3N \times 3N$ QMC correlation of the forces $\boldsymbol{\alpha}^{\mathrm{QMC}}$ from the above described correlation matrix $\bar{\boldsymbol{\alpha}}$ and obtain that:

$$\bar{\boldsymbol{\alpha}}_{\mathrm{ext}} = \bar{\boldsymbol{\alpha}} - \boldsymbol{\alpha}^{\mathrm{QMC}} \tag{4.33}$$

is the true external noise we have to add to the force components during the dynamics. Indeed the correlation matrix $\boldsymbol{\alpha}^{\mathrm{QMC}}$ can be independently evaluated during the dynamics and the computation of eq. (4.22) is possible with some statistical error. In this way, we can take into account that QMC forces are affected by a correlated noise, and obtain, in principle, an unbiased simulation following the correct SLD.

In order to confirm that our SLD scheme samples the exact partition function $Z = \int dx \exp(-V(x)/kT)$, we have also done a conventional Monte Carlo simulation with the fitted potential for the water monomer. In figure 4.1, both simulations show a perfect agreement in the internal energy at various temperatures consistent with the harmonic approximation up to 1000 K.

4.5 Role of the force covariance matrix

After several tests, we have discovered that the value of Δ_0, optimizing the efficiency of the calculation, is not necessarily the minimum one, i.e. $\Delta_0 = \tau$. Indeed much larger time steps and better performances are possible if $\Delta_0 \gg \tau$. In order to understand this behavior, it is important to realize that the covariance matrix $\boldsymbol{\alpha}^{\mathrm{QMC}}$ obtained with QMC is empirically proportional to the dynamical matrix (see figure 4.2). Therefore with a finite and large Δ_0 the high energy modes with high frequency vibrations can be systematically damped, and this clearly allows a faster propagation with larger time step τ.

According to F. Tassone et al. [79], the optimal choice of the friction used by damped dynamics is determined by the lowest frequency mode in order to have the least possible slowdown in that mode. This can be easily understood through a system consisting of several damped oscillators. The equation of

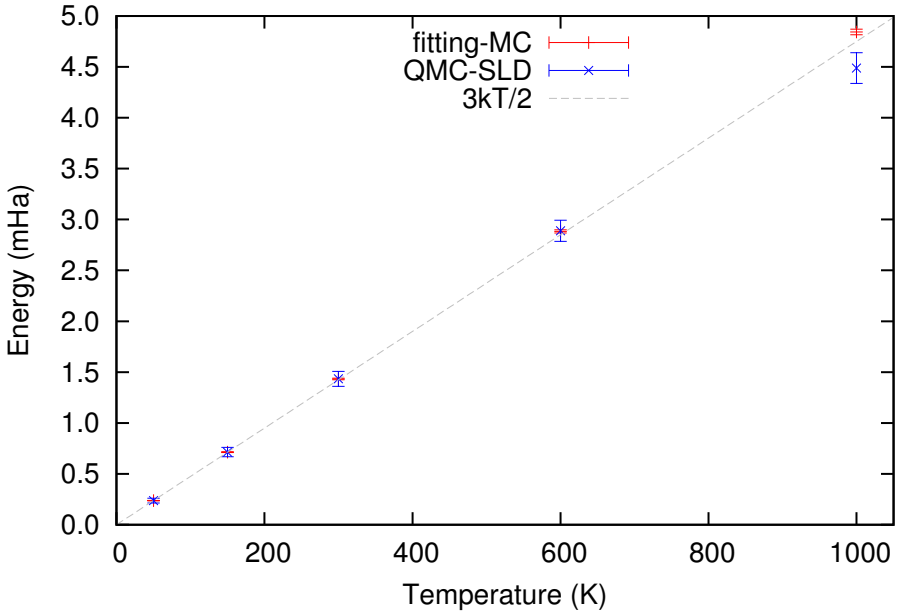

Figure 4.1: The internal energy of the water monomer in SLD simulations with QMC and MC simulations with the fitted potential at various temperatures. Since this potential expanded up to the fourth order is fitted with forces, the zeroth order coefficient is set to zero and the QMC internal energy, for comparison, is shifted by -17.24909(3) Ha which is the lowest possible variational energy with $\mathcal{N}_{\mathrm{QMC}} = 20480$ and $n_{\mathrm{opt}} = 20$. In these calculations, for simplicity, only coordinates x_{H_1}, x_{H_2} and y_{H_2} are free. Therefore, the internal energy should be $\frac{3}{2}kT$ if Harmonic approximation is assumed.

motion for this system reads

$$\ddot{\boldsymbol{R}} = -\boldsymbol{\gamma}\dot{\boldsymbol{R}} + \boldsymbol{f}(\boldsymbol{R}) \tag{4.34}$$

$$\boldsymbol{f}(\boldsymbol{R}) = -\boldsymbol{K}\boldsymbol{R} \tag{4.35}$$

$$\boldsymbol{\gamma} = 2\boldsymbol{\xi}\,, \tag{4.36}$$

where $\boldsymbol{K}$ is the dynamical matrix and $\boldsymbol{\gamma} = 2\boldsymbol{\xi}$ is the friction matrix. In the following, for simplicity, we assume that $[\boldsymbol{\gamma}, \boldsymbol{K}] = 0$. We can solve this equation by assuming a solution

$$\boldsymbol{R} = \exp(-\boldsymbol{W}t)\,, \tag{4.37}$$

where $\boldsymbol{W}$ is, in general, a matrix with complex elements. By substituting this assumed solution (4.37) back to eq. (4.34), we can obtain that

$$\boldsymbol{W} = \boldsymbol{\xi} \pm i\sqrt{\boldsymbol{K} - \boldsymbol{\xi}^2}\,. \tag{4.38}$$

In order to have the shortest correlation time, the optimal value of $\boldsymbol{\xi}$ is $\sqrt{\boldsymbol{K}}$ which maximizes the real part of $\boldsymbol{W}$ and achieves the fastest damping. In the

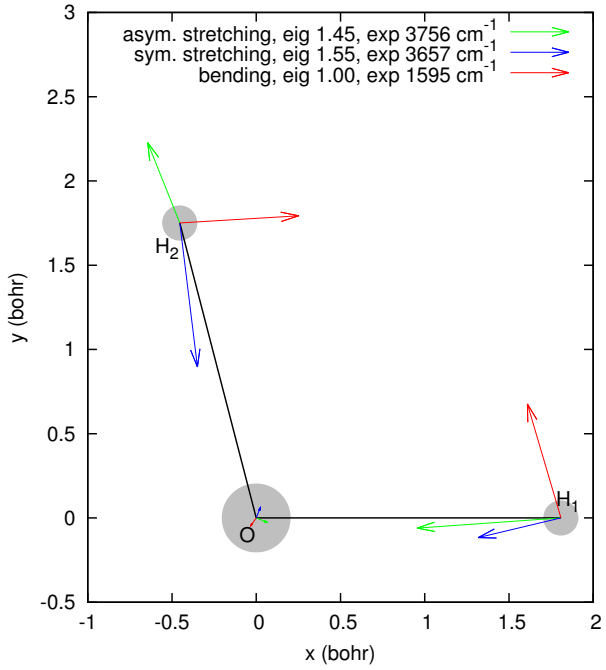

Figure 4.2: Eigenvectors of the 3×3 correlation matrix $\boldsymbol{\alpha}^{\mathrm{QMC}}$ which is constructed by $f^x_{\mathrm{H}_1}$, $f^x_{\mathrm{H}_2}$ and $f^y_{\mathrm{H}_2}$ (x_{H_1}, x_{H_2} and y_{H_2} are chosen as internal coordinates). These eigenvectors correspond to the three vibrational modes of the water monomer: bending (red), symmetrical (blue) and asymmetrical (green) stretching. The smaller eigenvalue of $\boldsymbol{\alpha}^{\mathrm{QMC}}$ corresponds to the lowest frequency vibrational mode. The eigenvalues in the plot are all rescaled by the lowest eigenvalue. The experimental values are fundamental frequencies.

case of [79], a friction matrix $\boldsymbol{\gamma}$ is replaced by a single value γ, the optimal value γ_{opt} is chosen as ω_{min} which is the square root of the smallest eigenvalue of the matrix $\boldsymbol{K}$.

In our MD scheme, the covariance matrix of the forces which implicitly includes the information of the dynamical matrix has already been utilized to build the friction matrix. When a large Δ_0 is used, the eigenvalues of $\boldsymbol{\gamma}$ spans a wide range. When these eigenvalues from the smallest to the largest ones map to the vibrational frequencies also from the smallest to the largest ones, the optimal choice of the frictions can be met by all the vibrational modes and much better performances is obtained compared with the standard method where only the lowest frequency mode is considered [79].

The largest possible time step τ is constrained by the highest frequency mode because it should be smaller than at least $1/2$ the time period of the highest frequency mode. Since the high frequency modes are damped very much, a larger time step τ can be used in our MD scheme. Meanwhile, since the low frequency modes are much less affected by the damping, increasing τ

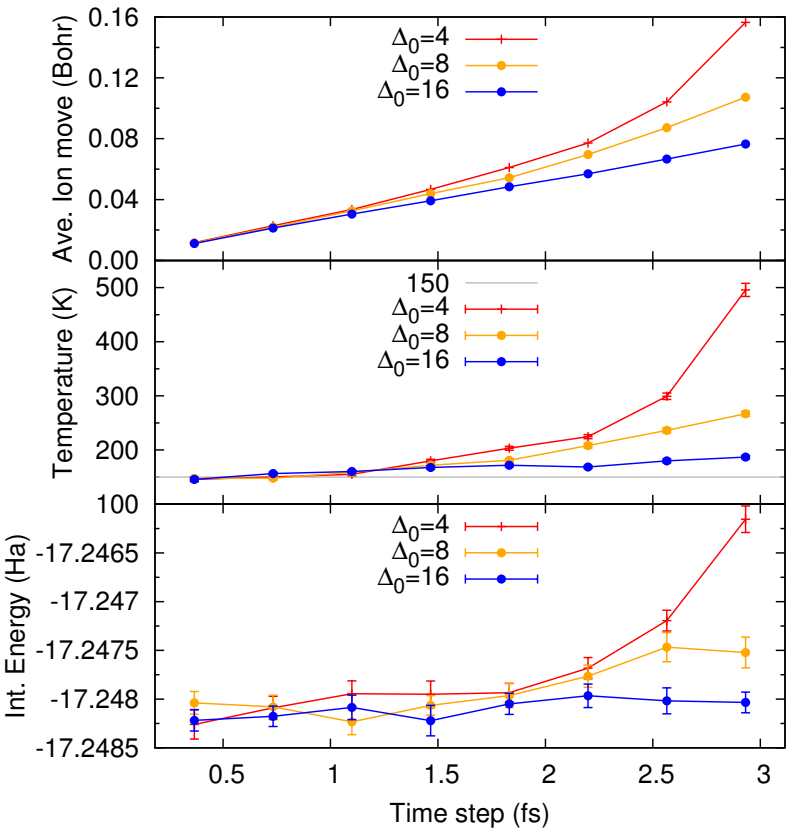

Figure 4.3: The average step length of collective ionic moves (upper), measured temperature (middle), internal energy (lower) as a function of the time step τ with $\Delta_0 = 4, 8, 16$ in test case water monomer.

also speed up more these modes than those high frequency modes.

In the figure 4.3, the curves with $\Delta_0 = 4, 8, 16$ are converged as $\tau < 1.5, 2.0, 3.0$ fs respectively. As Δ_0 increases, the length of the average ionic move decreases due to the increasing friction and thus the temperature and internal energy are stabilized at large τ. Therefore, large Δ_0 allows us to get benefits from large integration time step without hurting the accuracy. Let's consider the average length of each collective ionic move $||\boldsymbol{R}_n - \boldsymbol{R}_{n-1}||$ during the MD simulation as the dominant parameter for the integration error and make the plots. In the figure 4.4, the curves of internal energy with $\Delta_0 = 4, 8, 16, 64$ coincide with each other. This implies that the error of internal energy is directly related with the average ionic move $\overline{||\boldsymbol{R}_n - \boldsymbol{R}_{n-1}||}$ instead of τ. On the other hand, the curve of measured temperature with large Δ_0 choice shows a much better temperature control of our dynamics. In summary, large Δ_0 is recommended for both increasing the efficiency, accuracy and stability of our molecular dynamics simulation.

It has already been shown in ref. [35], that the present integration scheme of the Langevin equations is much better than the Euler integration method.

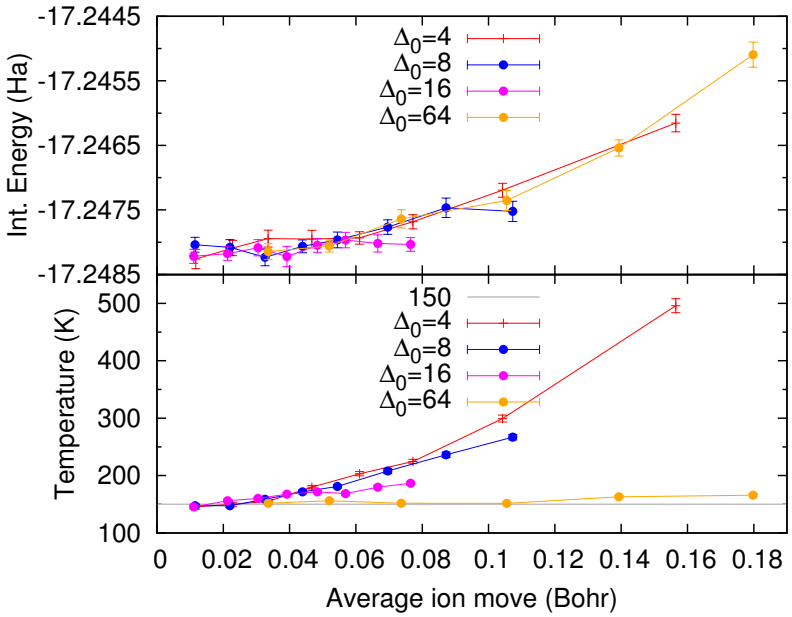

Figure 4.4: The internal energy $E_{\rm int}$ and measured temperature T as functions of average step length of collective ionic moves $\overline{||\boldsymbol{R}_n - \boldsymbol{R}_{n-1}||}$ in the MD simulations of the water monomer with various time steps τ. The four sets of simulations are performed with $\Delta_0 = 4, 8, 16, 128$ for water monomer at 150K.

In this thesis we also show that the present dynamics is also much more convenient within QMC because we can use a friction matrix proportional to the mentioned QMC covariance matrix ($\Delta_0 > 0$). To this purpose we have implemented exactly the same dynamics within the `Quantum ESPRESSO` [80] package, and show in figure 4.5 that, within the DFT dynamics, only quite smaller time steps τ are possible, just because in this case $\boldsymbol{\alpha}^{\rm QMC} = 0$, and it is not possible to damp the too high frequency vibrations.

Thanks to this important property of the covariance matrix $\boldsymbol{\alpha}^{\rm QMC}$, our SLD scheme driven by noisy QMC forces samples the phase space more efficiently than conventional SLD. Every coin has two sides and so does the noise.

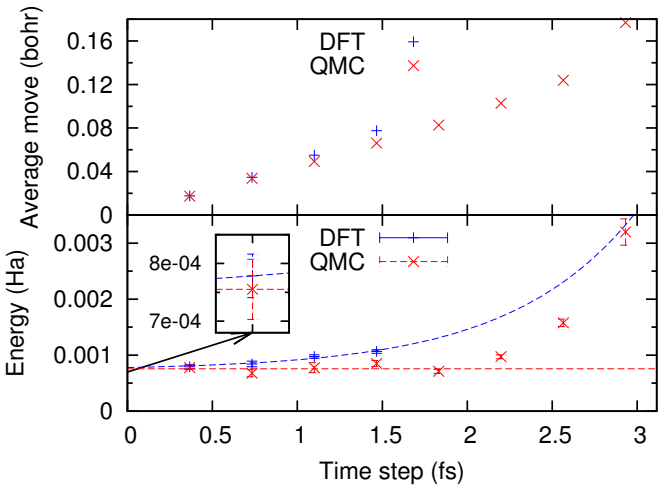

Figure 4.5: Convergence of the internal potential energy and average ion displacements as a function of the time step τ for the water dimer, obtained with MD at 50 K. The offset energy (minimum of the PES) values of DFT and QMC calculations are -34.410806 and -34.50405(4) Ha. The same $\alpha_0 = 0.03$ a.u. is used for both. Our simulation with DFT becomes unstable when $\tau \geq 1.8$ fs. Instead the QMC dynamics is always stable in the range studied because the friction matrix in this case contains also an important non-diagonal contribution proportional to QMC covariance matrix (see text and figure 4.2). In the top panel we compare the average distance that ions experience at each step in QMC and DFT dynamics. The increased stability achieved by using this covariance matrix in the friction is therefore obtained with an almost negligible slowing down of the QMC dynamics.

Chapter 5

Vibrational frequencies

5.1 Introduction

As we have discussed in the previous chapter, at present QMC allows us to compute forces with finite variance in an efficient way and we've already introduced the MD scheme by means of the QMC evaluation of forces. In this chapter, we'd like to systematically study how reliable this scheme is. In particular we want to understand:

- how the noise in the forces affects the reliability of the dynamics,
- how the systematic error due to the discretization in time affects the calculation in presence of noisy forces,
- how well the Born-Oppenheimer constraint is satisfied, namely how accurately it is possible to evolve the electronic wave function following the minimum energy condition. Indeed in a statistical method the variational parameters cannot be optimized with machine precision accuracy and the departure of the wave function from its minimum energy may represent an important bias requiring a careful study.

In order to answer to the above issues, we show the performance of QMC-based approaches (including MD simulations) by benchmark calculations of structural and vibrational properties of small molecules. These properties are often of interest in Chemistry and Materials Science because they help the interpretation of experiments, for instance of infrared and Raman spectroscopy [81].

In ab-initio approaches the vibrational properties are usually obtained within the Born-Oppenheimer approximation, that separates the electronic and nuclear degrees of freedom. Thus, their evaluation relies on the properties of the potential energy surface (PES) in the neighborhood of the structural minimum of the molecule [81, 82]. The simplest approach is to assume that the PES in the neighborhood of the minimum is well characterized within the harmonic approximation, so the frequencies are obtained from

the diagonalization of the mass-weighted Hessian matrix [81], which is calculated by performing static ab-initio computations in the minimum of the PES, or in its neighborhood. This approach neglects the anharmonicity of the PES, so "ad hoc" scaling factors [83, 84] have to be introduced in order to compare with the experimental frequencies. The most accurate approaches [85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96] go beyond the harmonic approximation, for instance taking the force fields of the PES around the configurational minimum up to the fourth order expansion and using the second order vibrational perturbation theory (PT2) [93, 94, 95, 96]. Other ab-initio approaches are based on ab initio molecular dynamics simulations, which directly includes finite temperature nuclear motions, and from the Linear Response Theory [97] and the Fermi Golden Rule, the Infrared and Raman spectra can be directly obtained from the Fourier transform of dipole and polarizability autocorrelation functions [98, 82].

In this chapter we have evaluated the structural and vibrational properties by using and comparing three different methods:

(*i*) with the Hessian, and in a few cases with higher order derivatives of the PES that are estimated by a careful fit of independent measurements of energy and forces. These quantities are calculated over a set of molecular configurations arranged on a grid [96, 50] around the equilibrium structure of the molecule;

(*ii*) same as *i* in the fitting by using the samples of energies and forces at molecular configurations generated instead automatically by a QMC-based MD simulation, at a given temperature T, with noisy forces [35];

(*iii*) by using time averaged correlations in a QMC-based MD simulation.

It is clear that, if the QMC-based MD simulation is consistent and the BO constraint is correctly satisfied, all different methods should provide consistent results, provided all sources of systematic errors can be removed in a controlled way, in order to converge to unbiased evaluations of the geometrical and vibrational properties.

In this chapter, we show that the method *ii* provides very accurate results with an efficiency comparable with the standard method *i*, whereas the method *iii* is computationally very demanding and is used therefore here only for testing the MD.The method *ii*, that we are proposing, is in our opinion better than the standard one *i* because it can be easily and systematically extended to complex systems containing several atoms. In such cases, it is very difficult to work with the standard method, because it relies on a careful choice of the grid of atomic positions that are used to fit the PES [96]. This method is difficult to be generalized to very complicated systems, and in particular the grid cannot be generated by a black box tool, as it depends instead on the user's choice. Instead we propose here the much more general

and flexible method *ii*, allowing a systematic and robust evaluation of harmonic frequencies. In this technique, only a single parameter has to be tuned, namely the target temperature of the MD simulation.

This chapter is organized as follows: in section 5.2 the three approaches of evaluating vibrational properties are explained in detail; in section 5.3 we describe the wave functions and the basis sets we use for all the molecules; whereas the discussion of all sources of systematic errors related to the present QMC dynamics is given in section 5.4; section 5.5 contains our results on several molecules with some discussion; finally in section 5.6 we draw our conclusions.

5.2 Calculation of vibrational properties

The standard method of calculating vibrational modes is well-known in the literature [99]. In this section, we summarize the main formulas and introduce the notations adopted in the rest of this chapter.

Within the Born-Oppenheimer approximation, the full Hamiltonian is separated into electronic and nuclear parts and only the latter is related to the calculation of vibrational modes. The nuclear Hamiltonian $\hat{H}$ is a summation of kinetic energy $\hat{T}$ and potential energy $\hat{V}$. Given a molecule with N atoms, $\hat{H}$ can be expressed in terms of $3N$-dimensional Cartesian coordinates $\boldsymbol{R}$:

$$\hat{H} = \hat{T} + \hat{V} = -\frac{1}{2}\sum_{\xi}^{N} \frac{1}{M_\xi}\nabla_\xi^2 + V(\boldsymbol{R})\,, \tag{5.1}$$

where M_ξ is the mass of the atom ξ. Since the potential energy $V(\boldsymbol{R})$ is generally assumed to be invariant under the translation and rotation of molecules, it is independent of both the molecule's center of mass and orientation. Therefore, it can be written in terms of $3N-6$ (or $3N-5$ for linear molecules) internal coordinates $\bar{\boldsymbol{s}}$.

Since the molecule is usually assumed to be semi-rigid, its potential energy can be simply recast in terms of the displacement from the equilibrium structure $\boldsymbol{R}_0$, corresponding to the (local) minimum of the PES. The Cartesian displacement is $\boldsymbol{X} = \boldsymbol{R} - \boldsymbol{R}_0$. Similarly, the displacement in the internal coordinates is $\boldsymbol{s} = \bar{\boldsymbol{s}} - \bar{\boldsymbol{s}}_0$.

In general, the mapping between Cartesian coordinates and internal coordinates is curvilinear. With Taylor expansion, $\boldsymbol{s}$ becomes

$$s_i = B_i^a X_a + \frac{1}{2!}B_i^{ab} X_a X_b + \frac{1}{3!}B_i^{abc} X_a X_b X_c + \cdots \tag{5.2}$$

where $i = 1, \ldots, 3N-6$ labels the internal coordinates and $a, b, c = 1, \ldots, 3N$ label the Cartesian coordinates. The Einstein summation notation of repeated indices is assumed hereafter. The coefficients in the series are the derivatives with respect to the Cartesian displacement: $B_i^a = \partial s_i / \partial X_a$,

$B_i^{ab} = \partial^2 s_i/(\partial X_a \partial X_b)$ and so on and so forth. The coefficients B_i^a in the linear term define the so called Wilson $\boldsymbol{B}$ matrix.

Hence, the potential energy can be expanded around the equilibrium structure in terms of internal coordinate displacements as

$$V(\boldsymbol{s}) = F^0 + F^i s_i + \frac{1}{2!} F^{ij} s_i s_j + \frac{1}{3!} F^{ijk} s_i s_j s_k + \cdots , \tag{5.3}$$

where the coefficients in the expansion are defined as $F^0 \equiv V$, $F^i \equiv \partial V/\partial s_i$, $F^{ij} \equiv \partial^2 V/(\partial s_i \partial s_j)$, etc., calculated at $s = 0$. Clearly, all the coefficients $F^i = 0$. Since F^0 is an irrelevant offset for any vibrational modes, we ignore it by putting $V(\boldsymbol{s}) - F^0$ instead of $V(\boldsymbol{s})$. In the rest of this chapter, V2 and V4 are used to indicate the potential energy surface expanded up to the second and the fourth order respectively.

In the standard method of calculating vibrational modes within the harmonic approximation, only the leading terms, i.e. the quadratic ones, are kept in both the potential and kinetic energies while all the rest is neglected:

$$V_{\mathrm{har}}(\boldsymbol{s}) = \frac{1}{2} F^{ij} s_i s_j = \frac{1}{2} \boldsymbol{s}^\dagger \boldsymbol{F} \boldsymbol{s} \tag{5.4}$$

$$T_{\mathrm{har}}(\dot{\boldsymbol{s}}) = \frac{1}{2} (\boldsymbol{G}^{-1})^{ij} \dot{s}_i \dot{s}_j = \frac{1}{2} \dot{\boldsymbol{s}}^\dagger \boldsymbol{G}^{-1} \dot{\boldsymbol{s}} , \tag{5.5}$$

where $\dot{\boldsymbol{s}}$ is the time derivative of $\boldsymbol{s}$, and the symbol † indicates the transpose. Meanwhile, the $(3N-6) \times (3N-6)$ matrix $\boldsymbol{G}$ is calculated as

$$G_{ij} = \sum_{\xi}^{N} \sum_{\alpha}^{3} \frac{1}{M_\xi} B_i^{\xi,\alpha} B_j^{\xi,\alpha} , \tag{5.6}$$

where $B_i^{\xi,\alpha}$ are the same linear terms defined in eq. (5.2), upon replacement of the index a with the pair (ξ, α), indicating more explicitly the component α corresponding to the atom ξ.

By introducing $3N-6$ normal coordinates $\boldsymbol{q}$, the potential and kinetic energies are recast as

$$V_{\mathrm{har}}(\boldsymbol{q}) = \frac{1}{2} \sum_{i}^{3N-6} \lambda_i q_i^2 \tag{5.7}$$

$$T_{\mathrm{har}}(\dot{\boldsymbol{q}}) = \frac{1}{2} \sum_{i}^{3N-6} \dot{q}_i^2 , \tag{5.8}$$

where $\lambda_i = \omega_i^2$ are the harmonic force constants corresponding to harmonic frequencies ω_i.[1] Assuming the transformation between internal coordinates

[1] We adopt in the chapter the conventional choice, that comes from spectroscopic conventions, to report the wavenumber $\bar{\nu}$, expressed in cm^{-1}, corresponding to the frequency ω. They are related by the equation $\omega = 2\pi c \bar{\nu}$, being c the speed of light in a vacuum.

and normal coordinates $\boldsymbol{s} = \boldsymbol{L}\boldsymbol{q}$, we replace $\boldsymbol{s}$ in eqs. (5.4) and (5.5) and compare them with eqs. (5.7) and (5.8). The final relations are written in matrix form as

$$\boldsymbol{L}^\dagger \boldsymbol{F} \boldsymbol{L} = \boldsymbol{\Phi}, \quad \boldsymbol{L}^\dagger \boldsymbol{G}^{-1} \boldsymbol{L} = \mathbf{I}\,, \tag{5.9}$$

where $\boldsymbol{\Phi}$ is a diagonal matrix with λ_i on the diagonal and $\mathbf{I}$ is a $3N-6$ dimensional identity matrix. With some very simple algebra, eqs. (5.9) turns into $\boldsymbol{G}\boldsymbol{F}\boldsymbol{L} = \boldsymbol{L}\boldsymbol{\Phi}$ which represents a standard generalized eigenvalue problem, where λ_r are the corresponding eigenvalues. This approach is also called Wilson's GF method [99].

Ab-initio methods can be used to calculate the Hessian matrix $\boldsymbol{F}$ in the potential energy, so that the application of the GF method is possible. In the standard method, it is necessary to perform a very accurate structural optimization of the molecule, and then to calculate the derivatives for the optimized geometry using analytic or finite-difference methods. A very tight structural optimization is computationally very demanding for QMC, thus alternative methods specifically engineered for stochastic-error affected approaches are preferable, as discussed in ref. [96] and summarized in section 5.2.1. We propose here other two possible approaches, described in sections 5.2.2 and 5.2.3.

We have reported also some results, labeled as fundamental frequencies, coming from second order perturbation theory (PT2), that uses also the third and forth order derivatives of $V(\boldsymbol{s})$, in order to take into account of the anharmonicity of the PES. The use and implementation of PT2 in presence of error affected PES have been widely discussed in Ref. 96, and we remand to this reference.

5.2.1 Simple fitting method

The conventional way to obtain the Hessian of V is to fit the parametrized Hessian matrix $\boldsymbol{F}$ with energies or forces computed at the chosen grid points of the $3N$ multidimensional space defined by the nuclear positions. In each of the $3N$ directions, at least 3 points are needed in the neighborhood of the equilibrium position in order to fit the Hessian. Obviously, this requires a tight (gradient $< 10^{-5}$ a.u., for the harmonic approach) or very tight (gradient $< 10^{-7}$ a.u., for PT2) structure optimization criteria [96] which can be easily achieved by self-consistent iterations in DFT or other deterministic methods.

However, these criteria are not feasible for QMC since all the energies and forces calculated by QMC are error-affected. The stochastic error σ_{QMC} is inversely proportional to the square root of the number of QMC samples $\mathcal{N}_{\mathrm{QMC}}$. Thus, in order to have an error 10 times smaller, a calculation 100 times more expensive is required. For this reason, the QMC stochastic errors are never pushed to very small values, especially for vibrational property calculations. Typically, the errors are $\sigma_E \sim 10^{-4}$ a.u. for energy and $\sigma_F \sim 10^{-3}$ a.u. for each

force component. In brief, both the PES and equilibrium structure are very much affected by the stochastic noise.

A. Zen et al. [96] proposed a multidimensional fitting scheme to obtain the accurate Hessian and equilibrium structure. They showed that the fitting with forces brings much less error than the fitting with energies. So we stick to forces for the fitting. Moreover, for our calculations with the water molecule we choose the "mesh-5" (see definition in ref. [96]), which consists of 59 independent grid points $\boldsymbol{R}_m$. The fitting was achieved by maximizing the likelihood function:

$$\mathcal{L}(\boldsymbol{k}|\boldsymbol{D}_F) = \prod_m^{\mathcal{N}_m} \frac{e^{-\frac{1}{2}\sum_{a,b}^{3N}({\boldsymbol{C}_m}^{-1})_{ab}\Delta\mathcal{F}_m^a(\boldsymbol{k})\Delta\mathcal{F}_m^b(\boldsymbol{k})}}{(2\pi)^{3N/2}\sqrt{\det(\boldsymbol{C}_m)}} \tag{5.10}$$

for the parameter set $\boldsymbol{k}$ to fit the $\mathcal{N}_m$ QMC samples $\boldsymbol{D}_F = \{\boldsymbol{R}_m, \boldsymbol{f}_m, \boldsymbol{C}_m\}_m$ where $\boldsymbol{f}_m$ is the force calculated at $\boldsymbol{R}_m$ and $\boldsymbol{C}_m$ is the $3N$-dimensional covariance matrix of the $\boldsymbol{f}_m$ (see definition eq. (4.22) and discussion in section 5.2.3), and

$$\Delta\mathcal{F}_m^a(\boldsymbol{k}) = \mathcal{F}^a(\boldsymbol{R}_m, \boldsymbol{k}) - f_m^a$$

the difference between the QMC force f_m^a of component a, calculated in $\boldsymbol{R}_m$, and the corresponding value of the parametrized force $\mathcal{F}^a(\boldsymbol{R}_m, \boldsymbol{k})$. The problem of maximizing $\mathcal{L}(\boldsymbol{k}|D_F)$ is equivalent to minimize the function:

$$\sum_m^{\mathcal{N}_m}\sum_{a,b}^{3N}({\boldsymbol{C}_m}^{-1})_{ab}\Delta\mathcal{F}_m^a(\tilde{\boldsymbol{k}})\Delta\mathcal{F}_m^a(\tilde{\boldsymbol{k}}) \tag{5.11}$$

and, as discussed in ref [96], in the case that we can neglect the covariance between QMC force evaluations of the different components (i.e., we can assume $\boldsymbol{C}_m$ diagonal, and the diagonal elements $(\boldsymbol{C}_m)_{aa} = (\sigma_m^a)^2$ are the variance of the QMC force evaluations of component a), the previous expression corresponds to the chi-squared-function:

$$\chi_F^2 = \sum_m^{\mathcal{N}_m}\sum_a^{3N}\left(\frac{\mathcal{F}^a(\boldsymbol{R}_m, \boldsymbol{k}) - f_m^a}{\sigma_m^a}\right)^2 . \tag{5.12}$$

We can quantify the quality of the fit by using the reduced-chi-squared function (goodness of fit):

$$\chi_{\text{red}}^2 = \frac{\chi_F^2}{(3N \times \mathcal{N}_m - N_k - 1)} , \tag{5.13}$$

where $3N$ is the number of force components, $\mathcal{N}_m$ is the number of molecular configurations considered and N_k is the number of fitted parameters. According to statistical theory, the closer $\chi_{\text{red}}^2 \simeq 1$, the better the fit is.

5.2.2 Fitting method with molecular dynamics

In the simple fitting method described in the previous section, the choice of the grid points, where energy and forces are evaluated, is crucial for accessing accurate vibrational frequencies. A good mesh should span a region neither too small, in order to be less affected by the stochastic noise, nor too large, to avoid strong anharmonicity (which cannot be well described by simple parametrization of the PES in a truncated Taylor expansion around the minimum). In order to reduce the systematic error, the best mesh should be expanded along the directions of the normal coordinates which are however known only after the fitting. An efficient compromise is to use internal coordinates based on certain conventional rules. After the region and expansion direction of the mesh is given, the density of the grid points should be also chosen properly. Too sparse mesh limits the accuracy while too dense mesh wastes computation. Since we did only once the evaluation of energy and forces on each grid point in the simple fitting method, good optimization of the wavefunction and accurate calculation of energy and forces are both necessary and therefore expensive. Usually, for each step during the optimization, relatively small statistics is used. In this way, the accuracy of the energy and forces evaluated in the last iteration of the optimization doesn't meet the necessary precision for the fitting. So a further much longer run at fixed optimal values of the variational parameters is required to compute the energy and forces precisely. This has also the drawback that the error in the optimization has to be negligible compared with the requested statistical error, a condition that is difficult to control systematically.

With molecular dynamics, things are instead much easier. By controlling only the temperature, a proper mesh is automatically decided by the trajectory of the moving ions. Since the dynamics follows the normal modes, the mesh has already been expanded around the best directions. The density and range of the mesh is directly tuned by the temperature. Unlike the simple fitting, now the same configuration will be approached several times during the dynamics, and, as a consequence, the stochastic error will be alleviated so that even quite low temperatures can be efficiently simulated. Since ions move very little for each iteration of MD, a heavy optimization is no longer necessary because the electronic wave function obtained in the previous iteration of the dynamics is a very good starting point for the current iteration, once only the positions of the atomic localized orbitals are consistently updated. This is actually a very remarkable advantage to expand our electronic variational wave function in terms of localized atomic orbitals. Indeed, after a few optimization steps, the wave function is usually converged within given statistical errors, and, as we will see later, the error in the optimization can be systematically controlled. Apart from the cheap optimization, energy and forces also require much less accuracy. In table 5.2, the error bars of the energy and forces during the dynamics are about 40 and 60 times larger than those

of the simple fitting. The values obtained in the last step of the optimization are already sufficient and thus, a substantial amount of computation is saved. The fitting procedure of the sampled configurations coming from molecular dynamics is exactly the same described in the previous section, and we can use eqs. (5.10)–(5.12).

In our tests on the water monomer, the anharmonic effects are quite strong if the MD is performed at high temperature. We have systematically studied the effect of the temperature in figure 5.1, where it is clear that anharmonic effects can be neglected only below 50K namely when the temperature corresponds to a frequency 50 times smaller than the lowest frequency of the system ($\simeq$ 2300K). This criterion cannot be easily extended to larger systems as the smallest frequency significantly drops, and a calculation at too small temperatures cannot provide enough information for the fit, yielding large statistical errors for the frequencies. For this reason it is important to include in the fit also the cubic and quartic terms, and, as it is also shown in figure 5.4, it is really remarkable that we can obtain a very reliable and converged estimate of the frequencies even at 1000K.

In order to improve further the fit, we generate more statistical samples by taking advantage of the molecular symmetry. For example, H_2O, H_2S and SO_2 have the C_{2v} symmetry while NH_3 and PH_3 have the C_{3v} symmetry. By simply swapping the positions and forces of each pair of H or O atoms in the molecules, we obtain once or twice more samples, without an extra computational effort. This procedure doesn't change much the frequencies and equilibrium geometry, as well as their statistical errors, but allows to enforce the symmetry of a molecule, namely recovering all equal X-H bond lengths and degenerate frequencies, if related by the mentioned symmetries. We have used a similar method also in the simple fitting (previous subsection), but in that case the purpose was mainly to reduce the number of points in the grid and to save computer resources.

5.2.3 Covariance matrix method

The previous two methods of evaluating vibrational frequencies give very accurate results but require an explicit parametrization of the PES in the neighborhood of the minimum, as well as the reduction of the number of parameters by using the symmetries of the molecule. To avoid this human overhead, we introduce another way of computing vibrational frequencies based on the evaluation of few appropriate covariance matrices, described in the following.

By employing a simple Gaussian integral over the statistical weight $\exp(-V_{\mathrm{har}}(\boldsymbol{s})/kT)$ where $V_{\mathrm{har}}(\boldsymbol{s})$ is defined in eq. (5.4) we easily obtain the relation

$$\boldsymbol{C}(\boldsymbol{s}) = kT\boldsymbol{F}^{-1} \tag{5.14}$$

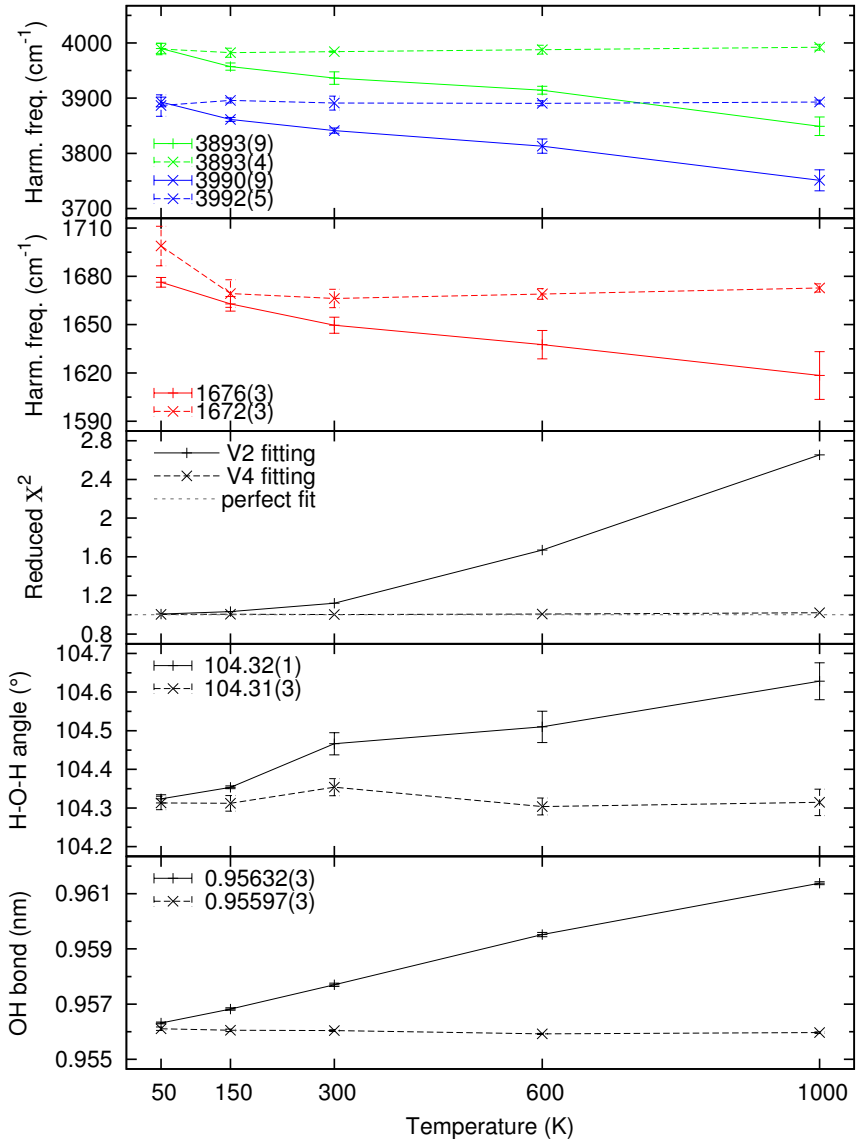

Figure 5.1: Harmonic vibrational frequencies and equilibrium geometry of water monomer obtained by fitting V2 (solid line) and V4 (dash line) as a function of temperature. The values reported in the keys are obtained at 50K for V2 and 1000K for V4.

where $\boldsymbol{C}(\boldsymbol{s})$ is the covariance matrix of the internal coordinates $\boldsymbol{s}$ according to the definition:

$$\boldsymbol{C}(\boldsymbol{s}) = \langle (\boldsymbol{s} - \langle \boldsymbol{s} \rangle)(\boldsymbol{s} - \langle \boldsymbol{s} \rangle)^{\dagger} \rangle \tag{5.15}$$

where $\langle \cdots \rangle$ refers to the ensemble average while in practice it is computed as the time average along the trajectory of MD. Therefore the matrix $\boldsymbol{F}$, necessary in the Wilson's GF method, can be obtained by computing $\boldsymbol{C}(\boldsymbol{s})$ with a simple algebraic inversion and a scaling by kT.

In a more direct method, the information of the matrix $\boldsymbol{F}$ can be obtained by computing the covariance matrix of the forces. Indeed, the forces $\boldsymbol{f}(\boldsymbol{s})$ defined as

$$f_i(\boldsymbol{s}) = -\frac{\partial V(\boldsymbol{s})}{\partial s_i} \tag{5.16}$$

have a very simple form if we can reliably work in harmonic approximation, namely $V(\boldsymbol{s}) \simeq V_{\mathrm{har}}(\boldsymbol{s})$, and are:

$$f_i(\boldsymbol{s}) \simeq -\boldsymbol{F}\boldsymbol{s}\,. \tag{5.17}$$

Therefore the covariance matrix of the forces is more simply related to the matrix $\boldsymbol{F}$ as

$$\boldsymbol{C}(\boldsymbol{f}) = kT\boldsymbol{F} \tag{5.18}$$

where $\boldsymbol{C}(\boldsymbol{f})$ is the covariance matrix of $\boldsymbol{f}$ similarly to eq. (5.15).

However, in QMC the forces are noisy and correlated since they are evaluated with the same Markov chains of finite length $\mathcal{N}_{\mathrm{QMC}}$, namely

$$\boldsymbol{f}_{\mathrm{noisy}} \equiv \langle \boldsymbol{f}_{\mathrm{local}} \rangle = \boldsymbol{f}_{\mathrm{exact}} + \boldsymbol{\delta} \tag{5.19}$$

where $\boldsymbol{f}_{\mathrm{local}}$ is the local force evaluated in each QMC sample and $\boldsymbol{\delta}$ is the statistical error associated to the QMC evaluation of the force. In order to obtain accurate frequencies, it is necessary to remove this bias for calculating the covariance matrix of forces, which we have done in the following way. The thermal average which is used to compute the covariance matrix can be divided into two steps — the average of all electronic realizations generated by quantum Monte Carlo at fixed ionic configuration $\boldsymbol{s}$ and the average of all sets of ionic configurations obtained during the Langevin dynamics. In the first step it is necessary to accumulate the covariance of the exact force components which however are known only with some statistical error ($\boldsymbol{f}_{\mathrm{noisy}}$). Therefore we can write $\boldsymbol{f}_{\mathrm{exact}} = \boldsymbol{f}_{\mathrm{noisy}} - \boldsymbol{\delta}$, where the error $\boldsymbol{\delta}$ depends on the quantum Monte Carlo statistics and vanishes only for $\mathcal{N}_{\mathrm{QMC}} \to \infty$. It follows therefore that:

$$\langle \boldsymbol{f}_{\mathrm{exact}} \boldsymbol{f}^{\dagger}_{\mathrm{exact}} \rangle_{\boldsymbol{s}} \;\approx\; \langle \boldsymbol{f}_{\mathrm{noisy}} \boldsymbol{f}^{\dagger}{}_{\mathrm{noisy}} \rangle_{\boldsymbol{s}} - \boldsymbol{C}_{\boldsymbol{s}}(\boldsymbol{\delta}) \tag{5.20}$$

where the subscript $\boldsymbol{s}$ refers to the restriction of a given ionic configuration $\boldsymbol{s}$, whereas $\boldsymbol{C}_{\boldsymbol{s}}(\boldsymbol{\delta})$ is the covariance of the noise, that can be in turn estimated by standard statistical methods using the finite number $\mathcal{N}_{\mathrm{QMC}}$ of Monte Carlo electronic samples that we have used to compute statistically the force components at a given ionic configuration $\boldsymbol{s}$, namely

$$\begin{aligned} \boldsymbol{C}_{\boldsymbol{s}}(\boldsymbol{\delta}) \;\approx\; & \frac{1}{\mathcal{N}_{\mathrm{QMC}}(\mathcal{N}_{\mathrm{QMC}}-1)} \\ & \times \sum_{j=1}^{\mathcal{N}_{\mathrm{QMC}}} (\boldsymbol{f}^{j,\boldsymbol{s}}_{\mathrm{local}} - \boldsymbol{f}^{\boldsymbol{s}}_{\mathrm{noisy}})(\boldsymbol{f}^{j,\boldsymbol{s}\dagger}_{\mathrm{local}} - \boldsymbol{f}^{\boldsymbol{s}\dagger}_{\mathrm{noisy}}) \end{aligned} \tag{5.21}$$

where $\boldsymbol{f}_{\text{local}}^{j,\boldsymbol{s}}$ are the force components corresponding to an independent electronic QMC sample j. With a finite number $\mathcal{N}_{\text{QMC}}$ of independent samples, $\boldsymbol{C_s}(\boldsymbol{\delta})$ scales as $1/\mathcal{N}_{\text{QMC}}$ and hence, the frequencies have a corresponding correction proportional to $\boldsymbol{C_s}(\boldsymbol{\delta})$, implying that the bias scales as $\frac{1}{\mathcal{N}_{\text{QMC}}}$ if the proposed noise correction is not applied. We will show the clear advantage to use this noise correction scheme in section 5.4.3.

5.3 Simulation setup

In all the following calculations we have used a variational wave function of a standard Jastrow-Slater form expanded on a localized basis set. According to our previous work [50] for the water monomer, we use (4s,5p,1d) primitive basis with 4 hybrid orbitals on oxygen (in short notation, O:(4s,5p,1d)/{4}) and H:(3s,1p)/{1} in the determinant part, whereas the Jastrow is expanded as a two-body part $\frac{1}{2b}(1-e^{-br})$ with one body rescaled and a three-body part with O:(3s,2p,1d)/{2} and H:(2s,2p)/{2} on hydrogen. Its VMC ground state energy at the equilibrium geometry is $-17.24927(3)$. During the dynamics, the exponents of the primitive basis in both determinant and three body Jastrow are all fixed and only 232 parameters in total are optimized on the fly. This basis is much more compact compared with the ones used in our previous work [50]. We have indeed verified that a larger basis does not improve much the inter-atomic description but decreases only the total energy. On the other hand, too many parameters make the optimization part too heavy and inefficient during the dynamics. Hence, we have to choose a compromise between accuracy and efficiency due to the available computational resources. Despite this limitation, we are generally working close to the Complete Basis Set limit as long as relevant chemical properties are concerned, thanks also to the rapid convergence in the basis set obtained within explicitly correlated wavefunctions, satisfying for instance all the electron-electron and electron-ion cusp conditions even with a finite basis set.

For H_2S, SO_2, NH_3 and PH_3 molecules, the basis sets used for the determinant part of the wave function are listed in table 5.3. The Jastrow has the same two body part as H_2O and its three-body part consists of (3s,2p,1d) on N/O/P/S and (2s,2p) on H.

In our calculation, energy-consistent pseudo-potentials (ECP) of Burkatzki et al. [100] are used to replace the core electrons of N, O, P and S atoms in order to have a reduced computational cost. Helium core is used for both N and O and Neon core is used for both P and S.

5.4 Control of all the sources of systematic errors in QMC dynamics

In the standard electronic structure calculation based on molecular dynamics, there are essentially three systematic errors to take into consideration: the time step τ used to integrate the SLD equations of motions, the accuracy in satisfying the Born-Oppenheimer approximation and the total time of simulation t_{TOT}. In the following we consider simple small molecules such that the simulation time is much larger than any reasonable correlation time of the system, so that this error can be safely neglected for simplicity. In QMC we have to take into account also that, at each step of the discretized dynamics in eqs. (4.28), only a statistical evaluation of the forces f_i with a finite number of samples $\mathcal{N}_{\mathrm{QMC}}$ is possible. This yields a statistical error $\propto \frac{1}{\sqrt{\mathcal{N}_{\mathrm{QMC}}}}$, that can be decreased very slowly with the computational time $\propto \mathcal{N}_{\mathrm{QMC}}$.

5.4.1 Time step error

As far as the time step τ error, this is simple to control, because unbiased solutions of the SLD equations of motion can be obtained by reducing τ to a sufficiently small value, within any reasonable integration scheme. In QMC we can perform the limit of $\tau \to 0$ for instance at fixed $\mathcal{N}_{\mathrm{QMC}}$. As long as there is no other source of bias (see next subsection) other than a finite $\mathcal{N}_{\mathrm{QMC}}$, we expect to have unbiased results for $\tau \to 0$ even within QMC, as explained in the following. After the time integration of the SLD equations in a small time interval τ the statistical noise associated to the forces is multiplied by the integration time τ in eq. (4.28), that is negligible compared to the stochastic part $\propto \sqrt{2T\tau}$ used to keep the temperature within the given target. In this way the systematic QMC error is expected to vanish *linearly* in τ and for $\tau \to 0$ the exact canonical distribution can be sampled,

$$\exp(-V(\boldsymbol{R})/kT) \tag{5.22}$$

where $V(\boldsymbol{R})$ is the BO-energy surface corresponding to a variational wavefunction $\psi_{\alpha,R}$ defined by several variational parameters α for given atomic positions:

$$V(\boldsymbol{R}) = \mathrm{Min}_{\alpha} \frac{\langle \Psi_{\alpha,\boldsymbol{R}} | H_{\boldsymbol{R}} | \Psi_{\alpha,\boldsymbol{R}} \rangle}{\langle \Psi_{\alpha,\boldsymbol{R}} | \Psi_{\alpha,\boldsymbol{R}} \rangle} \tag{5.23}$$

This error can be made in principle smaller by the "noise correction" scheme that was introduced in a previous work [47]. In practice, as it is shown in figure 5.2, the convergence in τ looks very well behaved and a reasonable accuracy is obtained also by using quite large time steps. In this case the mentioned noise correction scheme does not lead to a meaningful improvement probably because the use of a finite large $\Delta_0 = 8$ makes our dynamics more stable and less sensitive to the stochastic noise.

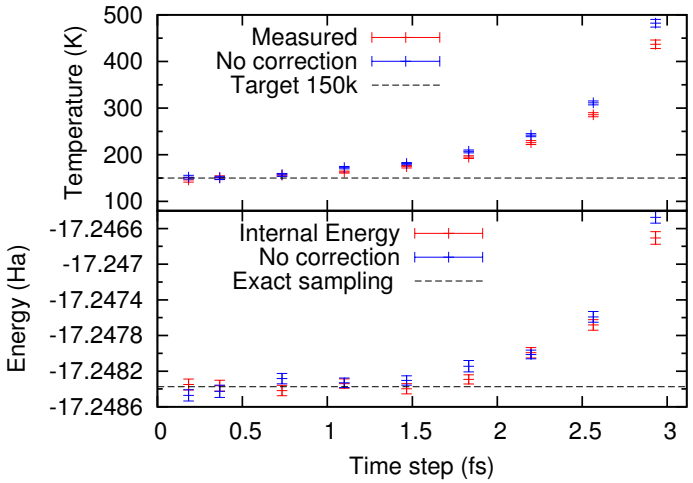

Figure 5.2: The measured temperature and internal potential energy versus the time step used in the dynamics of water monomer. The target temperature is set at 150K, the friction is 0.3 a.u. and $\Delta_0 = 8.0$. The minimum of the PES is $-17.24909(3)$ Ha. At each step of MD we perform $n_{\rm opt} \simeq 10$ steps of optimization, where all energy derivatives are estimated with $\mathcal{N}_{\rm QMC} = 20480$ samples generated by the Metropolis algorithm with 80 proposed attempts for each new sample (acceptance rate $\simeq 50\%$). The same plot, without using the noise correction (see eq. (4.33)) is also shown. The dashed lines indicate the "exact" results for the average temperature and the internal energy, the latter obtained by sampling exactly the fitted potential.

5.4.2 Error in sampling the BO energy surface

In the previous estimate of the error in τ we have to assume that, given the atomic positions, the energy derivatives of the BO energy surface $V(\boldsymbol{R})$ can be computed statistically, but without systematic bias. This means that the variational parameters α are exactly at the minimum energy condition that defines $V(\boldsymbol{R})$ in eq. (5.23), and only in this case the forces are unbiased. Unfortunately this condition is never met in a statistical optimization of the variational parameters and some approximation has to be done in practice. In the following we introduce the control parameter $n_{\rm opt}$. Each run of MD is obtained by performing several thousand iterations of the SLD discretized with a time interval τ. For each step of MD, we perform $n_{\rm opt}$ optimization steps of the electronic wave function with the SRH-CG method described in subsection 2.4.3. For $n_{\rm opt} \to \infty$ and fixed $\mathcal{N}_{\rm QMC}$, the optimized wavefunction converges to an approximate minimum of the BO energy surface where the energy derivatives, namely the atomic forces, differ at most by $\frac{1}{\sqrt{\mathcal{N}_{\rm QMC}}}$ from the exact BO ones. Therefore, we have found that it is convenient to study long, well equilibrated MD simulations at fixed statistical accuracy (i.e. $\mathcal{N}_{\rm QMC}$ fixed) and given τ, by increasing $n_{\rm opt}$ in a systematic way. In the optimization

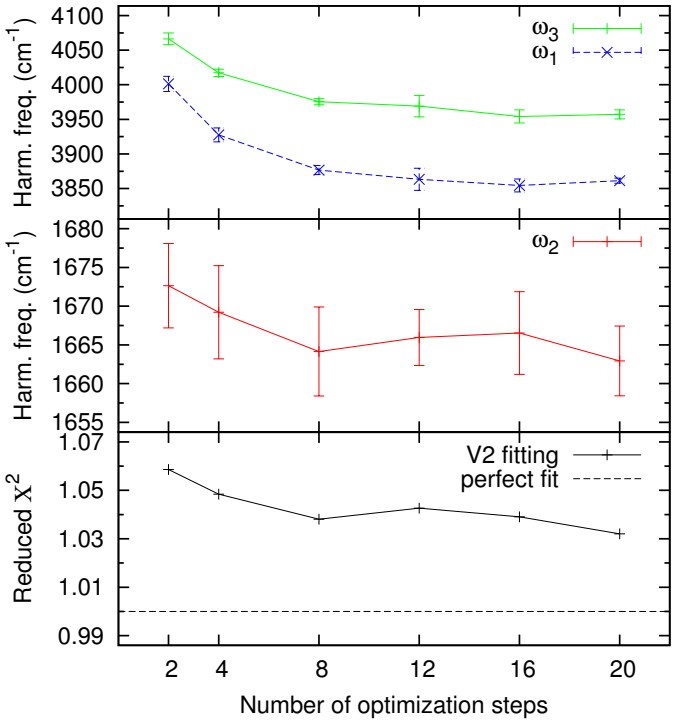

Figure 5.3: Harmonic vibrational frequencies of water monomer obtained by fitting V2 with samples generated by MD simulation at 150K as a function of QMC optimization steps $n_{\rm opt}$.

method we have used a given tolerance $\epsilon = 0.001$ in the inversion of the ill conditioned overlap matrix S corresponding to the chosen set of atomic orbitals used in the Jastrow and the determinantal part of our wave function. As now well established, the knowledge of this matrix S is extremely useful for an efficient optimization scheme (see ref. [26, 101]). Moreover, for the sake of a stable and systematic optimization technique we have also attenuated the wave function change predicted by the SRH-CG method by 50%. As it is shown in figures 5.3 and 5.4, the finite $n_{\rm opt}$ error is probably the most important one in QMC, because several optimization steps are necessary to achieve converged frequencies, especially the high frequency ones. Notice that, in this plot we use the fitting method, and the systematic error in τ, as defined in the previous subsection, is not present. Despite the slow convergence in $n_{\rm opt}$, it is quite evident that, by using $n_{\rm opt} \geq 10$ the simulations are still affordable and the error bars of the frequencies are quite small with reasonable computational resources, even for large $n_{\rm opt}$. In these plots the error bars have been evaluated by the standard Jackknife technique.

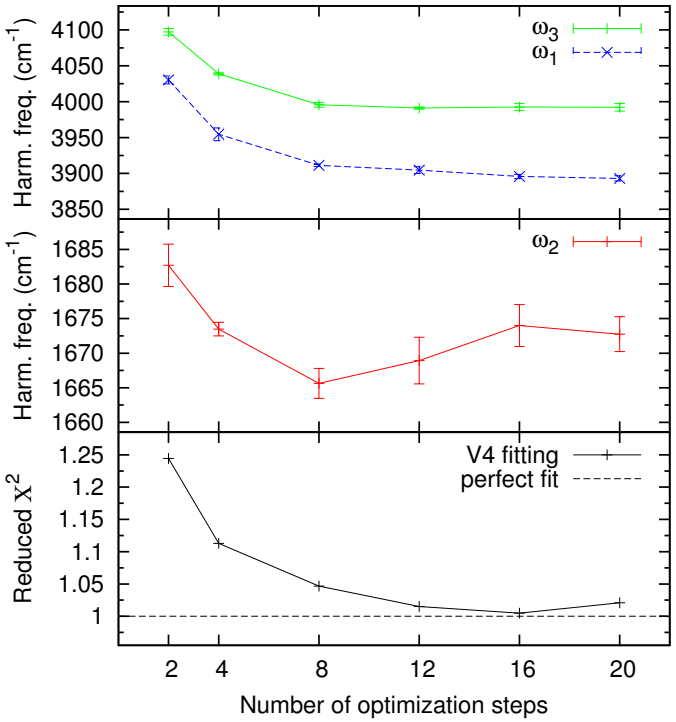

Figure 5.4: Harmonic vibrational frequencies of water monomer obtained by fitting V4 with samples generated by MD simulation at 1000K as a function of QMC optimization steps n_{opt}.

5.4.3 Residual QMC error due to a finite number of samples

Once all the above sources of error have been controlled, we are still left with the Monte Carlo statistical noise, namely the fact that we have to work with a finite number of samples $\mathcal{N}_{\text{QMC}}$ for each iteration of the dynamics. Among the various techniques considered in this chapter, this error affects mostly the method described in section 5.2.3.

As we have mentioned in the previous section this systematic error is affecting the evaluation of the energy derivatives by an error of order $\frac{1}{\sqrt{\mathcal{N}_{\text{QMC}}}}$. This means that the variational parameters α have a typical error of this magnitude $\frac{1}{\sqrt{\mathcal{N}_{\text{QMC}}}}$ during the MD simulation. However, since the energy at the minimum is affected quadratically by the error in the variational parameters, we can expect that all the frequency estimates, based only on energy expectation values, show a much smaller error inversely proportional to the number of sampling $\mathcal{N}_{\text{QMC}}$. This is readily seen in figure 5.5 where the calculation of frequencies is seen to converge linearly in $\frac{1}{\mathcal{N}_{\text{QMC}}}$. In this calculation it is also simple to identify the most important source of error, that is due to the stochastic estimation of the covariance matrix of the forces. Once we correct

this source of bias described in section 5.2.3, we see that this error is almost negligible (see figure 5.6).

Indeed, the other two fitting methods are also affected by the statistical correlation of the force components due to the finite QMC samples $\mathcal{N}_{\mathrm{QMC}}$. A better control of this bias is obtained by minimizing the function given in eq. (5.11) instead of the one corresponding to eq. (5.12) in section 5.2.1. In table 5.1, the two sets of frequencies labeled with 'cov' are done with the force covariance matrix, and differ very little from the simpler ones where different force components are assumed to be uncorrelated. This means that this approximation is almost correct in practice.

Finally, in order to quantify more clearly the statistical error, we have also performed a calculation with $\mathcal{N}_{\mathrm{QMC}}$ doubled and we have not been able to measure a sizable departure from the measured frequencies (see table 5.1, the harmonic frequencies at 150K, 'db' indicates double $\mathcal{N}_{\mathrm{QMC}}$).

In this thesis we have not shown how these different errors will affect larger systems, but it is clear that all the sources of errors, that we have described in detail in this section, can be systematically reduced by changing three parameters $\tau \to 0$, $n_{\mathrm{opt}}, \mathcal{N}_{\mathrm{QMC}} \to \infty$ in the same way we have done for the smaller system simulations.

5.5 Results

In this section we summarize our final results on the water monomer in table 5.1, as well as the calculation of frequencies for several systems of chemical interest within our Jastrow-Slater ansatz (see tables 5.3 and 5.4).

As far as the water monomer is concerned, the much smaller χ^2_{red} ($\chi^2_{\mathrm{red}} \sim 1$) obtained by fitting samples generated by our molecular dynamics is superior to the one obtained in the standard approach ($\chi^2_{\mathrm{red}} > 4$) with a fixed grid. In the data for the geometry, the H-O-H angle, a very sensitive parameter, has been improved significantly (0.11 ° closer to the experimental value) and the O-H bond length, a rather insensitive parameter, remains very good. The two sets of harmonic frequencies for the V_2 and V_4 fitting are both consistent within the error bars. According to our experience, the anharmonic fit including the V_4 term is preferred and more robust thanks to its more accurate parametrization of PES. Furthermore, for the same amount of computation time, the frequencies obtained by V4 fitting have statistical errors about one half smaller than those of the V2 fitting and even smaller than the ones of the simple fitting. In addition, the figures 5.3 and 5.4 show a very nice feature — the lower the frequency is, the smaller the corresponding error is. On the other hand, the calculation based on the covariance matrix method shows that this technique requires much more statistics than all the other methods because it requires a simulation much longer than the correlation time, that in turn, can be extremely large at low temperatures.

Method	temp. (K)	χ^2_{red}	equilibrium geom.		harmonic freq. (cm^{-1})			fundamental freq. (cm^{-1})		
			OH bond (nm)	H-O-H angle (°)	ω_2	ω_1	ω_3	ω_2	ω_1	ω_3
simple fitting V4	–	4.444	0.9565(1)	104.21(2)	1672(3)	3897(7)	3990(4)	1614(1)	3698(5)	3771(3)
MD fit V2	50	1.007	0.95632(3)	104.32(1)	1676(3)	3893(9)	3890(9)	–	–	–
MD fit V2 cov	50	0.672	0.95634(3)	104.32(1)	1677(3)	3892(9)	3989(8)	–	–	–
MD fit V4	150	1.005	0.95605(3)	104.31(2)	1669(8)	3895(11)	3982(12)	1614(90)	3738(178)	3628(210)
MD fit V4 db	150	1.002	0.95596(3)	104.34(2)	1666(6)	3892(5)	3986(10)	1670(38)	3697(99)	3661(174)
MD fit V4	1000	1.021	0.95597(3)	104.31(3)	1673(3)	3893(3)	3992(4)	1615(2)	3689(5)	3754(5)
MD fit V4 cov	1000	0.681	0.95598(2)	104.31(4)	1672(2)	3892(4)	3993(3)	–	–	–
covariance matrix	300	–	–	–	1673(24)	3765(28)	3912(20)	–	–	–
Experiment [102]	–	–	0.95721(30)	104.522(50)	1648.47	3832.17	3942.53	1594.59	3656.65	3755.79

Table 5.1: Equilibrium geometries and vibrational frequencies of the water monomer obtained by all the methods described in section 5.2.

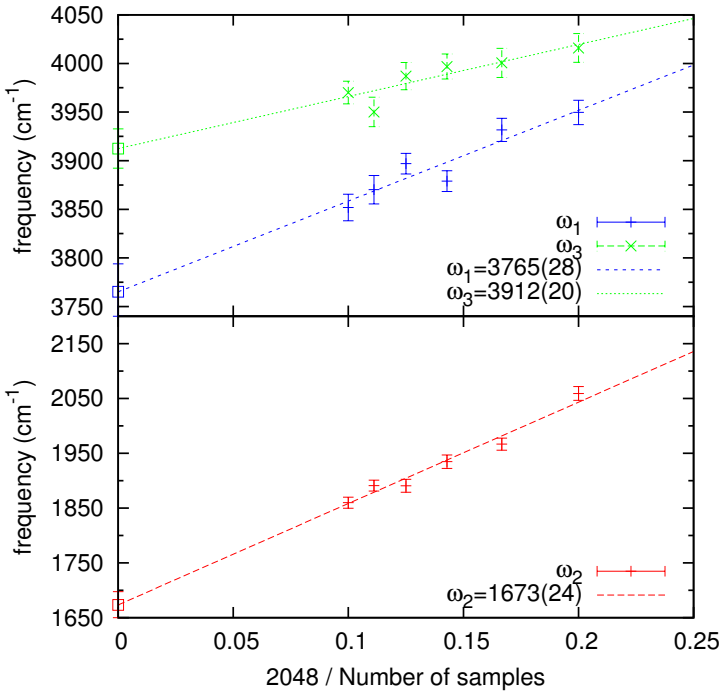

Figure 5.5: Vibrational frequencies of the water monomer obtained with the force-force covariance matrix without noise correction as a function of $1/\mathcal{N}_{\mathrm{QMC}}$ from MD simulation at 300K. The prefactor 2048 on the slope of the linear extrapolation doesn't change the intercept for the limit $1/\mathcal{N}_{\mathrm{QMC}} \to 0$.

Name	χ^2_{red}	Grid points	$\sigma(E)$ (Ha)	$\sigma(F)$ (a.u.)	Cpuh BG/Q
simple fitting V4	4.444	59	9.0E-5	1.5E-4	22.2k
MD fit V2 50K	1.007	13784	3.5E-3	6.2E-3	24.0k
MD fit V4 1000K	1.021	13784	3.7E-3	6.3E-3	23.5k

Table 5.2: Specifications of the fitting with manually chosen grid and MD.

The other four molecules we considered are divided into two groups — AB_2 and AB_3. In the first group, we use the same parametrization of the Hessian as in H_2O because this type of nonlinear AB_2 molecules is very similar to H_2O. The other group consists of two non-planar AB_3 molecules. The Hessian matrices used in both groups are simplified by using the molecular symmetries in order to improve the accuracy. We choose both JHF and JAGP types of wave function to compute the vibrational frequencies and compare them with CCSD(T) and experimental data from NIST database [103, 104].

H_2S molecule has 8 valence electrons which is exactly the same as the water monomer. In table 5.3, the equilibrium geometry obtained with JHF wavefunction but without the optimized exponents (JHF_nooptZ) gives the worst values. After optimizing the exponents, both geometry and frequencies

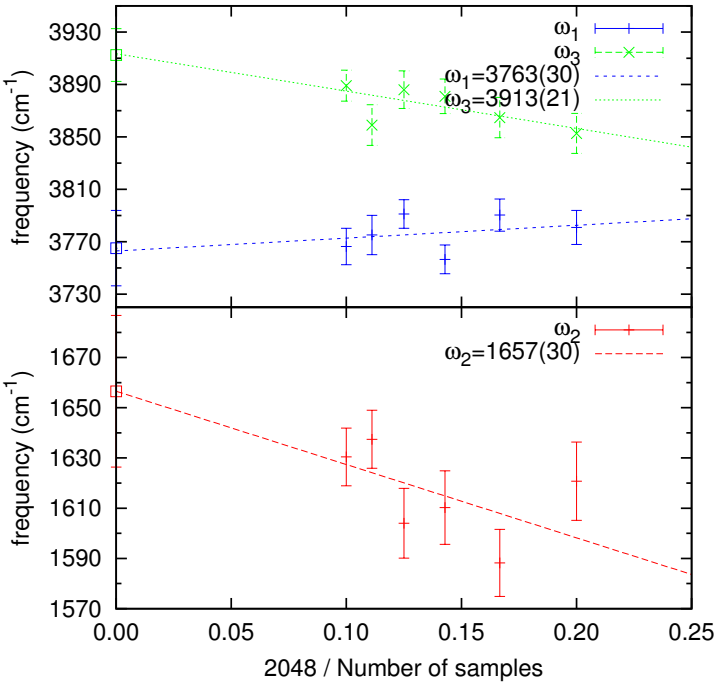

Figure 5.6: Vibrational frequencies of water monomer obtained with force-force covariance matrix plus noise correction as a function of $1/\mathcal{N}_{\mathrm{QMC}}$ from MD simulation at 300K.

are improved, in agreement with the conclusions of ref. [50]. For this reason, we have optimized all the exponents at equilibrium positions for all the remaining molecules studied. By using the JAGP wavefunction, the H-S-H angle of the equilibrium geometry is further improved as compared with experiments and the fundamental frequencies lower by $\sim$10 cm^{-1} for each mode in the best calculation reported in table 5.4. Compared with the experimental data, its RMS difference from fundamental frequencies is only 19 cm^{-1}.

SO_2 molecule has 18 valence electrons and requires a calculation much heavier than H_2S. Similar to H_2S, the use of JAGP wave function provides better results than those obtained with JHF wave function. Its RMS difference of fundamental frequencies from experimental values is 50 cm^{-1}. Even though CCSD(T) frequencies are much closer to the experiments, we should notice that they are harmonic frequencies rather than fundamental ones and the equilibrium geometry of CCSD(T) (+0.023 nm for S-O bond and -1.1 ° for O-S-O angle) is much worse than our values (-0.013 nm for S-O bond and +0.5 ° for O-S-O angle).

Both NH_3 and PH_3 have 8 valence electrons. Since we do not include the anharmonic correction for the fit, we compare their harmonic frequencies with the corresponding ones calculated with CCSD(T). For both molecules, the frequencies obtained with JAGP wavefunction are again better than those

name	Basis set	equilibrium geom.		GS energy
		bond (nm)	angle (°)	Hartree
H_2S		S-H	H-S-H	
JHF_nooptZ	S:(7s,8p,1d)/{4} H:(3p,1d)/{1}	1.33216(6)	92.00(3)	-11.40043(2)
JHF	S:(5s,4p,1d)/{4} H:(3p,1d)/{1}	1.33180(5)	92.42(2)	-11.40902(2)
JAGP	S:(5s,4p,1d)/{5} H:(3p,1d)/{3}	1.33237(5)	92.36(2)	-11.41154(2)
CCSD(T)	aug-cc-pVTZ	1.3419	92.299	–
Exp.	–	1.328	92.2	–
SO_2		S-O	O-S-O	
JHF	S:(6s,6p,1d)/{6} O:(6s,7p,1d)/{5}	1.4180(2)	119.91(6)	-42.27474(8)
JAGP	S:(6s,6p,1d)/{6} O:(6s,7p,1d)/{5}	1.4193(2)	120.06(6)	-42.28255(7)
CCSD(T)	aug-cc-pVTZ	1.4553	118.367	–
Exp.	–	1.432	119.5	–
NH_3		N-H	H-N-H	
JHF	N:(6s,6p,1d)/{4} H:(3p,1d)/{1}	1.00886(3)	107.01(1)	-11.74967(2)
JAGP	N:(4s,4p,1d)/{4} H:(3p,1d)/{1}	1.01014(3)	106.53(2)	-11.7512(3)
CCSD(T)	aug-cc-pVQZ	1.0128	106.541	–
Exp.	–	1.012	106.67	–
PH_3		P-H	H-P-H	
JHF	P:(6s,7p,1d)/{4} H:(3p,1d)/{1}	1.40925(7)	93.72(1)	-8.34788(1)
JAGP	P:(6s,7p,1d)/{4} H:(3p,1d)/{1}	1.41067(5)	93.580(9)	-8.34899(1)
CCSD(T)	cc-pVTZ	1.4186	93.501	–
Exp.	–	1.421	93.3	–

Table 5.3: Total energies and geometries of the structural minimum of H_2S, SO_2, NH_3 and PH_3 molecules. Potential V4 is fitted for H_2S and SO_2 while only V2 are fitted for NH_3 and PH_3.

corresponding to JHF wavefunctions. Both NH_3 and PH_3 have equilibrium geometries very close ($<0.01\,$nm for N/P-H bond and $<0.1\,^\circ$ for H-N/P-H angle) to those obtained by CCSD(T). The RMS difference of harmonic frequencies from CCSD(T) values are $46\,\mathrm{cm}^{-1}$ and $26\,\mathrm{cm}^{-1}$ for NH_3 and PH_3, respectively.

5.6 Conclusions

In this chapter, we have studied the performances of a recently developed molecular dynamics scheme based on quantum Monte Carlo. We have considered particularly simple systems by targeting the vibrational properties of simple molecules, that are well studied and understood with well established quantum chemistry methods. In this way we have been able to identify and systematically control all possible sources of systematic error which may affect this molecular dynamics. The main conclusion of this chapter is that the statistical error (the finite number of samples $\mathcal{N}_{\mathrm{QMC}}$ used for each iteration

name	T(K)	$\chi^2_{\rm red}$	type	freq. (cm^{-1})			
H_2S				A_1	A_1	B_2	
JHF_nooptZ	1000	1.077	harm.	1248(5)	2774(6)	2789(6)	
			fund.	1217(4)	2652(4)	2656(7)	
JHF	1000	1.018	harm.	1246(1)	2756(3)	2772(4)	
			fund.	1215(1)	2639(2)	2654(4)	
JAGP	1000	1.049	harm.	1235(3)	2752(3)	2767(4)	
			fund.	1206(2)	2630(1)	2643(2)	
CCSD(T)	–	–	harm.	1206	2711	2727	
Expt.	–	–	–	1183.0	2615.0	2626.0	
SO_2				A_1	A_1	B_2	
JHF	800	1.024	harm.	570(4)	1214(8)	1441(11)	
			fund.	563(3)	1211(5)	1429(8)	
JAGP	800	1.006	harm.	559(2)	1204(8)	1445(13)	
			fund.	557(2)	1193(7)	1426(10)	
CCSD(T)	–	–	harm.	506	1136	1332	
Expt.	–	–	–	517.7	1151.4	1361.8	
NH_3				A_1	E	A_1	E
JHF	50	1.006	harm.	1064(3)	1712(4)	3523(8)	3651(5)
JAGP	50	1.021	harm.	1098(7)	1709(6)	3523(7)	3640(7)
CCSD(T)	–	–	harm.	1159	1673	3476	3607
Expt.	–	–	–	950.0	1627.0	3337.0	3444.0
PH_3				A_1	E	A_1	E
JHF	50	1.025	harm.	1048(6)	1181(3)	2445(8)	2461(13)
JAGP	50	1.022	harm.	1045(4)	1178(3)	2431(8)	2437(4)
CCSD(T)	–	–	harm.	1018	1142	2412	2421
Expt.	–	–	–	992.0	1118.0	2323.0	2328.0

Table 5.4: Vibrational frequencies of H_2S, SO_2, NH_3 and PH_3.

of MD) and the time discretization error due to finite τ can be easily pushed to negligible values. On the other hand, we have found that the most difficult bias comes from the requirement to satisfy the BO constraint along the dynamics. We have found that it is important to employ a sufficiently large number $n_{\rm opt} \gtrsim 10$ of energy optimization for each step of molecular dynamics, in order to satisfactorily fulfill the BO constraint. Since the computational time is proportional to $n_{\rm opt}$, in the present scheme this is probably the most difficult bias to remove. Despite this difficulty, the calculation remains still feasible and can be extended to large systems like the liquid water [105] which will be discussed in the next chapter by using massively parallel supercomputers.

This work is also relevant to establish vibrational frequencies in complex electronic systems. Among the three methods that we have used for evaluating vibrational properties, the fitting method with samples generated by molecular dynamics gives the best results for the same amount of computation cost. Compared with the standard fitting procedure of ref. [96], it is easier and more systematic to set up and use, and yields better distributions of the configurations around the equilibrium structure, thus improving the quality of fit as well as a geometry closer to the experiment. Even though our method

based on the force-force correlations is the most direct and simplest approach, it usually requires much more statistics. All methods, apart from the one containing the anharmonic corrections, have favorable scaling with the system size, and are in principle very promising because can be extended to large systems, as well as generalized to the calculation of phonons in solids. However we have seen that, in order to neglect anharmonic effects, we have to work with so small temperatures that it is already very difficult to simulate a slightly larger system (such as the water dimer). On the other hand, we expect that the method which includes in the fit also the anharmonic corrections, should work also for larger systems, despite the difficulty to represent the V4 term with a number of parameters scaling with the fourth power of the number of atoms. Also for this reason this method is very difficult to implement in practice for large systems, and therefore we have limited our study to molecules containing at most four atoms.

A very interesting feature that we have noted in the estimation of vibrational frequencies by QMC, is that the small frequencies are much less biased by the systematic errors in our tests. This is really promising because small frequencies are often more interesting as they characterize the intermolecular interactions, whereas the high frequency modes are determined by the well understood intra-molecular properties. Moreover, we have systematically found that the use of the JAGP wave function in place of the more commonly-adopted Jastrow-Slater paradigm, improves significantly the calculation of both equilibrium structures and vibrational frequencies, basically without extra computational effort.

As well known QMC scales very well with system size and, once the problem of including anharmonic effects will be solved at least in an approximate way, say by self-consistent harmonic approximation [106], the computation of vibrational frequencies in large systems will be possible with a reasonable cost. In addition, we have shown that the present molecular dynamics can be extended to large systems already at present [35, 105], provided the few systematic errors are removed and especially the BO constraint are satisfied in the way we have carefully described in this chapter.

Chapter 6

Liquid water

6.1 Introduction

Water is one of the most important chemical substances on Earth. For its crucial role in many biological, chemical and environmental processes, understanding its microscopic structure is an issue of long-standing interest. Though it has been extensively studied for decades both at the theoretical and experimental levels under a wide variety of conditions, its microscopic nature still remains elusive. In 2005, the Science magazine included the problem "What is the structure of water?" on the list of the 125 most important questions [107] of modern science.

Water molecule is a polar molecule consisting of an oxygen and two hydrogen atoms. Since it has a large dipole moment, one might expect that the interaction between water molecules will be dominated by electrostatics and induction (polarization) with the dispersion interaction (van der Waals (vdW) interaction) playing a minor role. In liquid water, a proton is exposed to a high electronegative oxygen belonging to the next water monomer and thus forms a so called *hydrogen bond* (HB) which is not a true bond but a particularly strong dipole-dipole attraction. Water's many anomalous properties, which arise from the unique hydrogen bonding pattern and long range dispersion interaction, make it one of the most challenging liquids to model theoretically and a complete understanding of its structure and dynamics is still lacking.

Complementary to experiments, computer simulations have shown its power in quantitative characterizations and advancing qualitative understandings of water at ambient and extreme conditions. As was mentioned in the thesis introduction that a compromise between simulation scale and accuracy is always present, various theoretical descriptions of liquid water are employed for different target problems. For example, in order to demonstrate the process of protein folding in the water solution, tens of thousands up to millions of water molecules should be simultaneously simulated for a very long period of time. Nowadays, such simulations are accessible only with force fields,

given a reasonable computational cost. Well-established empirical force fields based on two-body interactions are able to reproduce properties of water in a wide range of pressures and temperatures. However, many of them are parametrized against experimental data or ab initio simulations [108], which raises the question of their predictive power for situations different from those experiments used by the parametrization. Moreover, since the many-body interactions are represented effectively by altering the two-body interaction, a direct physical interpretation of the liquid water at the molecular level is not straightforward. Those cooperative effects can be included to a certain extent by parametrizing many-body potentials against ab initio potential energy surfaces of small water clusters, but the inability to describe reactions remains and such parametrization substantially increases the difficulty of fitting, as well as the computational cost for many-body interactions.

In order to avoid the intrinsic drawbacks of force fields, ab initio approaches like DFT are also used to perform molecular dynamics simulation of liquid water. Nowadays, DFT-based MD simulations are routinely used to study several properties of condensed matter systems at ambient conditions up to extremely high pressures and temperatures [59, 60, 61, 62, 63, 64, 65, 66, 67], and represent a quite reliable tool to predict new materials, sometimes more effective or at least much cheaper than experiments. The first ab initio simulations of liquid water were attempted [109, 110] just a few years after the invention of Car-Parrinello molecular dynamics [58] (CPMD). Apart from the first excitement, it was soon clear that this substance was not so easy to understand with computer simulations. The oxygen-oxygen (O-O) radial distribution function (RDF), as far as the positions of the peaks were concerned, showed a pretty good agreement with the experiments available at that time but the overall shape given by the simulation was very overstructured. After these first attempts, many other standard DFT-based simulations have been reported, but they still lack a satisfactory agreement with the experiments. The equilibrium density given by simulations at ambient pressure (1 atm $\sim 10^{-4}$ GPa) is far from the expected 1 g/cm^3 and the simulated diffusion is one order of magnitude slower than those expected by experiments [111]. Moreover, it was shown that the evaluation of properties of liquid water in a periodic box of 32 molecules are affected by small but not negligible size effects [112, 113] and the structural properties [112, 114] are biased by too large values of the fictitious electronic mass in CPMD. Last but not the least, the solidification of water occurs at a temperature which is unrealistically large ($\sim$ 410 K), so that all the present DFT simulations of liquid water should be considered supercooled metastable phases [111, 115].

The DFT evaluations appear to be influenced not only by the choice of the functionals [113, 65], but also, within a given functional, by other setups of the calculations such as the pseudo-potential [111] and the basis set [116, 117]. The mostly used functionals for liquid water are those based on the generalized gradient approximation (GGA) to DFT (often PBE or BLYP density function-

als), yielding an over-structured water at ambient conditions. The accurate description of the exchange by using the computationally-more-expensive hybrid functionals was shown to improve only slightly the results [118, 119], probably due to their poor description of the long-range interaction forces. On the other hand, in order to overcome the well-known difficulty of DFT in describing long range interaction forces, the inclusion of empirical dispersion terms has been attempted either by using empirical pairwise interatomic potentials of the C_6R^{-6} form in the total energy [120, 121], or by adopting dispersion-corrected atom-centered potentials [122]. Even though these methods depend on external tunable parameters, all of them have provided improvements in some cases [123, 124, 125, 126, 127, 128] and are strongly dependent of the functional. Another way to improve DFT functionals is to add corrections obtained by performing many calculations of different configurations with accurate approaches like QMC or post-Hartree-Fock methods and then fitting the parameters with those results through machine-learning [129, 130]. Its application on the liquid water shows a significant improvement [131]. Finally, quantum effects have been shown to have an important role, as they lead to a more accurate description of the hydrogen-bond, improving the agreement with experimental data [132, 133, 134, 135, 136, 137] by broadening the RDF. Most of these achievements are very promising, but at this stage it is difficult to understand whether these improvements, mostly sacrificing the ab-initio approach, have a real predictive power.

A recent accurate experiment of X-ray diffraction [138] has raised again the reliability issue of present ab-initio molecular dynamics schemes, as it was found that, surprisingly, the position of the first peak was shifted towards larger distances. This observation is in excellent agreement with a recent extensive and independent review on the experimental structure of bulk water [139]. Indeed in ref. [139] a new methodology to interpret the experimental data is employed and the shifts of the intermolecular O-O, O-H and H-H peaks positions with respect to the old experimental references [140] are reported. These results are particularly embarrassing for ab-initio simulation because, the use of the PBE functional — until recently one of the most popular in this field — is being now questioned in favor of different functionals, like BLYP or B3LYP, that look clearly closer to present experiments [113], but less ab-initio from some point of view as their expression of the correlation energy is not universal and depends indeed on the external electron-ion potential, namely semi-empirical. In other words, we believe that, in order to make some progress for clarifying the present discrepancies between experiments and numerical simulations in this field, it is now timely to use a completely different approach for the following reasons:

- DFT — within its practical implementation with approximate functionals — is not systematically improvable, and cannot be validated by a variational principle as in the wave-function based approach. In particular it is not possible to judge the accuracy of a given functional without

knowing the experimental result, thus DFT lacks most of the predictive power of a usual wave function based approach. In our opinion this drawback is particularly limiting the present ab-initio functionals, because we believe that the main task of computer simulations should be also to predict the properties of materials when experiments are not possible (e.g. at very large pressures) and not only to reproduce them, with an appropriate and somehow arbitrary choice of the functional.

- the computing performances, especially in massively parallel architectures, are constantly growing with an impressive speed, as an exascale (10^{18} flops) supercomputer is expected much before 2020, and supercomputer architectures are becoming more and more suitable for statistical techniques rather than for deterministic methods such as DFT. Thus a wave function approach based on QMC is now becoming practical and competitive with DFT.

In this chapter we have employed the first ab-initio molecular dynamics simulation within a fully wave-function approach based on quantum Monte Carlo. We adopt the Born-Oppenheimer approximation, neglect the quantum effects on protons, and use the less accurate QMC approach — the variational Monte Carlo (VMC) — by means of a Jastrow Slater many-body wave function, to deal in the simplest way with the exact exchange and the electron correlation. Despite all the above approximations, a significant improvement in the description of liquid water has been achieved. In particular we have obtained that the O-O RDF, $g_{\mathrm{OO}}(r)$, is considerably less structured compared with DFT calculations of the same type (with no proton quantum effects). Moreover, it is also worth to emphasize that the position of the first peak is now in perfect agreement with the most recent and accurate experiments, a fact that was indeed found with a simulation dated before the new experimental data were distributed [141].

In this chapter, we first studied the basic setups of our simulation including wavefunction and MD simulation details in section 6.2. Then in section 6.3, we discussed the RDFs of liquid water in which the peak positions show perfect agreement with the experiments. In the following sections 6.4 and 6.5, the many-body interaction nature of the liquid water was displayed and the hydrogen bond related properties were analyzed within our VMC molecular dynamics simulation. Finally, in section 6.6, the discussion and conclusion are drawn.

6.2 Simulation specifications

In this section we first specify the wavefunction ansatz chosen for characterizing liquid water and then justify the basis set choice which is a compromise between accuracy and computational cost. In the last part, the setups of our MD simulation will be explained.

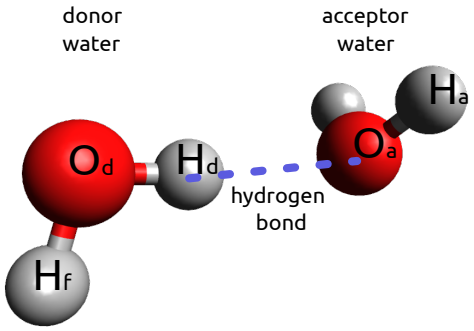

Figure 6.1: The geometry of the water dimer in its equilibrium configuration.

The many-body wavefunction function of the liquid water in our simulation is expressed by means of the following Jastrow Slater ansatz

$$\Psi(\boldsymbol{r}) = \Psi_{\mathrm{SD}}(\boldsymbol{r})J(\boldsymbol{r})\,. \tag{6.1}$$

similar to (3.1) which has an antisymmetric part now represented by a Slater determinant. The molecular orbitals which appear in the Slater determinant are expanded in a local atomic basis of the Gaussian/Slater type ((5s,4p,1d) contracted with 5 hybrid orbitals (introduced in section 3.3) on the oxygen and (3s,1p) contracted with 3 hybrid orbitals on the hydrogen). The two core electrons of the oxygen atoms have been described using the scalar-relativistic energy consistent pseudo-potential of Burkatzki et al. [100]. The Jastrow factor $J(\boldsymbol{r})$ (see section 3.5) consists of one-body, two-body and three-body terms only. The four-body terms are all excluded in order the save the number of parameters optimized during the simulation. The three body Jastrow is also expanded on an atomic basis set of Gaussian/Slater type with (3s,2p) on the oxygen and (2s,2p) on the hydrogen (no contraction used for the Jastrow). All the parameters except the exponents of the atomic orbitals are optimized with VMC energy minimization described in section 2.4.

In order to show the quality of the wavefunction ansatz and the accuracy of the chosen basis set, we have considered the water dimer, the simplest compound which owns the O-H hydrogen bond, as our benchmark system. In figure 6.2, the dissociation energy curve is faithfully reproduced, especially in the region of interest ($\sim$3 Å). The binding energy is not far from the experimental value ($\sim$5.0 Kcal/mol [142, 143]) and comparable with state-of-the-art post Hartree-Fock calculations (5.02 Kcal/mol for CCSD(T) [144, 145]). If the so called lattice regularized Diffusion Monte Carlo [11] (LRDMC) method is employed, this accuracy is further improved and the binding energy coincides with the best known theoretical results. LRDMC guarantees the full variational upper bound energy even in the case that pseudo-potentials are used [12].

In principle a more accurate MD simulation is possible with larger basis

set and LRDMC, but at present the computer resources required for such a simulation with tens of molecules remains out of reach. We'd like to remark here that, in molecular dynamics, the total energy is not important because what really matters is the relative forces between a couple of atoms, namely it suffices to compute accurate energy derivatives of the binding energy rather than the energy itself. As it can be seen in the inset of figure 6.2, by aligning the binding energy curve at their minimum (this shift doesn't affect its derivative), we obtain a rather good description of the binding shape in the region of interest $2 < R < 3.2$ with an acceptable error out of this region.

After establishing the wavefunction, the other details of our MD simulation are specified as follows. We apply the molecular dynamics driven by forces evaluated by quantum Monte Carlo shown in chapter 4. This MD scheme was introduced recently for the simulation of liquid hydrogen at high pressures [35] and systematically validated through the calculation of vibrational frequencies of small molecules [146]. We have performed a simulation of 32 waters at 300K (∼310K, the measured temperature) in the canonical ensemble (NVT) at experimental density $1\,\mathrm{g/cm^3}$, thus in a cubic cell with box side $L = 9.86\,\text{Å}$ and periodic boundary condition (PBC). Since the value of the mass is immaterial for static equilibrium properties, we have chosen the hydrogen mass equal to the oxygen one, and we have collected about 5000 iterations (that we can estimate to roughly correspond to more than 40 ps of simulation in a standard Newtonian MD simulation). At each iteration we optimize all variational parameters with 9 steps of efficient energy optimizations based on the SRH-SNF optimization scheme described in section 2.4 with 250 significant parameters in each iteration. At each optimization step, by applying the locality approximation (LA) (see section 3.3) with a cutoff $R_{\mathrm{MAX}} = 2.38\,\text{Å}$,[1] the total number of variational parameters decreases from ∼70000 to ∼12000, a huge reduction, without losing much accuracy (less than 0.1 mHa per water molecule). We have done several tests by further optimizing the wavefunction after 9 initial steps during the dynamics. The negligible energy, less than 0.1 mH per atom, gained by the extra optimization confirms that it is possible with this scheme to correctly follow the Born-Oppenheimer energy surface, namely the variational wave function remains at the minimum possible energy during the time evolution of the atomic positions.

In our simulation, we have started the simulation by using equilibrated configurations generated by DFT molecular dynamics with BLYP functional.

6.3 Radial distribution function

In statistical mechanics, the radial distribution function (RDF), also called pair correlation function, $g(r)$ in a system consisting of particles (atoms,

[1] This distance corresponds to the position of second valley in O-H RDF, namely all the hydrogen bond (second peak) are within this range.

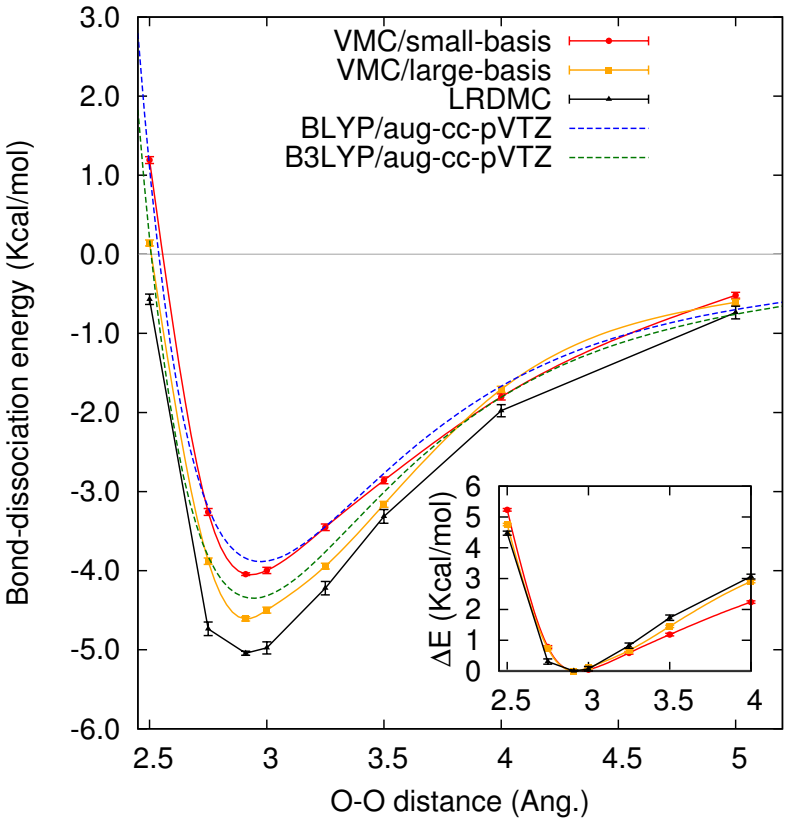

Figure 6.2: Dissociation energy of the water dimer, plotted as a function of the oxygen-oxygen distance, studied with VMC (the small-basis is the one used for the dynamics and the large basis is (9s,9p,3d,2f) contracted with 12 hybrid orbitals on the oxygen and (7s,6p,2d) contracted with 4 hybrid orbitals on the hydrogen for the determinant and (5s,4p,2d,1f) on the oxygen and (3s,2p,1d) on the hydrogen for the Jastrow), and LRDMC (that alleviates the basis set issue and is almost independent on the choice of the small or large basis). For a comparison, we report also the dissociation curve for DFT/BLYP and DFT/B3LYP, both with aug-cc-pVTZ basis set.

molecules, clusters, etc.), describes how density varies as a function of distance from a reference particle. This function is of fundamental importance in thermodynamics because various macroscopic thermodynamic quantities can usually be determined from $g(r)$.

The RDFs of liquid water obtained from our VMC-based and DFT/BLYP molecular dynamics simulation are reported in figure 6.3, in comparison with experimental results. Even though the BLYP functional describes the water dimer (the simplest system displaying the hydrogen bond) with a reasonable accuracy, comparable with the one obtained by our QMC scheme, as discussed in section 6.2, the peak positions and shapes of the RDFs are substantially different on the target 32 water system. Moreover we have verified that, within

this Langevin scheme, the correlation time estimated by the convergence of the RDF with much longer DFT calculations, is less than 2000 iterations, and therefore we are confident that our liquid is well equilibrated and the results are reliable. At variance with the Newtonian dynamics, our advanced method is capable of accelerating the slow modes while not affecting the fast modes by making use of the friction matrix (see section 4.5) and thus reduces a lot the auto-correlation time.

We see in figure 6.3 that these first results are very encouraging. Despite the noise, the outcome is very clear that the $g_{\mathrm{OO}}(r)$ is much closer to experiments than the corresponding DFT calculations. Not only the radial distribution function is much less overstructured (first valley and second peak) but also, as discussed in the introduction, the position of the first O-O peak is almost indistinguishable from the most recent experiments [138, 139] (ref. [139] introduced a novel method for removing the inelastic scattering from neutron data compared with the ref. [140]). The first peak of O-O RDF refers to all the nearest neighbor oxygen pairs. Between those two oxygen atoms, there are always an inner molecular O-H bond and an inter molecular O--H hydrogen bond which are represented by the first and second peaks of O-H RDFs. Since the O-H--O angle is distributed around 167° (see figure 6.6), all three atoms stay almost on a line. For this reason, the first peak position of the O-O RDF should be affected by the first and second peak positions of the O-H RDF. Indeed in the middle panel of figure 6.3, the second peak of the O-H RDF given by VMC simulation shifts to the right side of the corresponding DFT/BLYP peak while the first peaks of both O-H RDFs are very close. This indicates that in the liquid water, the strength of the O--H hydrogen bond has significant difference between DFT/BLYP and VMC. For the same reason, the second peak of the H-H RDF given by VMC also stands at the right side of the corresponding DFT/BLYP peak since it is the distance between the hydrogen atom on a hydrogen bond and the hydrogen atoms belonging to the water molecule whose oxygen participates the same hydrogen bond. It should be noted that the peak positions of the RDFs of DFT/BLYP calculations are always closer to the old experiment [140] while those given by VMC are closer or almost identical to the new experiment [139].

The improvement of RDFs implies a characterization of liquid water in VMC better than the usual DFT functionals due to a faithful description of the electronic correlation. In fact in ref. [147], it had been shown that the position of the first peak could be improved after better recovering the electronic correlation by employing the simplest (MP2) post-Hartree-Fock technique. Moreover, quantum effects should further improve the shapes of RDFs by broadening and lowering the peaks without shifting their positions, as it was shown before within DFT in ref. [132]. Despite this rather controversial fact [137, 148], because of the lack of long enough ab-initio simulations equipped with quantum effects, our results seem to support the claim made in ref. [132] about the relevance of proton quantum corrections in water. Indeed our RDFs

for classical protons remain sizably different from experiments, as far as the broadening and the heights of the first peaks are concerned, especially for what concerns the $g_{\mathrm{OH}}(r)$ and $g_{\mathrm{HH}}(r)$ radial distribution functions, where quantum effects are expected to be much more important.

In order to avoid possible size effects we have studied the position of the first peak with a much shorter simulation (∼600 steps, corresponding to about 5ps) with 64 molecules. Since our method equilibrates rather smoothly with the length of the simulation, it is possible to obtain this information also for the largest size by simple linear extrapolation. In this way we have found that the position of the first peak in $g_{\mathrm{OO}}(r)$ remains very close to the smaller 32 water molecule simulation, supporting the validity of our main finding.

6.4 Many water molecules' interaction

Within our method, we have obtained considerably better results, not only compared with the standard GGA functionals, such as BLYP and PBE. This is remarkable, because, within our approximation, the two-water interaction was basically dealt with the same degree of accuracy as the BLYP functional. This implies that the accurate description of the many water energy surface, and probably the long distance interactions — usually missed or incorporated by hands at the DFT level — should be important to close the gap with experiments.

In order to further support this issue we show in figure 6.4 the interaction between two water monomers

$$(\boldsymbol{f}^i_{\mathrm{H_2O}} - \boldsymbol{f}^j_{\mathrm{H_2O}}) \cdot (\boldsymbol{r}_{\mathrm{O_i}} - \boldsymbol{r}_{\mathrm{O_j}})/|\boldsymbol{r}_{\mathrm{O_i}} - \boldsymbol{r}_{\mathrm{O_j}}| \tag{6.2}$$

namely the repulsive force acting on the O-O axis, where the net molecular force reads

$$\boldsymbol{f}_{\mathrm{H_2O}} = \boldsymbol{f}_{\mathrm{H_1}} + \boldsymbol{f}_{\mathrm{H_2}} + \boldsymbol{f}_{\mathrm{O}}\,. \tag{6.3}$$

This interaction described by VMC is quite different from the BLYP prediction when two monomers are in the liquid water but almost the same in the vacuum. In the dimer panel of figure 6.4, namely dimer in the vacuum, even though DFT/B3LYP curve slightly differs (less than 4×10^{-4} a.u.) from the DFT/BLYP and VMC curves, the forces computed by all three methods are almost the same, which indicates that two-water interaction described by VMC has a similar level of accuracy as DFT which confirms our claim in section 6.2 based on the fact that the energy derivatives are important instead of the energy itself.

In figure 6.5, we show also the interaction between two ions A and B (oxygen or hydrogen) defined similarly to eq. (6.2) as

$$(\boldsymbol{f}_{\mathrm{A}} - \boldsymbol{f}_{\mathrm{B}}) \cdot (\boldsymbol{r}_{\mathrm{A}} - \boldsymbol{r}_{\mathrm{B}})/|\boldsymbol{r}_{\mathrm{A}} - \boldsymbol{r}_{\mathrm{B}}|\,. \tag{6.4}$$

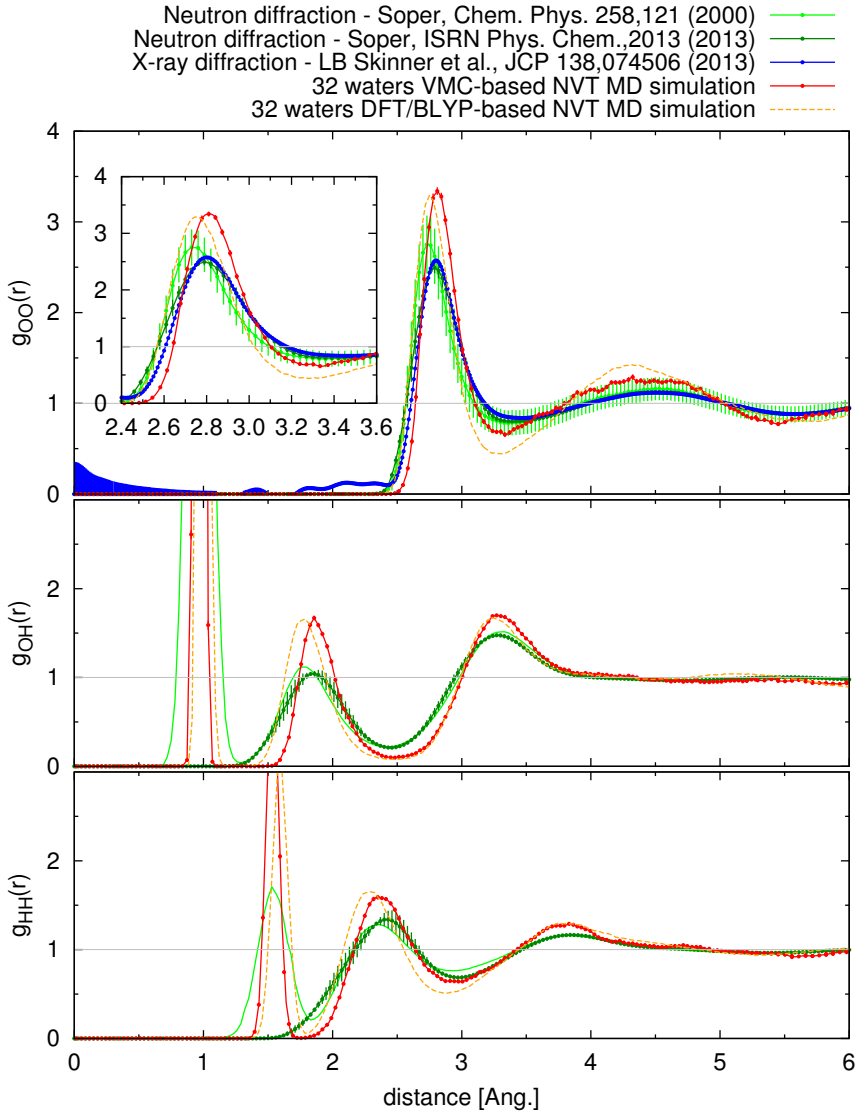

Figure 6.3: Radial distribution function obtained with 32 waters by a VMC-based dynamics in NVT ensemble (see text) as compared with DFT/BLYP and experiments (X-ray [138] or neutron diffraction [140, 139]) (a): Oxygen-Oxygen, (b) Oxygen-Hydrogen, (c) Hydrogen-Hydrogen.

The almost negligible difference between DFT/BLYP simulations with 32 and 64 water molecules indicates that the finite size effect is quite small in our DFT simulation. On the other hand, the curves of DFT/BLYP and VMC show substantial differences in their description of the interaction.

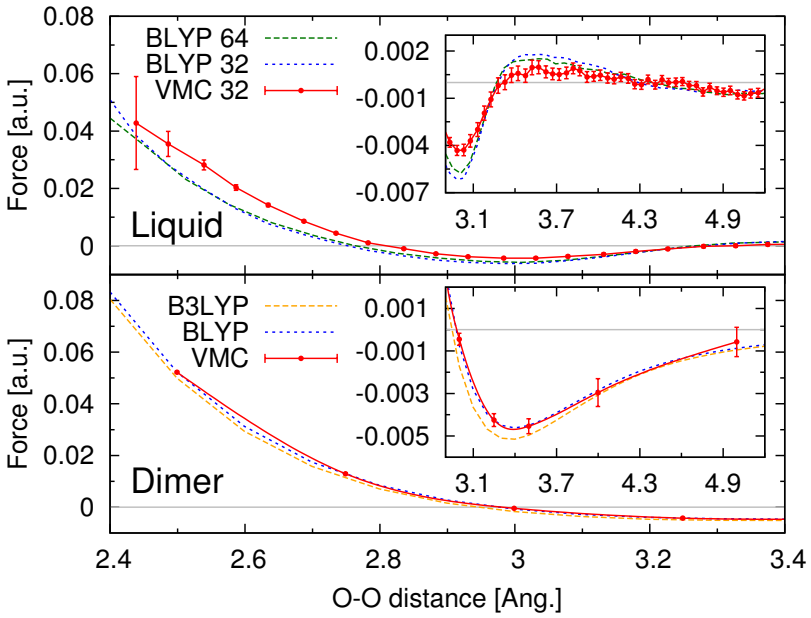

Figure 6.4: The net molecular force difference of a pair of water monomers (i, j) projected on their oxygen-oxygen (O-O) direction $(\boldsymbol{f}^i_{H_2O} - \boldsymbol{f}^j_{H_2O}) \cdot (\boldsymbol{r}_{O_i} - \boldsymbol{r}_{O_j})/r_{O_iO_j}$ as a function of O-O distance $r_{O_iO_j}$ where the net molecular force $\boldsymbol{f}_{H_2O} = \boldsymbol{f}_{H_1} + \boldsymbol{f}_{H_2} + \boldsymbol{f}_O$. The upper panel shows that in liquid water the VMC and DFT/BLYP forces have sizable difference at short range below 3Å and the difference between two DFT/BLYP calculations with 32 and 64 molecules in the unit cell is negligible. The lower panel shows that in the water dimer the two body behavior of VMC, DFT/BLYP and DFT/B3LYP results are almost the same both at short and long distance.

6.5 Role of the Hydrogen bond

As mentioned in the introduction of this chapter, hydrogen bonds play the key role for determining the structure of liquid water. It is always an unavoidable topic when liquid water is studied by various methods like force fields [149] and DFT calculations, especially when an improvement on the theoretical description of the dispersion interaction between water molecules is achieved [125, 127, 150]. Apart from ambient condition, it is also interesting to see how the structure changes by looking through the redistribution of hydrogen bonds when liquid water is under pressure [151]. Moreover when the nuclear quantum effects are taken into account, the behavior of hydrogen bonds changes dramatically due to the enhanced proton mobility and tunneling by the zero point motion [132, 136, 137].

We've measured the hydrogen bonds in our VMC simulation and they are identified in the following way. Since the hydrogen bonds exist only in the O-H--O structure displayed in figure 6.1. They are filtered out first by choosing oxygen atoms within a distance cutoff (3.32 Å) right at the position

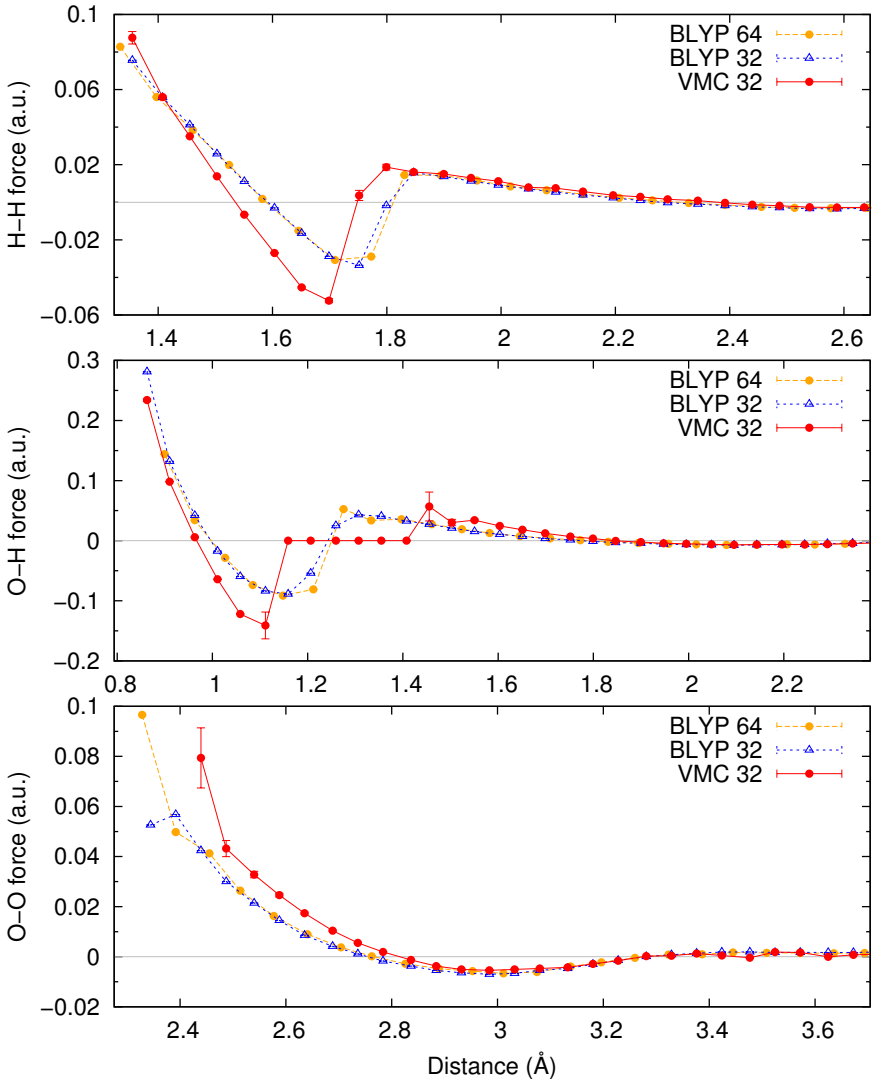

Figure 6.5: The atomic force difference between atom i and atom j projected on their direction $(\boldsymbol{f}_i - \boldsymbol{f}_j) \cdot (\boldsymbol{r}_i - \boldsymbol{r}_j)/r_{i,j}$ as a function of their distance r_{ij} for H-H (upper panel), O-H (middle panel) and O-O (lower panel) pairs in the liquid water. All three panels show that the VMC and DFT/BLYP forces have sizable difference at short range below 3Å and the difference between two BLYP calculations with 32 and 64 molecules in the unit cell is negligible.

of the first valley in O-O RDF. In the second step, the hydrogen (see H_d in figure 6.1) between the two oxygen in a O-H--O structure is selected as the one which provides the largest O-H--O angle among all the four hydrogen atoms belonging to these two oxygen atoms. In the last step, not all the H--O in the O-H--O structure are considered as hydrogen bonds but only those with their O-H--O angle larger than 140° are valid as ref. [151] did.

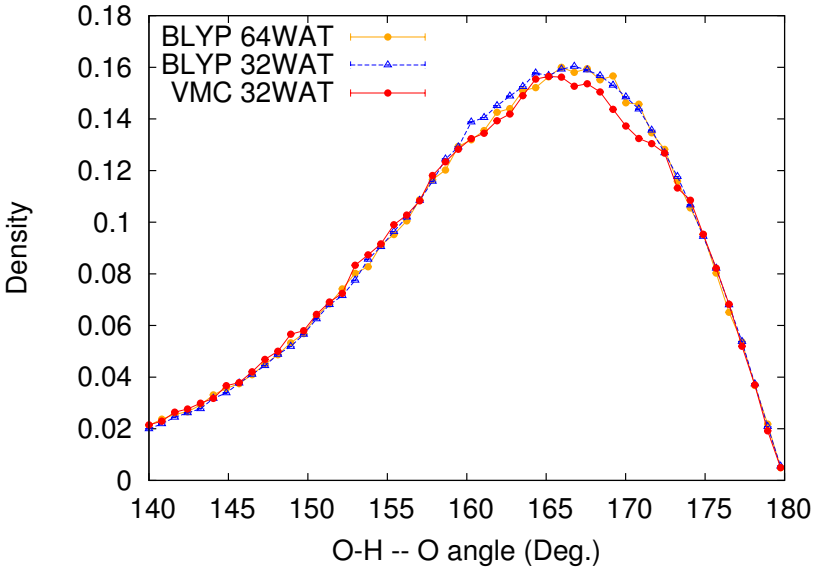

Figure 6.6: The distribution of the largest O-H--O angle between two water molecules.

Figure 6.6 shows the distribution of O-H--O in liquid water simulated by 32 and 64 water molecules with DFT/BLYP and 32 water molecules with VMC. Again, the curves given by DFT/BLYP simulations with 32 and 64 molecules are indistinguishable and the VMC curve, even with a little flaw at the peak due to limited statistics, almost coincide with the DFT/BLYP curves. This further strengthen our claim that the improvement of VMC O-O RDF over the DFT one is because of the change of the O--H hydrogen bond distance and not the O-H--O angle. As a comparison, when the DFT is corrected with empirical pair-wise potentials in order to recover the dispersion interaction, the O-H--O angle shows a significant change [124]. Our results question the validity of those methods.

We've also measured the distribution of proton-transfer coordinate ν, proposed in ref. [137], defined as

$$\nu = d(\mathrm{O_d\text{-}H_d}) - d(\mathrm{O_a\text{-}H_d})\,, \tag{6.5}$$

which the difference between the intra-molecular O-H distance d(O-H) and the inter-molecular O-H distance d(O--H) in an O-H--O structure. During our classical-ions simulation with VMC, we have not observed an auto-protolysis event (see figure 6.7). According to ref. [137] when the nuclear quantum correction is applied, liquid water shows a significant auto-protolysis events, namely positive ν. The fact that we haven't observed such proton-transfer indicates the nuclear quantum nature is clearly missing in our simulation.

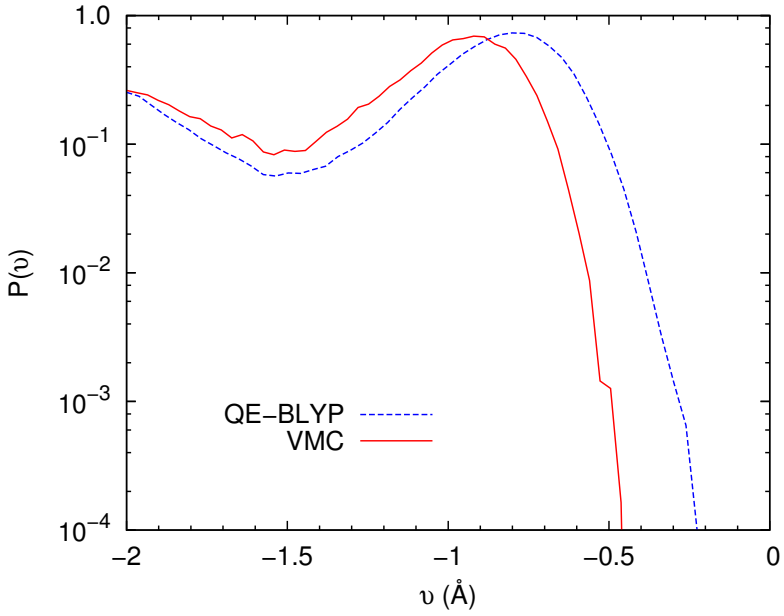

Figure 6.7: Distribution of the proton-transfer coordinate in ab initio dynamics of liquid water at 300k. This coordinate is defined as $\nu = d(\mathrm{O_d\text{-}H_d}) - d(\mathrm{O_a\text{-}H_d})$ on two water molecules connected by a hydrogen bond (see figure 6.1). We haven't observed one event that $\nu > 0$, namely no transfer among $\sim 1.8 \times 10^6$ and $\sim 3.6 \times 10^5$ hydrogen bonds in BLYP and VMC.

6.6 Discussion and conclusion

The first ab-initio simulation of liquid water by quantum Monte Carlo has been performed and yields so much excitement for us. The results show that several issues of ab-initio molecular dynamics, mainly DFT-based, can be solved within the present powerful approach. This opens a new frontier in the simulation of water because several questions about the mystery of water which are also difficult to answer by experiments could be solved with better accuracy by quantum Monte Carlo in the near future, especially by means of massively parallel supercomputers.

We have adopted the most simple quantum Monte Carlo method (the VMC) in a fully consistent and controlled way. Despite the roughness of this first attempt (as compared with the most recent DFT calculations), our study highlights a few clear results and also several perspectives described in the following.

The calculation by QMC is rather heavy, but possible, and there is room for considerable improvements along this fully ab-initio scheme. For instance it could be possible to work with a larger but more accurate basis with only a few times more computer resources, as our algorithm scales linearly with the basis dimension. Also we can add the four-body Jastrow which explicitly describes the dipole-dipole interaction to further improve the description of

liquid water. If even more computing power could be provided, projection methods like LRDMC would push higher accuracy on the evaluation of energy and forces, yielding better results.

Our agreement with experiments is rather satisfactory, and could be an excellent agreement if a larger system (64 waters could be sufficient) and nuclear quantum corrections, not included in this study, would be considered. In a DFT dynamics with 32 waters, the height of the first peak $g_{\mathrm{OO}}^{\mathrm{MAX}} = g_{\mathrm{OO}}(r_{\mathrm{MAX}})$ is expected to be overestimated by ~ 0.3 [112, 113] with respect to the converged value. If we assume that also QMC has similar size effects, the agreement with the experimental value should be substantially improved. Moreover, nuclear quantum corrections appear to further broaden the first peak and reduce its height by about ~ 0.4 (~ 0.24) if we consider as reference the path integral CPMD calculation reported in ref. [132] (ref. [148]). We can also probably capture substantial auto-protolysis events. Therefore, in future studies, by taking into account size and nuclear quantum effects, it may be possible to have a fully consistent ab-initio description of liquid water by QMC.

We finally remark that, thanks to good scaling properties of QMC algorithms with the system size and fast evolving supercomputers, this work opens promising perspectives for future application of such high-level ab initio molecular dynamics technique for the study of thermodynamic properties of complex clusters and liquids.

Chapter 7

Conclusion

In this thesis, the previously introduced molecular dynamics simulation with quantum Monte Carlo [47] is investigated and developed through meticulous and in-depth studies. Indeed this new advanced tool is becoming mature during these years [35, 146, 105].

The first part of this thesis introduces the basic knowledge of variational Monte Carlo and a wide range of optimization methods for the wavefunction within the framework of VMC. By applying all these methods on our test case Be_2, their performances are well displayed and their features are discussed. This part lays the foundation of the molecular dynamics since wavefunction optimization is intensively employed in order to fulfill the Born-Oppenheimer approximation when ions are moving during the molecular dynamics.

The first part also introduces the wavefunction ansatz chosen for characterizing the systems of interest. A good wavefunction should be accurate and should be evaluated efficiently. By assembling Jastrow factor and AGP represented by atomic hybrid orbitals and molecular orbitals, our wavefunction reaches a perfect balance between accuracy and compactness. In addition, through again Be_2 we see the importance of the four-body Jastrow in cases when the van der Waals interaction is particularly important.

In the second part, since we've already been equipped with efficient optimization methods and a good wavefunction, the ab initio molecular dynamics scheme with quantum Monte Carlo is proposed after introducing the methods for the accurate and efficient evaluation of forces [50]. In order to improve the accuracy and stability of this scheme, a better integration scheme is devised [146]. Now large time steps are allowed without losing accuracy, thus significantly reducing the number of steps during the equilibration. This part highlights the peculiar features of the covariance matrix of the forces which can be defined only within QMC and not in any other deterministic approach like DFT. It can be used to fine-tune the speeds of slow and fast modes of the dynamics and thus accelerate the simulation.

After that, our molecular dynamics scheme was validated by benchmarking calculations of the vibrational frequencies [146]. The calculations of the

frequencies of water and other small molecules show that many systematic biases in our MD scheme and QMC evaluation can be controlled and this further confirms the robustness of our MD scheme. We are now more confident to push this method forward to applications on large systems.

Since all the preparations are well done, it's time to hunt monsters. In the last chapter of the second part, the liquid water is simulated with our ab initio molecular dynamics [105]. The results are encouraging since we've obtained RDFs with peak positions almost identical to the most recent experiments and the shapes are much less structured than previous DFT-based ab initio simulations even if the two water molecule interaction is dealt with the same level of accuracy as the DFT/BLYP calculation. We have also studied the hydrogen bonds in the liquid water we have simulated.

Thanks to this advanced simulation tool, we've seen so much surprise and obtained a lot of encouraging results. However, there are still questions to be answered and improvements to be done. Here are the considerations of further improvements and applications

a) The advantages of the molecular dynamics scheme, using large time step and speeding up slow modes, introduced in this thesis rely on the force covariance matrix which has been empirically shown to be proportional to the dynamical matrix. Further theoretical studies are required to investigate how and why those two matrices are related. If we could extract the dynamical matrix from the force covariance matrix, we can get the vibrational modes and frequencies directly.

b) The vibrational frequencies of larger systems like water dimer are worth trying by our fitting method with molecular dynamics involved, especially if the anharmonic effects can be considered.

c) As for the simulation of liquid water, there is room for considerable improvements. On one hand, within our current ab initio scheme, it could be possible to work with a larger but more accurate basis set at a reasonable amount of extra cost. We could also add back the four-body Jastrow to have a better wavefunction ansatz to characterize the liquid water. On the other hand, we could compute forces with more advanced, however more costly, QMC techniques like LRDMC. In addition, longer simulations with larger system sizes (at least 64 molecules) could also reduce the finite size bias systematically.

d) We've seen from the simulation of liquid water that our simulation, even with substantial improvements, is still not exact compared with experiments. One key fact is that we have missed the quantum feature of ions which is required for a faithful description of liquid water. Further studies are necessary to describe ions within a fully consistent quantum mechanical approach.

Apart from a), all the other three points demand more computational resources. Actually, our simulation of liquid water has already hit the petascale performance. As the most powerful supercomputer on the Top500 list of high performance computer (HPC) now has already reached a measured performance of ~ 34 petaflops and the overall performance of supercomputers are growing at a rapid and steady speed, those improvements could be made possible in the very near future. In addition, the point d) also requires the implementation of new algorithms by employing path integral techniques.

However, there's another technical challenge that a big amount of work is required to adapt current programs to the new hardware architectures of supercomputers. The so called 'Distributed memory' era which mainly involves MPI[1] has passed and the 'Many core' era starts [152]. In this new era, the most heavy calculations are no longer carried by CPU but mainly by accelerators like GPU or MIC which integrate a huge number of compute units operating at lower frequency in order to perform with a much higher energy efficiency. This transition has been clearly shown in the most recent Top500 list announced in June 2014 that 4/10 of the top ten computers have already been equipped with accelerators. In order to achieve their decent performance, the current programming model needs to be revised. Fortunately, the new era is very promising for us since the new architectures favor statistical methods like quantum Monte Carlo rather than deterministic methods like DFT.

In conclusion, it is time to press the button to launch wide applications of the promising ab initio molecular dynamics by quantum Monte Carlo like what the DFT-based approach did two decades ago. I hope to see QMC continue expanding its territory in the following decades.

[1] MPI stands for 'Message Passing Interface'.

Appendix A

Solutions to the second order Langevin dynamics

A.1 Integration of second order Langevin equations

In this appendix we sketch how to integrate exactly the differential eq. (4.16):

$$\dot{\boldsymbol{v}}(s) = -\boldsymbol{\gamma}(\boldsymbol{R}) \cdot \boldsymbol{v}(s) + \boldsymbol{f}(\boldsymbol{R}) + \boldsymbol{\eta}(s) \tag{A.1}$$

in an arbitrary interval $\bar{t} \leq s \leq t$ within the assumption that the vector $\boldsymbol{R}$ is not changing much during the integration interval and that therefore it can be considered independent of s. In this appendix, in order to avoid confusion, we indicate by s the generic time defining the SLD dynamics, whereas with $\bar{t}$ and t, the initial and final time of the integration, respectively, so that the initial condition reads:

$$\boldsymbol{v}(\bar{t}) = \bar{\boldsymbol{v}} \,. \tag{A.2}$$

As well known this kind of equations can be solved in terms of the simple exponential solution $\boldsymbol{v}(s) = \exp\left[\boldsymbol{\gamma}(\boldsymbol{R})(\bar{t} - s)\right] \bar{\boldsymbol{v}}$, valid in absence of the external force and the noise (i.e. $\boldsymbol{f}(\boldsymbol{R}) + \boldsymbol{\eta}(s) = 0$). We search therefore a solution of the form:

$$\boldsymbol{v}(s) = \exp\left[-\boldsymbol{\gamma}(\boldsymbol{R})s\right] \boldsymbol{y}(s) \,. \tag{A.3}$$

By replacing the above equation in eq. (A.1) we easily obtain that:

$$\dot{\boldsymbol{y}}(s) = \exp\left[\boldsymbol{\gamma}(\boldsymbol{R})s\right] \left[\boldsymbol{f}(\boldsymbol{R}) + \boldsymbol{\eta}(s)\right] \tag{A.4}$$

with the initial condition given by inverting eq. (A.3) for $s = \bar{t}$:

$$\boldsymbol{y}(\bar{t}) = \exp\left[\boldsymbol{\gamma}(\boldsymbol{R})\bar{t}\right] \bar{\boldsymbol{v}} \,. \tag{A.5}$$

Eq. (A.4) can be integrated immediately from $\bar{t}$ to t, because its RHS is a known function of s:

$$\boldsymbol{y}(s) = \boldsymbol{y}(\bar{t}) + \int_{\bar{t}}^{s} \exp\left[\boldsymbol{\gamma}(\boldsymbol{R})t'\right] \left[\boldsymbol{f}(\boldsymbol{R}) + \boldsymbol{\eta}(t')\right] dt' \,. \tag{A.6}$$

We now go back to the original ansatz (eq. (A.3)), and by replacing the initial condition (eq. (A.5)) in the above equation, we obtain the final solution:

$$\begin{aligned}\boldsymbol{v}(t) = &\exp\left[\boldsymbol{\gamma}(\boldsymbol{R})(\bar{t}-t)\right]\boldsymbol{v}(\bar{t}) \\ &+\int_{\bar{t}}^{t}\exp\left[\boldsymbol{\gamma}(\boldsymbol{R})(t'-t)\right]\left[\boldsymbol{f}(\boldsymbol{R})+\boldsymbol{\eta}(t')\right]dt'\,. \end{aligned} \tag{A.7}$$

A.2 Better integration scheme

In this appendix we describe how to avoid the approximation in eq. (4.23) to integrate eq. (4.17), with a more involved method, that was already introduced in ref. [35]. However we have noted that in the proposed integrator it is not necessary to compute the velocities at half-integer times because we perform the integration of eq. (4.17) in an exact unbiased way. In the following we describe this derivation and obtain expressions very similar to the ones introduced in ref. [35], with the main difference that here we use integer time both for velocities and positions:

$$\boldsymbol{v}_n \equiv \boldsymbol{v}(t_n) \tag{A.8}$$

$$\boldsymbol{R}_n \equiv \boldsymbol{R}(t_n)\,. \tag{A.9}$$

Having the general expression of the velocity by eq. (A.7), we can use to integrate eqs. (4.16) and (4.17) in the interval $t_n \le s \le t_{n+1}$ and obtain, with a little involved algebra, just a bit more than the original scheme [47] described in section 4.4:

$$\boldsymbol{v}_{n+1} = e^{-\gamma_n\tau}\boldsymbol{v}_n + \boldsymbol{\Gamma}_n\cdot(\boldsymbol{f}_n+\tilde{\boldsymbol{\eta}}) \tag{A.10}$$

$$\boldsymbol{R}_{n+1} = \boldsymbol{R}_n + \boldsymbol{\Gamma}_n\cdot\boldsymbol{v}_n + \boldsymbol{\Theta}_n\cdot(\boldsymbol{f}_n+\tilde{\tilde{\boldsymbol{\eta}}}) \tag{A.11}$$

where we have introduced the following matrices, mainly to single out, as before, the actual noisy terms $\tilde{\boldsymbol{\eta}}$ and $\tilde{\tilde{\boldsymbol{\eta}}}$, that have to be added to the force components in the above Markov iterations for the velocities and coordinates,

respectively:

$$\boldsymbol{\Gamma}_n = \boldsymbol{\gamma}_n^{-1}(\mathbf{I} - e^{-\boldsymbol{\gamma}_n\tau}) \tag{A.12}$$

$$\boldsymbol{\Theta}_n = \boldsymbol{\gamma}_n^{-1}(\tau\mathbf{I} - \boldsymbol{\Gamma}_n)\,, \tag{A.13}$$

$$\begin{aligned}\tilde{\boldsymbol{\eta}} &= \boldsymbol{\Gamma}_n^{-1}e^{-\boldsymbol{\gamma}_n\tau}\int_{t_n}^{t_{n+1}} e^{\boldsymbol{\gamma}_n(t-t_n)}\boldsymbol{\eta}(t)\,dt \\ &= \boldsymbol{\Gamma}_n^{-1}e^{-\boldsymbol{\gamma}_n\tau}\int_0^{\tau} e^{\boldsymbol{\gamma}_n t}\boldsymbol{\eta}(t)\,dt\,, \end{aligned} \tag{A.14}$$

$$\begin{aligned}\tilde{\tilde{\boldsymbol{\eta}}} &= \boldsymbol{\Theta}_n^{-1}\int_{t_n}^{t_{n+1}} dt\int_{t_n}^{t} dt' e^{\boldsymbol{\gamma}_n(t'-t)}\boldsymbol{\eta}(t') \\ &= \boldsymbol{\Theta}_n^{-1}\int_0^{\tau} dt e^{-\boldsymbol{\gamma}_n t}\int_0^{t} dt' e^{\boldsymbol{\gamma}_n t'}\boldsymbol{\eta}(t')\,. \end{aligned} \tag{A.15}$$

In order to define the Markov process, it is enough to compute the correlation of the previously mentioned noisy terms, which we define as follows:

$$\langle\tilde{\boldsymbol{\eta}}_i\tilde{\boldsymbol{\eta}}_j\rangle \equiv \bar{\boldsymbol{\alpha}}_{ij}^{\mathbf{1,1}} \tag{A.16}$$

$$\left\langle\tilde{\tilde{\boldsymbol{\eta}}}_i\tilde{\tilde{\boldsymbol{\eta}}}_j\right\rangle \equiv \bar{\boldsymbol{\alpha}}_{ij}^{\mathbf{2,2}} \tag{A.17}$$

$$\left\langle\tilde{\boldsymbol{\eta}}_i\tilde{\tilde{\boldsymbol{\eta}}}_j\right\rangle \equiv \bar{\boldsymbol{\alpha}}_{ij}^{\mathbf{1,2}} \tag{A.18}$$

$$\left\langle\tilde{\tilde{\boldsymbol{\eta}}}_i\tilde{\boldsymbol{\eta}}_j\right\rangle \equiv \bar{\boldsymbol{\alpha}}_{ij}^{\mathbf{2,1}} = \bar{\boldsymbol{\alpha}}_{ij}^{\mathbf{1,2}}\,. \tag{A.19}$$

Then a straightforward integration in time, by using that the assumed correlation is given by eq. (4.18) and that, by eq. (4.19) the corresponding matrix $\bar{\boldsymbol{\alpha}} = 2T\boldsymbol{\gamma}$, we obtain:

$$\begin{aligned}\bar{\boldsymbol{\alpha}}^{\mathbf{1,1}} &= T\boldsymbol{\gamma}_n^2\coth(\boldsymbol{\gamma}_n\tau/2) \\ \bar{\boldsymbol{\alpha}}^{\mathbf{2,2}} &= T(2\boldsymbol{\Theta}_n - \boldsymbol{\Gamma}_n^2)\cdot\boldsymbol{\Theta}_n^{-2} \\ \bar{\boldsymbol{\alpha}}^{\mathbf{1,2}} &= T\boldsymbol{\gamma}_n\boldsymbol{\Gamma}_n\cdot\boldsymbol{\Theta}_n^{-1}\,. \end{aligned}$$

The above Markov process can be straightforwardly implemented, as well as the very similar one described in ref. [35]. However we have tested that all methods, including the simplest one described in section 4.4, behave equally well, with comparable performances, probably because a too high accurate integration scheme is not necessary for the available accuracy, possible at present with QMC.

As discussed also in ref. [47] (see also section 5.2.3), all the QMC force evaluations $\boldsymbol{f}$ are affected by an intrinsic stochastic noise, which usually determines an effective temperature higher than the target one. This problem

can be avoided, by generalizing the method of the *noise correction* described in section 4.4 to this specific case. Indeed we can follow the correct dynamics by adding to the QMC noise of the forces the two external noises $\tilde{\boldsymbol{\eta}}_{\text{ext}}$ and $\tilde{\tilde{\boldsymbol{\eta}}}_{\text{ext}}$ so that the total noises $\tilde{\boldsymbol{\eta}}$ and $\tilde{\tilde{\boldsymbol{\eta}}}$ satisfy the correct expressions in eqs. (A.16) to (A.19). In this way, we have to subtract the $3N \times 3N$ QMC correlation of the forces $\boldsymbol{\alpha}^{\text{QMC}}$ to each of the four sub-matrices, namely

$$\bar{\boldsymbol{\alpha}}_{\text{ext}}^{\boldsymbol{a},\boldsymbol{b}} = \bar{\boldsymbol{\alpha}}^{\boldsymbol{a},\boldsymbol{b}} - \boldsymbol{\alpha}^{\text{QMC}} , \tag{A.20}$$

is the true external noise we have to add to the system, to take into account that QMC forces contain already a correlated noise, that is independently evaluated during the dynamics. It can be shown that the resulting matrix $\bar{\boldsymbol{\alpha}}_{\text{ext}}$ is indeed positive definite provided Δ_0 is large enough, so that $\bar{\boldsymbol{\alpha}}_{\text{ext}}$ is a well defined correlation for an external noise.

Appendix B

Efficient calculation of $\hat{S}^2$

In this appendix we describe how to compute the expectation value of the total spin square $\hat{S}^2$ over the variational wavefunction Ψ_{JAGP} in the paired AGP case, namely with vanishing spin projection S_z along the z-axis and N electrons [55].

As is well known, within VMC, we need to compute the so called local estimator of the spin square:

$$\frac{\langle \boldsymbol{x}|\hat{S}^2|\Psi\rangle}{\langle \boldsymbol{x}|\Psi\rangle}$$

where $\boldsymbol{x} = \{\boldsymbol{r}_1^{\uparrow}, \ldots, \boldsymbol{r}_{N/2}^{\uparrow}, \boldsymbol{r}_1^{\downarrow}, \ldots, \boldsymbol{r}_{N/2}^{\downarrow}\}$ is a many body configuration where the electron positions and the spin projection along the z-axis $\sigma_i = \pm 1/2$ are defined. The application of $\hat{S}^2$ to a given configuration can be written as:

$$\hat{S}^2|\boldsymbol{x}\rangle = S_z^2|\boldsymbol{x}\rangle + \frac{1}{2}\sum_{\boldsymbol{r}_i,\boldsymbol{r}_j}(S^+_{\boldsymbol{r}_i}S^-_{\boldsymbol{r}_j} + \text{h.c.})|\boldsymbol{x}\rangle \tag{B.1}$$

where i and j label all electron positions, regardless of their spins. The above expression can be recast in the following way:

$$\hat{S}^2|\boldsymbol{x}\rangle = -\sum_{k,l}^{N/2}|\boldsymbol{x}_{kl}\rangle + \frac{N}{2}|\boldsymbol{x}\rangle \tag{B.2}$$

which generates $(N/2)^2$ new configurations

$$|\boldsymbol{x}_{kl}\rangle = -S^-_{\boldsymbol{r}_k^{\uparrow}}S^+_{\boldsymbol{r}_l^{\downarrow}}|\boldsymbol{x}\rangle \qquad (k,l = 1,\ldots,N/2) \tag{B.3}$$

where $k(l)$ labels only the spin-up(down) electrons, namely the new configuration $\boldsymbol{x}_{kl}$ is obtained by swapping the positions of the (k,l) electron pair with opposite spins. The minus sign in the above expression takes into account the Fermi statistics, in order to recast a spin-flip with a position exchange. Similarly the rightmost term in eq. (B.2) takes into account the local term $i = j$ in eq. (B.1), which is obtained by applying the spin-flip operator to each individual electron, leading to a trivial constant $(N/2)$ times $|\boldsymbol{x}\rangle$.

Therefore, the main problem is to compute the $(N/2)^2$ wavefunction ratios:

$$r_{kl} = \frac{\langle \boldsymbol{x}_{kl}|\Psi\rangle}{\langle \boldsymbol{x}|\Psi\rangle} = \frac{\det \boldsymbol{A}'}{\det \boldsymbol{A}}\frac{J'}{J} \tag{B.4}$$

$$\boldsymbol{A}_{ij} = \mathcal{G}(\boldsymbol{r}_i^{\uparrow}, \boldsymbol{r}_j^{\downarrow}) \tag{B.5}$$

which contain a determinant factor and a Jastrow factor.

B.1 Determinant part

For the determinant factor, the swapping of the (k, l) electron pair implies a change in the determinant $\boldsymbol{A} \to \boldsymbol{A}'$ given by:

$$\begin{aligned}
\boldsymbol{A}'_{ij} &= \boldsymbol{A}_{ij} + \delta_{ik}[\mathcal{G}(\boldsymbol{r}_l^{\downarrow}, \boldsymbol{r}_j^{\downarrow}) - \mathcal{G}(\boldsymbol{r}_k^{\uparrow}, \boldsymbol{r}_j^{\downarrow})] + \delta_{jl}[\mathcal{G}(\boldsymbol{r}_i^{\uparrow}, \boldsymbol{r}_k^{\uparrow}) - \mathcal{G}(\boldsymbol{r}_i^{\uparrow}, \boldsymbol{r}_l^{\downarrow})] \\
&\quad + \delta_{ik}\delta_{jl}\theta_{kl} \\
\theta_{kl} &= [\mathcal{G}(\boldsymbol{r}_l^{\downarrow}, \boldsymbol{r}_k^{\uparrow}) + \mathcal{G}(\boldsymbol{r}_k^{\uparrow}, \boldsymbol{r}_l^{\downarrow}) - \mathcal{G}(\boldsymbol{r}_l^{\downarrow}, \boldsymbol{r}_l^{\downarrow}) - \mathcal{G}(\boldsymbol{r}_k^{\uparrow}, \boldsymbol{r}_k^{\uparrow})]
\end{aligned}$$

We rewrite $\boldsymbol{A}'$ as

$$\begin{aligned}
\boldsymbol{A}' &= \boldsymbol{A}(I + \Delta) \\
\Delta_{ij} &= \boldsymbol{A}_{ik}^{-1} W_j + \delta_{jl} U_i \\
W_j &= \mathcal{G}(\boldsymbol{r}_l^{\downarrow}, \boldsymbol{r}_j^{\downarrow}) - \mathcal{G}(\boldsymbol{r}_k^{\uparrow}, \boldsymbol{r}_j^{\downarrow}) \\
U_i &= B_{i,k}^{\uparrow,\uparrow} - B_{i,l}^{\uparrow,\downarrow} + \boldsymbol{A}_{ik}^{-1}\theta_{kl}
\end{aligned}$$

where we have defined the $B^{\uparrow,\sigma}$ matrices as follows:

$$B_{i,j}^{\uparrow,\sigma} = \sum_z \boldsymbol{A}_{iz}^{-1}\mathcal{G}(\boldsymbol{r}_z^{\uparrow}, \boldsymbol{r}_j^{\sigma}) \tag{B.6}$$

Notice that these matrices can be computed only once for all spin flip ratios, amounting to $2(N/2)^3$ operations.

Then, by employing the Sherman-Morrison algebra, the ratio of the two determinants $\det \boldsymbol{A}'/\det \boldsymbol{A}$ is given by a determinant of a much simpler 2×2 matrix $\boldsymbol{M}$:

$$\boldsymbol{M} = \begin{pmatrix} 1 + \sum_i \boldsymbol{A}_{ik}^{-1} W_i & \boldsymbol{A}_{lk}^{-1} \\ \sum_i U_i W_i & 1 + U_l \end{pmatrix} \tag{B.7}$$

In this way, the number of operations necessary to obtain all the $(N/2)^2$ ratios scales as N^3 namely, with a computational time similar to the calculation of the energy.

B.2 Jastrow part

If the wavefunction is defined in terms of a spin dependent Jastrow, the total Jastrow factor can be generally written as

$$J = \exp\left(\sum_{i<j}^{N} V(i,j)\right) \tag{B.8}$$

where the summation over i and j are now over all the electrons, regardless of their spins, and V is defined as

$$V(i,j) = V_{\text{2B}}(i,j) + 4\sigma_i\sigma_j V_{\text{34B-Sz}}(i,j) \tag{B.9}$$

$$V_{\text{2B}}(i,j) = \frac{1}{2}[V_{\parallel}(i,j)(1+4\sigma_i\sigma_j) + V_{\perp}(i,j)(1-4\sigma_i\sigma_j)] \tag{B.10}$$

where V_{2B} is the two-body Jastrow consisting of spin parallel $V_{\parallel}$ and spin anti-parallel $V_{\parallel}$, and $V_{\text{34B-Sz}}$ is the three- and four- body spin Jastrow. The spin dependent part of V is rewritten as $V_{sd}(i,j)\sigma_i\sigma_j$, $V_{sd}(i,j) = 2(V_{\parallel}-V_{\perp})+V_{\text{34B-Sz}}$. Each time we swap the electron k with spin up and the electron l with spin down, we only need to flip the corresponding spins σ_k and σ_l in eq. (B.9). It is clear therefore that, by computing the auxiliary vector $\bar{V}(l) = \sum_{i\neq l} V_{sd}(i,l)\sigma_i$ once for all, all the Jastrow ratios J'/J can be easily computed as

$$\frac{J'}{J} = \exp\left\{\left[\bar{V}(l) - V_{sd}(k,l)/2\right] - \left[\bar{V}(k) + V_{sd}(l,k)/2\right]\right\},$$

namely also this with an irrelevant number of operations ($\simeq N^2$ operations).

List of abbreviations

HF	Hartree Fock
SD	Slater Determinant
AGP	Antisymmetrized Geminal Power
JHF	Jastrow correlated Hartree Fock wavefunction
JSD	Jastrow correlated Slater Determinant wavefunction
JAGP	Jastrow correlated AGP wavefunction
STO	Slater Type Orbital
GTO	Gaussian Type Orbital
MO	Molecular Orbitals
LA	Locality Approximation
SR	Stochastic Reconfiguration
CCSD	Coupled Cluster Single and Double excitation
DFT	Density Functional Theory
QMC	Quantum Monte Carlo
VMC	Variational Monte Carlo
DMC	Diffusion Monte Carlo
LRDMC	Lattice Regularized Diffusion Monte Carlo
BO	Born-Oppenheimer
MD	Molecular Dyanmics
CPMD	Car-Parrinello Molecular Dyanmics
LD	Langevin Dyanmics
SLD	Second order Langevin Dyanmics
RDF	Radial Distribution Function

List of publications

1. Andrea Zen, Ye Luo, Sandro Sorella, and Leonardo Guidoni. Molecular Properties by Quantum Monte Carlo: An Investigation on the Role of the Wave Function Ansatz and the Basis Set in the Water Molecule. *J. Chem. Theory Comput.*, 9(10):4332–4350, October 2013.

2. Andrea Zen, Emanuele Coccia, Ye Luo, Sandro Sorella, and Leonardo Guidoni. Static and Dynamical Correlation in Diradical Molecules by Quantum Monte Carlo Using the Jastrow Antisymmetrized Geminal Power Ansatz. *J. Chem. Theory Comput.*, 10(3):1048–1061, March 2014.

3. Ye Luo, Andrea Zen, and Sandro Sorella. Ab-initio molecular dynamics with noisy forces: validating the quantum monte carlo approach with benchmark calculations of molecular vibrational properties. *submitted to J. Chem. Phys.*, 2014.

4. Andrea Zen, Ye Luo, Guglielmo Mazzola, Leonardo Guidoni, and Sandro Sorella. Ab-initio molecular dynamics simulation of liquid water by Quantum Monte Carlo. *submitted*, 2014.

Bibliography

[1] David Philip Miller. *Discovering water: James Watt, Henry Cavendish, and the nineteenth century 'Water Controversy'*. Ashgate, 2004.

[2] E. Schrödinger. An undulatory theory of the mechanics of atoms and molecules. *Phys. Rev.*, 28:1049–1070, Dec 1926.

[3] V. Fock. Näherungsmethode zur losung des quanten-mechanischen mehrkörperprobleme. *Z. Phys.*, 61:126, 1930.

[4] Per-Olov Löwdin. Quantum theory of many-particle systems. iii. extension of the hartree-fock scheme to include degenerate systems and correlation effects. *Phys. Rev.*, 97:1509–1520, Mar 1955.

[5] W. M. C. Foulkes, L. Mitas, R. J. Needs, and G. Rajagopal. Quantum monte carlo simulations of solids. *Rev. Mod. Phys.*, 73:33–83, Jan 2001.

[6] D. M. Ceperley. Path integrals in the theory of condensed helium. *Rev. Mod. Phys.*, 67:279–355, Apr 1995.

[7] W. L. McMillan. Ground state of liquid he^4. *Phys. Rev.*, 138:A442–A451, Apr 1965.

[8] M. H. Kalos, D. Levesque, and L. Verlet. Helium at zero temperature with hard-sphere and other forces. *Phys. Rev. A*, 9:2178–2195, May 1974.

[9] R.C Grimm and R.G Storer. Monte-carlo solution of schrödinger's equation. *Journal of Computational Physics*, 7(1):134 – 156, 1971.

[10] Stefano Baroni and Saverio Moroni. Reptation quantum monte carlo: A method for unbiased ground-state averages and imaginary-time correlations. *Phys. Rev. Lett.*, 82:4745–4748, Jun 1999.

[11] M Casula, C Filippi, and S Sorella. Diffusion Monte Carlo method with lattice regularization. *Phys. Rev. Lett.*, 95(10):100201, 2005.

[12] Michele Casula, Saverio Moroni, Sandro Sorella, and Claudia Filippi. Size-consistent variational approaches to nonlocal pseudopotentials: Standard and lattice regularized diffusion Monte Carlo methods revisited. *J. Chem. Phys.*, 132(15):154113, 2010.

[13] George H. Booth, Alex J. W. Thom, and Ali Alavi. Fermion monte carlo without fixed nodes: A game of life, death, and annihilation in slater determinant space. *The Journal of Chemical Physics*, 131(5):–, 2009.

[14] Shiwei Zhang and Henry Krakauer. Quantum monte carlo method using phase-free random walks with slater determinants. *Phys. Rev. Lett.*, 90:136401, Apr 2003.

[15] W. A. Al-Saidi, Shiwei Zhang, and Henry Krakauer. Auxiliary-field quantum monte carlo calculations of molecular systems with a gaussian basis. *The Journal of Chemical Physics*, 124(22):–, 2006.

[16] C David Sherrill and Henry F Schaefer, III. The Configuration Interaction Method: Advances in Highly Correlated Approaches. pages 143–269. Elsevier, 1999.

[17] Rodney Bartlett and Monika Musiał. Coupled-cluster theory in quantum chemistry. *Rev. Mod. Phys.*, 79(1):291–352, February 2007.

[18] D. Ceperley, G. V. Chester, and M. H. Kalos. Monte carlo simulation of a many-fermion study. *Phys. Rev. B*, 16:3081–3099, Oct 1977.

[19] Nicholas Metropolis, Arianna W. Rosenbluth, Marshall N. Rosenbluth, Augusta H. Teller, and Edward Teller. Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, 21(6):1087–1092, 1953.

[20] P. J. Rossky, J. D. Doll, and H. L. Friedman. Brownian dynamics as smart monte carlo simulation. *The Journal of Chemical Physics*, 69(10):4628–4633, 1978.

[21] C. J. Umrigar, K. G. Wilson, and J. W. Wilkins. Optimized trial wave functions for quantum monte carlo calculations. *Phys. Rev. Lett.*, 60:1719–1722, Apr 1988.

[22] Martin Snajdr and Stuart 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. *The Journal of Chemical Physics*, 112(11):4935–4941, 2000.

[23] Dario Bressanini, Gabriele Morosi, and Massimo Mella. Robust wave function optimization procedures in quantum monte carlo methods. *The Journal of Chemical Physics*, 116(13):5345–5350, 2002.

[24] Sandro Sorella. Generalized lanczos algorithm for variational quantum monte carlo. *Phys. Rev. B*, 64:024512, Jun 2001.

[25] Michele Casula, Claudio Attaccalite, and Sandro Sorella. Correlated geminal wave function for molecules: An efficient resonating valence bond approach. *The Journal of Chemical Physics*, 121(15):7110–7126, 2004.

[26] Sandro Sorella, Michele Casula, and Dario Rocca. Weak binding between two aromatic rings: Feeling the van der Waals attraction by quantum Monte Carlo methods. *J. Chem. Phys.*, 127(1):014105, 2007.

[27] S Sorella. Wave function optimization in the variational Monte Carlo method. *Phys. Rev. B*, 71(24):241103, 2005.

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

[29] C. J. Umrigar, J. Toulouse, C. Filippi, S. Sorella, and H. Rhenning. Alleviation of the Fermion-sign problem by optimization of many-body wave functions. *Phys. Rev. Lett*, 98:110201, 2007.

[30] Michele Casula and Sandro Sorella. Geminal wave functions with jastrow correlation: A first application to atoms. *The Journal of Chemical Physics*, 119(13):6500–6511, 2003.

[31] M. Bajdich, L. Mitas, G. Drobný, L. K. Wagner, and K. E. Schmidt. Pfaffian pairing wave functions in electronic-structure quantum monte carlo simulations. *Phys. Rev. Lett.*, 96:130201, Apr 2006.

[32] John Lennard-Jones A.C. Hurley and J.A. Pople. The molecular orbital theory of chemical valency. xvi. a theory of paired-electrons in polyatomic molecules. *Proc. R. Soc. Lond. A*, 220(1143):446, 1953.

[33] Mariapia Marchi, Sam Azadi, Michele Casula, and Sandro Sorella. Resonating valence bond wave function with molecular orbitals: Application to first-row molecules. *The Journal of Chemical Physics*, 131(15):–, 2009.

[34] Tosio Kato. On the eigenfunctions of many-particle systems in quantum mechanics. *Communications on Pure and Applied Mathematics*, 10(2):151–177, 1957.

[35] Guglielmo Mazzola, Seiji Yunoki, and Sandro Sorella. Unexpectedly high pressure for molecular dissociation in liquid hydrogen by electronic simulation. *Nature Communication*, 5:3487, mar 2014.

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

[37] S. Fahy, X. W. Wang, and Steven G. Louie. Variational quantum monte carlo nonlocal pseudopotential approach to solids: Formulation and application to diamond, graphite, and silicon. *Phys. Rev. B*, 42:3503–3522, Aug 1990.

[38] Claudia Filippi and C. J. Umrigar. Multiconfiguration wave functions for quantum monte carlo calculations of first-row diatomic molecules. *The Journal of Chemical Physics*, 105(1):213–226, 1996.

[39] Stefano Evangelisti, Gian Luigi Bendazzoli, Roberto Ansaloni, Francesca Durì, and Elda Rossi. A one billion determinant full ci benchmark on the cray t3d. *Chemical Physics Letters*, 252(5-6):437 – 446, 1996.

[40] Julien Toulouse and C J Umrigar. Full optimization of Jastrow-Slater wave functions with application to the first-row atoms and homonuclear diatomic molecules. *J. Chem. Phys.*, 128(17):174101, 2008.

[41] Sandeep Sharma, Takeshi Yanai, George H. Booth, C. J. Umrigar, and Garnet Kin-Lic Chan. Spectroscopic accuracy directly from quantum chemistry: Application to ground and excited states of beryllium dimer. *The Journal of Chemical Physics*, 140(10):104112, 2014.

[42] Jeremy M. Merritt, Vladimir E. Bondybey, and Michael C. Heaven. Beryllium dimer-caught in the act of bonding. *Science*, 324(5934):1548–1551, 2009.

[43] Konrad Patkowski, Vladimír Špirko, and Krzysztof Szalewicz. On the elusive twelfth vibrational state of beryllium dimer. *Science*, 326(5958):1382–1384, 2009.

[44] Eric Neuscamman, C. J. Umrigar, and Garnet Kin-Lic Chan. Optimizing large parameter sets in variational quantum Monte Carlo. *Phys. Rev. B*, 85:045103, 2012.

[45] R Assaraf and M Caffarel. Computing forces with quantum Monte Carlo. *J. Chem. Phys.*, 113(10):4028–4034, 2000.

[46] R Assaraf and M Caffarel. Zero-variance zero-bias principle for observables in quantum Monte Carlo: Application to forces. *J. Chem. Phys.*, 119(20):10536–10552, 2003.

[47] Claudio Attaccalite and Sandro Sorella. Stable liquid hydrogen at high pressure by a novel ab initio molecular-dynamics calculation. *Phys. Rev. Lett.*, 100(11):114501, 2008.

[48] Sandro Sorella and Luca Capriotti. Algorithmic differentiation and the calculation of forces by quantum Monte Carlo. *J. Chem. Phys.*, 133(23):234111, 2010.

[49] T. Poole, W.M.C. Foulkes, J. Spencer, and P. Haynes. Algorithmic differentiation of diffusion monte carlo. In *APS March meeting*, 2014.

[50] Andrea Zen, Ye Luo, Sandro Sorella, and Leonardo Guidoni. Molecular Properties by Quantum Monte Carlo: An Investigation on the Role of the Wave Function Ansatz and the Basis Set in the Water Molecule. *J. Chem. Theory Comput.*, 9(10):4332–4350, October 2013.

[51] Emanuele Coccia, Daniele Varsano, and Leonardo Guidoni. Protein Field Effect on the Dark State of 11- cisRetinal in Rhodopsin by Quantum Monte Carlo/Molecular Mechanics. *J. Chem. Theory Comput.*, 9(1):8–12, January 2013.

[52] Matteo Barborini and Leonardo Guidoni. Reaction pathways by quantum Monte Carlo: Insight on the torsion barrier of 1,3-butadiene, and the conrotatory ring opening of cyclobutene. *J. Chem. Phys.*, 137(22):224309, 2012.

[53] Emanuele Coccia, Daniele Varsano, and Leonardo Guidoni. Ab Initio Geometry and Bright Excitation of Carotenoids: Quantum Monte Carlo and Many Body Green's Function Theory Calculations on Peridinin. *J. Chem. Theory Comput.*, page 140114080315003, January 2014.

[54] Riccardo Guareschi and Claudia Filippi. Ground- and Excited-State Geometry Optimization of Small Organic Molecules with Quantum Monte Carlo. *J. Chem. Theory Comput.*, 9(12):5513–5525, December 2013.

[55] Andrea Zen, Emanuele Coccia, Ye Luo, Sandro Sorella, and Leonardo Guidoni. Static and Dynamical Correlation in Diradical Molecules by Quantum Monte Carlo Using the Jastrow Antisymmetrized Geminal Power Ansatz. *J. Chem. Theory Comput.*, 10(3):1048–1061, March 2014.

[56] J. C. Grossmann and L. Mitas. Efficient quantum monte carlo energies for molecular dynamics simulations. *Phys. Rev. Lett*, 94:056403, 2005.

[57] F. A. Reboredo and J. Kim. Generalizing the self-healing diffusion Monte Carlo approach to finite temperature: A path for the optimization of low-energy many-body bases. *J. Chem. Phys.*, 140:074103, January 2014.

[58] R Car and M Parrinello. Unified Approach for Molecular-Dynamics and Density-Functional Theory . *Phys. Rev. Lett.*, 55(22):2471–2474, 1985.

[59] A Laio, S Bernard, GL Chiarotti, S Scandolo, and E Tosatti. Physics of iron at Earth's core conditions. *Science*, 287(5455):1027–1030, 2000.

[60] S Scandolo. Liquid-liquid phase transition in compressed hydrogen from first-principles simulations. *Proc. Natl Acad. Sci. USA*, 100(6):3051–3053, March 2003.

[61] C Cavazzoni. Superionic and Metallic States of Water and Ammonia at Giant Planet Conditions. *Science*, 283(5398):44–46, January 1999.

[62] G Kresse and J Hafner. Ab initio molecular dynamics for liquid metals. *Phys. Rev. B*, 47(1):558–561, January 1993.

[63] G Kresse and J Hafner. Ab initio molecular-dynamics simulation of the liquid-metal–amorphous-semiconductor transition in germanium. *Phys. Rev. B*, 49(20):14251–14269, May 1994.

[64] Ursula Rothlisberger and Wanda Andreoni. Structural and electronic properties of sodium microclusters (n=2–20) at low and high temperatures: New insights from ab initio molecular dynamics studies. *J. Chem. Phys.*, 94(12):8129, 1991.

[65] Joost VandeVondele, Fawzi Mohamed, Matthias Krack, Jürg Hutter, Michiel Sprik, and Michele Parrinello. The influence of temperature and density functional models in ab initio molecular dynamics simulation of liquid water. *J. Chem. Phys.*, 122(1):014515, 2005.

[66] D Alfe, M J Gillan, and G D Price. Composition and temperature of the Earth's core constrained by combining ab initio calculations and seismic data. *Earth Planet. Sci. Lett.*, 195(1-2):91–98, January 2002.

[67] M Bernasconi, P Silvestrelli, and M Parrinello. Ab Initio Infrared Absorption Study of the Hydrogen-Bond Symmetrization in Ice. *Phys. Rev. Lett.*, 81(6):1235–1238, August 1998.

[68] DM Ceperley and M Dewing. The penalty method for random walks with uncertain energies. *J. Chem. Phys.*, 110(20):9812–9820, 1999.

[69] C. Pierleoni and D. M. Ceperley. Computer Simulations in Condensed Matter Systems: From Materials to Chemical Biology Volume 1. *Lecture Notes in Physics*, 703:641–683, 2006.

[70] Miguel A. Morales, Carlo Pierleoni, Eric Schwegler, and D M Ceperley. Evidence for a first-order liquid-liquid transition in high-pressure hydrogen from ab initio simulations. *Proceedings of the National Academy of Sciences of the United States of America*, 107(29):12799–803, July 2010.

[71] Miguel A. Morales, Carlo Pierleoni, and D. M. Ceperley. Equation of state of metallic hydrogen from coupled electron-ion Monte Carlo simulations. *Physical Review E*, 81(2):021202, February 2010.

[72] S Duane, AD Kennedy, BJ Pendleton, and D Roweth. Hybrid Monte Carlo. *Phys. Lett. B*, 195(2):216–222, 1987.

[73] S. Sorella. `TurboRVB` quantum monte carlo package (accessed date oct. 2014).

[74] C Filippi and C J Umrigar. Correlated sampling in quantum Monte Carlo: A route to forces. *Phys. Rev. B*, 61(24):R16291–R16294, 2000.

[75] C J Umrigar. Two Aspects of Quantum Monte Carlo: Determination of Accurate Wavefunctions and Determination of Potential Energy Surfaces of Molecules. *Int. J. Quantum Chem.*, pages 217–230, 1989.

[76] J R Trail. Heavy-tailed random error in quantum Monte Carlo. *Phys. Rev. E*, 77(1):016703, 2008.

[77] Thomas Kühne, Matthias Krack, Fawzi Mohamed, and Michele Parrinello. Efficient and Accurate Car-Parrinello-like Approach to Born-Oppenheimer Molecular Dynamics. *Phys. Rev. Lett.*, 98(6):066401, February 2007.

[78] FR Krajewski and M Parrinello. Linear scaling electronic structure calculations and accurate statistical mechanics sampling with noisy forces. *Phys. Rev. B*, 73(4):041105, 2006.

[79] F. Tassone, F. Mauri, and R. Car. Acceleration schemes for ab initio molecular-dynamics simulations and electronic-structure calculations. *Phys. Rev. B*, 50:10561–10573, Oct 1994.

[80] Paolo Giannozzi, Stefano Baroni, Nicola Bonini, Matteo Calandra, Roberto Car, Carlo Cavazzoni, Davide Ceresoli, Guido L. Chiarotti, Matteo Cococcioni, Ismaila Dabo, Andrea Dal Corso, Stefano de Gironcoli, Stefano Fabris, Guido Fratesi, Ralph Gebauer, Uwe Gerstmann, Christos Gougoussis, Anton Kokalj, Michele Lazzeri, Layla Martin-Samos, Nicola Marzari, Francesco Mauri, Riccardo Mazzarello, Stefano Paolini, Alfredo Pasquarello, Lorenzo Paulatto, Carlo Sbraccia, Sandro Scandolo, Gabriele Sclauzero, Ari P. Seitsonen, Alexander Smogunov, Paolo Umari, and Renata M. Wentzcovitch. QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials. *J. Phys. Cond. Matt.*, 21(39), SEP 30 2009.

[81] Edgar Bright Wilson, J C Decius, and Paul C Cross. *Molecular vibrations: the theory of infrared and Raman vibrational spectra*. Dover Pubns, 1955.

[82] Marie-Pierre Gaigeot, Michaël Martinez, and Rodolphe Vuilleumier. Infrared spectroscopy in the gas and liquid phase from first principle molecular dynamics simulations: application to small peptides. *Mol. Phys.*, 105(19-22):2857–2878, October 2007.

[83] B A Hess, L J Schaad, P Carsky, and R Zahradnik. Abinitio Calculations of Vibrational-Spectra and Their Use in the Identification of Unusual Molecules. *Chem. Rev.*, 86(4):709–730, 1986.

[84] A P Scott and L Radom. Harmonic vibrational frequencies: An evaluation of Hartree-Fock, Moller-Plesset, quadratic configuration interaction, density functional theory, and semiempirical scale factors. *J. Phys. Chem.*, 100(41):16502–16513, 1996.

[85] S Carter, S J Culik, and J M Bowman. Vibrational self-consistent field method for many-mode systems: A new approach and application to the vibrations of CO adsorbed on Cu(100). *J. Chem. Phys.*, 107(24):10458–10469, 1997.

[86] J Koput, S Carter, and N C Handy. Ab initio prediction of the vibrational-rotational energy levels of hydrogen peroxide and its isotopomers. *J. Chem. Phys.*, 115(18):8345–8350, 2001.

[87] P Cassam-Chenai and J Lievin. Alternative perturbation method for the molecular vibration-rotation problem. *Int. J. Quantum Chem.*, 93(3):245–264, 2003.

[88] J M Bowman. The Self-Consistent-Field Approach to Polyatomic Vibrations. *Acc. Chem. Res.*, 19(7):202–208, 1986.

[89] R B Gerber and M A Ratner. Self-Consistent-Field Methods for Vibrational Excitations in Polyatomic Systems. *Adv. Chem. Phys.*, 70:97–132, 1988.

[90] G M Chaban, J O Jung, and R B Gerber. Ab initio calculation of anharmonic vibrational states of polyatomic systems: Electronic structure combined with vibrational self-consistent field. *J. Chem. Phys.*, 111(5):1823–1829, 1999.

[91] G Czako, T Furtenbacher, A G Csaszar, and V Szalay. Variational vibrational calculations using high-order anharmonic force fields. *Mol. Phys.*, 102:2411–2423, 2004.

[92] Edit Matyus, Gabor Czako, Brian T. Sutcliffe, and Attila G Csaszar. Vibrational energy levels with arbitrary potentials using the Eckart-Watson Hamiltonians and the discrete variable representation. *J. Chem. Phys.*, 127(8):084102, 2007.

[93] W Schneider and W Thiel. Anharmonic-Force Fields From Analytic 2nd Derivatives - Method and Application to Methyl-Bromide. *Chem. Phys. Lett.*, 157(4):367–373, 1989.

[94] D A Clabo, W D Allen, R B Remington, Y Yamaguchi, and H F Schaefer. A Systematic Study of Molecular Vibrational Anharmonicity and Vibration-Rotation Interaction by Self-Consistent-Field Higher-Derivative Methods - Asymmetric-Top Molecules. *Chem. Phys.*, 123(2):187–239, 1988.

[95] V Barone. Anharmonic vibrational properties by a fully automated second-order perturbative approach. *J. Chem. Phys.*, 122(1):014108, 2005.

[96] Andrea Zen, Delyan Zhelyazov, and Leonardo Guidoni. Optimized Structure and Vibrational Properties by Error Affected Potential Energy Surfaces. *J. Chem. Theory Comput.*, 8(11):4204–4215, November 2012.

[97] Ryōgo Kubo, Morikazu Toda, and Natsuki Hashitsume. *Statistical Physics II*, volume 31 of *Springer Series in Solid-State Sciences*. Springer Berlin Heidelberg, Berlin, Heidelberg, 1985.

[98] Marie-Pierre Gaigeot and Michiel Sprik. Ab Initio Molecular Dynamics Computation of the Infrared Spectrum of Aqueous Uracil. *J. Phys. Chem. B*, 107(38):10344–10358, September 2003.

[99] E Bright Wilson. Some Mathematical Methods for the Study of Molecular Vibrations. *J. Chem. Phys.*, 9(1):76, 1941.

[100] M Burkatzki, C Filippi, and M Dolg. Energy-consistent pseudopotentials for quantum monte carlo calculations. *J. Chem. Phys.*, 126(23):234105, 2007.

[101] Sandro Sorella. Variational Monte Carlo and Markov Chains for Computational Physics. pages 207–236. Springer Berlin Heidelberg, Berlin, Heidelberg, April 2013.

[102] W S Benedict, N Gailar, and Earle K Plyler. Rotation-Vibration Spectra of Deuterated Water Vapor. *J. Chem. Phys.*, 24(6):1139–1165, 1956.

[103] Nist chemistry webbook.

[104] Computational chemistry comparison and benchmark database.

[105] Andrea Zen, Ye Luo, Guglielmo Mazzola, Leonardo Guidoni, and Sandro Sorella. Ab-initio molecular dynamics simulation of liquid water by Quantum Monte Carlo. *submitted*, 2014.

[106] I Errea, M Calandra, and F Mauri. Anharmonic free energies and phonon dispersions from the stochastic self-consistent harmonic approximation: Application to platinum and palladium hydrides. *Phys. Rev. B*, 89:064302, 2014.

[107] Donald Kennedy and Colin Norman. What don't we know? *Science*, 309(5731):75, 2005.

[108] Sergei Izvekov, Michele Parrinello, Christian J. Burnham, and Gregory A. Voth. Effective force fields for condensed phase systems from ab initio molecular dynamics simulation: A new method for force-matching. *The Journal of Chemical Physics*, 120(23):10896–10913, 2004.

[109] K Laasonen, M Sprik, M Parrinello, and R Car. "Ab initio" liquid water. *J. Chem. Phys.*, 99(11):9080–9089, 1993.

[110] Michiel Sprik, Jürg Hutter, and Michele Parrinello. Ab initio molecular dynamics simulation of liquid water: Comparison of three gradient-corrected density functionals. *J. Chem. Phys.*, 105(3):1142, 1996.

[111] P H L Sit and Nicola Marzari. Static and dynamical properties of heavy water at ambient conditions from first-principles molecular dynamics. *J Chem Phys*, 122(20):204510, 2005.

[112] Jeffrey C Grossman, Eric Schwegler, Erik W Draeger, François Gygi, and Giulia Galli. Towards an assessment of the accuracy of density functional theory for first principles simulations of water. *J Chem Phys*, 120(1):300, 2004.

[113] Thomas D. Kühne, Matthias Krack, and Michele Parrinello. Static and Dynamical Properties of Liquid Water from First Principles by a Novel Car-Parrinello-like Approach. *Journal of Chemical Theory and Computation*, 5(2):235–241, February 2009.

[114] IFW Kuo, CJ Mundy, MJ McGrath, JI Siepmann, J VandeVondele, M Sprik, J Hutter, B Chen, ML Klein, F Mohamed, M Krack, and M Parrinello. Liquid water from first principles: Investigation of different sampling approaches. *J. Phys. Chem. B*, 108(34):12990–12998, 2004.

[115] Soohaeng Yoo, Xiao Cheng Zeng, and Sotiris S Xantheas. On the phase diagram of water with density functional theory potentials: The melting temperature of ice I[sub h] with the Perdew–Burke–Ernzerhof and Becke–Lee–Yang–Parr functionals. *J. Chem. Phys.*, 130(22):221102, 2009.

[116] Hee-Seung Lee and Mark E Tuckerman. Structure of liquid water at ambient temperature from ab initio molecular dynamics performed in the complete basis set limit. *J Chem Phys*, 125(15):154507, 2006.

[117] Hee-Seung Lee and Mark E Tuckerman. Dynamical properties of liquid water from ab initio molecular dynamics performed in the complete basis set limit. *J Chem Phys*, 126(16):164501, 2007.

[118] Teodora Todorova, Ari P Seitsonen, Jürg Hutter, I-Feng W Kuo, and Christopher J Mundy. Molecular Dynamics Simulation of Liquid Water: Hybrid Density Functionals ? *J. Phys. Chem. B*, 110(8):3685–3691, March 2006.

[119] Manuel Guidon, Florian Schiffmann, Jürg Hutter, and Joost VandeVondele. Ab initio molecular dynamics using hybrid density functionals. *J. Chem. Phys.*, 128(21):214104, 2008.

[120] Stefan Grimme. Accurate description of van der Waals complexes by density functional theory including empirical corrections. *J Comput Chem*, 25(12):1463–1473, 2004.

[121] Stefan Grimme. Semiempirical GGA-type density functional constructed with a long-range dispersion correction. *J Comput Chem*, 27(15):1787–1799, 2006.

[122] O von Lilienfeld, Ivano Tavernelli, Ursula Rothlisberger, and Daniel Sebastiani. Optimization of Effective Atom Centered Potentials for London Dispersion Forces in Density Functional Theory. *Phys Rev Lett*, 93(15):153004, October 2004.

[123] Jochen Schmidt, Joost VandeVondele, I F William Kuo, Daniel Sebastiani, J Ilja Siepmann, Jürg Hutter, and Christopher J Mundy. Isobaric-Isothermal Molecular Dynamics Simulations Utilizing Density Functional Theory: An Assessment of the Structure and Density of Water at Near-Ambient Conditions. *J Phys Chem B*, 113(35):11959–11964, September 2009.

[124] I-Chun Lin, Ari P Seitsonen, Maurício D Coutinho-Neto, Ivano Tavernelli, and Ursula Rothlisberger. Importance of van der Waals Interactions in Liquid Water. *J. Phys. Chem. B*, 113(4):1127–1131, January 2009.

[125] I-Chun Lin, Ari P Seitsonen, Ivano Tavernelli, and Ursula Rothlisberger. Structure and Dynamics of Liquid Water from ab Initio Molecular Dynamics—Comparison of BLYP, PBE, and revPBE Density Functionals with and without van der Waals Corrections. *J. Chem. Theory Comput.*, 8(10):3902–3910, October 2012.

[126] Omololu Akin-Ojo and Feng Wang. Effects of the dispersion interaction in liquid water. *Chemical Physics Letters*, 513(1-3):59 – 62, 2011.

[127] Jue Wang, G Román-Pérez, Jose M Soler, Emilio Artacho, and M V Fernández-Serra. Density, structure, and dynamics of water: The effect of van der Waals interactions. *J. Chem. Phys.*, 134(2):024516, 2011.

[128] Cui Zhang, Jun Wu, Giulia Galli, and François Gygi. Structural and Vibrational Properties of Liquid Water from van der Waals Density Functionals. *J Chem Theory Comput*, 7(10):3054–3061, October 2011.

[129] M J Gillan, F R Manby, M D Towler, and D Alfe. Assessing the accuracy of quantum Monte Carlo and density functional theory for energetics of small water clusters. *J. Chem. Phys.*, 136(24):244105, 2012.

[130] M J Gillan, D Alfè, A P Bartók, and G Csányi. First-principles energetics of water clusters and ice: a many-body analysis. *The Journal of chemical physics*, 139(24):244504, December 2013.

[131] D Alfe, A P Bartók, G Csányi, and M J Gillan. Communication: Energy benchmarking with quantum Monte Carlo for water nano-droplets and bulk liquid water. *J. Chem. Phys.*, 138(22):221102, 2013.

[132] Joseph Morrone and Roberto Car. Nuclear Quantum Effects in Water. *Phys. Rev. Lett.*, 101(1):017801, July 2008.

[133] Scott Habershon, Thomas E Markland, and David E Manolopoulos. Competing quantum effects in the dynamics of a flexible water model. *J. Chem. Phys.*, 131(2):024501, 2009.

[134] Francesco Paesani and Gregory A Voth. The Properties of Water: Insights from Quantum Simulations ? *J. Phys. Chem. B*, 113(17):5702–5719, April 2009.

[135] Francesco Paesani, Soohaeng Yoo, Huib J Bakker, and Sotiris S Xantheas. Nuclear Quantum Effects in the Reorientation of Water. *J. Phys. Chem. Lett.*, 1(15):2316–2321, August 2010.

[136] X Z Li, B Walker, and A Michaelides. Quantum nature of the hydrogen bond. *Proc. Natl. Acad. Sci. U. S. A.*, 108(16):6369–6373, April 2011.

[137] M Ceriotti, J Cuny, M Parrinello, and D E Manolopoulos. Nuclear quantum effects and hydrogen bond fluctuations in water. *Proc. Natl. Acad. Sci. U. S. A.*, 110(39):15591–15596, September 2013.

[138] Lawrie B Skinner, Congcong Huang, Daniel Schlesinger, Lars G M Pettersson, Anders Nilsson, and Chris J Benmore. Benchmark oxygen-oxygen pair-distribution function of ambient water from x-ray diffraction measurements with a wide Q-range. *J. Chem. Phys.*, 138(7):074506, 2013.

[139] A K Soper. The Radial Distribution Functions of Water as Derived from Radiation Total Scattering Experiments: Is There Anything We Can Say for Sure? *ISRN Phys. Chem.*, 2013(11):1–67, 2013.

[140] A K Soper. The radial distribution functions of water and ice from 220 to 673 K and at pressures up to 400 MPa. *Chem. Phys.*, 258(2-3):121–137, August 2000.

[141] See s. sorella's talk at the "advances in quantum monte carlo techniques for non-relativistic many-body systems". http://www.int.washington.edu/talks/WorkShops/int_13_2a/, June 27, 2013.

[142] Shepard A Clough, Yardley Beers, Gerald P Klein, and Laurence S Rothman. Dipole moment of water from Stark measurements of H2O, HDO, and D2O. *J. Chem. Phys.*, 59(5):2254, 1973.

[143] J A Odutola and T R Dyke. Partially deuterated water dimers: Microwave spectra and structure. *J. Chem. Phys.*, 72(9):5062, 1980.

[144] S22-benchmark noncovalent complexes.

[145] W Klopper, J G C M van Duijneveldt-van de Rijdt, and F B van Duijneveldt. Computational determination of equilibrium geometry and dissociation energy of the water dimer. *Phys. Chem. Chem. Phys.*, 2(10):2227–2234, 2000.

[146] Ye Luo, Andrea Zen, and Sandro Sorella. Ab-initio molecular dynamics with noisy forces: validating the quantum monte carlo approach with benchmark calculations of molecular vibrational properties. *submitted*, 2014.

[147] Mauro Del Ben, Mandes Schönherr, Jürg Hutter, and Joost VandeVondele. Bulk Liquid Water at Ambient Temperature and Pressure from MP2 Theory. *J. Phys. Chem. Lett.*, 4(21):3753–3759, November 2013.

[148] Sebastian Fritsch, Raffaello Potestio, Davide Donadio, and Kurt Kremer. Nuclear Quantum Effects in Water: A Multiscale Study. *J. Chem. Theory Comput.*, page 140123151236004, January 2014.

[149] Paolo Raiteri, Alessandro Laio, and Michele Parrinello. Correlations among hydrogen bonds in liquid water. *Phys. Rev. Lett.*, 93:087801, Aug 2004.

[150] Zhonghua Ma, Yanli Zhang, and Mark E. Tuckerman. Ab initio molecular dynamics study of water at constant pressure using converged basis sets and empirical dispersion corrections. *The Journal of Chemical Physics*, 137(4):–, 2012.

[151] Eric Schwegler, Giulia Galli, and Fran çois Gygi. Water under pressure. *Phys. Rev. Lett.*, 84:2429–2432, Mar 2000.

[152] Rob Neely. Supercomputing 101: A history of platform evolution and future trends. In *DOE CSGF*, 2014.