

UNIVERSITÉ PIERRE ET MARIE CURIE
UNIVERSIDAD NACIONAL DE SAN MARTÍN

Spécialité :

Physique

présentée par

Mario DAGRADA

Pour obtenir le grade de:

Docteur de
l'Université Pierre et Marie Curie et
de l'Universidad Nacional de San Martín

Improved quantum Monte Carlo simulations: from open to extended systems

---

dirigée par Francesco Mauri et Veronica Vildosola.
Co-directeur: Michele Casula

Soutenue le 28 Septembre 2016 devant le jury composé de :

M. Matthew Foulkes	Imperial College of London (Royaume-Uni)	Rapporteur
M. Carlo Pierleoni	Universitá dell'Aquila (Italie)	Rapporteur
M. Omar Fojon	Instituto de Física de Rosario (Argentine)	Rapporteur
Mme. Silke Biermann	École Polytechnique - CPHT	Examinateur
M. Bernard Bernu	UPMC - LPTMC	Examinateur
M. Michele Casula	UPMC - IMPMC	Co-directeur
Mme. Ana Maria Llois	Universidad Nacional de San Martin (Argentine)	Examinateur
Mme. Valeria Ferrari	Universidad Nacional de San Martin (Argentine)	Examinateur

Institut de minéralogie, de physique des matériaux et de cosmochimie
Université Pierre et Marie Curie
4, place Jussieu
75252 Paris CEDEX 05

École doctorale: Physique en Île-de-France

Acknowledgements

These three years have been definitely the most intense yet enriching and rewarding years of my life. The first thank must obviously go to my several supervisors.
At first Francesco who has been a guide, both scientifically and personally. Even if he did not follow me directly, I knew that I could always count on him and he demonstrated this to me several times.
Michele, with his experience and his ideas, has been surely the most important driving force of my work during the PhD. I really enjoyed spending time with him discussing and learning and, even if I decided not to take the path of physics research for the time being, I really hope we can continue to have the same fruitful collaboration we had during these years. Besides physics, he also taught me, with his example, the importance of professionalism and hard-working, I will always keep this in mind.
I had the incredible chance of living for almost a year in Argentina during my PhD. This would have not been possible without the help, the understanding and the passion of Veronica; she has been so nice and helpful that I was immediately feeling comfortable when working there. I will be back soon to CNEA for a visit and maybe longer.
A special thanks goes to Sandro. I do not exaggerate by saying that without the collaboration with him my PhD would have not been the same. He is one of the most passionate, dedicated and expert people I have ever known and I learned very much from him as he was always ready for an explanation, even for the smallest thing. Although I must admit that a lot of the *TurboRVB* code is still magical to me! :-) I also would like to thank Marco very much since he likely gave me the best advices till the very end of my PhD and he was always there for a little chat despite his very busy schedule. Finally, I would like to thank also Ruben for all the discussions we had, not only about physics.

I have met so many people in different countries during these years that it is really impossible to list them all here. With Adrien we spent a great time speaking about unusual French expression and (trying to) smoking cigars! I know that our friendship will remain the same also after the PhD. The same with Sara, who has the awesome ability to make me seeing things in a much clearer way when discussing with her. David, one of the most funny people I have ever met; we share many aspects of our personality, that is probably why we are so close to each other. I should point out, however, that I do not share his "feeling"! Andrea has become in a short time one of the closest friends of mine, I can safely say that we have grown together during these years and I know we will continue like this. Then Guilherme, who is one of the kindest and most generous people I know; I would have like to spend more time with him, especially during the last months, but I am sure we will find a way to stay in touch. Francesca, it was quite incredible coincidence

to meet her since, coming from the same city and same high school, we ended up in the same laboratory 1000 km away. She saved my life once and then we had a wonderful "coloc". It will not be difficult to see each other often! I conclude the Paris side with my fellow QMC mates! Nicolas – l'homme de la pause – has created such a nice environment in the lab, we have become very good friends, but only until he will beat me at tennis (so never)! :) The other Nicolas, instead, taught me a lot about French culture (although a little bit old I must say) and we had infinite discussions about science. Henry has helped me a lot during our daily fights with QMC and the mysteries of *TurboRVB*. With Felix we share many nice experiences, our trip to Telluride was one of the best I had.

Thanks also to Nieli whose smile made me feel better, Edoardo who has been a perfect office mate, Ornella, Eric, Johan, Aisha, Fabien, Cyrille, Sadaf, ZheSheng, Maya, Nina, Melanie, Thibaut, Betul, Raffaello, Lorenzo, Matteo, Francesca, Giulio, Achille and Alessandro.

My Argentinian/Chilean friends and colleagues Dilson, Abraham, Santiago, Leandro, Ramiro, Pablo, Ilan, Alejandra, Marcelo, Emmanuel and many others, they always helped out and thank to them I discovered the rich and interesting Argentinian culture, from politics to how to play paddle! However, my visit in Argentina would have not been the same without the De Vedia-Vera family: Maria Martha, Lautaro, Sofi, Mechi, Fede, Hernan just to mention the closest ones. They have been my second family (in many ways) and I am looking forward to be back there for spending more time with them.

Finally, the most important ones. My parents, who are always there for me in any situation, they are simply the best. My brother who gave me so many good advises. And my girlfriend Teresita, who has listened and helped me countless times. My whole PhD experience has started from her and with her, and it was definitely worth it!

Résumé

Dans cette thèse nous présentons des progrès algorithmiques ainsi que plusieurs applications des méthodes de Monte Carlo quantique (QMC) pour les calculs de structure électronique à partir des premiers principes. En général, le QMC est une téchnique qui donne des résultats très précis même pour des systèmes à forte corrélation électronique. En dépit d'un coût computationnel élevé, ses atouts restent l'allure polynomiale N^4 en fonction de la taille (N) du système et une très efficace parallelisation algorithmique. Les améliorations que nous proposons permettent d'étudier par QMC des systèmes de plus grosse taille voire infiniment étendus (périodiques), avec l'ambition de faire du QMC une alternative valable à la théorie de la fonctionnelle de la densité (DFT). Tous les résultats ont été obtenus par le logiciel TurboRVB, dont nous avons contribué au développement récent. D'abord, nous présentons une implémentation du QMC basée sur la fonction d'onde Jastrow-Geminale qui combine une grande flexibilité avec une forme analytique compacte, tout en gardant un traitement précis des corrélations électroniques. Nous avons appliqué une technique originale de plongement de l'atome dans son propre environnement pour réduire la taille de la base atomique d'orbitales à la molécule d'eau ainsi qu'à un modèle simplifié du transfert de protons (TP) dans l'eau. Nos résultats ouvrent la voie à l'étude des phénomènes microscopiques tels que le TP directement par QMC. Ensuite, nous avons amélioré notre méthode afin de simuler les solides cristallins. Grâce à une nouvelle procédure pour choisir de manière appropriée les conditions aux limites de la fonction d'onde, nous avons pu réduire considérablement les erreurs de taille finie qui affectent les simulations QMC des solides. Sur la base des techniques développées, nous étudions enfin le FeSe, supraconducteur à base de fer. Le QMC fournit le meilleur résultat concernant la structure cristalline du FeSe ; via une étude systématique du paysage énergétique à différentes configurations magnétiques, nous montrons un lien fort entre la structure, le magnétisme et les mouvements de charge dans ce matériau, prélude à une compréhension quantitative de la supraconductivité à haute température critique des premiers principes.

Abstract

In this thesis we present algorithmic progresses as well as applications of continuum quantum Monte Carlo (QMC) methods for electronic structure calculations by first principles. QMC generally yields very accurate results even for systems characterized by strong electron correlation. Despite its high computational cost, its advantages are the polynomial N^4 scaling with the number of particles (N) in the system and the negligible parallelization overhead of its algorithms. The improvements we propose allow to tackle much larger molecular as well as extended (infinite-size) systems by QMC, with the ultimate goal of making QMC a valid alternative to density functional theory (DFT). All results have

been obtained with the TurboRVB software, which we contributed to develop. At first, we present a QMC framework based on the Jastrow-Geminal wavefunction which combines great flexibility with a compact analytical form, while providing at the same time an accurate treatment of electron correlations. We apply an original atomic embedding scheme for reducing the basis set size to the water molecule and to a simple model of proton transfer (PT) in aqueous systems. Our results pave the way to the study of microscopic phenomena such as PT directly by QMC. Afterwards, we extend our QMC framework in order to simulate crystalline solids. We propose a novel procedure to find special values of the boundary conditions which allow to greatly reduce the finite-size errors affecting solid state QMC simulations. Using the techniques previously developed, we study the iron-based superconductor FeSe. We show that QMC provides the best crystal structure predictions on this compound; by means of a systematic study of the energy landscape at different magnetic orderings, we show a strong link between structural, magnetic and charge degrees of freedom in FeSe. Our results represent an important step towards a quantitative understanding of high-temperature superconductivity by first-principles.

Contents

Chapter 1

Introduction

The complexity of the problems tackled by modern science is growing faster and faster over the years. Nowadays, it is virtually impossible to find an analytical, pen-and-paper solution to any of the models developed for these problems, except that in particular and, in general, non-interesting situations. This aspect is a common thread of all scientific domains investigating our reality, ranging from the study of macroscopic phenomena such as the merge of two galaxies (scale of 10^{20} m) to the interaction of infinitely small particles such as quarks (scale of 10^{-15} m). If one focuses on the quantum mechanical description of atomic physics and chemistry, this fact is well summarized by the famous quotation from P. M. Dirac appeared in 1929: "*The underlying physical laws necessary for the mathematical theory of a large part of physics and the whole of chemistry are thus completely known, and the difficulty is only that the exact application of these laws leads to equations much too complicated to be soluble.*"

This evidence, combined with the computer revolution of the last few decades, paved the path to develop a new broad scientific domain: computational science. More than a third way of doing science, as it is usually considered, computational science bridges the gap between the two traditional pillars, namely analytical theory and experimentation, to build appropriate algorithms which can solve the otherwise unsolvable models with the aid of a computer. The comparison and the validation of the outcome with experimental results is obviously unavoidable. In some sense, computational science is a brute force, number-crunching approach to the aforementioned challenge of modern science. And its success demonstrates that this approach works in most situations.

The Schrödinger equation, the cornerstone of quantum mechanics, has been one of the first problems to be fed to a computer in the early 60s, thus giving birth to the field of first principles or *ab initio* calculations, namely the branch of computational science which devises methods to find approximate solutions of the full electronic Schrödinger equation

of atom and molecules (computational quantum chemistry) and solids (computational condensed matter).

Among *ab initio* techniques, the most important and successful framework is certainly density functional theory (DFT). The progress in DFT have advanced so far that it is nowadays possible to predict, simply by computer simulation, new materials with the desired electronic properties which can then be realized experimentally in a sort of reverse-engineering process. However the accuracy of DFT cannot be determined *a priori* since it is founded on an uncontrollable approximation for mapping the full many-body problem into an effective single-particle one. Furthermore it is based on the local electronic density only and hence it does not allow to access to the most important mathematical object of a quantum system, i.e. the electronic wavefunction. This information is however unavoidable for the fundamental understanding of many exotic phenomena such as high-temperature superconductivity, superfluidity, quantum Hall effects and likely a lot of yet undiscovered others.

The main focus of this thesis is on an alternative strategy based on the *ab initio* (or continuum) quantum Monte Carlo (QMC) techniques. The idea behind QMC is fascinatingly simple: propose a suitable form of the Hamiltonian ground state eigenvector, i.e. the electronic many-body wavefunction, and approximately solve the corresponding *full* Schrödinger equation to get the energy and other observables out of the simulation. The most important strength of continuum QMC is that it ensures a direct access to the many-body wavefunction and thus it gets extremely close to a realistic quantum mechanical description of how nature works at the atomic scale. This aspect, combined with its stochastic implementation required by the high dimensionality of the quantum problem, allows QMC to reach flexibility and accuracy far beyond one-electron methods both in open systems and in crystalline solids, as we will demonstrate throughout this thesis.

Furthermore, the computer architecture is rapidly changing over the years. The next generation supercomputing machines will pack up to 72 cores in each of their CPUs; general purpose GPUs presently used in scientific computing allow to exploit up to 65000 threads, i.e. independent processes, at the same time. The future of computer science is in parallelization, therefore it is almost certain that the future of *ab initio* calculations lies in methods able to fully benefit from such level of massive parallelism. QMC is likely the sole first-principles approach which is naturally built for the next generation computing architectures, thanks to its “embarrassingly“ parallel algorithms based on a large number of independent processes used in the stochastic evaluation of electronic integrals.

Thesis motivation and outline

Despite being a very promising and robust framework, nowadays there are only few groups around the world that make an intensive use of QMC. This is due to the intrinsic computational cost of QMC algorithms with respect to DFT. This issue requires algorithmic improvements and more computational power accessible worldwide. Furthermore the available QMC codes are relatively few, often not well documented and a great deal of human time is needed in order to carry out meaningful and reliable calculations, at variance with the almost black box nature of DFT packages.

This thesis aims at addressing the twofold nature of this problem and at giving a contribution towards a more widespread use of QMC techniques in electronic structure calculations. On the one hand, we propose several improvements to present QMC algorithms as well as new techniques for reducing the computational time necessary to carry out QMC calculations of realistic quantum systems. On the other hand, we describe their implementation within the *TurboRVB* [1] package for quantum Monte Carlo simulations, with the ultimate goal of making this powerful software package available under an open source license.

The structure of this manuscript closely follows the evolution of this QMC software carried out during the thesis, starting from molecular systems and evolving towards an accurate QMC description of crystalline materials. We always complement new algorithmic developments and programming efforts with extensive applications to diverse physical problems, outlined in the following. These aspects make the present manuscript a genuine interdisciplinary work in between computational chemistry and condensed matter physics where QMC techniques play the role of common thread. This thesis is organized in two main parts, each composed by multiple self-contained Chapters which can be read independently:

Part 1: QMC methods for finite systems

1. *Chapter 2*. In this Chapter we present a general overview of the full QMC framework used in this thesis for studying finite, open shell systems. Many of the topics addressed in this Chapter are often reused or evolved in the rest of the manuscript. The main focus is on the powerful Jastrow-geminal wavefunction ansatz, which can include high-level electron correlation within QMC at a reasonable computation cost. Preliminary DFT calculations, wavefunction optimization methods and single-point QMC techniques for computing observables are also detailed.

2. *Chapter 3*. In this Chapter we present extensive applications of the QMC methods introduced before to an open problem in quantum chemistry: the proton transfer

reaction in liquid water and aqueous systems. At first we devise a novel procedure for reducing the number of parameters of the QMC wavefunction and thus ensuring a considerable speedup in the calculations. We test our algorithm on the simple water molecule and then we apply it to a model of proton transfer reactions composed by two H_2O and an excess proton. We show our QMC implementation to be a highly accurate and less expensive alternative to deterministic quantum chemistry approaches for proton transfer systems.

Part 2: QMC methods for periodic systems

1. *Chapter 4.* By keeping the same structure as Chapter 2, we present all the important steps carried out for evolving our QMC framework from open systems towards an accurate treatment of extended solids. This evolution is of paramount importance since QMC is one of the few highly correlated techniques able to perform realistic solid state calculations.

2. *Chapter 5.* This Chapter is devoted to the most important issue which plagues QMC calculations of extended systems: the systematic errors deriving from the finite, and generally small size of the simulation cell. We test the standard techniques against a new procedure devised during this thesis which allows to find special values of the phase in front of the wavefunction within periodic boundary conditions. They can greatly reduce size errors coming from the kinetic energy part of the Hamiltonian. We apply our procedure to several realistic systems used as benchmark.

3. *Chapter 6.* In the last Chapter, we apply the solid state QMC methods to a major open problem in condensed matter physics: high-temperature superconductivity. We study structural and magnetic properties of iron selenide (FeSe), a compound belonging to the family of iron-based superconductors, with two complementary QMC approaches. This work is done in collaboration with the group of Lucas K. Wagner at the University of Illinois, Urbana-Champaign. We show that QMC captures the physics of FeSe with an unprecedented level of accuracy and that it can be the *ab initio* method of choice to tackle high-temperature superconductivity.

The algorithmic progresses as well as the extensive applications of QMC presented in this thesis prove that this method is getting closer to become the next generation standard in *ab initio* calculations.

Chapter 2

Quantum Monte Carlo methods for finite systems

2.1 Introduction

The term quantum Monte Carlo (QMC) broadly refers to a set of techniques for computing electronic structure properties of quantum systems. The main motif underlining all QMC methods is their extensive use of Monte Carlo stochastic integration for finding approximate solutions to the many-body stationary Schrödinger equation:

$$\hat{H}\Psi = E\Psi. \tag{2.1}$$

In this thesis we will always assume the Born-Oppenheimer approximation to hold, hence the Hamiltonian operator $\hat{H}$ of a system containing N electrons and $N_{\rm at}$ atoms in its non relativistic form reads:

$$\begin{aligned} \hat{H} &= -\frac{1}{2}\sum_{i=1}^{N}\Delta_{\mathbf{r}_i} + \frac{1}{2}\sum_{i=1}^{N}\sum_{j\neq i}^{N}\frac{1}{|\mathbf{r}_i-\mathbf{r}_j|} - \sum_{a=1}^{N_{\rm at}}\sum_{i=1}^{N}\frac{Z}{|\mathbf{r}_i-\mathbf{R}_a|} \\ &= \hat{T}+\hat{W}+V_{\rm ext} \end{aligned} \tag{2.2}$$

where we denote as $\mathbf{r}$ ($\mathbf{R}$) the electronic (ionic) positions and Z the atomic charges. $\hat{T}$ is the kinetic energy, $\hat{W}$ the electron-electron interaction and $V_{\rm ext}$ the external potential generated by the atomic arrangement. In Eq. 2.2 we have used atomic units ($e=1,\ \hbar=1,\ m_e=1$). This convention is kept throughout this manuscript.

The focus of this work is on continuum QMC methods whose fundamental object is the many-body trial wavefunction $\Psi_T(\mathbf{r}_1,\cdots,\mathbf{r}_N)$. By exploiting the variational principle, these methods attempt to find the best approximate solution of Eq. 2.1 using Ψ_T as an ansatz for the true Hamiltonian ground state eigenvector Ψ_0. The first application of

QMC dates back to the work on liquid He by MacMillan in 1965 [2]. The QMC techniques routinely available nowadays belong to two main categories, depending on the way the ansatz is sampled for filtering out the ground state. Within the first category, the wavefunction is evaluated in the full $3N$-dimensional configurational space generated by electronic coordinates. This includes variational Monte Carlo (VMC), Green function Monte Carlo (GFMC) [3] which has evolved to modern diffusion Monte Carlo (DMC) [4] with its variances – lattice regularized DMC [5] and release-node DMC [6]. These numerous QMC flavors can be combined with the path integral formalism in order to perform finite temperature simulations and, recently, molecular dynamics simulations [7, 8]. The second category contains the full configuration interaction quantum Monte Carlo (FCIQMC) [9], which samples stochastically the full set of determinants generated by the chosen basis set and the auxiliary field quantum Monte Carlo (AFQMC) [10] which rewrites the many-body interaction in terms of independent particles described by determinants interacting with an external auxiliary field.

The direct competitors of QMC in the domain of highly correlated first principles calculations of atom and molecules are quantum chemical techniques such as coupled cluster (CC) methods, many-body perturbation theory (MB) and the configuration interaction (CI) approach in its single-double (CISD) or full (FCI) variants. However, compared with these deterministic approaches QMC has two crucial advantages.

The first, and probably most obvious one is the stochastic nature of its backbone algorithms used to cope with the high dimensionality of electronic integrals. In fact, the central limit theorem ensures that the MC integration error ϵ, namely the statistical error, is independent on the dimensionality of the problem and it can be written as:

$$\epsilon \sim \sigma \sqrt{\frac{N}{N_{\mathrm{MC}}}} \tag{2.3}$$

where σ^2 is the variance on a single measurement and N_{MC} is the number MC samples or integration points. The statistical error in Eq. 2.3 can be reduced at will just by increasing the number N_{MC} of samples and therefore it is, in principle, possible to reach any target accuracy simply with more computer time.

Another consequence of stochasticity is that QMC algorithms can be straightforwardly parallelized by assigning a different random points used for integration to each processor. Since each point evolves independently during a QMC run, there is a negligible overhead due to the communication between processors. Diffusion Monte Carlo represents a special case to this respect since its standard algorithm involves more communication, as we will see in Sec. 2.4.2.

The second superior feature of QMC methods is their flexibility, which mainly results from the choice of the wavefunction. The functional form of the ansatz can indeed be

chosen without basically any restriction, except from being square-integrable and computable in a finite amount of time. However, a good trial wavefunction must cope with some practical requirements. At first, it is typically evaluated thousands of times during a simulation in order to achieve the desired statistical accuracy. It is thus of paramount importance to devise a compact and easily tractable functional form. Furthermore, the ansatz must be able to describe electron correlation, i.e. include space-space correlators, in order to account for a relevant part of the correlation energy. This important quantity is defined as:

$$E_{\text{corr}} = E_{\text{exact}} - E_{\text{HF}}, \tag{2.4}$$

where E_{HF} is the Hartree-Fock energy obtained by keeping only the exchange contribution in the interacting part of the Hamiltonian, and E_{exact} is the exact non-relativistic ground state energy. While E_{corr} represents only a small fraction of the total energy, it is fundamental to achieve the chemical accuracy ($\leq$ 1 Kcal/mol equivalent to 0.01 eV per atom) necessary to describe the energetics of weakly bonded molecular systems, but also to correctly reproduce the phase diagram of many solids. Examples of situations where this precision is unavoidable are presented in Chap. 3, Chap. 5 and Chap. 6.

This aspect becomes particularly relevant in order to make QMC methods competitive against other highly correlated techniques. Indeed, at variance from widely used CC methods (scaling $\sim N^5$ up to $\sim N^7$ when including double and perturbative triple excitations) or the FCI approach (exponential scaling), QMC displays a very favorable scaling of N^3-N^4 with the number electrons. Thanks to Eq. 2.3, it is evident that using QMC, chemical accuracy can be attained not only on small molecular models, but on much larger clusters not feasible with the other methods. This will be extensively discussed in the next Chapter.

Finally, all the QMC algorithms used in this work obey the so-called *zero-variance* principle, which is a consequence of the general variational principle applied to the evaluation of quantum expectations. The zero-variance principle establishes an intimate link between the quality of the trial wavefunction and the efficiency of a QMC simulation. It states that the statistical error on the final outcome goes to zero in the ideal case of the trial wavefunction Ψ_T being an exact eigenstate Ψ of the Hamiltonian. Therefore, if $\langle \Psi_T | \Psi \rangle \ll 1$, QMC simulations yield unreliable results affected by a large statistical error. It is not surprising that the starting ansatz is crucial. Wavefunction optimization methods have triggered a lot of efforts. The main procedures developed over the past years are efficient generalizations of Newton and steepest descent techniques for minimizing multivariate functions and in general they exploit stochastic algorithms.

In conclusion, it is worth noting that QMC techniques remains considerably more expensive than single-particle approaches such as Hartree-Fock, or effectively indepen-

dent electron methods such as density functional theory (DFT). Moreover, whereas it is straightforward to build simple and unbiased statistical estimators for the total energy, it is not the case for other important quantities such as the electron density or the ionic forces; also excited states are difficult to obtain and full spectra are not yet available out of a QMC simulation. However, for all aforementioned features, we believe that this trend will soon change as more computational power and improved algorithms will be available over the next decades.

2.1.1 Organization of the Chapter

This Chapter focuses on the QMC methods implemented in the *TurboRVB* [1] software. It is organized as follows. In Sec. 2.2 we introduce our DFT implementation on a localized basis set necessary to obtain an appropriate starting point for correlated QMC calculations. In Sec. 2.3 we describe a powerful wavefunction ansatz formed by a Jastrow correlated factor applied to a multi-configurational antisymmetrized geminal power (AGP). This highly flexible ansatz includes all flavors of electron correlation and it is expanded over a localized basis set whereas core electrons of the heavier atoms are replaced by pseudopotentials specifically designed for QMC calculations. The two main QMC techniques used in this work for evaluating total energies and other observables, i.e. VMC and lattice regularized DMC (LRDMC), are extensively detailed in Sec. 2.4. Finally, we present an efficient optimization procedures for both the wavefunction and the nuclear positions in Sec. 2.5. This energy minimization technique is based on the stochastic reconfiguration method and, at variance with previous implementations, it extensively uses the adjoint algorithmic differentiation programming technique for computing wavefunction and local energy derivatives. Finally, we briefly draw the conclusions of this Chapter in Sec. 2.6.

2.2 Density functional theory on a localized basis set

Calculations with density functional theory represent an essential step towards more accurate QMC simulations. The reason is twofold.

At first place, extensive benchmarks [11, 12] have proven that, compared to Hartree-Fock or natural orbitals, the DFT eigenstates obtained either with plain local density approximation or with hybrid energy functionals are the best single-particle starting guess for building the determinantal part of the QMC wavefunction. Besides the wavefunction, also the relaxed geometrical structure of molecules and solids at a DFT level is often used as initial point for QMC calculations, in particular when a correspondent experimental result is not available.

Secondly, the physics extracted from DFT calculations represents an essential preliminary scenario for interpreting the results of more accurate many-body techniques such as QMC. Due to its higher computational cost, it is nowadays advantageous to use QMC only to tackle physical problems where the DFT description is not satisfactory and which does not require a too large number of particles in the simulation model. This is the case for all applications presented in Chap. 3 and Chap. 6 of this thesis. This trend will likely change in the near future. With the algorithms available today, DFT will not be able to benefit of the increased computational power of the next generation supercomputers (*exascale* computing). On the contrary, as already mentioned, QMC methods will naturally benefit from the very large number of available cores for performing realistic simulations on systems with up to 10^4 electrons, at least one order of magnitude more of the capability of present QMC programs.

At variance with most of the available QMC softwares which rely on external programs, the *TurboRVB* code used in this thesis exploits a built-in DFT package which seamlessly interfaces with the QMC software, thus using the same simulation setup and avoiding any external tool for wavefunction conversion.

2.2.1 The Kohn Sham equations

DFT is a theory of many-body quantum systems whose purpose is to evaluate electronic structure properties of atoms, molecules and solids by first principles.

The basic idea of DFT is fundamentally different from wavefunction-based approaches such as Hartree-Fock configuration interaction as well as QMC methods. Rather than the ground state many-body wavefunction $\Psi(\mathbf{r}_1, \cdots, \mathbf{r}_N)$, the main physical object within DFT is the local electronic density $\rho(\mathbf{r})$, namely the diagonal elements of the one-body density matrix. This basic quantum mechanical quantity is defined as:

$$\rho(\mathbf{r}) = N \int d\mathbf{r}_2 \dots d\mathbf{r}_N \Psi^*(\mathbf{r}, \mathbf{r}_2, \dots, \mathbf{r}_N)\Psi(\mathbf{r}, \mathbf{r}_2, \dots, \mathbf{r}_N) \tag{2.5}$$

where N is the number of electrons in the system. The constitutional tenet of DFT is that *any* property of a many-body system can be casted into a functional of the ground state electronic density $\rho_0(\mathbf{r})$. In this way the actual functional form of $\Psi(\mathbf{r}_1, \cdots, \mathbf{r}_N)$ becomes irrelevant and needs not to be explicitly specified.

It can be shown that $\rho_0(\mathbf{r})$ is uniquely defined by the external potential $V_{\text{ext}}(\mathbf{r})$ which, in absence of external fields, only depends on the considered atomic arrangement (this goes under the name of *V-representability* of the electronic density). The demonstration of this simple yet revolutionary theorem is presented in the original work of Hohenberg and Kohn in Ref. [13].

If one considers the part of the quantum Hamiltonian in Eq. 2.2 independent from the external potential, the above statement implies that it exists a *universal* density functional of the energy which can be written as:

$$\mathcal{F}[\rho] = \langle \Psi | \, \hat{T} + \hat{W} \, | \Psi \rangle \, . \tag{2.6}$$

This expression is valid for any external potential $V_{\text{ext}}(\mathbf{r})$. By means of a straightforward application of the variational principle, the DFT ground state energy $\mathcal{E}[\rho] = \langle \Psi | \, \mathcal{F}[\rho] \, | \Psi \rangle + \int d\mathbf{r} \rho(\mathbf{r}) V_{\text{ext}}(\mathbf{r})$ given a specific external potential $V_{\text{ext}}(\mathbf{r})$ is recovered by performing a global functional minimization of Eq. 2.6 with respect to the electronic density, i.e. $\rho(\mathbf{r}) \rightarrow \rho_0(\mathbf{r})$.

This density based approach for solving the Schrödinger equation is very appealing. However it has several caveats which would have confined DFT to a merely academic exercise. On one hand, there are no known recipes for devising the form of the functional $\mathcal{E}[\rho]$ which, in addition, is certainly non local. Furthermore, although the ground state density is in principle sufficient for solving the many-body problem, there are no meaningful ways for directly extracting physical observables from the electronic density.

The widespread success of DFT is due to the ansatz proposed by Kohn and Sham (KS). The basic assumptions of KS formalism introduced in their seminal paper in Ref. [14] are the following:

1. The many-body problem represented by the functional in Eq. 2.6 can be mapped *exactly* into a set of single particle equations. These particles are fictitious fermionic quasi-particles (usually called KS particles) with the sole constraint of possessing, collectively, the same electronic density $\rho(\mathbf{r})$ of the physical electrons.

2. The KS particles are assumed *non-interacting* or independent.

Due to their non-interacting nature, we can describe the KS particle with a set of single-particle orbitals $\phi_i(\mathbf{r})$ with $i \in [1, N_{\text{orb}}]$. Notice that these orbitals are defined on the whole system and thus they are molecular orbitals. Typically, they are expanded over a suitable basis set (Eq. 2.13). Possible choices for this basis are discussed more in detail in Sec. 2.2.2. Given these orbitals, the electronic density can be rewritten as:

$$\rho(\mathbf{r}) = 2 \sum_{i \in \text{occ}} |\phi_i(\mathbf{r})|^2, \tag{2.7}$$

where the summation goes over the occupied orbitals. Here, for sake of simplicity, we focus on the case of doubly occupied orbitals integrating out the electron spin (hence the factor of 2). The extension to spin polarized systems is straightforward. Since the KS

particles have the same density as the physical electrons, the functional in Eq. 2.6 can be written more explicitly as a function of the orbitals $\phi_i(\mathbf{r})$:

$$\begin{aligned}\mathcal{F}[\rho] &= \sum_i \langle\phi_i| -\frac{1}{2}\Delta_{\mathbf{r}} |\phi_i\rangle + \frac{1}{2}\int\int d\mathbf{r}d\mathbf{r}'\frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} + \mathcal{E}^{\mathrm{XC}}[\rho] \\ &= \mathcal{E}^{\mathrm{KE}} + \mathcal{E}^{\mathrm{Ha}}[\rho] + \mathcal{E}^{\mathrm{XC}}[\rho] \end{aligned} \tag{2.8}$$

The first term of Eq. 2.8 is the kinetic energy, the second is the electron-electron repulsive interaction (or Hartree term); the last term is, by definition, the exchange-correlation energy $\mathcal{E}^{\mathrm{XC}}[\rho] = \mathcal{E}^{\mathrm{X}} + \mathcal{E}^{\mathrm{corr}}$ which includes all the many-body effects not related to the simple Coulomb repulsion: the exchange contribution $\mathcal{E}^{\mathrm{X}}$ derives from quantum particles indistinguishability and the correlation energy $\mathcal{E}^{\mathrm{corr}}$ has been defined in Eq. 2.4. If one drops the $\mathcal{E}^{\mathrm{corr}}$ contribution and computes $\mathcal{E}^{\mathrm{X}}$ exactly, DFT reduces to the well-known Hartree-Fock approximation.

$\mathcal{F}[\rho] + \int d\mathbf{r}\rho(\mathbf{r})V_{\mathrm{ext}}(\mathbf{r})$ can be minimized by taking a variation of the density $\delta\rho(\mathbf{r})$ and imposing the conservation of the total number of particles, i.e. $\int d\mathbf{r}\delta\rho(\mathbf{r}) = 0$. This constraint can be taken into account by defining a Lagrange multiplier which fixes the chemical potential. This leads to the following set of single particle equations usually called KS equations:

$$\hat{H}^{\mathrm{KS}}\phi_i(\mathbf{r}) = \epsilon_i\phi_i(\mathbf{r}). \tag{2.9}$$

The full KS Hamiltonian matrix is defined as $\hat{H}^{\mathrm{KS}} = \frac{\delta\mathcal{F}[\rho]}{\delta\rho(\mathbf{r})} + V_{\mathrm{ext}}(\mathbf{r})$ and it reads:

$$\hat{H}^{\mathrm{KS}} = -\frac{1}{2}\Delta_{\mathbf{r}} + \int d\mathbf{r}'\frac{\rho(\mathbf{r})}{|\mathbf{r}-\mathbf{r}'|} + \frac{\delta\mathcal{E}^{\mathrm{XC}}[\rho]}{\delta\rho(\mathbf{r})} + V_{\mathrm{ext}}(\mathbf{r}) = -\frac{1}{2}\Delta_{\mathbf{r}} + V_{\mathrm{eff}}(\mathbf{r}).$$

The above equation can be rewritten equivalently as:

$$\hat{H}^{\mathrm{KS}} = H^{1b} + V_{\mathrm{Ha}}[\rho] + V_{\mathrm{XC}}[\rho], \tag{2.10}$$

where H^{1b} includes one-body contributions, namely the kinetic energy and the external ionic potential. The set of KS equations is the core of DFT and it is an *exact* mapping of the true many-body problem into a single-particle framework subjected to the effective potential $V_{\mathrm{eff}}(\mathbf{r})$. We notice that Eq. 2.9 is formulated as an eigenvalue problem as the Schrödinger equation of the physical electrons; it can be therefore solved using standard linear algebra methods for secular equations. The eigenvalues $\{\epsilon_i\}$ are the electronic energy levels of the system described by the Hamiltonian $\hat{H}^{\mathrm{KS}}$.

Since KS quasi-particles and real electrons are only linked by the electron ground state density, the KS spectrum has in principle no physical meaning. However, the true electronic properties of the system are often very well accounted for by the KS solution. No rigorous theoretical explanation of this success is available so far.

It is important to remark that Eqs. 2.9 must be solved self-consistently: at each step the diagonalization of $\hat{H}^{\mathrm{KS}}$ yields new orbitals $\phi_i(\mathbf{r})$ which, in turn, define a new density. This density is used to construct a new $\hat{H}^{\mathrm{KS}}$ and the procedure is repeated until the difference in total energy or density norm between two steps is below a certain threshold. Once the DFT procedure is converged, the full many-body wavefunction of the Kohn-Sham system can be obtain by constructing a single Slater determinant out of the KS orbitals, compactly written as:

$$\Psi_{\mathrm{SD}}(\mathbf{r}_1, \cdots, \mathbf{r}_N) = \det[\phi_i(\mathbf{r}_j)], \tag{2.11}$$

where $i \in [1, N_{\mathrm{orb}}]$ and $j \in [1, N]$.

The most important issue a practical application of KS formulation has to face is that an exact form of $\mathcal{E}^{\mathrm{XC}}[\rho(\mathbf{r})]$ functional is not known and thus an approximate expression must be devised. What $\mathcal{E}^{\mathrm{XC}}[\rho(\mathbf{r})]$ needs to include in order to capture all the many-body effects due to exchange and long-range Coulomb interactions is the "holy grail" [15] of electronic structure calculations.

In this work, we mostly employ the widely used local density approximation (LDA) of $\mathcal{E}^{\mathrm{XC}}[\rho(\mathbf{r})]$. LDA is based on a drastic approximation since it imposes that the dependence of $\mathcal{E}^{\mathrm{XC}}$ on the density has only a local character, which is wrong even at the simple exchange level. The LDA exchange correlation functional reads:

$$\mathcal{E}_{\mathrm{XC}}^{\mathrm{LDA}}[\rho] = \int d\mathbf{r}\rho(\mathbf{r})E_{\mathrm{loc}}(\rho(\mathbf{r})). \tag{2.12}$$

The local exchange-correlation functional E_{loc} is obtained by fitting a set of QMC calculations on the homogeneous electron gas model at different densities [16, 17]. This is actually one of the first and most successful applications of QMC in electronic structure calculations. LDA has given impressively good results on a wide range of systems and this justifies its popularity. However, there are many situations where more complicated expressions based on the gradient of the density $\rho(\mathbf{r})$ must be employed for achieving the best result.

2.2.2 Kohn-Sham problem on non-orthogonal basis set

The choice of the DFT basis set to expand KS eigenvectors should be very flexible. The most successful functional form is certainly the plane wave (PW) basis set [18, 19]. PWs do not depend on any input parameter and they are very cheap to evaluate, therefore the complete basis set (CBS) limit can be systematically reached. Furthermore, their accuracy can be arbitrarily tuned by cutting off the PW expansion beyond a certain cutoff energy. PWs are suitable for open systems and also for solids with small simulation cells, since they naturally satisfy the Bloch theorem (see Chap. 4 for more details). However, for

extended periodic systems they can become inefficient when using sizable simulation cells since, in this case, the number of PWs needed in the expansion grows proportionally to the system size.

In the case of correlated techniques such as QMC, the choice of the basis set is more delicate due to its higher computational cost. For instance, PWs expansion of single-particle orbitals is not convenient in QMC since its computational cost grows as $\mathcal{O}(N^3)$ with a prefactor which depends on the PW cutoff [20]. Despite recent improvements on the PWs side [21], the most common solution in QMC is the use of atomic localized orbitals. This form is often used in quantum chemistry since they yield an intuitive picture of the inter-atomic chemical bonds and they give a better description of the atomic limit. Modern QMC approaches typically use localized orbitals with Slater type (STOs), Gaussian type (GTOs) or more evolved schemes based on localized splines expansion [20, 22].

In our implementation, we use a localized basis set with GTOs. For seamlessly interfacing with the QMC software, we use the same basis set in the built-in DFT package. In order to cope with this choice, suitable algorithms must be devised to diagonalize the KS Hamiltonian within a non-orthogonal basis set, as presented in this Section.

Atom centered Gaussian basis set

We expand the KS single particle orbitals $\phi_i(\mathbf{r})$ (Eq. 2.9) onto a localized basis set $\chi_i(\mathbf{r})$:

$$\phi_i(\mathbf{r}) = \sum_{j}^{N_{\text{tot}}} c_{ij}\chi_j(\mathbf{r}) \quad \text{with } i \in [1, N_{\text{orb}}], \tag{2.13}$$

where $N_{\text{orb}} \leq N_{\text{tot}}$ is the number of KS orbitals which depends on the number of electronic energy levels to be computed, and N_{tot} is the total basis dimension. The generalized index j are defined as $j = (\mu, a)$ where $a \in [1, N_{\text{at}}]$ goes over the atomic sites and $\mu \in [1, N_{\text{bas}}]$ spans the atomic basis set dimension. We assume for simplicity that every atom has the same basis set. Hence the total basis dimension is given by: $N_{\text{tot}} = N_{\text{bas}} \times N_{\text{at}}$.

The functions $\{\chi_j(\mathbf{r})\}$ can be STOs or GTOs or mixed forms. We will focus on GTOs which have been thoroughly employed for the results presented in this thesis. They are centered on the atomic positions $\{\mathbf{R}\}$ and they have the following form:

$$\chi^a_\mu(\mathbf{r}) = \chi^a_{(l,m,n)} \propto |\mathbf{r} - \mathbf{R}_a|^l e^{-\zeta_{l,n}|\mathbf{r}-\mathbf{R}_a|^2} Y_{l,m}(\Omega_{\mathbf{r}-\mathbf{R}_a}), \tag{2.14}$$

where $m \in [-l, l]$ and $n \in [1, n_l]$. $\{l, m\}$ are the angular momentum quantum numbers, n_l identifies the number of Gaussians for each angular momentum shell and $Y_{l,m}(\Omega)$ are the cubic harmonics defined as real-valued linear combinations of spherical harmonics.

Thanks to this choice, the orbitals are eigenfunctions of the angular momentum operators $\hat{l}^2$, ensuring the rotational invariance of the resulting wavefunction around the nucleus. The basis set is thus fully determined by the number of Gaussian per angular momentum channel and their exponents ζ. These values can be taken from available databases [23, 24] or they can be built from scratch. In the latter case, the basis exponents and coefficients are initially optimized by DFT(LDA) energy minimization with a manual procedure before starting QMC calculations. It is therefore useful to express the basis set in an *even tempered* form, namely for each angular momentum l the exponents of the Gaussians are written as a power series: $\zeta_{l,n} = \alpha_l \beta_l^{n-1}$ for $n = 1, \ldots, n_l$. This allows a reduction of the number of parameters to optimize, as for each shell l the full $\zeta_{l,n}$ series is fixed just by three values: α_l, β_l, and n_l. This choice is adopted to define an optimized primitive basis set for water molecule and proton transfer systems presented in Chap. 3. In other cases (Chap. 5 and Chap. 6) the basis is taken from databases and often optimized directly at a QMC level.

The formalism presented in this Section is valid for open systems which require real-valued orbitals. The extension of this basis set for constructing complex-valued wavefunctions suitable for periodic solids is described in Sec. 4.2.

Generalized eigenvalue problem

The major difficulty to solve the KS equation within a localized basis is that the orbitals in Eq. 2.14 are non orthogonal. In this situation, we define the basis set overlap matrix of dimension $N_{\text{tot}} \times N_{\text{tot}}$ as:

$$S_{ij} = \int d\mathbf{r}\ \chi_i(\mathbf{r})\chi_j(\mathbf{r}) = \langle \chi_i |\ \chi_j \rangle\,, \tag{2.15}$$

where we used the fact that χ_i orbitals are real-valued. Being an overlap matrix, $S = \{S_{ij}\}$ is positive definite. Notice that in this equation the spin is integrated out; in the case of periodic systems, the spin dependence will be explicitly kept (see Sec. 4.2). Analogously, we can define the KS Hamiltonian (Eq. 2.10) matrix elements as:

$$H_{ij}^{\text{KS}} = \langle \chi_i |\, H^{1b} \,| \chi_j \rangle + \langle \chi_i |\, V_{\text{Ha}}[\rho] + V_{\text{XC}}[\rho] \,| \chi_j \rangle = S_{ij}^H + V_{ij}[\rho]. \tag{2.16}$$

Thanks to the properties of the Gaussian functions, the integrals involved in the definition of $S = \{S_{ij}\}$ and $H^{\text{KS}} = \{H_{ij}^{\text{KS}}\}$ matrices can be computed analytically or semi-analytically and their computation can be efficiently parallelized as, for instance, implemented in the Crystal code [25]. However, in our approach we opt for representing the electronic density $\rho(\mathbf{r})$ and the wavefunction on a real space grid. The integrals can then be directly evaluated by an appropriate summation over the mesh points, except for the

Hartree term V_{Ha} which is computed as a convolution in Fourier space [26]. While obviously less efficient than the analytical evaluation, the grid approach is more general since it can be readily applied to non Gaussian orbitals; furthermore the computational cost can be also greatly reduced by casting the integration procedure on few fast matrix-matrix multiplications; finally, since we use cubic meshes only, the parallelization on one grid direction is straightforward to implement. There is no need for more complicated parallelization schemes as each integration point is independent from the others due to the locality of the KS Hamiltonian.

We define a mesh with dimension $N_m = N_x \times N_y \times N_z$ and volume $L_m = L_x \times L_y \times L_z$ where x, y, z are the Cartesian directions. To avoid possible divergences in the integrals due to electron-ion coalescence at grid points, the origin of the mesh is chosen such that the minimum distance of grid points from ionic positions is maximized. The size N_m of the grid represent an important input parameter which provides a fine control on the accuracy of the DFT energy and which must be carefully assessed to find the best compromise between precision and computational cost. For large grids, a memory buffer is required to avoid storing the whole set of mesh values during the run. Notice that, since the basis set does not change during the self-consistent DFT cycle, the density independent matrices S and S_H can be computed once for all at the initialization stage.

If we define the vectors $C_i = \{c_{ij}\}$ containing the coefficients of the KS eigenvectors in the localized basis and we plug Eq. 2.13 into Eq. 2.9, the KS problem reduces to the following secular equation:

$$H^{\mathrm{KS}} C_i = \epsilon_i S C_i \quad \text{with } i \in [1, N_{\mathrm{orb}}] \tag{2.17}$$

Eq. 2.17 is a generalized eigenvalue problem. Differently from orthogonal basis sets where $S = \mathbb{I}$, the stability of the numerical solution of Eq. 2.17 is strongly dependent on the overlap matrix condition number defined as the ratio between its largest and its lowest eigenvalues (ordered in ascending order):

$$s_{\mathrm{cond}} = \frac{s_{N_{\mathrm{tot}}}}{s_1}.$$

If this ratio is too large the KS problem is ill-conditioned, thus a very small change in the initial conditions leads to large and unpredictable errors on the diagonalization result and to inaccurate or garbage eigenvectors. This issue is even more significant when large basis sets are used and therefore it often prevents from reaching the CBS limit. This represents a strong limitation on the precision of the DFT wavefunction and, by consequence, on the quality of the corresponding QMC calculation. In fact, in many physical situations the basis set superposition error arising from a small basis set size dramatically deteriorates the description of the chemical bonds [26, 27].

Diagonalization algorithm

The accuracy and stability of the numerical solution of Eq. 2.17 can be considerably enhanced by applying the diagonalization algorithm introduced in Ref. [26]. This is composed by the following steps:

1. the overlap matrix S (generally ill-conditioned, in particular with a large basis set) is initially diagonalized with standard linear algebra methods [28] without special requirements. If the ith eigenvalue in the overlap spectrum is small, i.e. the ratio $s_i/s_{N_{\rm tot}}$ is of the order of the machine precision (typically $\sim 10^{-16}$), it is disregarded. In this way we are neglecting singular directions in the space generated by the overlap matrix which would lead to garbage eigenvalues in the Hamiltonian diagonalization. The non-singular directions are defined by the transformation:

$$\mathbf{u}_1^i = \frac{1}{\sqrt{s_i}}\mathbf{v}_i \tag{2.18}$$

   where $\mathbf{v}_i$ are the eigenvectors of S and $s_i/s_{N_{\rm tot}} \geq 10^{-16}$.

2. The initial S is rewritten in the basis U_1 defined by the transformation in Eq. 2.18. The new overlap matrix $\tilde{S} = U_1^T S U_1$ is well-conditioned and easy to diagonalize.

3. The diagonalization of $\tilde{S}$ yields a new transformation which we denote as U_2 and which is constructed in the same way as Eq. 2.18:

$$\mathbf{u}_2^i = \frac{1}{\sqrt{\tilde{s}_i}}\tilde{\mathbf{v}}_i \tag{2.19}$$

4. The global transformation $U = U_1 U_2$ generated by the two consecutive diagonalizations of S and $\tilde{S}$ is then applied to the KS Hamiltonian as follows:

$$\tilde{H}^{\rm KS} = U^T H^{\rm KS} U. \tag{2.20}$$

   The final matrix $\tilde{H}^{\rm KS}$ possesses the same spectrum of the initial Hamiltonian $\hat{H}^{\rm KS}$, since the transformation U is unitary, but it is now well-conditioned since singular directions have been eliminated. Eigenvalues and eigenvectors of $\tilde{H}^{\rm KS}$ are the solution to the generalized KS problem in Eq. 2.17.

The above diagonalization algorithm is applied at each step of the DFT self-consistent cycle and it allows to safely and efficiently work with large basis sets, as extensively demonstrated in Ref. [26]. The KS eigenvectors obtained from the converged calculation are used to build the determinant in Eq. 2.11 which constitutes the starting guess for the determinantal part of the QMC ansatz. This is the subject of the next Section.

2.3 The JAGP variational wavefunction

Since the discovery of high-temperature superconductivity in copper oxides materials [29], it became apparent the deficiency of single-particle techniques in the description of strongly correlated materials. This evidence gave a new impetus in the research on explicitly correlated methods, at first for solving lattice models connected with superconductivity, such as the multi-band Hubbard model, and afterwards also for tackling realistic systems from first principles.

Within wavefunction based methods, the most important mathematical object is the Slater determinant (SD), introduced in this context for the first time to construct the well-known Hartree-Fock (HF) approximation and used as many-body wavefunction for any independent or single-particle electron theory. We introduce this functional form in Eq. 2.11 in the context of DFT calculations.

Using the SD as ansatz for solving the many-body Schrödinger equation is convenient since it naturally includes effects arising from the antisymmetry of the many-body wavefunction. However, electronic repulsion due to long-range Coulomb electron-electron interaction is taken at the HF level and the correlation energy is zero. There are many ways for improving the SD in order to cure this problem. One widely used solution is to build the wavefunction as a linear combination of Slater determinants such in the aforementioned CC and CI methods. This approach relies on the fact that the full set of determinants constitutes a complete basis set of the whole Hilbert space.

Here we describe an alternative route. Our strategy is based on the so-called resonating valence bond (RVB) approach. RVB has been introduced for the first time in aromatic molecules by Pauling [30] and then generalized to describe metallic systems [31]. The main concept behind RVB is that *singlet* electron pairs forming a chemical bond between two distinct atomic sites can be subjected to an unsynchronized resonance. The superposition of these resonating bonds propagates through the molecule or the crystal and it ensures the stabilization of the electronic state if compared with the simple, non resonating HF approximation, as shown by numerical estimations of the resonance or condensation energy. Anderson [32] provided a rigorous mathematical framework for RVB theory suggesting that it could explain the physics of strongly correlated Mott insulators. Furthermore, the RVB state can support charged bosonic excitations with zero spin. Since such excitations could exist also at high-temperature, Anderson speculated that they could be the cause of the high-temperature superconductivity displayed by copper-oxides compounds [33] whose parent compound is a Mott insulator. He hypothesized that the coupling between copper and oxygen atoms could be a realization of the RVB state, however an experimental evidence of this speculation has not been found yet.

Within the second quantization formalism, the compact representation of the RVB state devised by Anderson reads:

$$|\Psi_{\mathrm{RVB}}\rangle = \hat{P}_N |\Psi_{\mathrm{BCS}}\rangle = \hat{P}_N \prod_{\mathbf{k}} (u_{\mathbf{k}} + v_{\mathbf{k}} c^{\dagger}_{\mathbf{k},\uparrow} c^{\dagger}_{-\mathbf{k},\downarrow}) |0\rangle . \tag{2.21}$$

The function $|\Psi_{\mathrm{BCS}}\rangle$ is the variational ansatz proposed by Bardeen, Cooper and Schrieffer for the theory of conventional superconductivity [34]. The operator $\hat{P}_N$ is a Gutzwiller projector defined as:

$$\hat{P}_N = \prod_{\mathbf{k}} (1 - n_{\mathbf{k}\uparrow} n_{\mathbf{k}\downarrow}) , \tag{2.22}$$

where $n_{\mathbf{k}\sigma} = c^{\dagger}_{\mathbf{k},\sigma} c_{\mathbf{k},\sigma}$ is the electron number operator in Fourier space. The role of this operator is to keep a fixed number of particles and forbid double occupancies of sites, thus freezing the charge fluctuations in the system allowed by the BCS ansatz.

By performing the Fourier transform of Eq. 2.21 we obtain its real space representation in a Hilbert space with fixed number N of electrons:

$$|\Psi_{\mathrm{RVB}}\rangle = \hat{P}_N |\mathrm{BCS}\rangle = \hat{P}_N \left[\sum_{\mathbf{r},\mathbf{r}'} \Phi(\mathbf{r} - \mathbf{r}') c^{\dagger}_{\mathbf{r},\uparrow} c^{\dagger}_{\mathbf{r}',\downarrow} \right]^{N/2} |0\rangle . \tag{2.23}$$

The electron pairing function $\Phi(\mathbf{r} - \mathbf{r}')$, alternatively known as *geminal*, is the Fourier transform of $u_{\mathbf{k}}/v_{\mathbf{k}}$ and it correlates two electrons belonging to the spin singlet representing the valence bond linking two generic lattice sites. We notice that the ansatz proposed by Anderson is composed by two factors: a product of pairing functions and a Gutzwiller projector operator.

Let us now consider the typical form of the QMC ansatz. In the following, we describe the N electrons of the system with the generalized coordinates $\{\mathbf{x}_i = (\mathbf{r}_i, \sigma_i)\}$, grouping positions and spins. We restrict our discussion to a spin unpolarized ($N_{\downarrow} = N_{\uparrow} = N/2$) state. The extension of the formalism presented in this Section to the spin polarized case is not treated here, but the interested reader can refer to Refs.[35, 36, 37]. The standard functional form of a QMC trial wavefunction is defined as:

$$\Psi(\mathbf{x}_1, \cdots, \mathbf{x}_N) = \exp(-J(\mathbf{r}_1, \cdots, \mathbf{r}_N)) \times \Psi_D(\mathbf{x}_1, \cdots, \mathbf{x}_N). \tag{2.24}$$

where $J(\mathbf{r}_1, \cdots, \mathbf{r}_N)$ is a bosonic factor, the so-called Jastrow factor, whereas the function $\Psi_D(\mathbf{x}_1, \cdots, \mathbf{x}_N)$ is an antisymmetric fermionic factor which ensures the right wavefunction symmetry and it is generally referred as determinantal part. The Jastrow antisymmetrized geminal power (JAGP) wavefunction employed in this thesis is an elegant and computationally efficient implementation of Eq. 2.23 in the context of realistic *ab initio* QMC calculations based on the ansatz of Eq. 2.24. This analogy will be explained in detail in the remaining part of this Section.

2.3.1 Multi-configurational AGP ansatz

We first describe the determinantal part $\Psi_D(\mathbf{x}_1, \cdots, \mathbf{x}_N)$ of Eq. 2.24. The antisymmetric product of equal geminals, or antisymmetric geminal power (AGP) wavefunction, is formally written as:

$$\Psi_{\mathrm{AGP}}(\mathbf{x}_1, \cdots, \mathbf{x}_N) = \hat{A}\Big[\, \Phi(\mathbf{x}_1, \mathbf{x}_2)\; \Phi(\mathbf{x}_3, \mathbf{x}_4)\; \ldots\; \Phi(\mathbf{x}_{N-1}, \mathbf{x}_N)\Big] \tag{2.25}$$

where $\hat{A}$ is an operator that antisymmetrizes the pairing product. The pairing function $\Phi(\mathbf{x}_i, \mathbf{x}_j)$ of two generic electrons i, j can be factorized into a symmetric radial part correlating two electrons and an antisymmetric spin singlet factor:

$$\Phi(\mathbf{x}_i, \mathbf{x}_j) = \Phi(\mathbf{r}_i^{\uparrow}, \mathbf{r}_j^{\downarrow}) = \Phi(\mathbf{r}_i, \mathbf{r}_j)\frac{\delta(\sigma_i, \uparrow)\delta(\sigma_j, \downarrow) - \delta(\sigma_i, \downarrow)\delta(\sigma_j, \uparrow)}{\sqrt{2}}. \tag{2.26}$$

Within this representation, the AGP has a very compact and computationally convenient expression [38] which generalizes the concept of SD to a pair of electrons:

$$\Psi_{\mathrm{AGP}}(\mathbf{x}_1, \cdots, \mathbf{x}_N) = \det\Big[\Phi(\mathbf{x}_i, \mathbf{x}_j)\Big] = \det\left[A_{i,j}\right]. \tag{2.27}$$

Notice that here we set $\hat{A}$ equal to the determinant for applying the antisymmetrization, but for instance a Pfaffian operator could have been chosen instead [39]. It has been demonstrated [38] that the AGP wavefunction is equivalent to the particle-conserving version of the BCS ansatz introduced in Eq. 2.21. This is therefore the first building block of the aforementioned analogy between JAGP and RVB wavefunctions.

Within a finite basis set calculation, the AGP can be expressed in two equivalent forms. We call the first one valence bond (VB) representation since it focuses on the RVB picture underlying the physics of this ansatz. In this context, the radial pairing function is expanded over the atom centered basis set $\{\chi_\mu^a(\mathbf{r})\}$ introduced in Eq. 2.14:

$$\begin{aligned}\Phi(\mathbf{r}, \mathbf{r}') &= \sum_{i,j}^{N_{\mathrm{tot}}} \lambda_{ij}\; \chi_i(\mathbf{r})\; \chi_j(\mathbf{r}') = \\ &= \sum_{a,b}^{N_{\mathrm{at}}} \sum_{\mu,\nu}^{N_{\mathrm{bas}}} \lambda_{\mu,\nu}^{a,b}\; \chi_\mu^a(|\mathbf{r} - \mathbf{R}_a|)\; \chi_\nu^b(|\mathbf{r}' - \mathbf{R}_b|).\end{aligned} \tag{2.28}$$

The total number of variational parameters of the ansatz within the VB representation is thus given by: $P^{\mathrm{VB}} \propto N_{\mathrm{tot}} \times N_{\mathrm{tot}}$.

The matrix $\Lambda = \{\lambda_{\mu,\nu}^{a,b}\}$ defines the weights of each orbital in the expansion and it has a direct link to the RVB picture. In fact, the Λ matrix elements connect different atoms within the system, giving the strength of the valence bond between two sites restricted to a specific element of the basis set. Besides giving an accurate picture of the chemical bonds,

the Λ matrix allows to simply fulfill other symmetries of the system, such as inversion or reflection, by setting appropriate equalities among elements connecting different orbitals. In this work, we usually employ a symmetric Λ matrix, while the formalism presented here holds, with slight modifications, also for generic non-symmetric arising from the inclusion of unpaired electrons Λ [36, 37]. As explained in Sec. 4.3.3, the mathematical framework used for non-symmetric geminal matrices is exploited also for generalizing the AGP wavefunction to the study of periodic solids.

Another representation of the pairing function can be obtained by expanding $\Phi(\mathbf{r}_i^{\uparrow}, \mathbf{r}_j^{\downarrow})$ over molecular orbitals (MOs). Within the MO representation, Eq. 2.28 reads:

$$\Phi(\mathbf{r}, \mathbf{r}') = \sum_{j}^{N_{\mathrm{MO}}} \lambda_j^{\mathrm{MO}} \phi_j^{\mathrm{MO}}(\mathbf{r}) \phi_j^{\mathrm{MO}}(\mathbf{r}') \tag{2.29}$$

where the functions $\phi_j^{\mathrm{MO}}(\mathbf{r})$ are developed over the same localized basis set used for the VB representation:

$$\phi_j^{\mathrm{MO}}(\mathbf{r}) = \sum_{i}^{N_{\mathrm{tot}}} c_{ji} \chi_i(\mathbf{r}). \tag{2.30}$$

The total number of variational parameters of the ansatz in the MO representation is given by: $P^{\mathrm{MO}} \propto N_{\mathrm{tot}} \times N_{\mathrm{MO}}$.

One can easily pass from VB to MOs representations simply by diagonalizing the geminal matrix Λ. Vice-versa, one can retrieve the full Λ matrix by expanding the molecular orbitals in their local basis set components through the following transformation:

$$\lambda_{ij} = \left[C^{\dagger} \Lambda^{\mathrm{MO}} C \right]_{ij}, \tag{2.31}$$

where $C = \{c_{ji}\}$ is a matrix constructed with the MOs coefficients. The VB and MO representations are equivalent if one choose as many molecular orbitals as the number of basis functions, i.e. $N_{\mathrm{MO}} = N_{\mathrm{tot}}$, while in the case of $N_{\mathrm{MO}} < N_{\mathrm{tot}}$, Eq. 2.29 is just an approximation of the complete pairing function in Eq. 2.28.

However, retaining the largest eigenvalues λ_j^{MO} of the geminal matrix is typically sufficient to obtain very accurate pairing functions and at the same time reduce the number of MOs. This choice is very convenient since the number of variational parameters of the ansatz can be greatly diminished: $N_{\mathrm{MO}} \ll N_{\mathrm{tot}} \implies P^{\mathrm{MO}} \ll P^{\mathrm{VB}}$. It can be proven [36] that by keeping only the first $N_{\mathrm{MO}} = N/2$ eigenvalues, i.e. the lowest allowed rank of the Λ matrix, the determinant of Eq. 2.27 is equivalent to the SD wavefunction. The uncorrelated HF ansatz is therefore contained within the AGP wavefunction as a special case.

Hence, the formalism presented here allows to have a fine control on the level of electron correlation, by smoothly transitioning from the independent-particle wavefunction to the

fully multi-configurational case. In fact, by adding to the geminal expansion in Eq. 2.29 additional MOs, the lower excited states come into play and contribute to decrease the total variational energy. The system is thus "resonating" among different configurations: the Fermi liquid ground state taken from the DFT wavefunction in Eq. 2.11 and an appropriate number of excited states, represented by the additional MOs introduced with a small weight. This procedure is fundamental in order to obtain a reliable description of systems displaying nearly-degenerate energy levels around the HOMO/LUMO gap (or the Fermi level in case of solids), all contributing to the true ground state; in this case the AGP wavefunction is equivalent to a linear combinations of SDs, each containing one degenerate or near degenerate orbital [35]. The AGP ansatz is thus particularly suitable for dealing with static electron correlation. Last but not least, the AGP wavefunction is also very convenient from a computational point of view. In fact the cost of AGP remains equal to the evaluation of a single determinant ($\sim N^3$ scaling with the number electrons) as in Eq. 2.27 even in the multi-reference case, namely when $N_{\mathrm{MO}} > N/2$ molecular orbitals are included in the simulation.

2.3.2 The missing ingredient: Jastrow factor

All determinantal wavefunctions such as the HF Slater determinant suffers from the absence of a correlation hole for electron pairs. This can lead to large energy fluctuations due to the Coulomb repulsion. In order to damp these fluctuations, it is common practice to multiply the determinantal part by a bosonic Jastrow factor. The Jastrow plays the role of the Gutzwiller projector in Eq. 2.23 and it disfavors double occupations of the orbitals. The analogy between JAGP and the RVB representation proposed by Anderson is thus complete.

Complementary to AGP, the Jastrow term is constructed to include dynamical electron correlation, i.e. the fraction of the correlation energy derived from charge (and thus spin) spatial motion in the system. This ensures an accurate treatment of weak intermolecular forces as Van der Waals interactions [40]. Furthermore, the Jastrow is tailored to solve the problem of wavefunction cusp conditions [41] affecting all approximate wavefunctions. Cusp conditions are constraints imposed on the derivatives of the many-body wavefunction at electron-ion and electron-electron coalescence points (namely when inter-particle distance $r \to 0$). They neutralize the divergence of the electron-ion and Coulomb potentials at the coalescence points by forcing an equivalent and opposite divergence in the kinetic energy, as it is automatically realized with the exact wavefunction. When the correct cusps are enforced into the variational ansatz, the rate of convergence of JAGP in the basis is much increased. Therefore the inclusion of a Jastrow factor, although re-

sulting in a larger number of variational parameters, brings great benefits in term of the final cost of wavefunction evaluation. Furthermore, when selecting a Slater determinant as antisymmetric part of the ansatz, the Jastrow represents the only way to go beyond HF and include electron correlation in the system.

Since it represents an essential part of the ansatz, we choose to define the Jastrow factor in a compact but complete way which takes into account the most relevant electron interactions. The general functional form of our Jastrow reads:

$$J(\mathbf{r}_1, \cdots, \mathbf{r}_N) = \sum_{a}^{N_{\text{at}}} \sum_{i}^{N} g_a^{\text{1body}}(\mathbf{r}_i - \mathbf{R}_a) + \sum_{i \neq j}^{N} g^{\text{2body}}(\mathbf{r}_i, \mathbf{r}_j). \tag{2.32}$$

The first term $g^{\text{1body}}(\mathbf{r} - \mathbf{R})$ is usually referred as one-body term and it is divided into two separate contributions:

$$g_a^{1body}(\mathbf{r} - \mathbf{R}_a) = v_a(r_a) + \sum_{\mu}^{N_{\text{basj}}} G_a^{\mu} \chi_{\mu}^{a}(\mathbf{r} - \mathbf{R}_a), \tag{2.33}$$

where $r_a = |\mathbf{r} - \mathbf{R}_a|$ are the electron-nucleus distances. $v(r)$ is a homogeneous function designed to impose the right nuclear cusps for curing electron-ion potential divergences:

$$v(r) = Z \frac{1 - e^{-\beta \sqrt[4]{(2Z)} r}}{\beta \sqrt[4]{(2Z)}}. \tag{2.34}$$

In Eq. 2.34 β is a parameter which is variationally tuned during the energy minimization procedure (see Sec. 2.5).

The second term in the right hand side of Eq. 2.33 recovers most part of the energy coming from the electron-ion interaction. As for the pairing function introduced in the previous section, this factor is expanded over a localized basis set $\chi_{\mu}^{a}(\mathbf{r} - \mathbf{R}_a)$ of dimension N_{basj} possessing the functional form introduced in Eq. 2.14. Besides ensuring a great level of flexibility during the wavefunction optimization, this choice does not affect the cusp conditions introduced by Eq. 2.34 since the localized orbitals smoothly reach a constant value faster than linearly when approaching the nucleus.

The many-body $g^{2body}(\mathbf{r}, \mathbf{r}')$ factor is composed by a homogeneous and non-homogeneous parts:

$$g^{2body}(\mathbf{r}, \mathbf{r}') = u(r) + \tag{2.35}$$

$$+ \sum_{\mu,\nu}^{N_{\text{basj}}} \sum_{a,b}^{N_{\text{at}}} C_{\mu,\nu}^{a,b} \, \chi_{\mu}^{a}(\mathbf{r} - \mathbf{R}_a) \, \chi_{\nu}^{b}(\mathbf{r}' - \mathbf{R}_b) + \tag{2.36}$$

$$+ \sum_{\mu,\nu}^{N_{\text{basj}}} \sum_{a,b}^{N_{\text{at}}} S_{\mu,\nu}^{a,b} \, \chi_{\mu}^{a}(\mathbf{r}^{\uparrow} - \mathbf{R}_a) \, \chi_{\nu}^{b}(\mathbf{r}'^{\downarrow} - \mathbf{R}_b) \tag{2.37}$$

where $r = |\mathbf{r} - \mathbf{r}'|$ is the electron-electron distance. The homogeneous term in Eq. 2.35 cures the Coulomb potential divergences at $r \to 0$ by imposing the correct electron-electron cusp conditions. We choose the following simple form for the $u(r)$ function:

$$u(r) = \frac{1}{A\,\gamma}(1 - e^{-\gamma r}). \tag{2.38}$$

In this equation γ is again an optimizable variational parameter. The scaling factor A is analytically determined by the cusps. It is always equal to 1/2 when only the unlike-spin contribution is considered. If this is not the case, A takes the values of 1/2 for unlike spin and 1/4 for like spin electron pairs; however this could lead to an undesired mixing of different electronic spin states, an effect usually known as *spin contamination*.

The remaining terms constitute an essential part of the many-body wavefunction. They are non-homogeneous as they are explicitly built upon space-space correlators between different atomic sites. The first term (Eq. 2.36) accurately includes long-range dipole-dipole interactions within the QMC ansatz. Thanks to its development over the same atomic basis set of size N_{basj} as the one-body term, it correlates single-particle orbitals describing electrons on the same or on different atomic sites. The spin is integrated out from Eq. 2.36 and therefore this term deals only with the charge sector. When spin fluctuations are relevant for the physical situation under investigation, the inclusion of a spin dependent part to the many-body Jastrow term may considerably improve the variational energy of the system. This is the case, for example, of the iron-based superconductor FeSe presented in Chap. 6. This term is defined in Eq. 2.37 and it acts as correlator for electrons with opposite spin. In principle, it allows to recover the amount of correlation energy coming from short to long ranged spin-spin interactions. The implementation of this term is completely analogous to the charge sector one.

2.3.3 Functional forms of the ansatz

In conclusion, we report here the two many-body wavefunctions we employ throughout this thesis. They are the Jastrow-antisymmetrized geminal power (JAGP) ansatz (used for the results in Chap. 3 for open systems) which reads:

$$\Psi_{\text{JAGP}}(\mathbf{x}_1, \cdots, \mathbf{x}_N) = \exp\left[-J(\mathbf{r}_1, \cdots, \mathbf{r}_N)\right] \det\left[\Phi(\mathbf{x}_i, \mathbf{x}_j)\right] \tag{2.39}$$

$$\Psi_{\text{JAGP}} = \exp\left[-J(\mathbf{R}_{\text{el}})\right] \times \hat{A}\Big[\ \Phi(\mathbf{r}_1, \mathbf{r}_2)\ \Phi(\mathbf{r}_3, \mathbf{r}_4)\ \ldots\ \Phi(\mathbf{r}_{N-1}, \mathbf{r}_N)\Big] \tag{2.40}$$

where $(i, j) \in [1, N]$ and the pairing function $\Phi(\mathbf{x}, \mathbf{x}')$ has been defined in Eq. 2.26. The second form is the Jastrow-Slater determinant (JSD) wavefunction (used for the results in

Chaps. 5 and 6 for extended systems), obtained from JAGP by truncating the molecular orbitals expansion to $N_{\rm MO} = N/2$. JSD wavefunction can be compactly expressed as:

$$\Psi_{\rm JSD}(\mathbf{x}_1, \cdots, \mathbf{x}_N) = \exp\left[-J(\mathbf{r}_1, \cdots, \mathbf{r}_N)\right] \det[\phi_i(\mathbf{r}_j^{\uparrow})] \det[\phi_i(\mathbf{r}_j^{\downarrow})] \tag{2.41}$$

with $i \in [1, N_{\rm MO}]$ and $j \in [1, N]$.

2.4 Evaluating physical observables within QMC

2.4.1 Variational Monte Carlo

Basic idea

VMC is the simplest application of Monte Carlo integration to the quantum problem and its first application to fermionic systems dates back to the work of Ceperley *et al.* in 1977 [42]. We consider a real-valued antisymmetric trial wavefunction $\Psi_T(\mathbf{r}_1, \cdots, \mathbf{r}_N) = \langle \mathbf{r}_1, \cdots, \mathbf{r}_N | \, \Psi_T \rangle$ with a generic functional form. The quantum expectation of the many-body Hamiltonian $\hat{H}$ (Eq. 2.2) over the trial Ψ_T reads:

$$\langle \hat{H} \rangle = E_{\rm VMC} = \frac{\int d\mathbf{R}_{\rm el} \Psi_T(\mathbf{R}_{\rm el}) \hat{H} \Psi_T(\mathbf{R}_{\rm el})}{\int d\mathbf{R}_{\rm el} \Psi_T^2(\mathbf{R}_{\rm el})}, \tag{2.42}$$

where $\mathbf{R}_{\rm el} = (\mathbf{r}_1, \cdots, \mathbf{r}_N)$. For the sake of clearness, spin is omitted from the following discussion. The variational principle establishes that Eq. 2.42 is an upper bound to the true ground state energy, i.e. $E_{\rm VMC} \geq E_0$.

Elementary algebra allows to recast Eq. 2.42 in the following form:

$$\begin{aligned} E_{\rm VMC} &= \int d\mathbf{R}_{\rm el} \underbrace{\frac{\Psi_T^2(\mathbf{R}_{\rm el})}{\int d\mathbf{R}'_{\rm el} \Psi_T^2(\mathbf{R}'_{\rm el})}}_{\text{distribution: } \pi(\mathbf{R}_{\rm el})} \overbrace{\frac{\hat{H}\Psi_T(\mathbf{R}_{\rm el})}{\Psi_T \mathbf{R}_{\rm el}}}^{\text{local energy: } e_L(\mathbf{R}_{\rm el})} \\ &= \int d\mathbf{R}_{\rm el} \, \pi(\mathbf{R}_{\rm el}) e_L(\mathbf{R}_{\rm el}) = \langle e_L \rangle \geq E_0. \end{aligned} \tag{2.43}$$

Eq. 2.43 is very convenient since it is an *importance sampling* transformation of Eq. 2.42. Importance sampling is widely employed in Monte Carlo based techniques since it allows to dramatically decrease the number of stochastic steps needed to attain a certain accuracy by drawing configurations from an appropriately chosen probability distribution, the so-called importance sampling distribution.

In Eq. 2.43 the importance sampling distribution is represented by the function $\pi(\mathbf{R}_{\rm el})$, i.e. the normalized probability amplitude given by Ψ_T^2. The *local energy* $e_L(\mathbf{R}_{\rm el}) =$

$\frac{\hat{H}\Psi_T(\mathbf{R}_{\rm el})}{\Psi_T\mathbf{R}_{\rm el}}$ arising from the importance sampling is simple to evaluate and its computational cost only depends on the cost of evaluating the ansatz Ψ_T. Notice that if the trial wavefunction is an eigenstate of $\hat{H}$, the local energy is equal to the exact ground state energy independently of the chosen configuration. This demonstrates that the VMC method satisfies the zero-variance property mentioned in the introduction to this Chapter.

The probability distribution $\pi(\mathbf{R}_{\rm el})$ is sampled by means of the standard Metropolis-Hastings algorithm introduced in Ref. [43] in 1953. This procedure allows to sample any probability distribution by generating a Markov stochastic process – or Markov chain – in the $3N$-dimensional space of electronic positions. Thanks to the properties of Markov chains, a new random point in the configurational space only depends on the previous configuration adopted by the system, i.e. the stochastic process has *no memory*.

At each step of the Markov process, a new configuration is generated by performing a single electron move. This move is accepted or rejected according to the Metropolis rule [43]; then the VMC energy is evaluated using Eq. 2.43 accumulating statistics until the desired accuracy is attained. A single electron move is preferable with respect to global displacements since the operations needed to update the wavefunction can be cast into fast matrix-matrix multiplications and reduced by means of the Sherman-Morrison formula[a]. The efficiency of the sampling can be further increased if the amplitude of the single electron move is adjusted depending on the distance with the nearest nucleus: if the electron is far from the ion, it moves with a step larger than the one of an electron closer to a nucleus. Furthermore, when the system is composed by two distant fragments, such as in the computation of dissociation limits (see for instance Sec. 3.3.2), large moves from one fragment to the other are allowed after a certain amount of MC samples.

Estimators in VMC

Given a finite number M of Monte Carlo generations, i.e. drawn configurations along the Markov chain, the expectation value of any operator $\hat{O}$ representing a physical observable, in analogy with Eq. 2.42, can be evaluated as:

$$O_{\rm VMC} = \langle \hat{O}_L \rangle = \frac{1}{M} \sum_{i=1}^{M} \hat{O}_L(\mathbf{R}_{\rm el}^i). \tag{2.44}$$

[a]Given the invertible square matrix A representing the determinant and the two column vectors u and v constituting a rank-1 update of A (hence a single-electron move), the Sherman-Morrison formula reads:

$$(A + uv^T)^{-1} = A^{-1} + \frac{A^{-1}uv^TA^{-1}}{1 + A^{-1}uv^T}.$$

Notice that the inverse A^{-1} has already been computed in the previous MC generation.

where $\hat{O}_L = \hat{O}\Psi_T/\Psi_T$ is the local variant of the operator. Eq. 2.44 is the sampling average of $\hat{O}_L$ over the stochastic process and it is therefore an unbiased estimator of the true expectation value of $\hat{O}$ thanks to the central limit theorem. An unbiased estimator of the variance $\sigma^2[O_{\mathrm{VMC}}]$ is given by:

$$\sigma^2[O_{\mathrm{VMC}}] = \frac{\sigma^2[\hat{O}_L]\tau_c}{M} = \frac{\tau_c}{M}\left[\langle \hat{O}_L^2 \rangle - (\langle \hat{O}_L \rangle)^2\right], \tag{2.45}$$

where $\sigma^2[\hat{O}_L]$ is the variance on single measurement and τ_c is the autocorrelation time between two generic MC steps defined as:

$$\tau_c = 1 + \frac{2}{\sigma[\hat{O}_L]M}\sum_{i<j}\mathrm{Cov}\left[\hat{O}_L(\mathbf{R}_{\mathrm{el}}^i), \hat{O}_L(\mathbf{R}_{\mathrm{el}}^j)\right], \tag{2.46}$$

which is equal to 1 in the ideal case of uncorrelated measurements, namely when it holds: $\mathrm{Cov}\left[\hat{O}_L(\mathbf{R}_{\mathrm{el}}^i), \hat{O}_L(\mathbf{R}_{\mathrm{el}}^j)\right] = 0 \ \ \forall\ (i,j)$. As already mentioned, it is evident that the statistical error can be adjusted below the desired accuracy by increasing the length M of the simulation.

In practical calculations, the number of samples M over which the expectation values are computed is restricted to a fixed block size $M_b < M$ in order to reduce autocorrelation time. The estimation of the final value is obtained by averaging over the blocks (*blocking technique*). The bootstrap resampling method is used on the block averaging to ensure a further reduction of the serial correlation among MC samples.

2.4.2 Diffusion Monte Carlo

Basic idea

VMC method is elegant and simple to implement, however its outcome strongly depends on the quality of the variational wavefunction, namely on its functional form, on the basis set size and on the level of optimization of the whole set of variational parameters. In order to systematically improve the VMC energy and reduce at most the dependence of the final result on the trial wavefunction Ψ_T, one can rely on the projector Monte Carlo class of algorithms. Here we describe a particular implementation of projector MC which it is known as diffusion Monte Carlo technique [3].

Let us call Ψ_0 the ground state eigenstate of the QMC Hamiltonian $\hat{H}$. DMC starts from realizing that the exact ground state energy E_0 can be obtained from a *mixed* expectation value defined as follows:

$$E_0 = \frac{\langle \Psi_0 | \hat{H} | \Psi_T \rangle}{\langle \Psi_0 | \Psi_T \rangle}. \tag{2.47}$$

However, this expression requires to access the exact ground state Ψ_0. One way to achieve this is by evolving the trial wavefunction along the imaginary time ($t \to -it$) axis. This is realized by the repeated application of the time evolution operator to Ψ_T:

$$\Psi_G(\mathbf{R}_{\rm el}, t) = e^{-[\hat{H}-\Lambda(t)]t}\Psi_T(\mathbf{R}_{\rm el}). \tag{2.48}$$

$\Psi_G(\mathbf{R}_{\rm el}, t)$ is the formal solution of the imaginary time Schrödinger equation associated with the Hamiltonian $\hat{H}$. The time dependent parameter $\Lambda(t)$ is an energy offset whose role will be elucidated later on and the ansatz Ψ_T is usually referred to as *guiding wavefunction*. By expanding Eq. 2.48 over the Hamiltonian eigenstates, it is straightforward to show that the infinite time limit of Eq. 2.48 is dominated by the Hamiltonian lowest-energy eigenstate Ψ_0 with finite overlap with the initial Ψ_T.

Using Eq. 2.48, we can define the following mixed expectation:

$$\begin{aligned} E_{\rm DMC} &= \frac{\langle\Psi_G|\,\hat{H}\,|\Psi_T\rangle}{\langle\Psi_G|\,\Psi_T\rangle} = \int d\mathbf{R}_{\rm el}\, \underbrace{\frac{\Psi_G(\mathbf{R}_{\rm el}, t)\Psi_T(\mathbf{R}_{\rm el})}{\int d\mathbf{R}'_{\rm el}\Psi_G(\mathbf{R}'_{\rm el})\Psi_T(\mathbf{R}'_{\rm el})}}_{\text{distribution: } \tilde{\pi}(\mathbf{R}_{\rm el},t)} \overbrace{\frac{\hat{H}\Psi_T(\mathbf{R}_{\rm el})}{\Psi_T(\mathbf{R}_{\rm el})}}^{\text{local energy: } e_L(\mathbf{R}_{\rm el})} = \\ &= \int d\mathbf{R}_{\rm el}\, \tilde{\pi}(\mathbf{R}_{\rm el}, t) e_L(\mathbf{R}_{\rm el}) = \langle e_L \rangle \end{aligned} \tag{2.49}$$

which coincides with Eq. 2.47 in the infinite time limit. Notice that Eq. 2.49 is an importance sampling transformation for the Hamiltonian expectation value, obtained by operating the substitution: $\Psi_T^2 \to \Psi_G\Psi_T$. Eq. 2.49 shares the same properties of the VMC analog (Eq. 2.43) such as the zero-variance character, and it can be evaluated in exactly the same way (Eqs. 2.44, 2.45). The sole difference among the two methods is represented by the distribution $\tilde{\pi}(\mathbf{R}_{\rm el}, t)$ to be sampled. However, we show in the following that this aspect requires a distinct stochastic algorithm to evaluate Eq. 2.49.

The fixed node approximation

In order to interpret $\tilde{\pi}(\mathbf{R}_{\rm el}, t)$ as a probability distribution, it must be positive everywhere. However, this is not verified by a fermionic wavefunction as Ψ_T which must possess positive and negative regions in order to fulfill the antisymmetric property dictated by the Pauli principle. Phase space regions where the wavefunction has the same sign are generally called *nodal pockets* and their exact position is an unknown property of the exact ground state which depends on the many-body physics of the system.

This issue would make the stochastic evaluation of Eq. 2.49 impossible except for special cases where the correct nodal structure can be determined by symmetry considerations. A simple but effective workaround is the so called *fixed node* (FN) approximation [16]. For

a local $\hat{H}$, this procedure consists in constraining the importance sampling distribution inside a nodal pocket of the many body wavefunction which translates in the following inequality for the unormalized distribution $f(\mathbf{R}_{\rm el}, t)$ to be sampled:

$$f(\mathbf{R}_{\rm el}, t) = \Psi_G(\mathbf{R}_{\rm el}, t)\Psi_T(\mathbf{R}_{\rm el}) \geq 0 \quad \forall \ \ \mathbf{R}_{\rm el} \text{ and } t. \tag{2.50}$$

Under the FN constraint, the projection in Eq. 2.48 is approximated also in the infinite time limit, unless the trial wavefunction possesses the same nodal structure as the true ground state. The variational principle still holds within the FN approximation.

Notice that Eq. 2.50 is valid only for real-valued wavefunctions and in the case of extended systems the FN approximation must be generalized, as discussed in Sec. 4.4.

The standard DMC algorithm

Given the importance sampling transformation within the FN approximation, the imaginary time Schrödinger equation associated to $f(\mathbf{R}_{\rm el}, t)$ can be written as follows:

$$-\frac{\partial f}{\partial t} = \underbrace{-\frac{1}{2}\Delta_{\mathbf{R}_{\rm el}} f(\mathbf{R}_{\rm el})}_{\text{diffusion}} + \underbrace{\nabla_{\mathbf{R}_{\rm el}} \cdot [\mathbf{v}_D(\mathbf{R}_{\rm el})\ f(\mathbf{R}_{\rm el})]}_{\text{drift}} + \overbrace{\left[e_L(\mathbf{R}_{\rm el}) - (1 + \frac{\partial}{\partial t})\Lambda\right]}^{\text{branching}} \tag{2.51}$$

where the quantity $\mathbf{v}_D(\mathbf{R}_{\rm el}) = \nabla_{\mathbf{R}_{\rm el}} \ln \Psi_T$ is referred as drift velocity. Eq. 2.51 can be interpreted as a master equation for the probability distribution $f(\mathbf{R}_{\rm el}, t)$; in fact it can be directly linked with a physical process of particles diffusion: the first term is a pure *diffusion* term; the second one is a factor representing a *drift* in the particles trajectories characterized by the velocity $\mathbf{v}_D$; the last term models particles which enter (birth) and exit (death) the simulation and it is thus equivalent to a rate process which is usually referred as *branching*.

In order to simulate this drifted diffusion process through a stochastic evolution of the particle configurations, the standard fixed node DMC algorithm relies upon writing the time evolution of Eq. 2.51 in its integral form as:

$$f(\mathbf{R}_{\rm el}, t) = \int d\mathbf{R}'_{\rm el}\ \mathcal{G}(\mathbf{R}_{\rm el} \leftarrow \mathbf{R}'_{\rm el}, t)\ f(\mathbf{R}'_{\rm el}, 0) \tag{2.52}$$

where the propagator $\mathcal{G}(\mathbf{R}_{\rm el} \leftarrow \mathbf{R}'_{\rm el}, t)$ is the Green's function solution of Eq. 2.51 with the initial condition: $\mathcal{G}(\mathbf{R}_{\rm el}, \mathbf{R}'_{\rm el}, 0) = \delta(\mathbf{R}_{\rm el}, \mathbf{R}'_{\rm el})$. $\mathcal{G}(\mathbf{R}_{\rm el}, \mathbf{R}'_{\rm el}, t)$ is approximated with a product of short-time expression using the well-known Trotter-Suzuki decomposition [44, 45] of the time evolution exponential operator $\mathrm{e}^{-\tau\hat{H}}$. This procedure introduces a further approximation since it is exact only when the time step τ used in the Trotter-Suzuki formula goes to zero: $\tau \to 0$. The time step error is generally small compared to the FN error and can be cured by extrapolation procedures.

In practical implementations of FN-DMC, a population of "walkers", each representing a different trajectory $\mathbf{R}_{\text{el}}^i$ in the configurational space, is generated with a weight w_i such that $\sum_i w_i = 1$. The system is initially thermalized by performing a small amount of VMC steps realized by setting $f(\mathbf{R}_{\text{el}}, 0) \rightarrow \Psi_T^2(\mathbf{R}_{\text{el}})$ and all weights to 1. The diffusion process given by Eq. 2.51 is simulated by letting the walkers diffuse with random moves in the configurational space (generally drawn from a Gaussian distribution with spread τ) and applying a drift $\tau \mathbf{v}_D$ to the trajectory. After a certain number of new configurations in the drift-diffusion process, the branching term is applied and a certain amount of configurations are killed or replicated by an appropriate update of the weights. The energy shift $\Lambda(t)$ is modified in order to maintain a stable walker population and to keep the global random walk in the regions where the sampling distribution is larger.

For a more detailed description of the DMC algorithm and its variances we refer the interested reader to the reviews in Refs. [46] and [47].

Estimators in DMC

In the DMC method the evaluation of observables $\hat{O}$ is more delicate with respect to VMC. In the latter approach, the estimator O_{VMC} defined in Eq. 2.44 approximates to the true or "pure" average $O_0 = \frac{\langle \Psi_0 | \hat{O} | \Psi_0 \rangle}{\langle \Psi_0 | \Psi_0 \rangle}$. The quality of this approximation depends on the trial wavefunction Ψ_T. On the contrary, FN-DMC samples a different distribution, thus a direct application of Eq. 2.44 yields the mixed estimator:

$$O_{\text{mix}} = \frac{1}{M} \sum_{i=1}^{M} \hat{O}_L(\mathbf{R}_{\text{el}}^i), \tag{2.53}$$

where $\{\mathbf{R}_{\text{el}}^i\}$ are configurations drawn from the mixed distribution $f(\mathbf{R}_{\text{el}})$. It can be shown that O_{mix} coincides with O_0 if and only if the operator $\hat{O}$ commutes with the Hamiltonian. This directly follows from the definition of the Hamiltonian expectation value given in Eq. 2.47.

In other cases, it is possible to build an extrapolated estimator by noting that the true estimator is approximated by:

$$O_0 = O_{\text{mix}} + \langle \hat{O} \rangle_{\delta\Psi}, \tag{2.54}$$

where the second average in the right-hand-side of the above Equation is evaluated over $\delta\Psi$. The latter quantity is the linear difference between the trial wavefunction and the true ground state and it is defined as follows:

$$\Psi_0(\mathbf{R}_{\text{el}}) = \Psi_T(\mathbf{R}_{\text{el}}) + \delta\Psi + \mathcal{O}(\delta\Psi^2). \tag{2.55}$$

Using Eq. 2.55 and taking the expectation value of $\hat{O}$ with respect to Ψ_T (namely O_{VMC}) we get the following relation:

$$O_{\mathrm{mix}} = O_{\mathrm{VMC}} + \langle \hat{O} \rangle_{\delta\Psi} \implies \langle \hat{O} \rangle_{\delta\Psi} = O_{\mathrm{mix}} - O_{\mathrm{VMC}}. \tag{2.56}$$

From Eqs. 2.54 and 2.56 we notice that the linear variation can be cancelled out by choosing the following extrapolated estimator for O_0:

$$O_0 \simeq 2O_{\mathrm{mix}} - O_{\mathrm{VMC}} + \mathcal{O}(\delta\Psi^2). \tag{2.57}$$

The applicability of Eq. 2.57 is obviously limited by the quality of the trial wavefunction Ψ_T, thus the initial VMC ansatz must be well optimized in order for this method to give accurate results. In this thesis we use the mixed estimators formalism for computing correlation functions presented in Chap. 6. Alternative schemes for accessing the pure estimate O_0 have been devised. Among them, one of the most successful is the forward walking technique [48] which evolves the DMC weights through imaginary time such that, after the projection, one gets: $f(\mathbf{R}_{\mathrm{el}}, t) \to \Psi_0$. Forward walking must be applied as post-processing tool after a DMC simulation and it has approximately the same cost of the initial simulation; on the contrary, the reptation Monte Carlo method [49], a modification of the standard DMC algorithm to work with imaginary time dependent variables, allows to directly access the pure estimators during the QMC run.

2.4.3 Pseudopotentials and lattice regularized DMC

Energy consistent pseudopotentials

One of the main bottlenecks of the DMC algorithm is the bad scaling with the atomic number Z. This has been proven to be of the order of $\sim Z^{5.5-6.5}$ by simulating noble gases with large atomic number within DMC [50, 51]. This fact would limit the applicability of QMC methods to the study of light elements and small extended systems.

In order to overcome this issue, the core electrons need to be replaced with cheap, but accurate pseudopotentials. In this work we use pseudopotential functions specifically designed for QMC, developed by Burzatki, Filippi and Dolg (BFD) in Refs. [52, 53]. BFD pseudopotentials are freely available for many elements in Ref. [23].

The idea of a pseudopotential is to substitute the electronic part of the many-body Hamiltonian with an appropriately defined operator restricted only to $N_v < N$ valence electrons. This operator can be written as:

$$\hat{H}_{\mathrm{v}} = -\frac{1}{2}\sum_{i=1}^{N_v}\Delta_i + \sum_{i<j}^{N_v}\frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} + \sum_{a=1}^{N_{\mathrm{at}}}\sum_{i=1}^{N_v}\mathcal{V}_{\mathrm{pp}}^{a}(\mathbf{r}_i). \tag{2.58}$$

In Eq. 2.58 the term $\mathcal{V}^a_{\rm pp}(\mathbf{r}_i)$ is a semi-local and angular momentum dependent operator. Its main role is to reproduce the core-valence repulsion and the orthogonality among core and valence electronic states. The functional form of $\mathcal{V}^a_{\rm pp}(\mathbf{r}_i)$ could have nodes and it must possess the correct asymptotic limit for electron-nucleus separations $r = |\mathbf{r} - \mathbf{R}| \to 0$. The $\mathcal{V}^a_{\rm pp}(i)$ used in this work is parameterized as a sum of localized Gaussians functions multiplied by powers of the electron-ion distance. Defining N_l the number of Gaussians per angular momentum channel l, the pseudopotential operator is constructed as:

$$\hat{V}_{\rm pp}(r) = \underbrace{-\frac{Z_v}{r} + \sum_{k=1}^{N_{\rm loc}} \frac{A^k_{\rm loc}\exp(-a^k_{\rm loc}r^2)}{r^{2-n^k_{\rm loc}}}}_{\text{local part: } V_{\rm L}} + \underbrace{\sum_{l=0}^{l_{\rm max}}\sum_{k=0}^{N_l} \frac{A^k_l\exp(-a^k_l r^2)\hat{P}_l}{r^{2-n^k_l}}}_{\text{non-local part: } V_{\rm NL}}. \tag{2.59}$$

where A, a, N, n are adjustable parameters. In the previous expression, $Z_v < Z$ is the charge of the pseudized nucleus. $\hat{P}_l$ denotes a standard projector operator on the spherical harmonics of angular momentum l:

$$\hat{P}_l = \sum_{m=-l}^{l} |ml\rangle \langle ml| .$$

The maximum angular momentum $l_{\rm max}$ in the projection is typically chosen equal to the highest angular momentum occupied in the ground state atomic state. However, it has been recently argued [54] that higher angular momentum components can be beneficial for the final accuracy of the QMC calculation. The non-local angular integration required by the projector is performed numerically on a randomized mesh of points distributed according a given quadrature [55]. The local and non-local parameters involved in the pseudopotential definition in Eq. 2.59, namely A,a,N,n both local and non-local, are determined by fitting a number of all electron, scalar relativistic Hartree-Fock calculations performed at different electronic states of the considered atom (*energy consistent* scheme). Unless otherwise specified, BFD pseudopotential are employed for all QMC results presented in this work as well as for preparatory DFT calculations. When computationally feasible, such as in the case of the single water molecule and the protonated water dimer (see Sec. 3.3), we estimated the pseudopotential error by comparing selected pseudopotential energies with corresponding all-electron calculations.

Lattice regularization

The inclusion of non-local pseudopotentials into $\hat{H}$ brings serious issues to the FN approximation, since the non-local term cannot be constrained within nodal pockets. In fact, in Eq. 2.51, the pseudopotential operator applied to the mixed distribution $f(\mathbf{R}_{\rm el}, t)$ can

lead to a non-local branching, thus evolving the diffusion process outside the nodal pockets. This issue can be circumvented by neglecting the non-local terms in the imaginary time equation with the locality approximation [56]. However, this procedure sacrifices the variational character of the DMC algorithm, since the local Hamiltonian ground state can be lower in energy than the true ground state of the system.

The lattice regularized DMC (LRDMC) technique [5, 57] offers an alternative solution which keeps the variational character of the standard DMC algorithm. Instead of relying on the Trotter breakup for building the short-time approximation, LRDMC takes a different path based on a discretization of the many-body Hamiltonian $\hat{H}$. The non-local kinetic energy operator is discretized on a grid with lattice space a. At each update of the electronic position during the diffusion process, the direction of the lattice is chosen with random angles θ and ϕ drawn from a normal distribution. The randomization of the direction allows to explore the continuous space thus ensuring ergodicity with nearest neighbors lattice moves only. The full non-local part of the Hamiltonian now reads:

$$\hat{T}^a = -\frac{1}{2}\sum_{i=1}^{N} \Delta^a_{\mathbf{r}_i}(\theta_i, \phi_i) + \hat{V}_{\mathrm{NL}} + \mathcal{O}(a^2), \tag{2.60}$$

where $\Delta^a_{\mathbf{r}_i}(\theta_i, \phi_i)$ is the discretized Laplacian and we have included also the non-local pseudopotential operator previously defined.

A better accuracy can be achieved by discretizing also the local part of the pseudopotential – the electron-ion Coulomb interaction – by requiring the continuous and discretized version of the local energy to be the same: $e^a_L(\mathbf{R}_{\mathrm{el}}) = e_L(\mathbf{R}_{\mathrm{el}})$. This turns out in the following regularization for the local potential:

$$V^a_L(\mathbf{R}_{\mathrm{el}}) = V_L(\mathbf{r}) - \frac{(\Delta_{\mathbf{r}} - \Delta^a_{\mathbf{r}})\Psi_T(\mathbf{R}_{\mathrm{el}})}{2\Psi_T(\mathbf{R}_{\mathrm{el}})}, \tag{2.61}$$

where V_L has been defined in Eq. 2.59. Notice that the locality character of the one-body potential is lost and its discretized version is a purely many-body quantity since it depends on all electron positions. Finally, the full discretized Hamiltonian reads:

$$\hat{H}^a = \hat{T}^a + \hat{V}^a_L + \hat{V}_{\mathrm{ee}} + \mathcal{O}(a^2) \tag{2.62}$$

where only electron-electron interaction $\hat{V}_{\mathrm{ee}}$ is not discretized. Despite being defined on a continuous space, Eq. 2.62 is effectively equivalent to a lattice Hamiltonian. Therefore, the number of $\hat{H}$ matrix elements connecting two electron positions are a finite number. This is valid also in the presence of non-local pseudpotentials, which do not constitute a problem within this scheme.

Within LRDMC the time step error is substituted by the lattice space error; the two quantities are linked by the relation: $a = \sqrt{\tau}$. Benchmark calculations for the LRDMC

energy extrapolation to zero lattice space are presented Fig. 6.6 for the case of iron selenide (FeSe). The mixed average is evaluated using the discretized Hamiltonian:

$$E_{\mathrm{LRDMC}} = \frac{\langle \Psi_G | \, \hat{H}^a \, | \Psi_T \rangle}{\langle \Psi_G | \, \Psi_T \rangle}, \tag{2.63}$$

which converges to Eq. 2.49 when $a \to 0$. The imaginary time evolution exploited for evaluating Eq. 2.63 is carried out with the fixed node DMC algorithm suitably modified to work on a lattice [58] by operating the substitution: $e^{-[\hat{H}-\Lambda]} \to \Lambda - \hat{H}$.

2.5 Efficient wavefunction optimization

The flexibility of the QMC ansatz is extremely important in order to accurately include different physical ingredients depending on the system under consideration without changing the functional form. Flexibility can only be achieved by parameterizing the variational wavefunction with a sufficiently large set of tunable parameters ($\sim 10^3 - 10^4$). As we already mentioned, the degree of optimization of the whole set crucially affects the statistical efficiency of the VMC and, to a lesser extent, DMC algorithms.

The optimization of a function depending on a large set $\{\alpha\}$ of independent variables is an important branch of mathematical analysis and it is typically carried out by minimizing some *cost* function with respect to the parameters. The main choices for this cost function for QMC are:

- the variational energy $E = \langle e_L \rangle$ introduced in Eq. 2.42 [59, 60, 61, 62, 63].
- the variance of the variational energy defined as $\sigma^2 = \langle (e_L - \langle e_L \rangle)^2 \rangle$ and generally minimized through a correlated sampling approach [64, 65, 66].
- more rarely, some combination of the two, for instance $\frac{\sigma^2}{E^2_{\mathrm{VMC}}}$ [67] or a linear combination of the two with arbitrary coefficients [68]: $c_1 E + c_2 \sigma^2$.

Independently from this choice, the standard optimization methods, e.g. Newton method or linear method, must be modified to cope with the additional complication arising from the statistical noise affecting the evaluation of the cost function.

In this Section we present an energy minimization scheme based on the efficient stochastic reconfiguration (SR) technique. SR is able to optimize up to 10^4 variational parameters at the same time with a stable and robust procedure. Like VMC and DMC, SR is a variational method based on a MC approach for computing integrals. SR is used for all the results presented in this thesis as the optimization method of both Jastrow and determinantal parts of our ansatz.

2.5.1 Stochastic reconfiguration method

The stochastic reconfiguration method is an energy minimization technique for a generic trial wavefunction. Originally devised for lattice calculations [62], it has been successfully generalized and improved to work in the continuum [35, 63, 69] on realistic *ab initio* systems. SR is based upon the fact that the repeated application of the projector operator $\Lambda - \hat{H}$ – with Λ sufficiently large – to the ansatz Ψ_T allows to filter out the exact ground state of $\hat{H}$ in a variational fashion. This approach is similar to the imaginary time evolution used in the LRDMC method (see Subsec. 2.4.2). The following derivation is carried out for a real wavefunction, the generalization to complex arithmetic is simple and is briefly outlined in Chap. 4.

In the following, we consider a generic trial wavefunction $|\Psi_T(\alpha^0)\rangle$ which is parameterized with an initial set of $p+1$ variational parameters: $\boldsymbol{\alpha}^0 = \{\alpha_0^0, \ldots, \alpha_p^0\}$. These parameters can be *non-linear*, which is the case of the Gaussian exponents and of the Jastrow parameters. Within a JAGP calculation, the full set of optimizable parameters includes the Jastrow coefficients $C_{\mu,\nu}^{a,b}$, $S_{\mu,\nu}^{a,b}$ and homogeneous terms β, γ, the geminal matrix $\lambda_{\mu,\nu}^{a,b}$ (or the MOs coefficients λ_k^{MO} depending on the representation), the exponents ζ_i of the Gaussian orbitals in both determinant and Jastrow basis sets and also, within the Born-Oppenheimer approximation, the ionic positions $\mathbf{R}$.

Given a small collective variation of the parameters such that $\boldsymbol{\alpha} = \boldsymbol{\alpha}^0 + \delta\boldsymbol{\alpha}$, we can write the Taylor expansion of the updated wavefunction $\Psi_T'(\boldsymbol{\alpha})$ in the following way:

$$\Psi_T'(\boldsymbol{\alpha}) = \Psi_T(\boldsymbol{\alpha}^0) + \sum_{k=1}^{p} \frac{\partial \Psi_T(\boldsymbol{\alpha}^0)}{\partial \alpha_k^0} \delta\alpha_k + \mathcal{O}(\delta\boldsymbol{\alpha}^2). \tag{2.64}$$

A simple transformation allows to rewrite Eq. 2.64 as:

$$\begin{aligned} \Psi_T'(\boldsymbol{\alpha}) &= \Psi_T(\boldsymbol{\alpha}^0) + \sum_{k=1}^{p} \Psi_T(\boldsymbol{\alpha}^0) \frac{\partial \ln\left[\Psi_T(\boldsymbol{\alpha}^0)\right]}{\partial \alpha_k^0} \delta\alpha_k \\ &= \sum_{k=0}^{p} \delta\alpha_k \hat{O}_k \Psi_T(\boldsymbol{\alpha}^0), \end{aligned} \tag{2.65}$$

where we retain only the linear terms of the expansion.

The local operators $\{\hat{O}_k\}$ are defined as the derivatives of the logarithm of the wavefunction with respect to each variational parameter and we conventionally define $\hat{O}_0 = \mathbb{I}$ and $\delta\alpha_0 = 1$. In the case of $\delta\alpha_0 \neq 1$, it is sufficient to rescale the other parameters as:

$$\delta\alpha_k \to \frac{\delta\alpha_k}{\delta\alpha_0}, \tag{2.66}$$

for obtaining the sum in Eq. 2.65. Notice that the set of $\{\hat{O}_k\}$ spans a $(p+1)$-dimensional subspace of the whole Hilbert space defined by the wavefunction Ψ_T. The dimension of

the subspace is determined by the size of the wavefunction basis set. It is evident that the linear approximation in Eq. 2.65 becomes more accurate as the true ground state is approached, since the variations in the parameters tends to zero.

The SR method consists in an iterative application of the projector $\Lambda - \hat{H}$ to the restricted subspace spanned by the $\{\hat{O}_k\}$ operators which has the dimension of the basis set. One step of the SR procedure consists in solving the following Equation for each variational parameter [62]:

$$\langle \Psi_T | \hat{O}_k (\Lambda \mathbb{I} - \hat{H}) | \Psi_T \rangle = \langle \Psi_T | \hat{O}_k | \Psi_T' \rangle \quad \text{with } k = 0, \ldots, p. \tag{2.67}$$

Substituting the linear expansion defined in Eq. 2.65 into Eq. 2.67, one obtains a system of equations for the set of variations $\{\delta\alpha_k\}$:

$$\begin{cases} \Lambda - \langle \hat{H} \rangle = \delta\alpha_0 + \sum_{l=1}^{p} \langle \hat{O}_l \rangle \delta\alpha_l & \text{for } k = 0 \\ \Lambda \langle \hat{O}_k \rangle - \langle \hat{O}_k \hat{H} \rangle = \delta\alpha_0 \langle \hat{O}_k \rangle + \sum_{l=1}^{p} \delta\alpha_l \langle \hat{O}_k \hat{O}_l \rangle & \text{for } k = 1,\ldots,p, \end{cases}$$

where the expectation values $\langle \cdot \rangle$ are evaluated over the initial trial state $|\Psi_T(\boldsymbol{\alpha}_0)\rangle$. The above conditions can be cast into a more compact and computationally convenient matrix form:

$$\frac{1}{2} f_k = \langle \Psi_T | \hat{O}_k \sum_{l=0}^{p} \delta\alpha_l \hat{O}_l | \Psi_T \rangle = \sum_{l=0}^{p} \delta\alpha_l S_{kl}, \tag{2.68}$$

where $f_k = 2\left(\langle \hat{O}_k \rangle \langle \hat{H} \rangle - \langle \hat{O}_k \hat{H} \rangle\right)$ is the known term and $S_{kl} = \langle \Psi_T | \hat{O}_k \hat{O}_l | \Psi_T \rangle$ is a positive definite covariance matrix. The elements of both these matrices are evaluated with a standard Monte Carlo integration based on the Metropolis algorithm as VMC (see Sec. 2.4). The solution of Eq. 2.68 gives the variation of the wavefunction parameters along a direction selected by the projector $\Lambda - \hat{H}$ and it is found by inverting the covariance matrix:

$$\delta\boldsymbol{\alpha} = \frac{1}{2} S^{-1} \mathbf{f}. \tag{2.69}$$

The new wavefunction $\Psi_T(\alpha)$ after one SR step can be found by updating the parameters as:

$$\boldsymbol{\alpha} = \boldsymbol{\alpha}^0 + \Delta\delta\boldsymbol{\alpha} = \boldsymbol{\alpha}^0 + \Delta \frac{1}{2} S^{-1} \mathbf{f}, \tag{2.70}$$

where Δ is an appropriately small value related to the $\delta\boldsymbol{\alpha}$ normalization (Eq. 2.66); therefore $\Delta \propto 1/\Lambda$ and it is adjustable during the simulation for achieving a better convergence rate of the iterative scheme.

In fact, we show in the following that the energy gain at each SR iteration depends on the Δ chosen. This can be understood by noting that the derivative of the Hamiltonian expectation value with respect to a generic variational parameter can be written as:

$$\begin{aligned} -\frac{\partial \langle \hat{H} \rangle}{\partial \alpha_k} &= -\frac{\left\langle \frac{\partial \Psi_T}{\partial \alpha_k} \right| \hat{H} \left| \Psi_T \right\rangle + \left\langle \Psi_T \right| \hat{H} \left| \frac{\partial \Psi_T}{\partial \alpha_k} \right\rangle}{\left\langle \Psi_T \right| \left. \Psi_T \right\rangle} + 2 \frac{\left\langle \Psi_T \right| \hat{H} \left| \Psi_T \right\rangle \left\langle \frac{\partial \Psi_T}{\partial \alpha_k} \right| \left. \Psi_T \right\rangle}{\left\langle \Psi_T \right| \left. \Psi_T \right\rangle^2} \\ &= 2 \left(\langle \hat{O}_k \rangle \langle \hat{H} \rangle - \langle \hat{O}_k \hat{H} \rangle \right) = f_k \end{aligned} \tag{2.71}$$

where the second equality derives from the relation $\frac{\partial \Psi_T}{\partial \alpha_k} = \Psi_T \hat{O}_k$ and the hermiticity of the Hamiltonian. We also assume that $\hat{H}$ is independent on the variational parameter α_k. This is not the case if one considers the ionic positions as variational parameters for structural relaxation, as we will see later on in this Section.

Eq. 2.71 allows to interpret $\mathbf{f}$ as a generalized force with respect to the variational parameters $\boldsymbol{\alpha}$. By considering a sufficiently small Δ, the energy expectation value at one SR iteration can be expanded around the parameter variation. By retaining only the linear term, one obtains the energy gain at each SR step:

$$-\mathbf{f}^T(\boldsymbol{\alpha} - \boldsymbol{\alpha}^0) = -\frac{1}{2}\Delta \, \mathbf{f}^T S^{-1} \, \mathbf{f}, \tag{2.72}$$

which is always negative since the S matrix is always positive definite. As already mentioned, the Δ parameter controls the speed of convergence and the overall efficiency of the method. It must be an appropriate tradeoff between large values – breaking the approximation Eq. 2.72 and leading to unstable simulations – and too small values, leading to a very low convergence rate and thus long simulations [70].

2.5.2 Stochastic reconfiguration with ionic forces

Thanks to the Born-Oppenheimer approximation, the SR can be readily applied to structural minimization by considering the atomic positions $\{\mathbf{R}\}$ as additional variational parameters of the many-body wavefunction. At the VMC level, the force acting on an atom at position $\mathbf{R}$ can be written as a gradient of the local energy as follows:

$$f_{\mathbf{R}} = -\nabla_{\mathbf{R}} \langle \hat{H} \rangle = -\nabla_{\mathbf{R}} \langle e_L \rangle \tag{2.73}$$

where $e_L = \hat{H}\Psi_T/\Psi_T$ is, as usual, the local energy. In the definition of e_L both $\hat{H}$ and Ψ_T explicitly depend on the atomic positions. An implicit dependence in the wavefunction is also present through the electronic variational parameters: $\alpha_l = \alpha_l(\mathbf{R})$. However, it is easy to show that the gradient $\frac{\partial \langle \hat{H} \rangle}{\partial \alpha_l} \frac{\partial \alpha_l}{\partial \mathbf{R}}$ (obtained after applying the chain rule) vanishes when approaching the energy minimum for all variational parameters [71].

In VMC calculations, the well-known Hellmann-Feynman theorem is not directly applicable due to the approximate nature of the trial wavefunction. Therefore it is more convenient and efficient [72] to rely on the analytic expression of the gradient in Eq. 2.73. This quantity reads:

$$f_{\mathbf{R}} = -\frac{\int d\mathbf{R}_{\rm el} \frac{\partial e_L}{\partial \mathbf{R}} \Psi_T^2}{\int d\mathbf{R}_{\rm el} \Psi_T^2} - 2\frac{\int d\mathbf{R}_{\rm el} (e_L - \langle \hat{H} \rangle) \frac{\partial \ln |\Psi_T|}{\partial \mathbf{R}} \Psi_T^2}{\int d\mathbf{R}_{\rm el} \Psi_T^2}, \tag{2.74}$$

where the first term is an Hellmann-Feynman like expression, whereas the second term is usually referred as Pulay term and it vanishes when reaching the exact ground state.

The efficiency in the evaluation of Eq. 2.74 can be improved of ~ 1 order of magnitude by employing the differential *space warp coordinate transformations* (SWCTs) [66, 71, 72].

A SWCT is a differential coordinate transformation which is designed to mimic the displacement of an electron at position $\mathbf{r}_i$ when a chosen nucleus located at $\mathbf{R}_a$ is displaced of a vector $\mathbf{D}_a$ from its original position. In particular, if the electron is very far from the chosen atom, it does not feel the movement of the ion and hence it remains at the same position; on the other hand, if the electron is very close to the atom, due to the Born-Oppenheimer approximation it will approximately moves of the same amount $\sim \mathbf{D}_a$. This intuitive picture is achieved mathematically with the following transformation:

$$\begin{cases} \mathbf{R}_b \to \mathbf{R}'_b = \mathbf{R}_b + \mathbf{D}_a & \text{for } b \in [1, N_{\rm at}] \\ \mathbf{r}_i \to \mathbf{r}'_i = \mathbf{r}_i + \omega(|\mathbf{r}_i - \mathbf{R}_a|)\mathbf{D}_a & \text{for } i \in [1, N] \end{cases} \tag{2.75}$$

where

$$\omega(|\mathbf{r}_i - \mathbf{R}_a|) = \frac{|\mathbf{r}_i - \mathbf{R}_a|^{-4}}{\sum_{a=1}^{N_{\rm at}} |\mathbf{r}_i - \mathbf{R}_a|^{-4}}$$

is a weight function which adjusts the displacements according to the aforementioned reasoning. As detailed in Ref. [71], the analytical forces can be simply rewritten after the application of the above transformation. Alternative estimators for ionic forces have been devised [73, 74] which do not rely on the analytical expression in Eq. 2.74.

Eq. 2.74 combined with SWCT ensures, in principle, a very efficient evaluation of the forces $f_{\mathbf{R}}$. However, the main bottleneck in this computation is represented by the derivatives of the local energy $\frac{\partial e_L}{\partial \mathbf{R}}$ and of the variational wavefunction $\frac{\partial \ln |\Psi_T|}{\partial \mathbf{R}}$ appearing in Eq. 2.74. It is well-known [72] that the straightforward finite difference implementation is inefficient and affected by a larger statistical error with respect to the local energy computation. This fact often leads to unfeasible calculations especially for large systems. There have been a few attempts to overcome this issue such as, for instance, the one based on the correlated sampling technique for an accurate evaluation of energy differences [75]. In this work we rely on an alternative solution which is the subject of the next Section.

2.5.3 An alternative route to derivatives: adjoint algorithmic differentiation

It is evident that an efficient implementation of the SR method requires a fast computation of wavefunction derivatives with respect to a large set of variational parameters. As we previously mentioned, this is particularly important for evaluating ionic forces. Within our approach, this is achieved with the adjoint or automatic algorithmic differentiation (AAD), presented in this Section. AAD refers to a programming technique whose purpose is to calculate derivatives of an arbitrary complex function F implemented as a computer program. AAD is based on the idea that a program is always composed by a sequence of basic arithmetic operations (such as addition and multiplication) and intrinsic functions (such as exp, log, etc. etc.) which, individually, are elementary differentiable. It is therefore possible to compute the derivative of the output(s) with respect to the input(s) of the function by a mechanical and almost black-box application of the basic rules of differentiation. The resulting derivatives are therefore *analytical* with neither approximation nor truncation errors. Furthermore, the computational of AAD cost is proportional, with a small prefactor, to the cost of the initial program. Hence, it is sufficient to devise an efficient algorithm to compute the function F in order to obtain, with the same efficiency, all its derivatives with respect to the input(s).

For this reason, AAD is a very attractive approach in situations where there is no obvious or computationally efficient algorithm to compute the derivatives of a certain quantity which, on the contrary, can be evaluated fast. This is the case, for example, of the determination of derivative assets as well as portfolio risks in computational finance, for which AAD has been applied the first time outside computer science research [76, 77]. Within QMC, this approach has been introduced in Ref. [72] for an efficient computation of many-body ionic forces. In this thesis, this approach has been extended to all the derivatives of the many-body wavefunction Ψ_T with respect to the variational parameters needed by the SR optimization method.

Following Ref. [72], we consider a computer implemented function which, from a large number M_I of inputs collectively denoted as the vector $X = (X_1, \ldots, X_{M_I}) \in \mathbb{R}^{M_I}$, allows to evaluate a smaller number M_O of outputs referred as the vector $Y = (Y_1, \ldots, Y_{M_O}) \in \mathbb{R}^{M_O}$ ($M_I \gg M_O$) in two steps:

$$Y = F(X) \quad \text{with} \quad X \to U \to V \to Y. \tag{2.76}$$

In the implementation within QMC, the outputs are the local energy and the many-body wavefunction ($M_O = 2$), whereas the inputs are represented by all the electron coordinates and variational parameters (M_I up to $\sim 10^4$). The two intermediate steps U and V can be composed, in the simplest case, by a single code instruction or by an entire high-level

subroutine. In the illustrative case of Eq. 2.76, the derivative of the output with respect to the input is the result of the following application of the chain rule:

$$\frac{\partial Y}{\partial X} = \frac{\partial Y}{\partial V}\frac{\partial V}{\partial U}\frac{\partial U}{\partial X}. \tag{2.77}$$

If one focuses on an individual step, such as $V \to Y$, its corresponding contribution to the chain rule in Eq.2.77 is defined as the derivatives of an arbitrary linear combination of the output vector $\{Y_i\}$ with respect to the input as follows:

$$\bar{V_k} = \sum_{j=1}^{M_O} \bar{Y_k}\frac{\partial Y_j}{\partial V_k}. \tag{2.78}$$

The quantity $\bar{V_k}$ is called the *adjoint* of the input variable V referred to the instance $V \to Y$ of the function F. $\bar{Y}$ is a given vector in $\mathbb{R}^{M_I}$ defining the arbitrary linear combination, as we shall see in greater detail later on. By using the above definition of adjoint for each step, the chain rule in Eq. 2.77 can be rewritten as follows:

$$\bar{Y}\frac{\partial Y}{\partial X} = \bar{V}\frac{\partial V}{\partial U}\frac{\partial U}{\partial X} = \bar{U}\frac{\partial U}{\partial X} = \bar{X} \tag{2.79}$$

The final result $\bar{X}$ is the adjoint of the input and it represents a linear combination of the rows of the Jacobian $J_{jk} = \frac{\partial F_j(X)}{\partial X_k}$ of the function F, defined as:

$$\bar{X_k} = \sum_{j=1}^{M_O} \bar{Y_k}\frac{\partial Y_j}{\partial X_k} = \sum_{j=1}^{M_O} \bar{Y_j}J_{jk} \quad \text{with } k = 1, \ldots, M_I. \tag{2.80}$$

The weights of each component of the Jacobian are therefore given by the vector $\bar{Y}$, which is set at the beginning of the AAD procedure. For example, if one wants to calculate the derivative with respect to the first component only of the input X, the $\bar{Y}$ vector can be set to $(1, 0, \ldots, 0)$.

From Eq. 2.79 it is easy to realize that, once all adjoints are defined, the derivatives of the function F (Eq. 2.80) can be computed by simply following the original algorithm in the reversed sequential order as follows:

$$\bar{Y} \to \bar{V} \to \bar{U} \to \bar{X}. \tag{2.81}$$

Hence, the AAD procedure consists in rewriting the initial or *direct* program from its last instruction towards the beginning by propagating the adjoints of each step, i.e. by differentiating each instruction following the chain rule sketched in Eq. 2.77. The final result of the AAD is a new program $\bar{F}$ which implements the *reverse* of the original algorithm: each time the original program holds a variable `v`, the reverse holds a variable

Table 2.1 – Example of AAD implementation in Fortran90. Direct and reverse algorithms for computing the distance between two generic real arrays and its derivatives (see text for definition). The adjoints of each variable are denoted by the $\_b$ signature. The outcome of the procedure is represented by the vectors A_b and B_b containing the derivatives of the distance e with respect to each component of the input vectors A and B. The function `PUSH` (`POP`) are used to store in (retrieve from) memory buffers appropriately defined the intermediate variables of the direct program needed by the reverse algorithm. For the sake of simplicity, in this explanatory program the arbitrary vector $\bar{Y}$ possesses only one component which is obviously set to 1.

DIRECT PROGRAM	REVERSE PROGRAM

`e2 = 0.0`
`A(1:N) = 0.0`
`B(1:N) = 0.0`
`do` `i = 1,N`
`  e1 = A(i) - B(i)`
`  call PUSH(e1)`
`  e2 = e2 + e1**2`
`  call PUSH(e2)`
`end do`
`e = sqrt(e2)`
	`A_b(1:N) = 0.0`
	`B_b(1:N) = 0.0`
	`e_b = 0.0`
	`e1_b = 0.0`
	`e2_b = 0.0`
	`e2_b = e2_b + 0.5/sqrt(e2)*e_b`
	`e_b = 0.0`
	`do` `i = N,1,-1`
	`  call POP(e1)`
	`  e1_b = e1_b + 2*e1*e2_b`
	`  call POP(e2)`
	`  A_b(i) = A_b(i) + e1_b`
	`  B_b(i) = B_b(i) - e1_b`
	`  e1_b = 0.0`
	`end do`
	`e2_b = 0.0`

$\bar{\mathtt{v}}$ with the same shape, containing the derivative of `v` with respect to an arbitrary linear combination of the inputs.

A real program devised with the sequential programming logic is made of numerous steps of the kind in Eq. 2.76. Furthermore, these steps typically possess a tree-like connectivity, in the sense that the output of each step constitutes not only the input for the step following directly, but also for other parts of the program, therefore contributing to different adjoints in the reverse algorithm. It is thus possible to increase the efficiency of AAD by storing in memory, by appropriately defined buffers, the intermediate outputs of the original program (*Store-All* approach, see Tab. 2.1 for a simple example) before start-

ing the reverse, instead of recomputing them within the reverse program (*Recompute-All* approach). In the *TurboRVB* implementation it is always used the more efficient, but harder to code, *Store-All* approach.

The cost of AAD (Eq. 2.80) satisfies the following inequality [76]:

$$\frac{\text{Cost}\,[F' + F]}{\text{Cost}\,[F]} \leq \omega M_O \tag{2.82}$$

where $F' = \nabla F$ refers to Jacobian and F to the function itself, whereas $\omega \in [3, 4]$ is a small prefactor multiplying the small number M_O of dependent variables. Furthermore, the cost of AAD is also linear in the number of the components of F' needed in output. In order to obtain all the Jacobian components is indeed sufficient to repeat M_I times the procedure by setting the $\bar{Y}$ arbitrary vector as $(1, 0, 0, \ldots, 0)$ / $(0, 1, 0, \ldots, 0)$ / $(0, 0, 1, \ldots, 0)$ and so on.

The practical implementation of the reverse algorithm is mechanical in nature and therefore it can be performed itself by a program. For this purpose several tools have been developed. In the *TurboRVB* code, a few reversed routines among the most complicated ones have been created by the software Tapenade [78], while the remaining ones have been coded by hand. Notice that AAD can be simply extended to complex variables needed for treating periodic solids within QMC (see Chap. 4). We present the adjoint version of two of the most important operations among complex vectors in Appendix C.

An exhaustive description of the programming rules needed to implement AAD is beyond the scope of this section and the interested reader can refer to the original paper in Ref. [72]. Nevertheless, it is useful in conclusion to report an example of practical implementation of AAD. In Tab. 2.1 we show the direct and reverse algorithms of a function calculating the distance between two real vectors A and B of dimension N defined as: $e = \sqrt{\sum_{i=1}^{N}(A_i - B_i)^2}$. Despite its simplicity, this sample code presents all the main steps needed to differentiate an algorithm via AAD in the *Store-All* approach.

2.6 Conclusions

In this Chapter we presented a comprehensive theoretical framework based on quantum Monte Carlo techniques for studying the electronic structure of molecular, open shell systems.

We use the accurate AGP wavefunction built as an antisymmetric product of geminal or electron pairing functions as the determinantal part of our ansatz. Thanks to a DFT package built-in to the QMC software *TurboRVB* we are able to use the same setup (basis set and pseudopotential) for DFT and QMC calculations without relying on any parser

tool. A flexible representation of the AGP wavefunction based on a molecular orbital expansion allows us to go from the independent particle DFT wavefunction to the correlated AGP by keeping the same computational cost ($\sim N^3 - N^4$ with, however, a much larger prefactor with respect to the DFT calculations). The combination of the Jastrow factor with the pairing function yields the *ab initio* realization of the resonating valence bond theory developed by Pauling and Anderson. The JAGP wavefunction recovers most part of both dynamic and static electron correlation and allows a fine control on the strength of chemical bonds among different atomic orbitals.

The wavefunction optimization is crucial to obtain a good QMC variational energy. For this purpose, we have introduced the well-known stochastic reconfiguration energy minimization technique. At variance with previous implementations of this method, in our work we utilize a novel and efficient approach for computing wavefunction derivatives based on adjoint algorithmic differentiation. Besides its efficiency, AAD also allows a fast implementation of new functional forms of the wavefunction ansatz, such as for example the Pfaffian function [39], whose usefulness will be pointed out in the last Chapter of this thesis. In fact, it is only necessary to devise an efficient algorithm to evaluate the wavefunction and the local energy, then its derivatives are automatically obtained with low coding effort and with the same computational complexity as the direct algorithm.

Chapter 3

Applications to aqueous systems: proton transfer reactions

3.1 Introduction

In this Chapter we present extensive applications of the QMC framework previously developed to an open problem in quantum chemistry: the phenomenon of proton transfer (PT) reactions in aqueous systems. Our investigation involves also a thorough study of the water molecule, as a fundamental preliminary assessment for the study of PT systems.

Proton transfer is a ubiquitous phenomenon occurring along hydrogen bond networks which develops on energy scales lower than the chemical accuracy of 1 Kcal/mol. Its accurate *ab initio* description requires therefore advanced computational methods such as coupled cluster techniques or QMC. On the other hand, in order to extract quantitative properties of PT physics, such as the activation barrier of the reaction or the proton diffusion coefficient, large scale simulations are needed. We will show in this Chapter that the favorable scaling of our QMC framework combined with an efficient method for reducing the number of variational parameters introduced here, allows us to cope with both of these requirements.

Previous *ab initio* molecular dynamics studies have been mainly carried out using potential energy surfaces (PESs) obtained from density functional theory calculations [79, 80] due to its very cheap computational cost. However, DFT-based PESs fail to reproduce basic properties of liquid water such as the melting temperature [81] and the oxygen-oxygen radial distribution function g_{OO} [82]. Recent *ab initio* molecular dynamics on liquid water [7, 70] based on QMC forces have improved this scenario and they will likely shed new light on H_2O properties when next generation supercomputers will enable longer and more accurate simulations.

The DFT failure on liquid water affects also the DFT description of PT reactions. This process, whose microscopic mechanism is known as the Grotthuss mechanism [83], is extremely important in a wide range of biological systems and influences many dynamical processes in material science, biochemistry [84] and bioenergetics [85, 86]. It also plays a fundamental role in fuel cell development and fabrication.

The theoretical predictions based on DFT simulations [87, 88] are not in accordance with experiments [89, 90] even for a basic descriptor of the reaction such as the PT activation barrier. Nevertheless, these simulations have already given many insights into the physics of Grotthuss mechanism [91, 92, 93]. In particular, it has been found that proton transport in water is abnormally fast and is not driven by any ordinary diffusion process. The picture which is instead commonly accepted describes the translocation of excess positive charge as a structural diffusion of defects, involving both structural and dynamical rearrangements of the hydrogen-bonded network [93].

As already mentioned, one of the main issues preventing from a complete understanding of PT in water, is related to its very sensitive thermal behavior. The required PES precision is of the order of few tenths of Kcal/mol, down to the so-called *subchemical accuracy*, which has been reached only recently by computational methods beyond DFT, such as coupled cluster (CC), multi-reference configuration interaction (MRCI) and the full configuration interaction quantum Monte Carlo (FCIQMC) methods. However, CC scales as $N^5 - N^7$ with the number of particles, whereas MRCI and FCIQMC show a weak exponential scaling; this makes large scale realistic molecular dynamics simulations unfeasible with these methods.

Therefore, highly correlated investigations mainly focused, so far, on the two most important structural models devised to explain proton hydration. M. Eigen [94] considered as the core of PT the complex $H_9O_4^+$ in which a hydronium H_3O^+ is strongly bound to three water molecules; G. Zundel [95] proposed a different reaction which involves the simpler $H_5O_2^+$ complex (the "Zundel ion" or protonated water dimer) as the core of the transfer of a proton between two H_2O molecules. These two models occur only as limiting case of a more complex phenomenon [91, 93]; nevertheless, they represent a perfect testing ground for correlated computational methods.

In particular, the protonated water dimer $H_5O_2^+$ is the smallest system in which an excess proton is shared between water molecules; due to its simplicity, many studies have been carried out to elucidate its structure [96], energetics [97, 96] and vibrational properties [98]. Also, molecular dynamics simulations including quantum effects via Feynman path integrals have been performed [99]. Last but not least, an accurate PES has been recently produced [100] by fitting a large set of coupled cluster calculations including single, double and perturbative triple excitations (CCSD(T) method).

3.1.1 Organization of the Chapter

The aim of this Chapter is to set the stage for realistic PT calculations based on our QMC approach. In order to reach this goal, we focus our efforts on two tightly coupled directions.

On one hand, we address a major issue of QMC simulations: the rapid growth ($\sim N_{\rm bas}^2$) of the computational burden with the size of the primitive basis set. This issue could impose severe limitations on the large-scale simulations of PT systems needed to extract physical observables comparable with experiments (such as the proton transfer reaction barriers and diffusion coefficients), but also to take into account solvation effects with a sufficiently large number of H_2O molecules. We develop a technique based on a novel kind of contracted orbitals which we dub geminal embedded orbitals (GEOs). GEOs are contracted hybrid orbitals designed to possess optimal overlap with the initial uncontracted basis set. In the context of PT, they are first constructed for the H_2O molecule alone and then directly transferred to the more complex Zundel model. The theoretical foundations of the GEOs method are presented in Secs. 3.2.1 and 5.5.1 and the method is compared to standard natural orbitals in Sec. 3.2.3. We work out the optimal contracted basis for the water monomer in Sec. 3.2.4.

On the other hand, we present in this Chapter an extensive analysis of structural properties of the $H_5O_2^+$ complex in order to assess the accuracy of our QMC framework for more complex PT systems. In particular, we address in Sec. 3.3.1 the ground state geometrical structure of the Zundel ion as modeled by QMC. In Sec. 3.3.2 we work out its energy landscape along the oxygen-oxygen distance, the natural coordinate of the PT reaction. The transition between the symmetric configuration – where the proton is evenly shared among the two oxygens – and the asymmetric regime – where the excess proton is more bound to one H_2O molecule – is critical to understand the PT mechanism. In Secs. 3.3.3, 3.3.4 we demonstrate that QMC can describe this crossover with an agreement down to subchemical accuracy ($\sim$ 0.2 Kcal/mol) with the quantum chemistry "golden standard" CCSD(T) method [100], and we discuss the connections of our results with realistic finite temperature simulations. Finally in Sec. 3.3.5 we test the ability of our approach to tackle much larger protonated water clusters in an affordable computational time. This is achieved by performing test calculations on a simulation model containing 6 water molecules and an excess proton. We draw the conclusions to this Chapter in Sec. 3.4

3.2 Geminal embedded orbitals

In this Section we present a procedure to reduce the primitive basis set size of the QMC ansatz (Eq. 2.28) in an effective and optimal way. This is achieved by defining contracted hybrid orbitals, dubbed geminal embedded orbitals (GEOs), by means of an original density matrix embedding procedure published in Ref. [101]. We apply this procedure to the H_2O molecule and devise an optimal contracted basis set, transferable to more complex water clusters such as the protonated water dimer presented in Sec. 3.3.

Reducing the number of localized orbitals over which the wavefunction is expanded has several advantages and it is crucial for tackling large molecular clusters or sizable supercell calculations with QMC. At first it obviously decreases the computational effort for the wavefunction evaluation. Furthermore, it also guarantees a more robust energy minimization procedure (see Sec. 2.5). Indeed, given the fact that QMC energy derivatives are noisy, using a more compact basis set reduces redundancy and helps find more quickly the global minimum since the number of the effective directions in the Hilbert space is smaller.

3.2.1 Embedding scheme

The scheme presented in this section is based on a density matrix embedding of the determinantal part of the wavefunction. The Jastrow factor is not considered in the following discussion since its convergence with basis set is much quicker and therefore it generally requires few basis functions. This method is an application of the concept of quantum entanglement between a part of a system (A) interacting with the environment (B), where one represents A (B) with a set of M_A (M_B) orthogonal states labeled by the index i (j) and writes down the wavefunction of the universe $U = A \cup B$ as:

$$|U\rangle = \sum_{i=1,j=1}^{M_A,M_B} \psi_{ij} |i \otimes j\rangle. \tag{3.1}$$

It is straightforward to show [102] that the optimal way to describe the universe by using only a few $p << M_A$ states of the system embedded in the environment is obtained by using the p eigenvectors corresponding to the largest eigenvalues of the density matrix:

$$D_{ii'} = \sum_j \psi^*_{ij} \psi_{i'j}. \tag{3.2}$$

This approach is extremely simple and general, and it has been successfully applied in a variety of embedding schemes such as, for instance, the popular density matrix renormalization group (DMRG) [102].

Within the implementation presented in this section, the universe is restricted to the AGP wavefunction defined in Eq. 2.27: the many-body coefficients ψ_{ij} of Eq. 3.1 are replaced by the matrix $\Lambda = \{\lambda_{i,j}\}$ (Eq. 2.28), the states $\{|i\rangle\}$ are the one-electron basis orbitals (Eq. 2.14) and the universe size $M_A + M_B$ is the total basis set size $N_{\rm tot}$.

3.2.2 Detailed procedure

Let us now consider the MOs representation of the geminal reported in Eq. 2.29. For the sake of simplicity, we assume a real-valued geminal in the following discussion. The generalization to complex arithmetic is straightforward. As already mentioned, the geminal is also suited to describe the Slater determinant as a limiting case when $N_{\rm MO} = N/2$. In this situation, the pairing function becomes equal to the standard one-body density matrix. The embedding proposed here is done at the geminal level, by left projecting Eq. 2.29 over a single atom centered at $\mathbf{R}$:

$$U^{\mathbf{R}}_{\rm proj}(\mathbf{r},\mathbf{r}') = \sum_{k}^{N_{\rm MO}} \lambda^{\rm MO}_k \psi^{\rm proj}_{k,\mathbf{R}}(\mathbf{r})\, \phi^{MO}_k(\mathbf{r}') = \tag{3.3}$$

$$= \sum_{k}^{N_{\rm MO}} \lambda^{\rm MO}_k \sum_{i|\mathbf{R}_i=\mathbf{R}}^{N_{\rm tot}} \sum_{j}^{N_{\rm tot}} d_{ki}\, d_{kj}\, \chi_i(\mathbf{r})\, \chi_j(\mathbf{r}') \tag{3.4}$$

where the indices (i,j) run over the total basis set. The $\psi^{\rm proj}_{k,\mathbf{R}}(\mathbf{r}) = \sum_{j|\mathbf{R}_j=\mathbf{R}} d_{kj}\chi_j(\mathbf{r})$ are the molecular orbitals restricted only to those components centered on the atom chosen for the projection. Unless otherwise stated, hereafter we are going to omit the symbol $\mathbf{R}$ in the projected quantities, for readability. The term defined in Eq. 3.4 carries information on the intra-atomic electronic structure affected by inter-atomic interactions between the site $\mathbf{R}$ and its environment. The inter-atomic interactions are explicitly kept by the left-partial projection of the full density matrix.

For the next step of the procedure, we choose to represent the left-projected geminal of Eq. 3.4 in terms of $N_{\rm geo} \ll N_{\rm MO}$ atomic hybrid orbitals $\phi^{\rm GEO}_k$, dubbed geminal embedded orbitals (GEOs). The procedure explained in this Section aims at finding the coefficients for this reduced basis set which best reproduce the initial uncontracted localized basis. Using $\phi^{\rm GEO}_k$, left-projected on the atom located at $\mathbf{R}$, and a set of auxiliary molecular orbitals $\bar{\psi}_k$ spanning all the basis set, we can rewrite Eq. 3.4 as follows:

$$\bar{U}_{\rm proj}(\mathbf{r},\mathbf{r}') = \sum_{k=1}^{N_{\rm geo}} \phi^{\rm GEO}_k(\mathbf{r})\bar{\psi}_k(\mathbf{r}'). \tag{3.5}$$

Notice that the auxiliary orbitals $\bar{\psi}_k$ are defined, exactly as the standard $\phi^{MO}_k(\mathbf{r})$, as a linear combination of the primitive Gaussian basis set. In order to find the best GEOs, we

minimize the Euclidean distance $L = |U_{\rm proj} - \bar{U}_{\rm proj}|$ between the original and the truncated geminal functions. These functions are defined in $\mathcal{R}^3 \times \mathcal{R}^3$ in such a way that:

$$\begin{aligned} L^2 &= |U_{\rm proj}|^2 - 2\sum_k \int d\mathbf{r} d\mathbf{r}' U_{\rm proj}(\mathbf{r},\mathbf{r}')\phi_k^{\rm GEO}(\mathbf{r})\bar{\psi}_k(\mathbf{r}') \\ &+ \sum_k \int d\mathbf{r}\ \bar{\psi}_k^2(\mathbf{r}), \end{aligned} \tag{3.6}$$

where $|U_{\rm proj}|^2 = \int d\mathbf{r} d\mathbf{r}' U_{\rm proj}^2(\mathbf{r},\mathbf{r}')$, and we assumed that the optimal atomic orbitals are orthonormal. This assumption is valid without loss of generality, because - whatever is the solution for the minimum - we can always orthogonalize the corresponding optimal orbitals $\phi_i^{\rm GEO}$ and get a solution written in the same form as in Eq. 3.5. Taking the variation over all possible unconstrained functions $\bar{\psi}(\mathbf{r})$, one can show that the steady condition $\frac{\delta L^2}{\delta \psi_k(\mathbf{r})} = 0$ implies:

$$\bar{\psi}_k(\mathbf{r}) = \int d\mathbf{r}' U_{\rm proj}(\mathbf{r}',\mathbf{r})\phi_k^{\rm GEO}(\mathbf{r}'). \tag{3.7}$$

Replacing Eq. 3.7 into Eq. 3.6 yields:

$$L^2 = |U_{\rm proj}|^2 - \sum_k \int d\mathbf{r} d\mathbf{r}' D_{\rm proj}(\mathbf{r},\mathbf{r}')\phi_k^{\rm GEO}(\mathbf{r})\phi_k^{\rm GEO}(\mathbf{r}'), \tag{3.8}$$

where $D_{\rm proj}$ is the density matrix kernel defined, in analogy with Eq. 3.2, as:

$$D_{\rm proj}(\mathbf{r},\bar{\mathbf{r}}) = \int d\mathbf{r}' U_{\rm proj}(\mathbf{r},\mathbf{r}')U_{\rm proj}(\bar{\mathbf{r}},\mathbf{r}'). \tag{3.9}$$

Thus, in order to minimize L^2 one needs to maximize the quadratic form involving $D_{\rm proj}$, with the constraint that the orbitals $\phi_k^{\rm GEO}(\mathbf{r})$ are orthonormal. $D_{\rm proj}$ is a density matrix kernel and therefore it is a positive definite symmetric matrix which satisfy the minimum/maximum property of symmetric operators. It is thus clear that L^2 is minimized just when the optimal GEOs coincide with the $N_{\rm geo}$ eigenvectors of the density matrix $D_{\rm proj}$ with maximum eigenvalues w_i. Indeed, all the eigenvalues w_i must be positive, and the corresponding eigenvectors are obviously an orthonormal set of states, consistently with the assumption.

From Eq. 3.9 and the choice of the atomic projectors used in its definition, it follows that the density matrix kernel $D_{\rm proj}$ can be expressed in terms of the atomic basis $\{\chi_i\}$ restricted around a given atom at the selected position $\mathbf{R}_i = \mathbf{R}$. As a consequence, also the optimal GEOs can be expanded on the same local basis:

$$\phi_i^{\rm GEO}(\mathbf{r}) = \sum_{j|\mathbf{R}_j=\mathbf{R}} d_{ij}^{\rm GEO}\chi_j(\mathbf{r}). \tag{3.10}$$

In the non-orthogonal finite basis $\{\chi_j(\mathbf{r})\}$, this turns into the generalized eigenvalue equation:

$$\sum_{j|\mathbf{R}_j=\mathbf{R}} \Big[(\lambda s \lambda^\dagger) s\Big]_{ij} d^{\mathrm{GEO}}_{kj} = w_k d^{\mathrm{GEO}}_{ki} \quad \text{for } i \text{ s.t. } \mathbf{R}_i = \mathbf{R}, \text{ and } j \text{ s.t. } \mathbf{R}_j = \mathbf{R} \ , \tag{3.11}$$

where the matrix λ has been defined through the transformation in Eq. 2.31 by restricting the indices i, j to span only the basis elements centered around the atom located at $\mathbf{R}$ in order to fulfill the left-projection. $s_{ij} = \langle \chi_i | \, \chi_j \rangle$ is the positive definite basis set overlap matrix. Eq. 3.11 can be immediately solved by standard linear algebra packages[a], by considering that the overlap matrix s is symmetric and positive definite. After the diagonalization, the eigenvector coefficients satisfy the orthogonality requirement $d^{GEO} s (d^{GEO})^\dagger = I$, that we have previously assumed. Moreover, the truncation error, i.e. the residual distance, is $\epsilon_L = |U_{\mathrm{proj}}|^2 - \sum_{i=1}^{p} w_i$.

Thanks to the embedding scheme, the GEOs in Eq. 3.10 are the best compromise between basis set size and accuracy while they also carry physical information on the most representative atomic states. Furthermore, GEOs are expanded over the whole set of atomic angular momenta, therefore they are hybrid orbitals. Thanks to this feature they can automatically account for non-trivial chemical hybridization and for the crystal field effect in solids [103].

Optimal hybrid orbitals are typically determined within a small cluster (supercell) and then ported to describe larger molecular (solid) models. The optimal GEOs describing each atom can be evaluated by applying the procedure outlined in this section to previous DFT calculations (SD wavefunction), or fully optimized JAGP wavefunctions after removing the Jastrow factor.

Once the optimal GEOs are constructed, the AGP wavefunction of Eq. 2.28 can be rewritten as:

$$\tilde{\Phi}(\mathbf{r}, \mathbf{r}') = \sum_{a,b}^{N_{\mathrm{at}}} \sum_{i,j}^{N_{\mathrm{geo}}} \tilde{\lambda}^{a,b}_{i,j} \, \phi^{\mathrm{GEO}}_{i,a}(\mathbf{r}) \, \phi^{\mathrm{GEO}}_{j,b}(\mathbf{r}'). \tag{3.12}$$

where the $\tilde{\lambda}^{a,b}_{i,j}$ are given by the overlap maximization of the latter $\tilde{\Phi}$ with the original Φ in the $\mathbb{R}^3 \times \mathbb{R}^3$ space, and we explicitly write the GEOs dependence on the atomic index. Recalling that P^{VB} represents the number of parameters in the uncontracted representation, it is evident that Eq. 3.12 greatly reduces the dimension of the basis as the number of variational parameters amounts now to:

$$P^{\mathrm{GEO}} \propto (N_{\mathrm{tot\ geo}})^2 + N_{\mathrm{tot\ geo}} \times N_{\mathrm{bas}} \ll P^{\mathrm{VB}}, \tag{3.13}$$

[a]This linear problem corresponds to a generalized eigensystem equation of the type $ABx = \lambda x$ with $A = \lambda s \lambda^\dagger$ a symmetric matrix and $B = s$ a symmetric and positive definite one, x and λ being the eigenvector and the corresponding eigenvalue, respectively.

where $N_{\text{tot geo}} = N_{\text{geo}} \times N_{\text{at}}$ is the total number of GEOs in the optimal basis set.

Table 3.1 – Overlap $\langle\tilde{\Phi}|\Phi\rangle^2/(\langle\Phi|\Phi\rangle\langle\tilde{\Phi}|\tilde{\Phi}\rangle)$ between the geminal Φ of the fully optimized Jastrow-single determinant wave function of the water molecule in the primitive basis set (see Sec. 3.2.4 for more details) and the best $\tilde{\Phi}$ developed on the GEO/ANO basis set reported in the first column for oxygen and hydrogen; for instance, (8O,5H) means that we used 8 GEOs/ANOs for the oxygen atom and 5 GEOs/ANOs for the hydrogen atom. Given the basis set size, the contracted atomic orbitals are determined in the "standard" way (third column, Eq. 3.16) and by the geminal embedding scheme described in Sec. 3.2 (second column). The embedding scheme presented here systematically gives better overlaps and it converges to full overlap already for a (4O, 4H) basis set. The last line corresponds to the complete basis set limit for the contractions with respect to the space spanned by the primitive basis set, where all methods have to converge by definition (as it is actually found numerically).

GEO/ANO basis set	GEOs overlap (%)	"standard" ANOs overlap (%)
(4O,1H)	99.8390	99.0696
(8O, 2H)	99.9511	99.1846
(4O, 4H)	100.0000	99.1929
(4O, 5H)	100.0000	99.1933
(8O, 5H)	100.0000	99.2188
(12O, 6H)	100.0000	99.8305
(20O, 8H)	100.0000	99.9458
(30O, 10H)	100.0000	100.0000

3.2.3 Comparison with standard natural orbitals

The concept of atomic natural orbitals (ANO), introduced by Löwdin in 1956 [104], has been widely developed in quantum chemistry. Efficient schemes to find optimal hybrid basis set contractions have been devised [105, 106] and efforts have been devoted to extract clear chemical pictures from ANOs [107]. However, the typical scheme for building ANOs differs from the embedding procedure introduced in this section. Standard ANOs are defined as eigenstates of the local atomic density matrix left *and* right projected on the given site. Given $D = D(\mathbf{r}, \mathbf{r}')$ the density matrix generally defined in Eq. 3.2, the atomic projected version read:

$$D^{\mathbf{R}}_{\text{atom}} = \mathcal{P}^{\text{at}}_{\mathbf{R}} D \mathcal{P}^{\text{at}}_{\mathbf{R}}, \tag{3.14}$$

where $\mathcal{P}^{\text{at}}_{\mathbf{R}}$ are the standard atomic projectors within a non-orthogonal basis set:

$$\mathcal{P}^{\text{at}}_{\mathbf{R}} = \sum_{i|\mathbf{R}_i=\mathbf{R}} \sum_{j|\mathbf{R}_j=\mathbf{R}} |\chi_i\rangle \, s^{-1}_{ij} \, \langle\chi_j| \,. \tag{3.15}$$

This density matrix is clearly different from the left-projected one defined in Eq. 3.9 used in our embedding scheme.

The optimal ANOs are obtained as the p eigenvectors with largest eigenvalue w_i of the following linear system:

$$\sum_j D^{ij}_{\text{atom}} d^{\text{ANO}}_{kj} = w_k s_{ij} d^{\text{ANO}}_{kj} \quad \text{for } i \text{ s.t. } \mathbf{R}_i = \mathbf{R}, \text{ and } j \text{ s.t. } \mathbf{R}_j = \mathbf{R}. \tag{3.16}$$

We carried out a comparison of the performances of the two techniques for effectively reducing the basis set. For measuring the efficiency of GEOs and ANOs, we constructed both basis sets with the same dimension $N_{\text{totgeo}} \ll N_{\text{tot}}$ and then we computed the normalized overlap between the pairing function expressed within the full basis set (Φ) and the one expanded over the hybrid orbitals ($\tilde{\Phi}$): $\left\langle \tilde{\Phi} \middle| \Phi \right\rangle^2 / \langle \Phi | \Phi \rangle \left\langle \tilde{\Phi} \middle| \tilde{\Phi} \right\rangle$. We found that using GEOs yields a better overlap with respect to standard ANOs no matter the chosen basis set size, as benchmarked with the water molecule in Tab. 3.1.

3.2.4 Application to the water monomer

The purpose of this section is to apply our geminal embedding scheme to a simple, but challenging system: the single water molecule.

H_2O is ubiquitous on Earth and is a key element for life. The single water molecule is the main building block of the hydrogen-bonded network, but its intra-molecular degrees of freedom will also significantly affect the intermolecular environment due to its large dipole moment and the strong directionality of the H bond. Despite its importance, its exotic properties arising from its unique structure are not fully understood yet. It is thus not surprising that a vast amount of experimental as well as theoretical works has been published on this subject (see the introduction to this Chapter for more details).

The importance of the basis set in the description of the properties of water has been thoroughly investigated by the quantum chemistry community [108, 109]. Moreover, a recent QMC study on H_2O has appeared [71], which uses the same JAGP ansatz as ours. H_2O represents therefore a perfect laboratory for testing our embedding scheme. In addition, an optimal reduction of the number of variational parameters in the ansatz is crucial to tackle larger water clusters such as the protonated water dimer, subject of Sec. 3.3.

In this Section, we assess first the accuracy of our QMC framework in predicting the structure and energetics of water, with a particular focus on the impact of static electron

[a]Here the number of parameters is the same as the one in the JAGP wave function since in the JSD ansatz we rewrite the corresponding geminal (of rank $N/2$) on the uncontracted basis in order to optimize the MOs, as explained by Ref. [110].

Table 3.2 – VMC energies of the water molecule and number of variational parameters in the QMC wave functions. The geometry is the experimental one in the pseudopotential calculations, while it is the QMC relaxed one in all-electron calculations. The total number of parameters (last column) and the wave function quality vary depending on the contraction level of the GEOs used in the determinantal part. The Jastrow functional form has been kept fixed in all calculations. This gives a number of 195 and 418 Jastrow parameters for the pseudopotential and all-electron calculations, respectively. The other parameters are in the determinant, coming from both $\lambda^{a,b}$ (third to last column) and the basis set, i.e., $\chi^a_\mu(\mathbf{r})$ (Eq. 2.14) for the primitive Gaussian basis and $\phi^{\mathrm{GEO}}_{i,a}(\mathbf{r})$ (Eq. 3.10) for the GEOs (second to last column).

Wave function ansatz	VMC energies			number of parameters		
	Energy E_x (Ha)	Variance (Ha^2)	$E_x - E_{\mathrm{JSD}}$ (mHa)	$\lambda^{a,b}_{\mu,\nu}$	$\chi^a_\mu, \phi^{\mathrm{GEO}}_{i,a}$	total
pseudopotential calculations						
JSD: primitive GTOs	-17.24821(7)	0.2655(6)	0.0	682	18	895[a]
JAGP: (4O,1H) GEOs	-17.25013(8)	0.2635(12)	-1.91(11)	21	158	374
JAGP: (4O,5H) GEOs	-17.25183(6)	0.2510(6)	-3.62(10)	105	238	538
JAGP: (8O,2H) GEOs	-17.25267(7)	0.2426(18)	-4.46(10)	78	298	571
JAGP: (8O,5H) GEOs	-17.25302(6)	0.2412(34)	-4.89(10)	171	358	724
JAGP: primitive GTOs	-17.25389(6)	0.2296(5)	-5.68(10)	682	18	895
all-electron calculations						
JSD: primitive GTOs	-76.40025(8)	1.412(3)	0.0	1383	19	1820[a]
JAGP: (9O,2H) GEOs	-76.40504(9)	1.399(6)	-4.79(12)	91	361	870
JAGP: primitive GTOs	-76.40660(7)	1.374(3)	-6.35(11)	1383	19	1820

correlation on this system. Starting from this reference JAGP wavefunction, we construct several GEO basis sets and select the one which yields the best compromise between accuracy and computational efficiency. This basis will be then transferred to the $H_5O_2^+$ model (Sec. 3.3). A complementary investigation on the impact of the pseudopotential is also carried out in this section. Its outcome is valid also in the case of the protonated water dimer.

For all QMC calculations presented here we use a primitive Gaussian basis of $O[5s, 5p, 2d]$ and $H[4s, 2p]$ for the determinantal part, while for the Jastrow factor it is $O[3s, 2p, 1d]$ and $H[2s, 1p]$. The first guess for the determinantal basis set is obtained by performing DFT(LDA) energy minimization starting from the even tempered form, as explained in Sec. 2.2.2. Afterwards, the exponents of the localized orbitals are fully optimized with the SR method. No spin-dependent term is included in the many-body Jastrow factor by avoiding spin contamination in the homogeneous part and by setting to zero all the matrix elements $S^{a,b}_{\mu,\nu}$ in Eq. 2.35. Note that the latter basis set has been recently claimed by Ref. [71] to be one of the most accurate. If not otherwise specified, the core electrons of the oxygen atom are replaced with BFD pseudopotentials (see Sec. 2.3), whereas hydrogen is always treated all-electron.

At first, we test the energetics yielded by the JAGP wavefunction against the simpler JSD. These two functional forms are defined in Sec. 2.3.3. This comparison is crucial in order to select the best reference wavefunction from which the geminal embedded orbitals are constructed.

The energy difference is reported in Tab. 3.2. This Table highlights the size of static correlations in the system, which amounts to 5-6 mHa. This leads also to better geometrical properties, as seen in Tab. 3.3. The JAGP geometry is closer to experiment than the JSD one, in both the OH distance and the HOH angle. The structural effects of static correlations in the water molecule have been already pointed out in Ref. [71], where they were attributed mainly to a change in the local description of the oxygen atom. Thanks to the flexibility of the MOs representation (Eq. 2.29), the AGP wavefunction allows us to analyze in deeper detail the role of static correlation. If one goes beyond the SD wavefunction – for the H_2O case this is achieved by taking $N_{\rm MO} > N/2 = 4$ for pseudopotential calculations and 5 for all-electron calculations – the impact of correlation is linked to the modulus of the AGP eigenvalues $\lambda_k^{\rm MO}$, obtained by diagonalizing the full Λ matrix (Eq. 2.28). When the geminal embedded orbitals are used for representing AGP (Eq. 3.12), the physical implications are more subtle. Indeed, in this case the resulting AGP wavefunction (without the Jastrow factor) is expanded over a hybrid basis set (the GEOs) which mixes several angular momentum channels, and it is directly related to the one-body density matrix of the system. The AGP eigenvalues for the water monomer are plotted in Fig. 3.1 for different basis sets.

The Figure shows that indeed the orbitals above the HOMO (which corresponds to the highest occupied molecular orbital in the SD representation) have a sizable, i.e. $\sim 10^{-2}$, eigenvalue, with a distribution which falls abruptly to zero only after the 40-th orbital (gray area in the plot). This reflects the multi-determinant character of the water molecule, taken into account by the AGP ansatz. This proves that, although the entanglement of quantum levels at the origin of static correlations can come from the oxygen atom, its impact in water has a genuine molecular character. Last but not least, the multi-determinant AGP representation leads to a better description of the nodal surface of the true ground state, with a gain of about 2.5 mHa in the fixed node LRDMC energy with respect to the one obtained by using the JSD as trial wave function, as reported in Tab. 3.4.

It is important to remark that by using the full primitive basis the total number of JAGP parameters reaches the value of 895 (last column of Tab. 3.2). Whereas this value remains affordable with small molecular clusters, it likely leads to computationally unfeasible calculations when the dimension of the simulation model is increased. In fact, without using any basis set contraction, the number of variational parameters of the

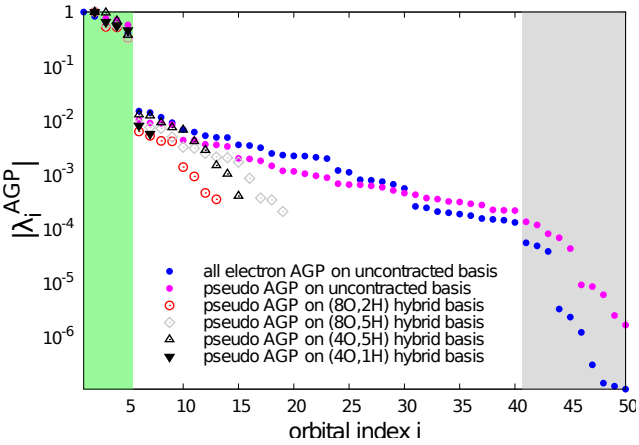


Figure 3.1 – Semilog plot of the modulus of the AGP eigenvalues versus the molecular orbital index for different basis sets and calculations. The orbital indexes include always the oxygen 1s electrons, replaced in the pseudopotential calculations. The green area represents the exactly occupied molecular orbitals in the SD representation, with $\lambda_i = 1$ for $i \in \{1, \ldots, \text{HOMO}\}$ and $\lambda_i = 0$ for $i \geq \text{LUMO}$. In the AGP representation, by diagonalizing the geminal we obtain the corresponding molecular orbitals (eigenvectors) and their occupations λ_i^{AGP} (eigenvalues). In the AGP, also the orbitals above the HOMO are occupied, with a weight which jumps across the HOMO to LUMO transition (in going from the green to white region). The gray area shows when the MO occupation tail falls rapidly to zero in the full AGP (expanded on a primitive basis set), signalling that the MOs above that threshold start to be irrelevant to describe the static correlations in the system.

Table 3.3 – Geometrical properties of the global minimum of the water molecule. We report a comparison between different QMC wave functions and experimental results[111].

	$\overline{OH}$ (Å)	$\angle HOH$ (°)
pseudo JSD	0.9542(4)	104.730
pseudo JAGP	0.9549(4)	104.549
all-electron JSD	0.9539(4)	105.187
all-electron JAGP	0.9557(4)	105.101
experiment[111]	0.95721(30)	104.522(50)

determinant is fixed by the size of the matrix $\{\lambda_{\mu,\nu}^{a,b}\}$ of Eq. 2.28. This dimension of this matrix clearly increases as the square of the primitive basis set size.

In order to alleviate this problem, we use the geminal embedding scheme to create the optimal hybrid contractions (Eq. 3.12). Since we demonstrated the reliability of JAGP to describe water properties, we use this wavefunction as the starting point for constructing the GEOs.

The results are reported in Tab. 3.2. The smallest basis set which includes the 1s for H and the 2s and 2p orbitals for O, thus taking into account the 2s2p near degeneracy

Table 3.4 – LRDMC energy results extrapolated to the zero lattice space limit. The LRDMC calculations are performed in the fixed-node approximation. In the last row, we compute the energy gain due to a better nodal description provided by the JAGP wave function with respect to the JSD one. Note that the agreement between the pseudopotential and all-electron calculations has an accuracy of the order of 0.1 mHa, despite their sharp difference in total energies.

	pseudo	all-electron
E_{JSD} (Ha)	-17.26280(6)	-76.42475(15)
E_{JAGP} (Ha)	-17.26528(6)	-76.42690(14)
$E_{\mathrm{JAGP}} - E_{\mathrm{JSD}}$ (mHa)	-2.48(9)	-2.15(21)

at the atomic O level, is the 4O 1H hybrid set (in self-explanatory notations). It gives the poorest energy and variance among the GEO basis sets considered, though, being lower than the JSD ansatz. The best energy is obtained with the largest hybrid basis tried here, namely the 8O 5H set. It recovers a large fraction of the static correlation and its energy is less than 1 mHa above the uncontracted JAGP one. However, the price to pay is that the parameters reduction is weak, the total number of parameters being close to the one of the full JAGP expansion (see last column of Tab. 3.2). Indeed, while the number of $\lambda^{a,b}_{\mu,\nu}$ is still significantly lower than the one for the uncontracted basis set, the number of parameters required to build the GEOs grows too much. The best compromise between efficiency, i.e. total number of variational parameters, and accuracy, i.e., variational energy, is provided by the 8O 2H basis, as it yields a significant gain in energy with a small/moderate number of parameters. This advantage will be remarkable for large numbers of atoms, as the number of variational parameters corresponding to the GEOs grows only *linearly* with the number of atoms (see Eq. 3.10); on the other hand, the number of parameters corresponding to $\lambda^{a,b}_{\mu,\nu}$, grows instead quadratically, but it remains still affordable since it is dramatically reduced by this approach (see Tab. 3.2).

Finally, we study how the AGP spectrum changes with the contracted hybrid basis sets. Fig. 3.1 shows that, after a complete wave function optimization, the modulus of the GEO eigenvalues covers the $10^{-2} - 10^{-4}$ range of the fully uncontracted AGP expansion, except for the smallest 4O 1H basis, which clearly spans a too restricted Hilbert space. Moreover, we checked that the JAGP expanded on the optimal 8O 2H basis gives the same fixed node LRDMC energy as the full JAGP, signalling that the nodal surface is properly described even by the hybrid 8O 2H contraction. Therefore we are going to use it in the study of the protonated water dimer presented in the next Section. The gain in efficiency of the GEOs basis set is expected to be larger and larger as the system size increases, as the quadratic growth of the number of parameters in the atomic basis set with the number of atoms is strongly dependent on the primitive basis set size.

In conclusion to this section, we assess the accuracy of the pseudopotential approximation by presenting all-electron calculations on H_2O. With respect to the pseudopotential case, the primitive and contracted basis sets for oxygen have been extended in order to account for the additional 1s electrons. The primitive basis set is then O($6s6p2d$) and O($3s3p1d$) for the determinant and Jastrow term respectively, while the optimal contracted hybrid basis set is 9O 2H. In Tab. 3.2 we report the variational energies for different wave functions. The energy gain provided by the all-electron JAGP wavefunction is very close to the one in the pseudopotential calculations. The substantial agreement between the two calculations is apparent also in Fig. 3.1, where the eigenvalues of the higher energy molecular natural orbitals in the AGP behave similarly. The LRDMC energy difference between the JAGP and JSD trial wave functions calculated using all electrons coincides within the error bars with the difference calculated with pseudopotentials, as shown in Tab. 3.4. The nodal contribution to the fixed-node energy is the same. The JAGP LRDMC energy is one of the best ever published, in statistical agreement with the value computed by Lüchow and Fink [112] ($-76.429(1)$ Ha), who used 300 determinants in the trial wave function, and with the value from Ref. [71] ($-76.42660(2)$ Ha), where the same ansatz as ours for the trial wavefunction has been used. The JAGP LRDMC projected energy is only 11 mHa higher than the extrapolated exact result of -76.438 Ha [113].

The all-electron calculations confirm the importance of including static correlations to have a better description of the geometry, as shown in Tab. 3.3. However the HOH angle turns out to be less accurate than the one obtained with pseudopotentials if compared to experiment, most probably because it is a quantity very sensitive to basis set convergence, which is harder to reach in all-electron calculations. Another drawback of all-electron calculations is of course the larger variance for an equivalent wave function ansatz, due to the 1s electron fluctuations, as one can easily evince from comparing the values presented in Tab. 3.2.

Due to the larger primitive basis set required in all-electron calculations, the parameter reduction allowed by the GEOs contraction (9O 2H) has a great impact on the efficiency. The total number of parameters is reduced by almost a factor of 4 in the determinantal part of the single molecule, without any significant loss of accuracy in the JAGP total energy (see Tab. 3.2).

The use of the BFD pseudopotential for oxygen, and the JAGP ansatz together with the optimal GEO basis set obtained for the water molecule, is transferred to the protonated water dimer studied in the next section.

3.3 Proton transfer reactions made simple: the Zundel model

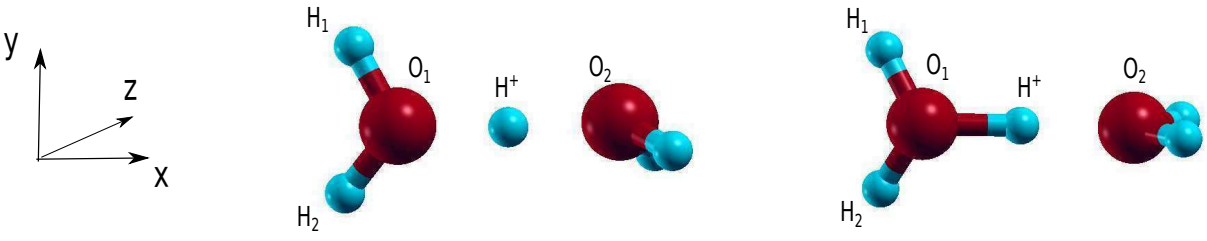
The protonated water dimer, or Zundel model, represents the simplest model for PT in aqueous systems. We study the energetics of $H_5O_2^+$ by choosing a suitable reaction coordinate (RC). Selecting a RC allows to reduce the complexity of the full-dimensional PES and project it onto a single-dimensional subset which retains the desired physical features of the full hypersurface.

As modeled by our system, the PT reaction takes place in three different steps. First, the excess proton is bound to one water molecule forming an $H_3O^+ + H_2O$ complex. By thermal fluctuations the oxygen-oxygen separation can approach the optimal distance of the Zundel complex (around 2.39 Å); at this stage the system assumes a "Zundel configuration" with the proton equally shared between the two oxygens. A further stretch of $\overline{OO}$ destabilizes the Zundel configuration and a new $H_2O + H_3O^+$ complex is produced. The overall effect of this process is a transfer of a proton along the hydrogen bond between two oxygen atoms. This mechanism suggests the choice of the $\overline{OO}$ separation as RC for the protonated dimer potential energy curve.

In the following we present the geometry and energetics of the protonated water dimer at the VMC and LRDMC levels and we compare our results with the quantum chemistry "golden standard" CCSD(T) approach. All QMC calculations are performed with the JAGP wave function ansatz with the optimal GEOs obtained for the water molecule in Sec. 3.2.4. The results presented in this Section are published in Ref. [114]

3.3.1 Properties of the symmetric global minimum

Figure 3.2 – QMC optimized geometries for global C_2 minimum (left) and for C_s local minimum (right).



The minimum energy structure of $H_5O_2^+$ has been debated in literature, as there are two candidates with competing energies: a C_2 symmetric structure, commonly known as the Zundel configuration, with the proton evenly shared between the two oxygens, and a C_s-Inv one with the proton slightly closer to one H_2O molecule (see Fig. 3.2). Several *ab initio* investigations [97, 96] have shown that a better treatment of electron correlation results in a change of the ground state (GS) geometry from the C_s to the C_2 configuration. Accurate highly-correlated studies [115, 116, 96, 100] have eventually confirmed that the global minimum is C_2-symmetric.

At a QMC level of theory, the $H_5O_2^+$ GS has a C_2 Zundel configuration with centrosymmetric excess proton (left-hand side of Fig. 3.2), in agreement with previous studies. The main geometrical parameters of the global minimum are presented in Tab. 3.5, for both the pseudopotential and all-electron calculations.

Table 3.5 – Geometrical properties (distances in Å, angles $\angle$ and dihedral angles ω are in °) of the C_2-symmetry minimum of protonated water dimer, comparison between different computational methods. See Fig. 3.2 (left-hand side) for atomic notation.

	DFT(PBE)	DFT(B3LYP) [115]	QMC - with pseudo	QMC - all-electron	CCSD(T) [100]
$\overline{O_1O_2}$	2.4111		2.3847(5)	2.3905(4)	2.3864
$\overline{O_1H^+}$	1.2074	1.2172	1.1942(8)	1.1989(5)	1.1950
$\overline{H^+O_2}$	1.2074		1.1930(5)	1.1944(6)	1.1950
$\angle O_1H^+O_2$	173.661	173.6	174.71(7)	174.43(9)	173.730
$\overline{O_1H_1}$	0.9697	0.9784	0.9605(8)	0.9630(7)	0.9686
$\overline{O_1H_2}$	0.9691	0.9778	0.9650(8)	0.9628(6)	0.9682
$\angle H_1O_1H_2$	109.161		109.16(9)	109.40(7)	108.8
ω-$H_1O_1H^+O_2$	295.690		293.5	296.6	295.3
ω-$H_2O_1H^+O_2$	163.809		158.5	162.0	163.6

Table 3.6 – Geometrical properties (distances in Å, angles in °) of the C_s-Inv minimum of protonated water dimer. See Fig. 3.2 (right-hand side) for atomic notation.

	DFT - B3LYP[115]	QMC - with pseudo	QMC - all-electron	CCSD(T)[100]
$\overline{O_1O_2}$		2.3996(6)	2.3913(3)	2.3989
$\overline{O_1H^+}$		1.1154(8)	1.1285(5)	1.1233
$\overline{H^+O_2}$	1.2507	1.2852(4)	1.2648(4)	1.2720
$\angle O_1H^+O_2$	175.4	176.5(1)	175.29(6)	175.646
$\overline{O_1H_1}$	0.9746	0.9641(7)	0.9635(4)	0.9641
$\overline{O_1H_2}$	0.9741	0.9625(4)	0.9616(5)	0.9645
$\angle H_1O_1H_2$		110.56(8)	110.66(6)	110.153

When compared with CCSD(T), the QMC ground state geometries agree within an atomic separation of 0.005 Å. The agreement in the intra-molecular angles is within 0.3°,

while the largest discrepancy between QMC and CCSD(T) is in the dihedral angles ω between two H_2O planes, which are related to soft vibrational modes and therefore affected by a larger statistical bias during the relaxation. Nevertheless, these differences do not affect the overall description of the GS and the energetics of the system.

In the present study, also the C_s-Inv structure has been taken into account (right-hand side of Fig. 3.2). Tab. 3.6, which reports the VMC optimized C_s-geometries, confirms the trend seen in Tab. 3.5 for the GS, although the discrepancies in the bond lengths are slightly larger between different methods, the CCSD(T) values being in between the all-electron and pseudopotential VMC results.

In the pseudopotential calculations, the energy difference between the C_s configuration and the C_2-symmetric global minimum turns out to be 0.25(8) Kcal/mol at the VMC level, and 0.23(8) Kcal/mol at the LRDMC level. This is in satisfactory agreement with previous results carried out with Møller-Plesset (MP2) perturbation theory [116] (0.28 Kcal/mol) and CCSD(T) [100] (0.46 Kcal/mol) techniques.

3.3.2 Stretching the $\overline{OO}$ distance

From the C_2-sym global minimum we stretch the $\overline{OO}$ distance in order to study the potential energy curve and elucidate the PT character in the dimer. We verified that an analogous stretch from the C_s local minimum yields a higher energy over the entire $\overline{OO}$ range and thus we disregard this possibility.

The structural relaxation at the VMC level for fixed $\overline{OO}$ separation requires a careful procedure due to the flatness of the PES. Starting from a DFT(PBE) optimized geometry and a JAGP variational wave function fully optimized in the electronic part, the atomic coordinates are relaxed with the stochastic reconfiguration method employing the AAD technique for ionic forces evaluation (see Sec. 2.5 for technical details).

For the sake of comparison, we minimize a parameterized full-dimensional PES fitted from CCSD(T) calculations [100] to find the best coupled cluster estimates of energy and geometry. By means of the *downhill simplex* minimization technique, we find the configuration of lowest CCSD(T) energy at the same constrained $\overline{OO}$ distance as the corresponding QMC and DFT calculations.

In Fig. 3.3 we plot the energy landscape along the RC for DFT(PBE), CCSD(T), VMC and LRDMC, the latter computed at the VMC geometry. For the VMC technique, we also report all-electron calculations for some $\overline{OO}$ separations. We notice a good agreement among all techniques in the region at the left of the global minimum of the curve, except for a constant shift by 0.02 Å between the DFT(PBE) results and the others. The PBE $\overline{OO}$ minimum is indeed located at 2.41 Å, while the minimum of the other methods turns

Figure 3.3 – Potential energy curve (Kcal/mol) of the protonated water dimer projected on the $\overline{OO}$ distance. Comparison between different computational methods. Structural relaxation is performed at each level of theory, except for the LRDMC landscape which uses VMC relaxed structure. Each curve has its minimum as reference point.

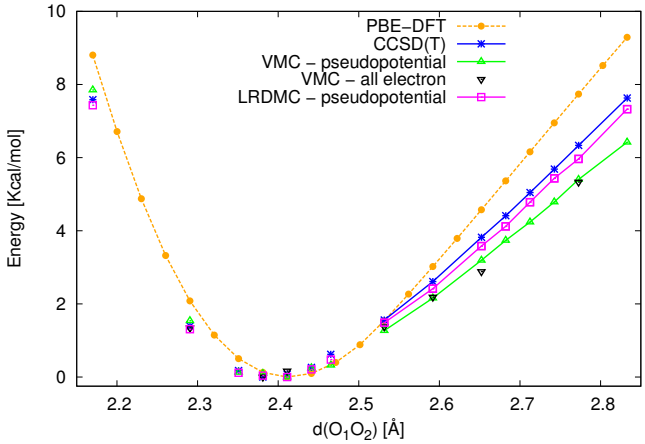
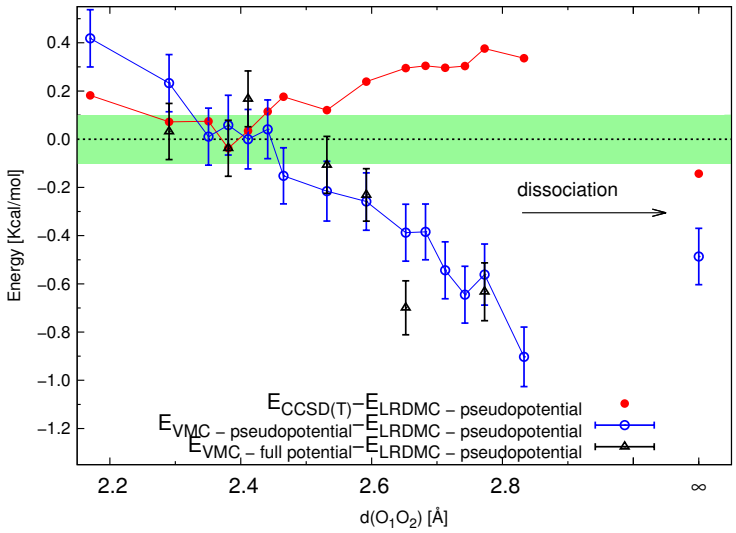


Figure 3.4 – CCSD(T) and VMC energy differences (Kcal/mol) as a function of the $\overline{OO}$ distance (Å) reported in Fig. 3.3 plotted with respect to the LRDMC energies differences (the zero of the y-axis). Full potential VMC results are reported for a few points along with dissociation energies. The green area represents the statistical error bar achieved in a typical LRDMC run, i.e. $\sim$ 0.1 Kcal/mol.



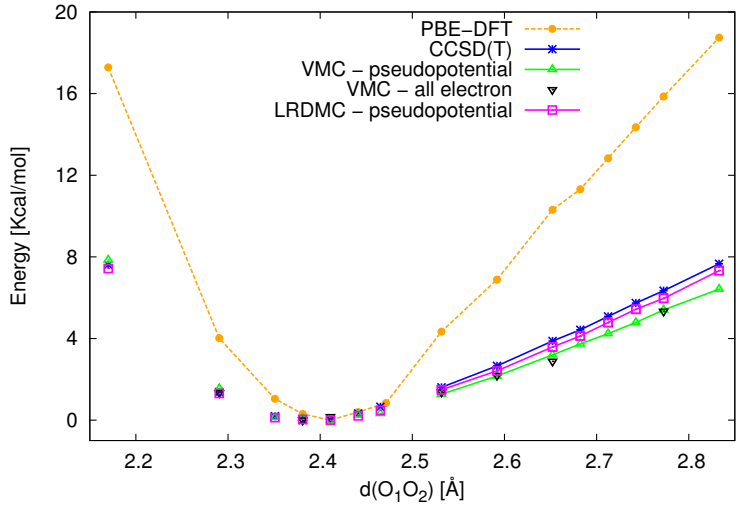
out to be at $\sim$ 2.39 Å. Hence this part of the energy curve is only slightly influenced by a better treatment of correlations.

On the other hand, the region at the right of the minimum, at intermediate $\overline{OO}$ distances ($\geq$ 2.55 Å), displays a different behavior. The DFT(PBE) calculations overestimate the slope of the curve with respect to the most accurate techniques. LRDMC, which yields the best QMC correlation energy, shows a remarkably good agreement with

the state-of-the-art CCSD(T) results. In particular all CCSD(T) values are in the range of ~ 0.3 Kcal/mol, three times the statistical error of the LRDMC calculations, as shown in Fig. 3.4. This Figure also shows that pseudopotential and all-electron VMC calculations are in statistical agreement for the reported $\overline{OO}$ separations. This demonstrates the quality of the BFD pseudopotentials in the Zundel ion, verified also in Tab. 3.5 and Tab. 3.6 for the geometry of the C_2-sym and C_s-Inv minima, respectively (see Sec. 3.3.1).

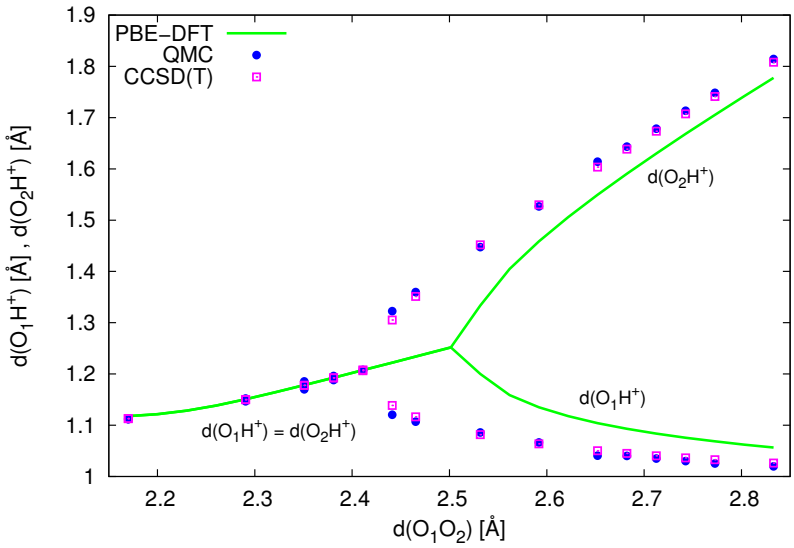
As mentioned before, the curves in Fig. 3.3 are obtained with the minimum energy geometry at each level of theory (except for the LRDMC curve, whose geometry is set at the VMC level). In order to have a more reliable comparison and avoid the bias from the use of different geometries, we carried out the same calculations employing the VMC-optimized structures for every technique. The result is reported in Fig. 3.5. The trend displayed in Fig. 3.3 for method-optimized geometries is remarkably enhanced when the same configuration of the dimer is considered. Away from the minimum, the DFT(PBE) energies show a larger overestimation of the slope.

Figure 3.5 – Protonated water dimer energy landscape (Kcal/mol) as a function of the $\overline{OO}$ distance (Å). All the calculations are performed at the VMC-optimized geometry. The zero energy reference point corresponds to the minimum of each curve.



The slope of the dimer potential energy curve is related to the behavior of the excess proton in the system. In Fig. 3.6 this property is elucidated. We report the separations between the excess proton and each of the two oxygens in the same plot: $\overline{O_1H^+}$ and $\overline{H^+O_2}$. The structure of the dimer is relaxed at different levels of theory, with the same procedure used to obtain the potential energy curve in Fig. 3.3. The plot clearly shows the appearance of two distinct regimes of the dimer. One is characterized by a symmetric Zundel configuration with the proton evenly bonded to the two oxygens; from the point of view of proton transfer physics, it is basically equivalent to the GS configuration.

Figure 3.6 – Separations (Å) between the two oxygens and the excess proton as a function of the reaction coordinate for different computational methods. The QMC results are obtained by minimizing the VMC ionic forces.



Stretching the $\overline{OO}$ distance results in the formation of a $H_3O^+ + H_2O$ complex with the proton localized on one water molecule. These configurations belong to the *asymmetric regime* of the dimer. Within this regime, the initial C_2 point symmetry of the GS geometry is broken due to proton localization.

The distances obtained by the VMC relaxation of the atomic coordinates are in excellent agreement with the CCSD(T) calculations; in particular the root-mean-square distance between the two data sets over the whole $\overline{OO}$ range is ~ 0.007 Å for both $\overline{O_1H^+}$ and $\overline{O_2H^+}$. Electron correlation plays a key role in determining the stability of the symmetric configuration of the dimer. We define d_c as the critical $\overline{OO}$ distance at which the symmetric configuration of the excess proton is broken. The overestimation of the PES slope by DFT(PBE) in Fig. 3.3 corresponds to an overestimation of d_c by $\sim +0.13$ Å with respect to highly correlated *ab initio* methods beyond DFT. Furthermore DFT(PBE) poorly describes the geometry in the broken-symmetry region close to d_c, where the discrepancy in the $\overline{OH^+}$ distances is the largest (going up to 0.15 Å).

3.3.3 Implications for more realistic PT models

The zero-temperature potential energy curves reported in Fig. 3.3 and Fig. 3.5 seem to conflict with the proton transfer mechanism discussed in the introduction of this Section, since the configuration with centrosymmetric proton is energetically favored and therefore it represents a stable state rather than a transition state between two asymmetric

configurations with localized proton.

However, this contradiction can be easily explained. Indeed, it has been shown in Ref. [91] that the introduction of thermal and polarization effects due to the environment, favors the asymmetric regime of the complex. At finite temperature the *free energy* landscape displays a global minimum shifted towards the asymmetric regime and the Zundel-like structure does not represent the energetically favoured configuration any more. Moreover, a recent experimental result showed that the average $\overline{OO}$ distance in liquid water is 2.81 Å [117] which clearly corresponds to a symmetry-broken configuration of the dimer (see Fig. 3.6). In any case, in order to jump from a water molecule to one of its neighbors, the proton must pass through a Zundel configuration. Correctly describing the energetics and geometry of the protonated water dimer in the symmetry-breaking transition region is therefore of paramount importance to get an accurate description of PT in more realistic models.

As mentioned before, Fig. 3.6 highlights that at 0 K a better treatment of electron correlation yields the asymmetric regime stable over a considerably wider range of $\overline{OO}$ distances with respect to DFT results. As a consequence, we expect the calculated proton diffusion in water to be very sensitive to the level of theory [91, 88, 87].

3.3.4 Properties of the symmetry-broken configurations

A quantity which has been extensively studied over the past years [118, 119, 87] is the *static proton transfer barrier*, i.e. the barrier that the H^+ has to overcome in order to jump from one H_2O molecule to the other at a fixed $\overline{OO}$ distance in the symmetry-broken regime. This quantity does not provide a realistic comparison with the experimental activation barrier for PT, as the $\overline{OO}$ distance will shorten during the proton hopping. Nevertheless, it is relevant in order to provide a further check of the accuracy of our QMC approach. Fixing the $\overline{OO}$ distance, the barriers are obtained as energy differences between the asymmetric configuration with localized proton and a structure with the excess proton at equal distance from the two oxygens. Calculations are performed in three representative $\overline{OO}$ separations; the results are shown in Tab. 3.7, where they are compared with existing data in literature.

The first $\overline{OO}$ distance is 2.47 Å, very close to d_c displayed by highly correlated approaches (2.43 Å). In contrast, at a DFT level the configuration with centrosymmetric proton is still energetically favored, as shown by Fig. 3.6. It has been noted [96] that the PES in this region is particularly flat. Furthermore the potential energy curve of the dimer develops on very tiny energy differences around d_c. These issues make the calculations in this region of the PES extremely delicate since the stochastic noise can considerably

affect the quality of the QMC predictions.

Tab. 3.7 shows that the QMC and CCSD(T) results display a vanishing energy barrier of the order of $0.2-0.3$ Kcal/mol at a $\overline{OO}$ distance close to d_c. The height of this barrier is slightly above the attained statistical error in our typical QMC run. However, despite the very sensitive behavior of the dimer PES around d_c, the accuracy of our force minimization algorithm allows us to account for the tiny energy differences involved. Thus it ensures the necessary precision to describe the PT physics in the dimer. As we discuss in Sec. 3.3.5, a similar accuracy can be achieved in larger molecular clusters with a reasonable amount of computational time. Therefore our QMC framework guarantees a reliable description of the PT physics also for more realistic models.

Table 3.7 – Static proton transfer barriers (Kcal/mol) at fixed $\overline{OO}$ separations. Comparison between different level of theory at different geometries. If not specified, the structure is optimized with the same method as the corresponding energy calculation.

Method		$O-O$ distance		
Energy	Geometry	2.47 Å	2.6 Å	2.7 Å
DFT(PBE)		-0.74	0.19	1.94
VMC		0.28(6)	2.99(8)	5.99(8)
LRDMC	VMC	0.37(8)	2.64(7)	5.57(7)
CCSD(T)		0.22	2.37	5.24
CCSD(T)	VMC	0.28	2.29	5.32
CCSD(T) [119]	MP2		2.08	4.85
QCISD(T) [119]	MP2		2.06	4.82
MS-EVB [87]			2.05	5.11
MP2 [119]			1.77	4.39

The other results in Tab. 3.7 are obtained at larger oxygen separations, further away from d_c. They confirm the general behavior already seen along the $\overline{OO}$ reaction coordinate. The LRDMC and CCSD(T) results are in a agreement to within 0.3 Kcal/mol, whereas the VMC overestimates the barrier of about 0.6 Kcal/mol. As is known from previous work, DFT substantially underestimates the barrier with respect to post-HF methods, which provide a better treatment of correlation.

Finally, let us analyze the extreme limit of the asymmetric Zundel configuration, namely when $\overline{OO} \to \infty$, with the formation of one H_2O and one hydronium. The dissociation energy D_e of $H_5O_2^+$ is computed by setting the distance between the two oxygens to 14 Å. With the CCSD(T) PES, we checked that this is already in the large distance plateau. We get a D_e of 33.02(9) Kcal/mol with VMC, and 32.54(8) Kcal/mol with LRDMC, to be compared with the CCSD(T) value of 32.68 Kcal/mol computed in Ref. [100],

while DFT(PBE) gives $D_e = 29.55$ Kcal/mol. The agreement between the LRDMC and CCSD(T) is impressive, while it is good for the VMC estimate. This is mainly due to the size consistency of the JAGP ansatz, obtained once the Jastrow factor is close to the complete basis set limit [120]. The VMC and CCSD(T) dissociation energies are plotted in Fig. 3.4 with respect to the LRDMC values.

3.3.5 Perspectives on realistic simulations

QMC methods present a favorable scalability with the number of particles with respect to other highly correlated approaches such as CC. With the aim of demonstrating this feature within our framework, we performed a benchmark calculation on a more realistic PT model composed of 6 water molecules and one excess proton, which will be the subject of a further study. In Tab. 3.8 we report a comparison of the computational time required to carry out typical VMC and LRDMC runs for different sizes of the protonated water cluster. These calculations are performed at fixed optimized variational parameters and geometry (single-point calculations).

The VMC structural relaxations are more expensive than single-point runs since they involve variational parameter optimization and evaluation of the atomic forces. Within our calculations, the computational cost of a typical structural relaxation run in order to obtain well-converged parameters and forces is about 5 times larger than the VMC single-point calculations presented in Tab. 3.8. Note that this cost strictly depends on the quality of the starting wave function and geometry of the system.

Table 3.8 – Total computational wall time in hours of typical VMC and LRDMC runs at fixed optimized variational parameters and geometry. Calculations have been performed with the program TurboRVB [1] on 32 thin nodes of the Curie HCP machine (2.7 GHz, 16 core Intel Sandy Bridge processors). The target statistical error reached is of 0.06 Kcal/mol in total energies. A single water molecule, a protonated dimer and a larger cluster of 6 water molecules with one excess proton are compared. The LRDMC is carried out at a lattice space of 0.125 a_0. CCSD(T) computational times have been evaluated by assuming a perfect N^7 scaling with the number of particles and no parallelization overhead.

	Total wall time (h) on 512 CPUs		
# of water molecules	VMC	LRDMC	CCSD(T)
1	0.05	0.15	0.0001
2	0.24	2.13	0.01
6	6.49	164.35	21.87

All calculations in Tab. 3.8 have been carried out on the HPC Curie thin nodes (2.7 GHz 16 core Intel Sandy Bridge processors), and the QMC target statistical error has been set to 0.06 Kcal/mol for the total energy. By means of a simple polynomial fit to the data in

Tab. 3.8, we see that the LRDMC method, carried out at a lattice space of $a = 0.125a_0$, displays an almost perfect N^4 scaling. The simpler variational Monte Carlo technique shows an N^3 scaling. As the variance scales linearly with N (see Eq. 2.3), and the QMC all-electron move costs N^3, the resulting theoretical scaling is N^4. An apparent scaling faster than N^4 is probably due to the BLAS-based implementation [28] of the program TurboRVB, which becomes more efficient as the system size increases, by reducing the N^4 prefactor the most. It turns out that LRDMC calculations are feasible in a reasonable computational time for the 6 H_2O cluster, while VMC calculations are still cheap at that cluster size.

In order to contextualize the values reported in Tab. 3.8, we performed a benchmark CCSD(T) calculation of the small Zundel ion on the same machine. We use Dunning's correlation-consistent *aug-cc-pVTZ* basis set, the same employed in Ref. [100] for fitting the Zundel PES. Calculations have been done with the Gamess [121] program. A single-point run on 128 CPUs costs 2.25 minutes, approximately one order of magnitude less than the corresponding QMC calculation, since at this cluster size the large QMC prefactor dominates over its favorable scaling.

According to the theoretical N^7 scaling of CCSD(T), an analogous calculation on the 6-molecule cluster would have a slightly larger cost with respect to the corresponding VMC calculation (see Tab. 3.8). Therefore, the 6-molecule complex represents a crossing point in the relative efficiency of the VMC and CCSD(T) methods; it is obvious that a further increase in the cluster size would make the QMC approach considerably favoured in terms of computational demand. Moreover, the coupled-cluster estimated cost has been obtained by assuming no parallelization overhead and no constraints of memory allocation, criteria which are hard to meet in coupled-cluster calculations. On the other hand, quantum Monte Carlo has less demanding memory requirements and an almost perfect parallel scaling with the number of cores.

3.4 Conclusions

In this Chapter we have presented extensive applications of QMC calculations with the JAGP ansatz to the description of water and its PT-related phenomenon.

At first, we introduced the geminal embedded orbitals. They are optimally constructed hybrid orbitals which allow to be drastically reduced the number of variational parameters within the JAGP wavefunction without appreciable loss of accuracy. This is demonstrated by constructing the optimal GEOs basis set for the single H_2O molecule and comparing results obtained using the GEO basis with previously published QMC calculations.

The GEO basis is directly portable to the Zundel ion, a simple model for proton trans-

fer reactions, consisting of two H_2O moleculs and an excess proton. We study its energy landscape as a function of the oxygen-oxygen distance with the VMC and LRDMC methods. LRDMC is shown to be in excellent agreement with the quantum chemistry "golden standard" CCSD(T) (within 0.3 kcal/mol), whereas minor error (up to about 1 Kcal/mol) are shown by VMC. On the contrary, DFT(PBE) strongly overestimates the slope of the energy landscape away from the minimum. In contrast with previous works, where the geometry was either taken from experiment or drawn from force field calculations, we are able to optimize both the electronic and structural parts at the VMC level thanks to the algorithmic differentiation technique explained in Chap. 2.

VMC structural relaxation provides geometries remarkably close to the ones obtained from a CCSD(T) fitted PES. We show the presence of two distinct regimes of the dimer depending on the $\overline{OO}$ distance: one with a centrosymmetric excess proton and the other with the proton localized on one of the water molecules. The stability of these configurations crucially depends on the level of theory: a better treatment of electron correlation results in the stability of the asymmetric proton geometry over a wider range of $\overline{OO}$ distances.

These results, together with the proton transfer static barrier and the dissociation energy, show that our QMC approach has a global accuracy comparable with the most advanced quantum chemistry methods in both geometry and energetics of the dimer. This finding, combined with a better scaling with the system size demonstrated on a larger water cluster, makes QMC a promising candidate to obtain a fully *ab initio* description of the microscopic mechanism of proton transfer in water.

This Chapter concludes the first part of this thesis, devoted to the theory as well as applications of QMC in the domain of electronic structure of molecular systems. We will show in the next part of this manuscript that, at variance with its direct competitors, *ab initio* QMC can be extended to the study of crystalline solids keeping the same precision and scalability and thus opening the way to study a vast amount of new systems and phenomena with an unprecedent level of accuracy.

Chapter 4

QMC simulations of extended systems

4.1 Introduction

In this Chapter we present an overview on the theoretical and technical aspects necessary to extend the continuum QMC framework to investigate the electronic structure of solids. This task has taken a significant part of this thesis work.

The main outcome of the previous Chapters is that QMC techniques, among correlated methods, are capable of reaching a precision down to $\sim$ 0.25 Kcal/mol $\approx$ 0.01 eV/atom whilst keeping at the same time an affordable computational cost. Coupled cluster methods (CC), many-body perturbation theory (MBPT) and quadratic configuration interaction[a] (QCI) are highly-correlated alternatives which possess the crucial property of being size-consistent and have been extended to tackle solids [123]. However, as we will discuss more in detail, the explicit inclusion of electron correlation requires to perform calculations at several sizes and then extrapolate the final result to the limit of the infinite solid. On one hand, extrapolation generally requires to simulate a larger number of atoms with respect to open systems. Furthermore, it is well known that the infinite-size limit can yield divergent energies in perturbation theories such as MP2, MBPT or CC methods [124, 125]. Only very recently, this problem has been overcome by combining CC with the random-phase approximation [126] or by extending a variant of the method called the "equation of motion" CC approach to periodic systems [127]. However, so far, applications of quantum chemical methods to solids have been limited to the model level

[a]Quadratic configuration interaction [122] is a suitable modification of the CI algorithm which introduces additional quadratic terms in the determinant coefficients. These new terms correct for size-consistency errors in the total energy for single and double excitations CI methods.

or to very simple compounds [128]. QMC does not suffer from these limitations and it has a favorable scaling with the number of particles. Therefore, nowadays it represents one of the few available choices for performing realistic *ab initio* investigations of solids.

There is a vast amount of physical situations where the mainstream electronic structure methods based on effective independent electron theories fail.

The first and most obvious scenario where explicit introduction of electron correlation is beneficial is the study of solids presenting strong electron correlation such as rare earth materials and Mott insulators. Electron correlation arising from localized d and f orbitals cannot be, in general, well accounted either by plain DFT or by hybrid DFT; the outcome is often functional dependent and the DFT+U method, besides sacrificing a truly first principle approach by introducing the screened Coulomb parameter U, has been proven reliable in particular cases only. QMC offers an explicit treatment of electron correlation, a purely first principle approach and the possibility to directly access the many-body wavefunction of the whole crystal. Successful applications of QMC range from rare-earth compounds [103], to transition metal oxides [129, 130] and, more recently, to the vast domain of high-temperature superconductivity (see for instance Ref. [131]). For more details on the last category we refer the reader to the introduction of Chap. 6 where we present an extensive QMC investigation of the high-temperature superconductor FeSe.

As in the case of open systems, solids dominated by weak non-covalent forces such as Van Der Waals interactions are not well accounted by DFT. Again, explicit inclusion of dispersion effects in the DFT Hamiltonian has been proposed, but it often requires phenomenological input parameters [132]. QMC can describe, with a unified *ab initio* framework, both covalent and non-covalent bindings in an accurate way as demonstrated by benchmark calculations [40] as well as applications to real systems such as solid neon [133], glass polymorphs [134] and layered materials [135].

These features opened the way for investigating a variety of low-energy phenomena within QMC. For instance, phase transitions on several kind of systems have been addressed [103, 136] and also the prediction of binding properties and diffusion coefficients of molecules on different substrates represents a relevant application [137, 138], which can be potentially applied to commercially important fields such as battery development.

Although a fully first principles study of solids by QMC is still at an early stage, it is of paramount importance to develop reliable techniques and algorithms for extended systems since in the near future large-scale QMC simulations will be feasible and this will allow to deeply investigate emergent quantum phenomena with an unprecedent level of accuracy. In the introduction to Chap. 6 we will discuss one of these open possibilities: a coherent *ab initio* description of the microscopic pairing mechanism responsible for high-temperature superconductivity.

4.1.1 Organization of the Chapter

The main difficulty arising from the extension of the theoretical framework developed in Chap. 2 to solids is the simulation of the infinite periodic structure of a crystal with a finite and very small number of electrons and atoms. The structure of this Chapter closely follows the one of Chap. 2 as well as the theoretical and programming efforts spent during this thesis work for this purpose.

The first tool which has been modified is the built-in DFT code. In Sec. 4.2 we briefly revise the Bloch theorem, the cornerstone of mean-field band structure theory. We will see that the application of the Bloch theorem prescribes several important modifications to the DFT algorithms ranging from a new single-particle basis set involving complex-valued orbitals to the implementation of efficient schemes for sampling the Brillouin zone.

In Sec. 4.3 we show how to generalize the Bloch theorem to QMC calculations and how solids must be described within a correlated approach. This involves the application of a phase factor, or twist condition to the many-body wavefunction. In Sec. 4.4 we explain how to extend the VMC and DMC methods to the case of periodic solids and thus complex-valued trial wavefunctions. Finally, in Sec. 4.5, we provide a detailed overview on the SR method for optimizing the wavefunction of periodic systems and we present an original scheme allowing to efficiently optimize the ansatz using several twist conditions at the same time. We draw the conclusions to this Chapter in Sec. 4.6.

4.2 Localized basis DFT calculations for periodic systems

4.2.1 Primitive cell calculations

A typical three dimensional crystal is composed by $\sim 10^{23}$ atoms. If one assumes that this crystal has no defects this macroscopic structure can be fully described by the periodic repetition along the axes of a single building block, which is usually called unit cell. Since all unit cells are equivalent, they form a lattice in the real space – the Bravais lattice – which is fully specified by a set of basis vectors $\{\mathbf{A}_i\}$ with $i \in [1,3]$. A unit cell constructed to contain only one lattice point is called *primitive cell* (PC). Given the PC and the basis vectors, all lattice points can be written as: $\mathbf{L}_m = m_1\mathbf{A}_1 + m_2\mathbf{A}_2 + m_3\mathbf{A}_3$ where $\{m_i\}$ are integers. Notice that there is an infinit number of equivalent choices of primitive vectors generating the same lattice. Therefore the crystal structure of a solid is fully determined by specifying both the basis vectors and the Bravais lattice.

We focus here on the DFT description of a solid, but it is important to bear in mind that

the following discussion is valid for any effectively independent electron theory. Consistent with the idea of an infinite solid, the crystal is usually modeled by a single PC with N electrons. The PC is subjected to periodic boundary conditions (PBC) to mimic the repetition along the Cartesian directions.

At variance with the open system case, the electronic density $\rho(\mathbf{r})$ is now a periodic quantity satisfying the relation: $\rho(\mathbf{r}+\mathbf{L}_m) = \rho(\mathbf{r})$ for all lattice vectors. Therefore the KS Hamiltonian introduced in Eq. 2.10 must have the following invariance property:

$$\begin{aligned} \hat{H}^{\mathrm{KS}}(\mathbf{r}+\mathbf{L}_m) &= \\ H^{1b}(\mathbf{r}+\mathbf{L}_m) &+ V_{\mathrm{Ha}}[\rho(\mathbf{r}+\mathbf{L}_m)] + V_{\mathrm{XC}}[\rho(\mathbf{r}+\mathbf{L}_m)] = \hat{H}^{\mathrm{KS}}(\mathbf{r}) \end{aligned} \tag{4.1}$$

where we use the fact that the external electron-ion potential is obviously periodic due to the repetition of the same atomic arrangement.

4.2.2 The Bloch theorem

From Eq. 4.1, we notice that the Hamiltonian commutes with the single electron translational operator $\hat{T}_{\mathbf{L}}$ of any primitive lattice vector, hence: $[\hat{H}_{\mathrm{KS}}, \hat{T}_{\mathbf{L}}] = 0$. The two operators possess the same eigenfunctions. It is a textbook demonstration that this leads to the following condition on the whole set of KS orbitals $\{\phi_i\}$:

$$\phi_i^{\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}} u_{\mathbf{k}}(\mathbf{r}) \quad \forall\, i \in [1, N_{\mathrm{orb}}], \tag{4.2}$$

where $\mathbf{k}$ is a wavevector in units of $2\pi/\mathbf{L}$ usually called $\mathbf{k}$-point and $u^{\mathbf{k}}(\mathbf{r})$ is a periodic function such that $u^{\mathbf{k}}(\mathbf{r}+\mathbf{L}) = u^{\mathbf{k}}(\mathbf{r})$ for any lattice vector. This relation must hold for *any* single-particle orbitals describing a periodic solid. If PBC are applied to the primitive cell, the inequivalent $\mathbf{k}$-points can be chosen within the first Brillouin zone (1BZ) of the reciprocal lattice corresponding to $\{\mathbf{L}_m\}$.

From Eq. 4.2 it derives that the orbitals change only in a phase factor when an electron is displaced of a lattice vector:

$$\phi_i^{\mathbf{k}}(\mathbf{r}+\mathbf{L}) = e^{i\mathbf{k}\cdot\mathbf{L}} \phi_i^{\mathbf{k}}(\mathbf{r}) \quad \forall\, i \in [1, N_{\mathrm{orb}}]. \tag{4.3}$$

The wavevector $\mathbf{k}$, expressed in $2\pi/\mathbf{L}$ units, can be interpreted as a generic boundary condition applied to the single particle wavefunction when it wraps around cell boundaries. $\mathbf{k} = (0,0,0)$ is equivalent to simple PBC, while a $\mathbf{k} = (\frac{1}{2}\mathbf{B}_1, \frac{1}{2}\mathbf{B}_2, \frac{1}{2}\mathbf{B}_3)$ corresponds to fully anti-periodic boundary conditions, where $\mathbf{B}_1, \mathbf{B}_2, \mathbf{B}_3$ are the reciprocal lattice vectors. Eq. 4.2 and 4.3 are two equivalent formulations of the celebrated Bloch theorem.

The independent electron description of solids is tremendously simplified by this result. In fact, the Bloch theorem implies that a $u_{\mathbf{k}}(\mathbf{r})$ defined in the primitive cell is an eigenstate of all translational operators on the Bravais lattice simultaneously.

In principle, in order to extract intrinsic properties of the periodic system such as the total energy or generic susceptibilities, one has to sum over all the single-particle orbitals describing the infinite number of electrons in the system. Thanks to the Bloch theorem, this formidable task reduces to solving the KS problem within a single PC for a given set of wavevectors $\mathbf{k}$ lying in the 1BZ of the system. Therefore, for any periodic observable f defined on the crystal lattice, its value for the infinite system, i.e. its thermodynamic limit, is given by the following integral:

$$f_\infty(\mathbf{r}) = \frac{\Omega}{(2\pi)^3} \int_{\Omega_{\mathrm{BZ}}} d\mathbf{k} \sum_{i \in \mathrm{occ}} \left\langle \phi_i^{\mathbf{k}}(\mathbf{r}) \right| f_i(\mathbf{r}) \left| \phi_i^{\mathbf{k}}(\mathbf{r}) \right\rangle , \tag{4.4}$$

where the index n goes over occupied KS orbitals, Ω is the PC volume and $\Omega_{\mathrm{BZ}} = \frac{(2\pi)^3}{\Omega}$ the 1BZ volume. Eq. 4.4 gives in principle a way to compute *exact* properties of the infinite solid within an independent electron framework. However, in practical implementations the integral must be approximated by an appropriate summation, as we shall see later on in this section.

Notice that the Bloch wavefunctions defined in Eq. 4.2 are complex-valued. The basis set over which they are expanded (see Eq. 2.13) must cope with this requirement. In the following we present the technical aspects of the basis set for periodic systems implemented in the *TurboRVB* code.

4.2.3 Gaussian basis set for periodic systems

The functional form of the localized orbitals used in both DFT and QMC calculations for open systems has been discussed in detail in Sec. 2.2.2. The generalization to periodic systems can be achieved following different routes presented in this Section.

Let us consider, for the sake of simplicity, a simulation box of lengths (L_x, L_y, L_z) subjected to periodic boundary conditions. We denote $\mathbf{L}_m = m_1\mathbf{A}_1 + m_2\mathbf{A}_2 + m_3\mathbf{A}_3$ the lattice vectors associated with this simulation cell which can be either the primitive cell or a larger box constructed using several primitive cells. The following discussion is independent of this choice.

We require the orbitals to fulfill the single particle Bloch theorem (Eq. 4.2). In this way, the KS eigenvectors expanded over this basis possess the same property and they are hence suitable to describe the infinite solid. As already mentioned, in the case of a generic wavevector $\mathbf{k}^\sigma$ they must also be complex-valued. In the following, we explicitly keep the dependence on the electron spin σ since we will leave the freedom to assign different boundary conditions to $\uparrow$ / $\downarrow$ spin electrons.

We tested two possible choices for constructing these orbitals. The first one directly plugs a phase factor in front of the localized functions introduced in Eq. 2.14 for open

systems. A suitable coordinate transformation $\hat{T}$ is applied to the electron-ion distances as follows:

$$X_j(\mathbf{r}, \mathbf{k}^\sigma) = e^{i\mathbf{k}^\sigma \cdot \mathbf{r}} \chi_j(\hat{T}(\mathbf{r} - \mathbf{R}_a)) \quad \forall\ j \in [1, N_{\text{tot}}]. \tag{4.5}$$

The operator $\hat{T}$ makes the electron-ion distances in the argument of the localized orbitals periodic functions of the simulation box. We choose this transformation in the following way [139, 136]:

$$\hat{T}(\mathbf{r} - \mathbf{R}_a) = \Bigg[\left(\frac{L_x}{\pi} \sin\left(\frac{\pi}{L_x}(r_x - L_x) \right) \right)^2 + \\ + \left(\frac{L_y}{\pi} \sin\left(\frac{\pi}{L_y}(r_y - L_y) \right) \right)^2 + \left(\frac{L_z}{\pi} \sin\left(\frac{\pi}{L_z}(r_z - L_z) \right) \right)^2 \Bigg]^{1/2}. \tag{4.6}$$

This kind of transformation is very efficient since the derivatives of the periodic orbitals can be evaluated *analytically* by simply applying the elementary chain rule. Initially, it has been introduced [140] to avoid discontinuities in the wavefunction derivatives when a particle crosses the boundaries between periodic images of the simulation box. It is well suited for our purposes since it is apparent that the way Eq. 4.5 is constructed naturally satisfies the Bloch theorem for single particle orbitals (Eq. 4.2).

We tested also a second functional form for the localized orbitals which is inspired by the definition of Wannier functions [141]. These are localized orbitals obtained by performing the Fourier transform of appropriately rotated Bloch eigenstates (Eq. 4.2). We consider a Wannier orbital centered around the atomic position $\mathbf{R}$ and referenced with respect to a generic lattice vector $\mathbf{L}_m$. This reads:

$$w(\mathbf{r} - \mathbf{R}_a - \mathbf{L}_m) = \frac{\Omega}{(2\pi)^3} \int d\mathbf{k}^\sigma\ e^{-i\mathbf{k}^\sigma \cdot \mathbf{L}_m} \phi(\mathbf{r} - \mathbf{R}_a, \mathbf{k}^\sigma), \tag{4.7}$$

where Ω is the volume of the simulation box. The Bloch functions are retrieved by inverting the Fourier transform in Eq. 4.7 as follows:

$$\phi_i(\mathbf{r}, \mathbf{k}^\sigma) = \sum_m\ e^{-i\mathbf{k}^\sigma \cdot \mathbf{L}_m} w(\mathbf{r} - \mathbf{R}_a - \mathbf{L}_m). \tag{4.8}$$

If we substitute the localized Wannier functions $w(\mathbf{r} - \mathbf{R})$ with our localized orbitals, we obtain the final functional form of our single-particle orbitals:

$$X_i(\mathbf{r}, \mathbf{k}^\sigma) = \sum_m\ e^{-i\mathbf{k}^\sigma \cdot \mathbf{L}_m} \chi_i(\mathbf{r} - \mathbf{R}_a - \mathbf{L}_m). \tag{4.9}$$

The sum goes over the infinite set of direct lattice vectors $\{\mathbf{L}_m\}$. We truncate the summation for a given term m^* which ensures the following condition to be fulfilled for all orbitals:

$$\zeta_i \left[(m_1^* L_x)^2 + (m_2^* L_y)^2 + (m_3^* L_z)^2 \right]^\alpha \geq \epsilon_{\text{cut}} \quad \text{with } i \in [1, N_{\text{tot}}], \tag{4.10}$$

where the $\alpha = 1$ for Gaussian type orbitals and $\alpha = 1/2$ for Slater type orbitals. $\{\zeta_i\}$ are the exponents of the localized functions (Eq. 2.14) and $\epsilon_{\rm cut}$ is a suitable input cutoff typically defined such as: $e^{-\epsilon_{\rm cut}} \sim 10^{-6} - 10^{-8}$ a.u. Notice that the optimal value of this cutoff is strongly system dependent and must be assessed for each simulation. We realized empirically that the DFT calculations usually require a larger cutoff to obtain convergent results with respect to the corresponding correlated QMC runs. Eq. 4.9 does not need any coordinate transformation, thus the localized functions $\{\chi_i(\mathbf{r})\}$ can be evaluated with the same routines employed for Eq. 2.14 already exploited for open systems.

The orbitals in Eq. 4.9 satisfy two important properties which are listed here and which will be useful later on:

- when an electron is displaced of any lattice vector ($\mathbf{r} \rightarrow \mathbf{r} + \mathbf{L}_n$) it picks only a phase factor:

$$\begin{aligned} X_i(\mathbf{r} + \mathbf{L}_n, \mathbf{k}^\sigma) &= \sum_m e^{-i\mathbf{k}^\sigma \cdot \mathbf{L}_m} \chi_i((\mathbf{r} + \mathbf{L}_n) - \mathbf{R}_a - \mathbf{L}_m) \\ &= e^{-i\mathbf{k}^\sigma \cdot \mathbf{L}_n} X_i(\mathbf{r}, \mathbf{k}^\sigma) \quad \forall\ n \end{aligned} \tag{4.11}$$

  This is nothing else than the Bloch theorem and it directly derives from the Wannier-like construction of our orbitals.

- a similar relation holds when the atomic position $\mathbf{R}_a$ is translated by any lattice vector ($\mathbf{R}_a \rightarrow \mathbf{R}_a + \mathbf{L}_n$):

$$\begin{aligned} X_i^{\mathbf{R}_a + \mathbf{L}_n}(\mathbf{r}, \mathbf{k}^\sigma) &= \sum_m e^{-i\mathbf{k}^\sigma \cdot \mathbf{L}_m} \chi_i(\mathbf{r} - (\mathbf{R}_a + \mathbf{L}_n) - \mathbf{L}_m) \\ &= e^{i\mathbf{k}^\sigma \cdot \mathbf{L}_n} X_i(\mathbf{r}, \mathbf{k}^\sigma) \quad \forall\ n \end{aligned} \tag{4.12}$$

The main drawback with respect to the first choice is that the computational cost needed by the summation over lattice vectors is higher. However, as we will see in the next Section, accurate many-body QMC calculations are typically performed on large supercells, thus only one or two terms in the summation are necessary to satisfy the cutoff condition in Eq. 4.10.

We carry out benchmark calculations on the DFT(LDA) determinant developed on both basis sets in order to assess their accuracy. As apparent from Tab. 4.1, the second implementation in Eq. 4.9 gives a lower DFT(LDA) energy for all tested boundary conditions. Despite the non-variational nature of the DFT method when changing the basis set functional form, a lower energy usually provides a better starting point for the QMC wavefunction optimization procedure. In view of these results, we adopted Eq. 4.9 as the basis set form for the determinantal part of the QMC wavefunction.

Table 4.1 – Benchmark calculations on the solid bcc phase of hydrogen. Calculations are performed within the primitive cell and a complex-valued wavefunction is used consistently for all boundaries. We report the value of the DFT(LDA) energy in Ha/atom using the two different implementations of localized orbitals discussed in the text.

k-point	Energy (Eq. 4.5)	Energy (Eq. 4.9)
(0.500, 0.500, 0.000)	-0.449646	-0.460164
(0.500, 0.500, 0.500)	-0.563586	-0.623700
(0.250, 0.250, 0.000)	-0.459667	-0.461574
(0.250, 0.250, 0.250)	-0.486947	-0.501187
(0.166, 0.166, 0.500)	-0.475306	-0.479029

4.2.4 Generic boundary conditions

In several situations presented in this manuscript, it is useful to impose a boundary condition $\mathbf{k}^{\uparrow}$ for spin up particles different from $\mathbf{k}^{\downarrow}$ applied to spin down electrons.

One scenario where this feature becomes unavoidable is to obtain an efficient QMC determinant optimization with complex wavefunctions, as explained in Sec. 4.5.2. Another technique which exploits the generic boundary conditions framework here introduced is the flavor twist method for reducing finite-size effects, detailed in Sec. 5.7 of the next Chapter.

Given the single-particle basis previously described, we can define new spin dependent overlap matrices for basis set and density independent part of the Hamiltonian:

$$S_{ij}^{\sigma} = \int d\mathbf{r}\ X_i^*(\mathbf{r},\mathbf{k}^{\sigma}) X_j(\mathbf{r},\mathbf{k}^{\sigma}), \tag{4.13}$$

$$S_{ij}^{H,\sigma} = \int d\mathbf{r}\ X_i^*(\mathbf{r},\mathbf{k}^{\sigma}) \hat{H}^{\mathrm{KS}} X_j(\mathbf{r},\mathbf{k}^{\sigma}), \tag{4.14}$$

where both S and S^H are Hermitian matrices by definition and the overlap S is also positive definite. A generalization to non collinear spins is possible, but it has not been implemented.

Given the above definition of the overlap matrices, the exchange correlation part of $\hat{H}^{\mathrm{KS}}$ is then computed using standard LDA or LSDA approximations. Since both $V_{\mathrm{Ha}}[\rho(\mathbf{r})]$ and $V_{\mathrm{XC}}[\rho(\mathbf{r})]$ are purely density dependent quantities, they are real-valued, thus keeping the hermiticity of the full Hamiltonian. Ewald summations are used for dealing with the divergence in the $1/r$ contribution to the integrated Coulomb potential due to the interaction between periodic images. It is important to remark that, within our generic boundary implementation, the Hamiltonian and thus the electronic density $\rho(\mathbf{r})$ must be always evaluated separately for $\uparrow$ and $\downarrow$ spin electrons at each self-consistent cycle.

This leads to a double memory consumption for the complex DFT algorithm, making a proper use of buffering more relevant than in the open systems case. Notice that, when $\mathbf{k}^{\downarrow} = -\mathbf{k}^{\uparrow}$, the eigenvectors of the two spin sectors are Hermitian conjugates of each other. This ensures that the resulting determinant matrix Λ (Eqs. 2.28 and 4.29) possesses the Hermitian property. In this case, suitable optimizations to the algorithm reduce both memory requirements and computational cost.

The diagonalization procedure sketched in Sec. 2.2 can be extended to the case of Hermitian matrices using the complex diagonalization routines provided by the standard linear algebra library LAPACK [28]. The diagonalization is applied separately to the $\uparrow$ and $\downarrow$ spin KS problems and it yields a double set of eigenvectors $\{\phi_i(\mathbf{r}, \mathbf{k}^{\uparrow})\}$ and $\{\phi_i(\mathbf{r}, \mathbf{k}^{\downarrow})\}$ which are molecular orbitals. The DFT single determinant wavefunction is now compactly expressed as:

$$\Psi_{\mathrm{SD}}(\mathbf{r}_1, \cdots, \mathbf{r}_N) = \det[\phi_i(\mathbf{r}_j, \mathbf{k}^{\uparrow})] \det[\phi_i(\mathbf{r}_j, \mathbf{k}^{\downarrow})], \tag{4.15}$$

where $i \in [1, N_{\mathrm{tot}}/2]$ and $j \in [1, N/2]$. Eq. 4.15 can be used as initial guess for QMC periodic calculations after building the proper interface.

4.2.5 Numerical Brillouin zone integration

The continuous integral in Eq. 4.4 for evaluating thermodynamically converged observables in DFT is approximated with a summation over a finite number N_k of $\mathbf{k}$-points.

A special case of Eq. 4.4 is the evaluation of the total DFT energy for the infinite crystal. The way this quantity is practically evaluated deserves better attention since it will be crucial for understanding the special twist method for reducing finite-size effects introduced in the next Chapter. In the standard DFT framework, the total KS energy reads:

$$\begin{aligned} \mathcal{E}_{\infty}[\rho_{\infty}] &= \sum_{\sigma} \sum_{j=1}^{N_k} w_j \sum_{n \in \mathrm{occ}} \left\langle \phi_n(\mathbf{r}, \mathbf{k}_j^{\sigma}) \right| \hat{H}^{\mathrm{KS}}[\rho_{\infty}] \left| \phi_n(\mathbf{r}, \mathbf{k}_j^{\sigma}) \right\rangle = \\ &= \sum_{\sigma} \sum_{j=1}^{N_k} w_j \sum_{n} F_{\beta}(\mathcal{E}_n^{\mathbf{k}_j^{\sigma}}[\rho_{\infty}] - \mu_{\infty})\, \mathcal{E}_n^{\mathbf{k}_j^{\sigma}}[\rho_{\infty}]. \end{aligned} \tag{4.16}$$

where the index n goes over occupied electronic bands and the function F_{β} is defined later on (Eq. 4.18). . The electronic density $\rho_{\infty}(\mathbf{r})$ of the infinite crystal is defined in a similar fashion:

$$\rho_{\infty}(\mathbf{r}) = \sum_{\sigma} \sum_{j=1}^{N_k} w_j \sum_{n} F_{\beta}(\mathcal{E}_n^{\mathbf{k}_j^{\sigma}}[\rho_{\infty}] - \mu_{\infty})\ |\phi_n(\mathbf{r}, \mathbf{k}_j^{\sigma})|^2. \tag{4.17}$$

Eqs. 4.16 and 4.17 contain several crucial tricks required for accurate evaluation of crystal properties which have subtle physical implications. These features of the DFT code have been implemented in the *TurboRVB* package during this thesis.

First of all, one can considerably improve the accuracy of the approximation in Eq. 4.16 whilst keeping, at the same time, an affordable number of **k**-points, by choosing an uniformly spaced grid in the 1BZ following the well-known Monkhorst-Pack method [142]. To each grid point it can be assigned an appropriate weight w_i determined by exploiting the point group symmetries of the Bravais lattice associated with the simulation box. In order for the averaged electronic density to fulfill the right crystal symmetries, this quantity must be symmetrized after each self-consistent cycle by summing over the stars of equivalent lattice vectors; in this case, shifting the origin of the grid can be harmful since it might break the underlying crystal symmetries thus hindering a correct symmetrization. We do not impose any shift to the real mesh in the case of the weighted **k**-points summation.

Secondly, the occupation of each band is computed using the smearing function F_β which, in our implementation, we take with Fermi-Dirac form:

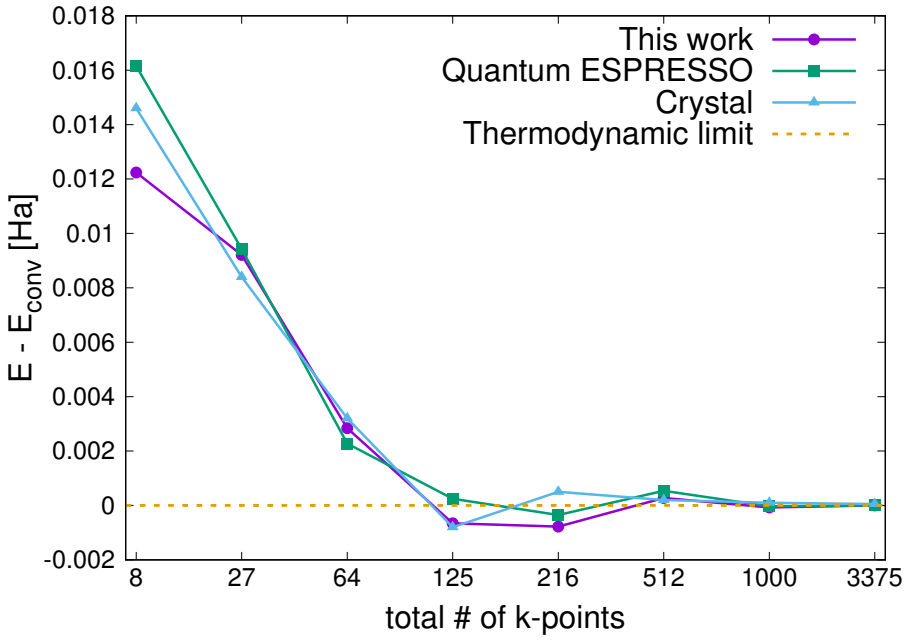
$$F_\beta(\mathcal{E} - \mu_\infty) = \frac{1}{1 + \mathrm{e}^{\beta(\mathcal{E}-\mu_\infty)}}. \tag{4.18}$$

β is the smearing parameter which determines the steepness of F_β and which is equivalent to the application of a fictitious temperature to the electronic degrees of freedom in the system. The smearing is crucial for the convergence of the summation in Eqs. 5.9,4.17 in the case of metallic systems possessing discontinuous electronic occupations around the Fermi energy. Notice that the effect of the electronic temperature is to introduce fractional electron occupations for each boundary. It is thus clear that this method works within the *grand canonical* ensemble. The variable number of particles per each boundary is set by the chemical potential μ_∞ updated at each self-consistent cycle. Notice that at zero temperature μ_∞ coincides with the single-particle Fermi energy ϵ_F.

We present in Fig. 4.1 a benchmark **k**-point energy convergence on the lithium bcc phase with our implementation and we compare it to the Quantum ESPRESSO [18] software, which uses a plane wave basis set, and to the Crystal [25] program, which employs the same Gaussian basis set as ours, but with an analytical procedure for the evaluation of two-electron integrals. As expected, the three methods display a similar efficiency in converging to the $\mathcal{E}_\infty[\rho_\infty]$. This confirms the quality of our **k**-point sampling implementation.

In conclusion, it is worth mentioning that the KS problem can be solved independently for each value of **k**. This feature naturally leads to build an additional level of parallelism on top of the splitting on the x direction of the integration mesh (see Sec. 2.2.2). To this purpose, we exploit the features of the message passage interface (MPI) library for parallel calculations to create a two dimensional grid of processors where each row deals with a single **k**-point and the results are then averaged over columns, i.e. **k**-points, at

Figure 4.1 – Benchmark calculations on our implementation of DFT within a localized basis set. We show the convergence with the number of **k**-points used to integrate over the 1BZ of the bcc structure of lithium. We use uniformly spaced grid of points [142] for sampling the 1BZ. On the x axis the total number of **k**-points in the grid is reported whereas on the y axis we show the total energy difference in Ha between the chosen grid and the thermodynamic limit calculated with a $20 \times 20 \times 20$ mesh. We compare our DFT software with two widely used codes (Quantum ESPRESSO [18] and Crystal [25]) implementing different DFT frameworks.



the end of each DFT cycle, for evaluating Eqs. 4.16 and 4.17. The basic algorithm for creating the two dimensional grid of processors is presented in Appendix B. The same parallel scheme is also applied to QMC calculations on periodic systems which is the topic of the next section.

4.3 Periodic systems with QMC

4.3.1 Many-body Hamiltonian in supercell calculations

Correlated wavefunction methods such as QMC explicitly treat the full quantum many-body problem through the ansatz $\Psi_T(\mathbf{r}_1, \cdots, \mathbf{r}_N)$. The full Coulomb electron-electron interaction in a solid is long-ranged and non-periodic. In the following we denote as $\mathbf{L}_P$ the lattice vectors defined in the primitive cell. By defining the many-body Hamiltonian within a single PC, the non-periodic Coulomb interaction breaks the following symmetry:

$$\hat{H}(\mathbf{r}_1, \ldots, \mathbf{r}_i + \mathbf{L}_P, \ldots, \mathbf{r}_N) = \hat{H}(\mathbf{r}_1, \cdots, \mathbf{r}_N), \tag{4.19}$$

which is instead valid for the effective single-particle Hamiltonian (Eq. 4.1). The sole invariance left is the simultaneous translation of *all* electrons of a primitive lattice vector:

$$\hat{H}(\mathbf{r}_1 + \mathbf{L}_P, \ldots, \mathbf{r}_i + \mathbf{L}_P, \ldots, \mathbf{r}_N + \mathbf{L}_P) = \hat{H}(\mathbf{r}_1, \cdots, \mathbf{r}_N). \tag{4.20}$$

Eq. 4.20 is a direct consequence of the periodic nature of the total electron-ion potential. However, this condition is not sufficient to describe the infinite crystal using a single PC. In explicitly correlated calculation one is forced to extend the simulation cell to a larger box, usually called *supercell* (SC). Notice that SCs are extensively used also in independent electron methods in order to study non-periodic structure or solids with defects. In correlated calculations, supercells are unavoidable in order to capture the long-range decay of correlation functions and of the Coulomb potential itself.

We denote as $\mathbf{L}_S$ the SC lattice vectors and as $N_S > N$ the number of electrons in the SC. The SC is typically chosen as $n_x \times n_y \times n_z$ integer multiple of the primitive cell. In this way the supercell lattice is compatible with the translational symmetries given by the primitive cell lattice spanned by the vectors $\mathbf{L}_P$ since $\{\mathbf{L}_S\} \subset \{\mathbf{L}_P\}$. As for the primitive cell case, periodic images are generated by subjecting the SC to periodic boundary conditions and the Ewald summation formalism can be exploited for curing the divergent $1/r$ series of the Coulomb potential. The many-body wavefunction of a $n_x \times n_y \times n_z$ supercell corresponds to a set of single-particle wavefunction defined on a $n_x \times n_y \times n_z$ grid of $\mathbf{k}$-points in the 1BZ of the primitive cell. Fig. 4.2a shows a representation of a $\mathbf{k}$-point mesh corresponding to a 4×4 supercell within the Brillouin zone of a 2D square lattice.

The electron many-body Hamiltonian (Eq. 2.2) within a supercell can be rewritten as follows:

$$\hat{H} = -\frac{1}{2}\sum_{i=1}^{N} \Delta_{\mathbf{r}_i} + \frac{1}{2}\sum_{\mathbf{L}_S}\sum_{i<j}^{N} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j - \mathbf{L}_S|} - \sum_{\mathbf{L}_S}\sum_{a=1}^{N_{\text{at}}}\sum_{i=1}^{N} \frac{Z}{|\mathbf{r}_i - \mathbf{R}_a - \mathbf{L}_S|} \tag{4.21}$$

where the summation $\sum_{\mathbf{L}_S}$ goes over all the lattice vectors of the supercell.

4.3.2 Twisted boundary conditions

It is simple to generalize the single particle Bloch theorem to the many-body eigenstates of the Hamiltonian in Eq. 4.21 [143, 144]. In fact, the SC approach introduces an *artificial* translational symmetry in $\hat{H}$ given by:

$$\mathcal{H}(\mathbf{r}_1, \ldots, \mathbf{r}_i + \mathbf{L}_S, \ldots, \mathbf{r}_{N_S}) = \mathcal{H}(\mathbf{r}_1, \cdots, \mathbf{r}_{N_S}) \quad \forall\, i \in [1, N_S]. \tag{4.22}$$

Consider now a many-body wavefunction $\Psi(\mathbf{r}_1, \cdots, \mathbf{r}_{N_S})$, assumed to be an eigenstate of $\hat{H}$. If one picks an arbitrary couple of electrons identified by the positions $\mathbf{r}_i$ and $\mathbf{r}_{i+1}$

and displaces the first particle by a SC lattice vector, the one-particle Bloch theorem must hold:

$$\Psi(\mathbf{r}_1, \ldots, \mathbf{r}_i + \mathbf{L}_S, \ldots, \mathbf{r}_{N_S}) = e^{i\boldsymbol{\theta}_i \cdot \mathbf{L}_S} \Psi(\mathbf{r}_1, \cdots, \mathbf{r}_{N_S}) \tag{4.23}$$

where we use the symbol $\boldsymbol{\theta}$ in order to distinguish from the single-particle boundary $\mathbf{k}$. One can apply the same shift to the second electron as:

$$\Psi(\mathbf{r}_1, \ldots, \mathbf{r}_i, \mathbf{r}_{i+1} + \mathbf{L}_S, \ldots, \mathbf{r}_{N_S}) = e^{i\boldsymbol{\theta}_{i+1} \cdot \mathbf{L}_S} \Psi(\mathbf{r}_1, \cdots, \mathbf{r}_{N_S}) \tag{4.24}$$

where for now $\boldsymbol{\theta}_i \neq \boldsymbol{\theta}_{i+1}$. In Eq. 4.24 we can exchange the two electrons; using the antisymmetric property of Ψ one obtains:

$$\Psi(\mathbf{r}_1, \ldots, \mathbf{r}_{i+1} + \mathbf{L}_S, \mathbf{r}_i, \ldots, \mathbf{r}_{N_S}) = -e^{i\boldsymbol{\theta}_{i+1} \cdot \mathbf{L}_S} \Psi(\mathbf{r}_1, \cdots, \mathbf{r}_{N_S}) \tag{4.25}$$

By applying the transformation: $\mathbf{r}_{i+1} \to \mathbf{r}_{i+1} - \mathbf{L}_S$ and exchanging again the two particles, one retrieves the following identity:

$$\Psi(\mathbf{r}_1, \ldots, \mathbf{r}_{N_S}) = e^{i(\boldsymbol{\theta}_{i+1} - \boldsymbol{\theta}_i) \cdot \mathbf{L}_S} \Psi(\mathbf{r}_1, \cdots, \mathbf{r}_{N_S}) \tag{4.26}$$

Eq. 4.26 holds only if $\boldsymbol{\theta}_{i+1} - \boldsymbol{\theta}_i \in \mathbf{G}_S$, where $\{\mathbf{G}_S\}$ are the supercell reciprocal lattice vectors. This follows directly from the definition of the reciprocal lattice: $e^{i\mathbf{G}_S \cdot \mathbf{L}_S} = 1$ holds for all lattice vectors. Since we applied PBC to the supercell, the supercell wavevector $\boldsymbol{\theta}$ remains unchanged if translated within the SC reciprocal lattice. $\boldsymbol{\theta}_i$ can be reduced into the 1BZ of the supercell and we can set $\boldsymbol{\theta}_i = \boldsymbol{\theta}_{i+1}$ without loss of generality. The electron couple has been arbitrarily chosen at the beginning of this derivation, therefore the phase of the wavefunction must be the same for *any* electron.

In conclusion, we can formulate the many-body Bloch theorem with two equivalent relations, which generalize Eqs. 4.2 and 4.3. When any electron wraps around the SC boundaries it picks the same phase factor $\boldsymbol{\theta}$:

$$\Psi(\mathbf{r}_1, \ldots, \mathbf{r}_i + \mathbf{L}_S, \ldots, \mathbf{r}_{N_S}) = e^{i\boldsymbol{\theta} \cdot \mathbf{L}_S} \Psi(\mathbf{r}_1, \cdots, \mathbf{r}_{N_S}) \quad \forall\ i \in [1, N_S]. \tag{4.27}$$

This is equivalent to requiring Ψ to possess the following functional form:

$$\Psi^{\boldsymbol{\theta}}(\mathbf{r}_1, \cdots, \mathbf{r}_{N_S}) = W_{\boldsymbol{\theta}}(\mathbf{r}_1, \cdots, \mathbf{r}_{N_S}) \exp\left(i\boldsymbol{\theta} \cdot \sum_{i=1}^{N_S} \mathbf{r}_i \right) \tag{4.28}$$

where the function W is periodic on the SC lattice.

The wavevector $\boldsymbol{\theta}$ is usually called *twist condition* and it is an extension of the concept of $\mathbf{k}$-point introduced for mean-field theories. In Fig. 4.2 we present a graphical visualization of the action of a twist condition on the primitive cell BZ. The contour plot is a generic Fermi surface slice. On the left panel we show the $\mathbf{k}$-points in the primitive BZ

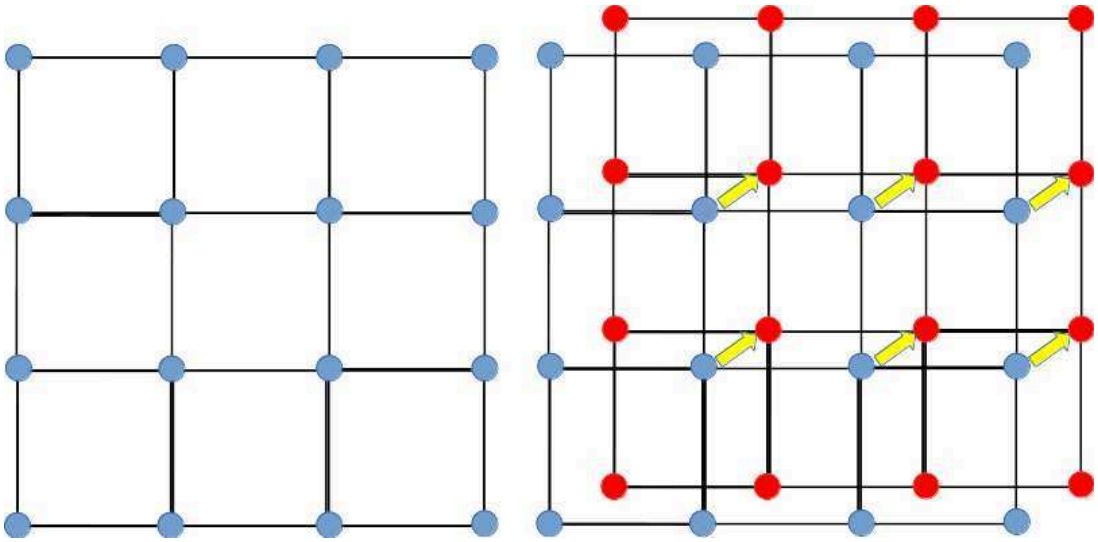


Figure 4.2 – Graphical visualization on the difference between supercell twist and **k**-point in the first BZ of the primitive cell. In the left panel we represent the **k**-points mesh in the first BZ of the primitive cell associated with a 4×4 supercell. In the right panel, the action of a twist condition as an offset to this grid is shown. An integration over the whole set of **k**-points allowed in the first BZ is equivalent to integrate over all twist conditions allowed in the supercell reciprocal lattice.

corresponding to the supercell wavefunction. When a twist $\boldsymbol{\theta}$ is applied (right panel), an offset proportional to the twist components along the Cartesian directions is applied to this grid which now explores different points in the first BZ. From Fig. 4.2 it is evident that the BZ integration for single particle systems is equivalent, in a many-body framework, to an integration over all the twist conditions allowed in the SC Brillouin zone. As in the independent particle case, the discrete set of allowed values is determined by the PBC imposed on the supercell. Therefore, an accurate determination of the properties of the infinite crystal within QMC usually requires us to perform calculations with wavevectors $\boldsymbol{\theta}$ different from zero.

As in the DFT case, the main computational difficult arising from a generic $\boldsymbol{\theta}$ is that the wavefunction must be complex-valued. Complex arithmetic can be avoided only for twist values which correspond to a **k**-point mesh with inversion symmetry, such as $\boldsymbol{\theta} = 0$ or $\boldsymbol{\theta} = (\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$. However, as we will see in the next Chapter, such twist conditions are usually not sufficient to obtain thermodynamically converged properties within QMC. The last two sections of this Chapter are devoted to the generalization of the main QMC algorithms introduced in Chap. 2 to the complex formalism.

4.3.3 Complex JAGP ansatz

Using the single-particle orbitals defined in Eq. 4.9, we can rewrite the electron pairing function for extended systems. In the valence bond representation for a spin unpolarized system – introduced for open systems in Eq. 2.28 – this quantity reads:

$$\Phi(\mathbf{r}^{\uparrow}, \mathbf{r}'^{\downarrow}) = \sum_{a,b}^{N_{\text{at}}} \sum_{\mu,\nu}^{N_{\text{bas}}} \lambda_{\mu,\nu}^{a,b} \; X_{\mu}^{a}(\mathbf{r}, \boldsymbol{\theta}^{\uparrow}) \; X_{\nu}^{b}(\mathbf{r}', \boldsymbol{\theta}^{\downarrow}). \tag{4.29}$$

The way single-particle orbitals are defined allows us to impose a generic twist condition for $\uparrow$ and $\downarrow$ spin electrons in the determinantal part. In contrast to the open systems case, the matrix $\Lambda = \{\lambda_{\mu,\nu}^{a,b}\}$ now has complex-valued elements. In the alternative MOs representation (see Eq. 2.29), we expand the geminal as:

$$\Phi(\mathbf{r}^{\uparrow}, \mathbf{r}'^{\downarrow}) = \sum_{j}^{N_{\text{MO}}} \lambda_{j}^{\text{MO}} \phi_{j}^{\text{MO}}(\mathbf{r}^{\uparrow}) \phi_{j}^{\text{MO}}(\mathbf{r}'^{\downarrow}). \tag{4.30}$$

In Eq. 4.30, both the eigenvalues λ_j^{MO} - obtained through the diagonalization of the Λ geminal matrix - as well as the coefficients $C^{\sigma} = \{c_{ji}^{\sigma}\}$ - defining the MO expansion over the localized basis set - are phase dependent.

without loss of generality we can impose the coefficients λ_j^{MO} independently of the chosen twist condition. On the contrary, the coefficients $C^{\sigma} = \{c_{ji}^{\sigma}\}$ defining the MOs expansion over the localized basis set are phase dependent.

When using the generic twist condition framework previously introduced, particular care must be taken when passing from the VB to the MOs representation. Indeed, since unlike spin particles are no longer equivalent, in this situation the Λ matrix diagonalization procedure must be generalized as is usually done in the case of spin polarized systems [37]. We extend the geminal matrix by constructing a $2N_{\text{tot}} \times 2N_{\text{tot}}$ matrix $\tilde{\Lambda}$ defined as:

$$\tilde{\Lambda} = \begin{pmatrix} 0 & \Lambda \\ \Lambda^{\dagger} & 0 \end{pmatrix} \tag{4.31}$$

where $\Lambda^{\dagger}$ is the Hermitian conjugate of Λ. The basis set overlap matrix, since the basis orbitals are different between the two spin sectors, must also be extended, with the same dimension, as follows:

$$\tilde{S} = \begin{pmatrix} S^{\uparrow} & 0 \\ 0 & S^{\downarrow} \end{pmatrix} \tag{4.32}$$

where the matrix S^{σ} has been defined in Eq. 4.14. The complex MO coefficients and eigenvalues λ_j^{MO} are obtained by solving the generalized eigenvalue problem associated with $\tilde{\Lambda}$ which corresponds to the following set of secular equations:

$$\begin{cases} S^{\uparrow} \Lambda S^{\downarrow} C^{\downarrow} = \lambda_j^{\text{MO}} S^{\uparrow} C^{\uparrow} \\ S^{\downarrow} \Lambda^{\dagger} S^{\uparrow} C^{\uparrow} = \lambda_j^{\text{MO}} S^{\downarrow} C^{\downarrow} \end{cases} \tag{4.33}$$

Notice that Eqs. 4.33 couple the two spin sectors, thus they must be solved at the same time by means of an algorithm similar to the one used for the KS Hamiltonian diagonalization. All the properties of the AGP wavefunction established for open systems are valid also in the case of solid calculations. The starting $N/2$ molecular orbitals are taken from the DFT(LDA) wavefunction in Eq. 4.15.

The Jastrow factor is only constructed by means of density-density correlators. Since the electronic density is an intrinsically phase independent quantity, due to its periodicity over the atomic lattice, in the case of periodic systems the Jastrow factor can be chosen twist-independent and therefore real-valued as in the case of atoms and molecules. Notice that this quantity is defined on the same periodic box as the determinant, therefore it must satisfy periodic boundary conditions when a particle wraps around cell boundaries. In order to fulfill this condition, we expand the short-range part of the Jastrow factor (Eqs. 2.33 and 2.35) over a set of localized orbitals with the usual functional form as for open systems. However, in our current implementation periodic boundary conditions are enforced on these one-particle functions by using the coordinate transformation on the electron-ion distances defined in Eq. 4.6. We plan to extend also the Jastrow basis set to the functional form defined in Eq. 4.9.

4.4 Single-point QMC calculations with complex wavefunctions

In this Section we present the variational and diffusion Monte Carlo algorithms suitably modified to handle complex-valued wavefunctions. The evolution of the VMC algorithm is straightforward. The DMC case is instead more delicate since the standard fixed-node approximation introduced in Eq. 2.50 cannot be applied and must be reformulated to cope with complex trial and guiding wavefunctions.

4.4.1 Variational Monte Carlo

The Hamiltonian expectation value introduced in Eq. 2.42 is modified as follows:

$$E_{\mathrm{VMC}} = \frac{\int d\mathbf{R}_{\mathrm{el}} \Psi_T^*(\mathbf{R}_{\mathrm{el}}) \hat{H} \Psi_T(\mathbf{R}_{\mathrm{el}})}{\int d\mathbf{R}_{\mathrm{el}} |\Psi_T(\mathbf{R}_{\mathrm{el}})|^2}, \tag{4.34}$$

where the local energy $e_L(\mathbf{R}_{\mathrm{el}}) = \hat{H}\Psi_T/\Psi_T$ is now a complex-valued function of the electronic positions. The importance sampling transformation for the efficient evaluation of Eq. 4.34 is analogous to the real case and the Markov chain generation process remains

the same. The only relevant difference is that the computational cost of the algorithm is approximately 2/2.5 times the real case due to the need of complex arithmetic.

Unlike the local energy, the Hamiltonian expectation value and any other physical observables $\langle \hat{O} \rangle$ extracted from the QMC simulation are real quantities. The VMC operator estimator must be modified accordingly:

$$\hat{O}_{\mathrm{VMC}} = \langle \hat{O}_L \rangle = \frac{1}{M} \sum_{i=1}^{M} \mathrm{Re}[\hat{O}_L(\mathbf{R}_{\mathrm{el}}^i)]. \tag{4.35}$$

where M is the number of MC generation on a single block of stochastic measurements and $\hat{O}_L(\mathbf{R}_{\mathrm{el}}) = \hat{H}\Psi_T/\Psi_T$ is the local version of the operator. We evaluate the block variance in the following way:

$$\begin{aligned} \sigma^2[\langle \hat{O}_L \rangle] &= \frac{1}{M-1} \sum_{i=1}^{M} \left[\hat{O}_L(\mathbf{R}_{\mathrm{el}}^i) - \langle \hat{O}_L \rangle \right]^* \left[\hat{O}_L(\mathbf{R}_{\mathrm{el}}^i) - \langle \hat{O}_L \rangle \right] \\ &= \frac{1}{M-1} \sum_{i=1}^{M} \left[\hat{O}_L(\mathbf{R}_{\mathrm{el}}^i)\ \hat{O}_L^*(\mathbf{R}_{\mathrm{el}}^i) - \langle \hat{O}_L \rangle^2 \right]. \end{aligned} \tag{4.36}$$

Block averaging and the bootstrap technique [b] are then used to decrease the serial correlation among the MC measurements.

A particular attention must be paid when multiple twist conditions are included in the simulation. Twist calculations are uncorrelated, so we can compute the expectation value by a simple arithmetic average at the end of the simulation with no need of block regrouping:

$$\langle \hat{O}_L \rangle = \frac{1}{N_t} \sum_{i=1}^{N_t} \langle \hat{O}_L^{\boldsymbol{\theta}_i} \rangle \tag{4.37}$$

where N_t is the number of twists and $\langle \hat{O}_L^{\boldsymbol{\theta}} \rangle = \frac{\int d\mathbf{R}_{\mathrm{el}} \Psi_T^*(\mathbf{R}_{\mathrm{el}},\boldsymbol{\theta}) \hat{O} \Psi_T(\mathbf{R}_{\mathrm{el}},\boldsymbol{\theta})}{\int d\mathbf{R}_{\mathrm{el}} |\Psi_T(\mathbf{R}_{\mathrm{el}},\boldsymbol{\theta})|^2}$ is the operator expectation value for a single twist evaluated in the standard way. We can write the variance on a multiple twist calculation as:

$$\sigma^2[\langle \hat{O}_L \rangle] = \frac{1}{N_t} \left(\frac{1}{N_t - 1} \sum_{i=1}^{N_t} \sigma^2[\langle \hat{O}_L^{\boldsymbol{\theta}_i} \rangle] \right) \simeq \frac{1}{N_t^2} \sum_{i=1}^{N_t} \sigma^2[\langle \hat{O}_L^{\boldsymbol{\theta}_i} \rangle] \tag{4.38}$$

where $\sigma^2[\langle \hat{O}_L^{\boldsymbol{\theta}} \rangle]$ is the single twist variance (Eq. 4.36) and we used the usual expression for the unbiased estimator of the variance. Notice that with this definition, if ideally the wavefunction is exact for each twist, the variance goes to zero and the zero-variance

[b] Boostrap resampling is a statistical method which consists in resampling the obtained data by following the empirical distribution of the data themselves or, in other words, constructing a number of resamples with replacement of the observed statistical data and then recover the desired average value from them.

property is maintained as expected. It is important to remark that multiple twists do not necessarily lead to a lower efficiency of the algorithm with respect to the single twist calculation as twists are uncorrelated. In fact, the statistics for each boundary condition can be accumulated and, keeping the same number of MC samples, the statistical noise when passing from a calculation with a single twist condition to one including N_t boundaries is reduced according to:

$$\sigma[\langle \hat{O}_L \rangle] \rightarrow \frac{1}{\sqrt{(N_t)}} \sigma[\langle \hat{O}_L \rangle], \tag{4.39}$$

which directly derives from Eq. 4.38. As we will see later on, this feature is particularly relevant to achieve efficient wavefunction optimization with multiple twists (Sec. 4.5), as well as to effectively reduce finite size errors by averaging over different boundary conditions (Sec. 5.3).

4.4.2 Fixed-phase diffusion Monte Carlo

A diffusion Monte Carlo algorithm able to handle complex-valued wavefunctions allows the treatment of systems where time-reversal symmetry is broken. This feature is needed in many physical situations not only because it enables the application of generic boundary conditions to the wavefunction, but also since it opens the way to treat Hamiltonians containing external magnetic fields within an accurate correlated technique as DMC.

A complex trial wavefunction hinders a probabilistic interpretation of the important sampling transformation $f(\mathbf{R}_{\rm el}, t)$ (Eqs. 2.49, 2.50). A simple but effective workaround, known as the fixed phase approximation, was introduced in Refs. [145, 146] in the context of the 2D electron gas under an external magnetic field. For complex-valued trial and guiding wavefunctions, we can rewrite the importance sampling transformation in the following way:

$$\begin{aligned} f(\mathbf{R}_{\rm el}, t) &= \Psi_G^*(\mathbf{R}_{\rm el}, t)\Psi_T(\mathbf{R}_{\rm el}), \\ &= |\Psi_G(\mathbf{R}_{\rm el}, t)|\ |\Psi_T(\mathbf{R}_{\rm el})| e^{i(\phi_T(\mathbf{R}_{\rm el}) - \phi_G(\mathbf{R}_{\rm el}, t))}, \end{aligned} \tag{4.40}$$

where ϕ_T (ϕ_G) is the phase of the trial (guiding) wavefunction.

The simplest way to retrieve a real-valued distribution from the above definition is to *fix* the phase of the guiding wavefunction during the imaginary time evolution to be equal to the one of the initial ansatz Ψ_T. This yields the following relation, usually called fixed-phase approximation:

$$\phi_G(\mathbf{R}_{\rm el}, t) = \phi_T(\mathbf{R}_{\rm el}) \quad \forall \ \mathbf{R}_{\rm el} \text{ and } t \tag{4.41}$$

By enforcing the above condition in Eq. 4.40, we immediately notice that the resulting

importance sampled distribution:

$$\tilde{f}(\mathbf{R}_{\text{el}}, t) = |\Psi_G(\mathbf{R}_{\text{el}}, t)|\ |\Psi_T(\mathbf{R}_{\text{el}})|, \tag{4.42}$$

is real-valued and positive definite, since it is defined as a product of the wavefunction moduli. The FP constraint is thus analogous to the FN approximation (Eq. 2.50), where the role of the nodal surface of Ψ_T is substituted by the 3N-dimension phase function $\phi_T(\mathbf{R}_{\text{el}})$. It can be easiliy demonstrated that within the FP approximation the variational character of the fixed-node approximation - a special case of the FP approximation when a real trial wavefunction is used - is maintained.

The FP constraint is very convenient since the time-dependent Schrödinger equation for $\tilde{f}(\mathbf{R}_{\text{el}}, t)$ (Eq. 2.51) is only slightly modified with respect to its real-valued counterpart. In fact, only the drift term shows a phase dependency:

$$\begin{aligned}
\nabla_{\mathbf{R}_{\text{el}}} \ln \Psi_T(\mathbf{R}_{\text{el}}) &= \nabla_{\mathbf{R}_{\text{el}}} \ln \left[|\Psi_T(\mathbf{R}_{\text{el}})| e^{i\phi_T(\mathbf{R}_{\text{el}})} \right] = \\
&= \nabla_{\mathbf{R}_{\text{el}}} \ln |\Psi_T(\mathbf{R}_{\text{el}})| + \nabla_{\mathbf{R}_{\text{el}}} \ln e^{i\phi_T(\mathbf{R}_{\text{el}})} = \\
&= \underbrace{\nabla_{\mathbf{R}_{\text{el}}} \ln |\Psi_T(\mathbf{R}_{\text{el}})|}_{\text{standard drift term}} + i \nabla_{\mathbf{R}_{\text{el}}} \phi_T(\mathbf{R}_{\text{el}}).
\end{aligned} \tag{4.43}$$

When performing the mixed average with the real function $\tilde{f}(\mathbf{R}_{\text{el}}, t)$ within the FP constraint, the expectation value of the purely imaginary term in the right-hand-side of Eq. 4.43 gives a vanishing contribution. Hence the master equation associated with $\tilde{f}(\mathbf{R}_{\text{el}}, t)$ can be solved using the very same algorithm outlined in Sec. 2.4.2 for the real case. The lattice regularization of the Hamiltonian is also maintained for introducing non-local pseudopotentials. As usual, the FP algorithm has a cost which is $\sim 2/2.5$ that of its real-valued counterpart due to the need for complex arithmetic. The FP approach outlined in this Section has been implemented in the *TurboRVB* package during this thesis. The FP method can be generalized when a vector potential $\mathbf{A}$ is included in the Hamiltonian for modeling the action of a magnetic field. This extension is discussed in Ref. [145], but has not been implemented here.

An evolution of the FP algorithm has been introduced by Ref. [147] under the name release-phase diffusion Monte Carlo. This method relaxes the FP constraint by allowing the phase of Ψ_G to evolve during the drift-diffusion process until branching occurs; at this stage the total phase is accumulated among all walkers and the walker weights are updated accordingly. The high computational cost of this method has limited – to our knowledge – its application to isolated atoms and as a benchmark for the simpler FP approximation [147, 148].

4.5 Wavefunction optimization with periodic systems

In standard QMC calculations on periodic systems, the determinantal part is kept fixed at the DFT level (Eq. 4.15) whereas only the Jastrow factor is optimized. The supposedly weak dependence of the DMC method on the trial Ψ_T makes this procedure typically sufficient to attain the desired accuracy and this is the methodology followed by most part of QMC codes available on the market. Furthermore, this scheme is convenient since the Jastrow factor is a twist independent quantity which can thus be chosen equal for all boundaries. It is evident that if one considers several uncorrelated twist conditions in the same simulation (for example for reducing systematic finite size effects with the methods explained in Secs. 5.3 and 5.7) the statistical noise on the minimization procedure is effectively reduced by averaging over the twists at every optimization step or for each twist separately at the end of the simulation (see Eq. 4.38).

The determinantal part is instead a $\boldsymbol{\theta}$-dependent quantity and its full QMC minimization would require, in principle, separate calculations for each twist. On the other hand, an accurate orbital optimization allows one to greatly reduce the size of the atomic basis set with a negligible loss in accuracy. Moreover, this possibility opens the way, using our JAGP wavefunction, to extract QMC electron pairing functions resolved in reciprocal space. This offers invaluable information for investigating, for example, the mechanism of high-temperature superconductivity, as demonstrated in Ref. [149] and discussed more in detail in Chap. 6.

In this Section we aim at extending the wavefunction optimization tools we introduced in Chap. 2 to the case of periodic systems. At first in Sec. 4.5.1 we briefly discuss the evolution of the SR method to the case of complex variational parameters. Afterwards, in Sec. 4.5.2, we introduce an original parameterization of the pairing function which brings a solution to the problem of determinant optimization in the presence of several twist conditions.

4.5.1 Stochastic reconfiguration with complex wavefunction

In Sec. 2.5 we introduced the stochastic reconfiguration algorithm in the context of real wavefunction optimization. The trial wavefunction $\Psi_T(\boldsymbol{\alpha})$ is parameterized with a set of now *complex* variational parameters $\{\boldsymbol{\alpha}\}$. We can write a collective variation of these parameters as:

$$\boldsymbol{\alpha} + \delta\boldsymbol{\alpha} = (\boldsymbol{\alpha}^{\mathrm{Re}} + \delta\boldsymbol{\alpha}^{\mathrm{Re}}) + i(\boldsymbol{\alpha}^{\mathrm{Im}} + \delta\boldsymbol{\alpha}^{\mathrm{Im}}). \tag{4.44}$$

When performing the linear expansion around the parameter variations introduced in Eq. 2.64, we can decouple the resulting problem into two real-valued equations for the

real and imaginary parts of the variational parameters. We then apply the same algorithm developed in Sec. 2.5.1 for the molecular system case. This gives rise to a couple of real-valued equations for the real and imaginary parts of the variational parameters, which are solved using the standard SR algorithm In our implementation, we promoted to complex the variational parameters in the determinant, namely the Λ geminal matrix and the coefficients of the atomic or GEO contractions.

The derivatives necessary for the SR complex algorithm are computed via algorithmic differentiation, introduced in Sec. 2.5.3. The adjoints of the two most important algebraic operations extended to complex-valued matrices are presented in Appendix C.

Since the complex pairing function in Eq. 4.29 is an analytical function of the variational parameters $\{\lambda_{\mu,\nu}^{a,b}\}$, we can make use of the well known Cauchy-Riemann relations. They link the derivatives of the real and imaginary parts of an analytic function $\Phi = \Phi^R + \mathrm{i}\Phi^\mathrm{I}$ with respect to the real and imaginary part of a generic complex-valued variational parameter $\lambda = \lambda^R + \mathrm{i}\lambda^\mathrm{I}$:

$$\begin{aligned} \frac{\partial \Phi^R}{\partial \lambda^R} &= \frac{\partial \Phi^I}{\partial \lambda^I} \\ \frac{\partial \Phi^R}{\partial \lambda^I} &= -\frac{\partial \Phi^I}{\partial \lambda^R} \end{aligned} \tag{4.45}$$

Eqs. 4.45 allows to avoid the computation of the derivatives of the imaginary part of the pairing function, thus leading to a considerable reduction (factor ~ 2) in the final computational cost of some parts of the complex algorithm.

4.5.2 Twist independent parameterization of the pairing function

For the following discussion, we consider the electron pairing function introduced in Eq. 4.29. In order to simplify the notation, we denote as $\lambda^{\mathbf{R}_a,\mathbf{R}_b}$ the geminal matrix elements connecting atoms $\mathbf{R}_a$ and $\mathbf{R}_b$ for *all* atomic basis set components. Analogously, $X_\sigma^{\mathbf{R}_a}$ refers to all basis set elements localized on the atom $\mathbf{R}_a$ for particles with spin σ, defined in Eq. 4.9.

The pairing function of a quantum system is intimately connected with its density matrix (see Eq. A.2), which is a physical observable. It is therefore natural to impose its variational parameters $\{\lambda_{\mu,\nu}^{a,b}\}$ invariant under any a supercell lattice vector translation:

$$\lambda^{\mathbf{R}+\mathbf{L}_m,\mathbf{R}'+\mathbf{L}_n}\, X_\uparrow^{\mathbf{R}+\mathbf{L}_m} X_\downarrow^{\mathbf{R}'+\mathbf{L}_n} = \lambda^{\mathbf{R},\mathbf{R}'}\, X_\uparrow^{\mathbf{R}} X_\downarrow^{\mathbf{R}'}. \tag{4.46}$$

Eq. 4.46 is valid for any pair of atoms $(\mathbf{R}, \mathbf{R}')$ in the simulation supercell.

In parallel, it follows from Eq. 4.12 that, when the atomic coordinates are translated

on the supercell lattice, the following transformation also holds:

$$\lambda^{\mathbf{R}+\mathbf{L}_m,\mathbf{R}'+\mathbf{L}_n}\, X_{\uparrow}^{\mathbf{R}+\mathbf{L}_m} X_{\downarrow}^{\mathbf{R}'+\mathbf{L}_n} = e^{-i\left[\boldsymbol{\theta}^{\uparrow}\cdot\mathbf{L}_m+\boldsymbol{\theta}^{\downarrow}\cdot\mathbf{L}_n\right]}\lambda^{\mathbf{R}+\mathbf{L}_m,\mathbf{R}'+\mathbf{L}_n}\, X_{\uparrow}^{\mathbf{R}} X_{\downarrow}^{\mathbf{R}'}. \tag{4.47}$$

If we combine Eqs. 4.46 and 4.47, we obtain the following transformation for the Λ matrix elements:

$$\lambda^{\mathbf{R}+\mathbf{L}_m,\mathbf{R}'+\mathbf{L}_n} = e^{\mathrm{i}\left[\boldsymbol{\theta}^{\uparrow}\cdot\mathbf{L}_\mathrm{m}+\boldsymbol{\theta}^{\downarrow}\cdot\mathbf{L}_\mathrm{n}\right]}\lambda^{\mathbf{R},\mathbf{R}'}, \tag{4.48}$$

which must hold for any pair of supercell translational vectors $(\mathbf{L}_m, \mathbf{L}_n)$. Eq. 4.48 provides a consistent definition of the pairing function matrix on the supercell lattice. Indeed, if this condition is verified, one can take an arbitrary pair of atomic positions $(\mathbf{R}+\mathbf{L}_m, \mathbf{R}'+\mathbf{L}_n)$ on the supercell lattice equivalent to the original ones $(\mathbf{R}, \mathbf{R}')$ for periodic boundary conditions, and obtain the same pairing function up to a phase factor. The supercell translation operations are compatible with the primitive cell ones, since the supercell is usually chosen as integer multiple of the primitive cell. This implies that Eq. 4.48 must be always satisfied for the real pairing function of the crystal and not only within the supercell.

In the special case of $\mathbf{L}_m = \mathbf{L}_n$, i.e. when translating the couple of atoms into another equivalent supercell, it directly follows from Eq. 4.48 that the translational invariance of the geminal function is respected if one chooses the twist conditions as:

$$\boldsymbol{\theta}_{\downarrow} = -\boldsymbol{\theta}^{\uparrow}, \tag{4.49}$$

which means that down spin particles have opposite phase with respect to up spin ones. This relation highlights the importance of having a general boundary condition framework as introduced already within DFT (Sec. 4.2).

Considering $\boldsymbol{\theta}_{\downarrow} = -\boldsymbol{\theta}^{\uparrow}$, we can rewrite the condition of Eq. 4.48 as:

$$\lambda^{\mathbf{R}+\mathbf{L}_m,\mathbf{R}'+\mathbf{L}_n} = e^{i\boldsymbol{\theta}^{\uparrow}\cdot(\mathbf{L}_m-\mathbf{L}_n)}\lambda^{\mathbf{R},\mathbf{R}'} = e^{i\boldsymbol{\theta}^{\uparrow}\cdot\mathbf{L}_l}\lambda^{\mathbf{R},\mathbf{R}'}, \tag{4.50}$$

where $\mathbf{L}_l = \mathbf{L}_m - \mathbf{L}_n$ is a vector which still belongs to the supercell lattice. For each pair of atomic positions $\mathbf{R}, \mathbf{R}'$ on the lattice it is thus possible to find a suitable value of $\mathbf{L}_l$ such that $|\mathbf{R}-\mathbf{R}'|$ is mapped into a reference supercell as follows:

$$|\mathbf{R}-\mathbf{R}'+\mathbf{L}_l(\mathbf{R},\mathbf{R}')| \leq \frac{L_i}{2} \quad \text{for } i \in (x,y,z). \tag{4.51}$$

Eq. 4.51 ensures that any couple of atoms $(\mathbf{R}, \mathbf{R}')$ are always translated within, at most, nearest neighbor supercell images of the chosen reference. This is achieved thanks to an appropriately chosen translational vector $\mathbf{L}_l(\mathbf{R}, \mathbf{R}')$ which actually depends on the considered pair of atoms. $\mathbf{L}_l(\mathbf{R}, \mathbf{R}')$ indirectly depends also on the electronic positions. In fact, in the valence bond picture, they can change the connections between atoms by

breaking and forming valence bonds connecting different $\mathbf{R}, \mathbf{R}'$ atoms with respect to the previous configuration. Therefore the translational vector $\mathbf{L}_l(\mathbf{R}, \mathbf{R}')$ must be recomputed after each single electron move.

If one enforces the parameterization of Eq. 4.51 on a generic phase-dependent pairing function (Eq. 4.48), the result is an effective geminal matrix which is independent of the chosen boundary condition. This key outcome derives from the fact that periodic nearest neighbor images do not have any gauge-dependent phase attached by definition and they are therefore purely translational invariant quantities.

Table 4.2 – Benchmark calculations on the solid bcc phase of hydrogen. We used a 64 atoms supercell and 8 twist conditions. The wavefunction Ψ is computed by separate optimization of each twist, while the wavefunction $\tilde{\Psi}$ is extracted from a calculation including all twist with the procedure explained in the text. Total energy difference $E_\Psi - E_{\tilde{\Psi}}$ in Ha, the overlap $\langle\Psi|\tilde{\Psi}\rangle$ and the corresponding statistical errors $\delta(E_\Psi - E_{\tilde{\Psi}})$ and $\delta\left[\langle\Psi|\tilde{\Psi}\rangle\right]$ have been computed using the correlated sampling technique. For all twists, the VMC total energy difference given by the two wavefunctions is lower than 2 mHa.

$\boldsymbol{\theta}$	$E_\Psi - E_{\tilde{\Psi}}$	$\delta(E_\Psi - E_{\tilde{\Psi}})$	$\langle\Psi\|\tilde{\Psi}\rangle$	$\delta\left[\langle\Psi\|\tilde{\Psi}\rangle\right]$
1	-0.00063	0.00024	0.98688	0.00012
2	-0.00096	0.00026	0.98509	0.00016
3	-0.00194	0.00030	0.98457	0.00016
4	-0.00162	0.00041	0.98260	0.00017
5	-0.00220	0.00040	0.98296	0.00034
6	-0.00154	0.00027	0.98514	0.00023
7	-0.00104	0.00044	0.98732	0.00022
8	-0.00099	0.00045	0.98787	0.00018

Using Eqs. 4.50 and 4.51 we can define a set of effective variational parameters for the pairing function as:

$$\lambda_{\text{eff}}^{\mathbf{R},\mathbf{R}'} = \lambda^{\mathbf{R}+\mathbf{L}_m(\mathbf{R},\mathbf{R}'),\mathbf{R}'}. \tag{4.52}$$

The derivatives required by the SR optimization procedure (Eq. 2.65) can be now written as a function of $\{\lambda_{\text{eff}}^{\mathbf{R},\mathbf{R}'}\}$ by applying the chain rule:

$$\begin{aligned} \frac{\partial \Psi_T(\mathbf{R}_{\text{el}})}{\partial \lambda_{\text{eff}}^{\mathbf{R},\mathbf{R}'}} &= \frac{\partial \Psi_T(\mathbf{R}_{\text{el}})}{\partial \lambda^{\mathbf{R},\mathbf{R}'}} \frac{\partial \lambda^{\mathbf{R},\mathbf{R}'}}{\partial \lambda_{\text{eff}}^{\mathbf{R},\mathbf{R}'}} = \\ &= \frac{\partial \Psi_T(\mathbf{R}_{\text{el}})}{\partial \lambda^{\mathbf{R},\mathbf{R}'}}\, e^{i\boldsymbol{\theta}^\uparrow \cdot \mathbf{L}_m} \end{aligned} \tag{4.53}$$

The left-hand-side of Eq. 4.53 is by definition a phase independent quantity. Therefore its computation strongly benefits from the statistics accumulated for each twist condition;

the statistical error on its evaluation is effectively reduced in this way, exactly as for the Jastrow variational parameters.

This leads to an optimized SR reconfiguration algorithm for dealing with single or multiple twist conditions with the same efficiency. As before, by including the atomic positions as variational parameters, this algorithm can be extended also to geometry relaxation. The main steps of the procedure are:

- The phase independent derivatives are constructed by attaching an appropriately chosen phase factor $e^{i\boldsymbol{\theta}^{\uparrow}\cdot\mathbf{L}_m}$ to the original derivatives for each twist condition.

- One step of the SR algorithm, modified for complex-valued pairing functions, is applied using the derivatives $\frac{\partial\Psi_T(\mathbf{R}_{\rm el})}{\partial\lambda_{\rm eff}^{\mathbf{R},\mathbf{R}'}}$. The statistical noise is effectively decreased by accumulating statistics per each boundary condition.

- The new optimized parameters for each twist condition are obtained by simply inverting Eq. 4.53. Now we can construct a new pairing function with the definition in Eq. 4.29 and perform the next SR step.

This algorithm is immediately applicable to the single determinant wavefunction by casting Eq. 4.29 in the MO representation and truncating the expansion to the first $N/2$ molecular orbitals.

We tested the efficiency of this effective parameterization by performing benchmark calculations on 64 atoms supercell of bcc hydrogen. We consider 8 twist conditions. We start from a previously optimized Jastrow factor and from the DFT(LDA) determinant for all twists. The primitive basis set used is $J[1s1p]D[2s]$ where D stands for determinant and J for the Jastrow factor.

We optimized the determinant orbitals for each twist condition *separately (8 calculations)* and also, using the same statistics, all boundary conditions *together (single calculation)* employing the phase-independent parameterization outlined in this Section. In Tab. 4.2, we present for each twist condition the VMC energy difference and the overlap between the wavefunctions Ψ and $\tilde{\Psi}$ obtained with the first and second methods respectively. The efficiency of our technique is apparent since the difference in the total VMC energy is lower than 2 mHa for all twists.

In conclusion to this Section it is worth mentioning some limitations of our parameterization. This method is not applicable in the presence of spin polarized systems with $N_\uparrow \neq N_\downarrow$. The unpaired electrons breaks the crystal lattice symmetries and it is evident that the basic geminal transformation derived in Eq. 4.48 is no longer valid. For similar reasons the gauge independent parameterization we proposed cannot be directly used if the electron pairing is described with the more complicated Pfaffian functional form [39].

The Pfaffian naturally accommodates a triplet state with an unpaired electron, thus it hampers us again from exploiting crystal lattice symmetries to derive Eq. 4.48. As we shall see in Chap. 6, the use of a Pfaffian can be crucial to properly account for spin fluctuations in the context of high temperature superconductivity, thus further investigations are required to overcome this inconvenience of the approach here outlined.

4.6 Conclusions

In this Chapter we presented the theoretical background and practical implementations of our QMC framework to the study of periodic solids.

We showed how to extend the single-particle localized orbitals to the case of periodic systems by testing a couple of distinct functional forms fulfilling the well-known Bloch theorem of one-electron functions. Furthermore, we implemented a generic boundary conditions framework which allows to apply a different boundary condition to $\uparrow$ and $\downarrow$ spin particles, a feature which turns out to be of paramount importance for efficient wavefunction optimization within QMC.

These schemes are extended to the case of QMC calculations. Due to the long-range Coulomb interaction, QMC calculations must be performed within supercells composed of several primitive cells. We discussed how the Bloch theorem can be extended to this situation and we defined the action of a phase factor (or twist condition) when applied to the many-body wavefunction. This has to be distinguished from the similar concept of a $\mathbf{k}$-point in effective single-particle theories.

In the last two sections, we focused on the evolution of the QMC algorithms within the *TurboRVB* code developed during this thesis work. We showed that the main complications arise from the need of a complex wavefunction to properly describe solid systems. The VMC algorithm is straightforwardly extended to this case. The backbone LRDMC algorithm can also be maintained with complex wavefunctions, but in order to project the procedure to the real axis not the sign, but the phase of the guiding wavefunction must be equal to the phase of the trial wavefunction.

Finally, we have presented the SR algorithm for wavefunction optimization as well as an original scheme to obtain optimized determinants in the presence of several twist conditions. By enforcing translational invariance with a proper choice of twist condition for $\uparrow$ and $\downarrow$ spin particles, we were able to parameterize the pairing function, the main building block of our ansatz, in a boundary independent way. This allows us to compute the derivatives over the variational parameters in the determinant by accumulating the statistics over all twist conditions included in the simulation, thus yielding the same efficiency as a single-twist calculation.

The theoretical and programming efforts presented in this Chapter allowed us to extend the capabilities of the *TurboRVB* code towards an accurate treatment of periodic solids and paved the way to the application of several techniques for reducing the systematic error arising from the finite size of the simulation supercell. Novel methodologies for curing these errors as well as several applications of the approaches presented here are the subject of the next Chapter.

Chapter 5

Alternative approaches to reduce finite size effects in QMC solid state calculations

5.1 Introduction

The simulation of an infinite periodic crystal using a finite number of particles introduces systematic errors into the results of the numerical experiment. These errors strongly affect all observables which do not converge rapidly with the system size, among them the total energy, structural parameters and correlation functions. Since in *ab initio* QMC calculations the statistical error can be reduced down to the desired precision if enough computer time is spent, finite size (FS) effects typically represent the most important source of error for solid state calculations.

The *ab initio* determination of different phases in a crystalline material typically requires the sub-chemical accuracy (< 0.01 eV/atom) in many-body solid state calculations. This has been verified in a vast amount of solids ranging from the simple crystalline benzene [150] to silicon [136] and polymorphic compounds [134]. FS errors are typically larger than this energy difference and they therefore prevent a reliable application of QMC to these problems. In addition to crystalline phases, magnetic phases can also be hard to distinguish without a proper treatment of FS effects; we will see an example in Chap. 6 where we present a careful investigation of FS effects on the high-temperature superconductor FeSe.

Effectively independent electron theories such as HF or DFT rely on the Bloch theorem to reduce FS effects thanks to a suitable integration of the Brillouin zone. As we explained in the previous Chapter, in principle this procedure allows to suppress all one-body FS

errors. Within a correlated approach such as QMC, the FS problem is more involved. The explicit presence of electron-electron correlation results in a classification of FS errors into two main categories which are distinguished by their physical origin.

The first one derives from the single-particle terms in the quantum Hamiltonian and it also appears within mean-field techniques. The one-body part of $\hat{H}$ is constituted by:

$$\hat{H}^{1b} = \hat{T} + \hat{V}_{ext}$$

where $\hat{T}$ is the kinetic energy and $\hat{V}_{ext}$ the external ionic potential. The latter is periodic for a perfect crystal and thus does not contribute to FS errors. Errors instead come mainly from fluctuations in the kinetic energy $\hat{T}$, as we will explain more in detail later on in this Chapter.

The many-body part of $\hat{H}$, namely the electron-electron interaction, constitutes a source of erro peculiar to correlated methods. The electron-electron interaction can be always decomposed as follows:

$$\hat{V}_{ee} = \hat{V}_{Ha} + \hat{V}_{XC}.$$

The Hartree term $\hat{V}_{Ha} = \int d\mathbf{r}' \frac{\rho(\mathbf{r})}{|\mathbf{r}-\mathbf{r}'|}$ only depends on the local electronic density $\rho(\mathbf{r})$ which is also a periodic quantity and rapidly convergent with the system size. Many-body FS errors come from the exact non-local exchange correlation potential $\hat{V}_{XC}$ whose dependence on the simulation cell size cannot be known *a priori*. The standard procedure for curing many-body FS effects consists in simulating the system at different supercell sizes and extrapolating the energy to infinite size, i.e., to the thermodynamic limit. In 3 dimensions, the leading order of the error as a function of the system size is $1/N$, while in lower dimensional systems different powers are more appropriate [151].

However, in the largest supercell the many-body ansatz is typically more difficult to optimize and the energy is not, in general, a simple analytic function of $1/N$. Beyond-leading-order corrections are necessary to alleviate this problem and improve the quality of the extrapolation procedure [152]. Numerous schemes have been devised to this purpose. Some of them are based on the random phase approximation (RPA) of the electronic structure factor [153, 151, 152]. In fact, at small $\mathbf{k}$ values it can be substituted with its RPA formulation to get a suitable correction applicable to finite-size systems. The rational behind this choice is that in the high-density regime (small $\mathbf{k}$ values) the RPA becomes exact. Another scheme is based on the generation of an *ad-hoc* density functional which can account for many-body errors directly at a DFT level and provide a simple post-processing correction to QMC energy data [154]. This is the procedure we used in this work and it will be elucidated more in detail later on.

Finally, it is worth mentioning also that in supercell simulations subjected to PBC the $1/r$ term of the Coulomb potential must be substituted with a different interaction to

obtain a convergent sum over periodic images. The model Coulomb potential (MPC) introduced in Refs. [155, 156] has been found to produce results which converge more rapidly with the system size with respect to the standard Ewald summations [157]. The other side of the coin is that MPC introduces an undesired distortion in the exchange correlation hole [151] which can affect the outcome of QMC simulations at small system sizes.

In this Chapter, our focus is on single-particle errors related to the kinetic energy part of the Hamiltonian. The other sources of FS effects are treated with standard methods which are discussed. In the last Section we will introduce and test a method which utilizes strategies similar to the one-body correction methods, but is designed to fully account for both single-particle and many-body effects. This approach is called flavor twisted boundary conditions (FTBC) and is particularly relevant since it has never been used before in realistic calculations. The common thread of all error correcting techniques presented in this Chapter is the use of twisted boundary conditions applied to the many-body wavefunction (see Sec. 4.3.2 for the definition of twist).

5.1.1 Organization of the Chapter

The first part of this Chapter is devoted to the theoretical foundations of single-particle error correcting methods. A detailed discussion on one-body shell effects is presented in Sec. 5.2. In Sec. 5.3 we describe the mainstream correction methodologies based on averaging over several twist conditions. In Sec. 5.4 an alternative method constructed with special values of the twist condition in the supercell Brillouin zone is presented. These special values are devised to reproduce the thermodynamically converged energy at a mean-field level, thus yielding the correct independent particle limit of the many-body wavefunction. We propose here an original procedure for extracting special twist values with an adjustable numerical accuracy which we dub the exact special twist (EST) method. We test it on the 3D homogeneous electron gas, the most important model for metallic systems. In Sec. 5.6 we apply EST procedure to realistic simulations of metals. We present benchmark calculations on both energy extrapolation within the VMC and LRDMC methods, and, a special case of correlation functions, on ionic forces. We show that our procedure is the best single-point solution to the one-body FS errors problem and is competitive with twist averaging techniques at the cost of a single twist condition. This is especially favorable for diffusion Monte Carlo calculations where most of the simulation time is spent in equilibration of the initial walker configurations. The main body of this Chapter concludes in Sec. 5.7, which is devoted to the flavor twist boundary conditions method. This approach is applicable thanks to our generic boundary conditions

implementation introduced in the previous Chapter. In fact, the application of FTBC involves averaging over a set of twists distinct for ↑ and ↓ spin particles. We apply this method to two paradigmatic metallic systems studied with the special twist method in order to test its efficiency against the standard twist averaging technique. We conclude this Chapter in Sec. 5.8.

5.2 Kinetic energy shell fluctuations

The simulation of a solid within a supercell subjected to PBC is equivalent to sampling the primitive cell BZ with a grid of **k**-points whose spacing is inversely proportional to the supercell dimension.

When increasing the supercell size in a correlated calculation in order to extrapolate to infinite size, the underlined **k**-point mesh in the primitive BZ becomes denser. The Bloch theorem states that the occupations of single-particle orbitals depend on the **k** vectors; this leads to the formation of a Fermi surface for metallic systems. In that case the orbital filling can change abruptly when varying the system size. Typically, this can be related to a change from a closed-shell to an open-shell character and vice-versa. This discontinuous change in orbital occupations gives raise to large fluctuations in the single-particle energy operators, namely in the kinetic energy (KE) part of the many-body Hamiltonian and, to a lesser extent, the one-body potential energy.

Due to their one-electron nature, these fluctuations are present already at the mean-field level and they are thus related only to the determinantal part of the many-body wavefunction. An alternative way of looking at the shell filling problem is to note that, within a single-particle theory, the KE contribution to the energy per particle is given by:

$$\mathcal{T} = \frac{1}{N}\frac{1}{N_k}\sum_{\mathbf{k}}\sum_{n=1}^{N_{\text{bands}}^{\text{occ}(\mathbf{k})}} \epsilon_{n,\mathbf{k}} - \sum_{i=1}^{N} V_{\text{eff}}(\mathbf{r}_i) \tag{5.1}$$

where N_k is the number of boundary conditions used and $\epsilon_{n,\mathbf{k}}$ are the electronic bands. If we focus on the DFT framework, the effective single-particle potential V_{eff} has been defined in Eq. 2.10. In the infinite size limit, the summation over discrete single particle levels is substituted by an integration as follows:

$$\frac{1}{N}\frac{1}{N_k}\sum_{\mathbf{k}}\sum_{n=1}^{N_{\text{bands}}^{\text{occ}(\mathbf{k})}\epsilon_{n,\mathbf{k}}\to\frac{\Omega}{N}\int_0^{\epsilon_F} d\mathcal{E}\ G(\mathcal{E})\ \mathcal{E},} \tag{5.2}$$

where Ω is the primitive cell volume and ϵ_F is the Fermi energy and the density of states

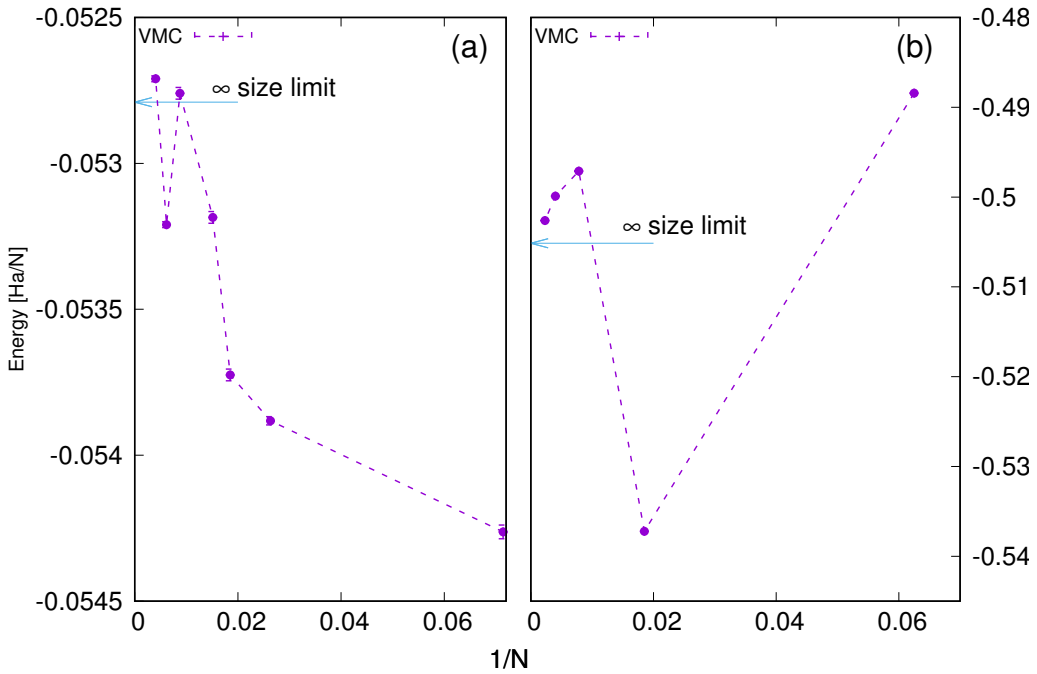
per unit volume $G(\mathcal{E})$ is defined as:

$$G(\mathcal{E}) = \frac{1}{(2\pi)^3} \sum_{n=1}^{N_{\text{bands}}^{\text{occ}(\mathbf{k})} \int d\mathbf{k}\ \delta(\mathcal{E}-\epsilon_{n,\mathbf{k}})}. \quad (5.3)$$

For metallic systems $G(\mathcal{E})$ keeps a finite value when approaching the Fermi energy ϵ_F due to the absence of gap, hence resulting in a sharp Fermi surface. This gives rise to large and uncontrollable integration errors when replacing the exact density of states with a simple summation over energy levels in the region close to ϵ_F. The shell filling effects on the KE are thus related to an incorrect sampling of the Fermi surface of the simulated material.

It is worth noting that additional FS errors on the KE arise from the long-range part of the many-body wavefunction associated with the Jastrow factor. This issue can be partially cured with the aforementioned RPA corrections [153], but its detailed discussion is beyond the scope of this thesis.

Figure 5.1 – PBC energy at different supercell sizes in the 3D electron gas (panel a, taken from Ref. [158]) and in the hydrogen bcc structure (panel b). On the x axis we report the inverse of the number of particles whereas the y axis shows the energy per particle in Ha. We notice in both cases a strongly fluctuating behavior when approaching the infinite size limit. This is mainly due to the change in shell filling at different simulation cell sizes.



We conclude this Section by showing the impact of shell filling effects through a practical example. In Fig. 5.1 we present energy versus $1/N$ curves on the 3D-HEG [158] (left

panel) and on the bcc structure of hydrogen (right panel) obtained at the VMC level with the Jastrow-single determinant wavefunction and without the application of any twist condition. We notice that in both cases shell filling effects give rise to a highly discontinuous behavior in the energy versus size curve. It is therefore impossible to perform any kind of extrapolation to estimate the infinite size limit of the energy.

5.3 Twist averaging techniques

Shell fluctuations can be considerably reduced by averaging over several twisted boundary conditions. The concept of twist has been introduced in Sec. 4.3.2; if we restrict, for simplicity, to a cubic supercell of size L, the allowed twist conditions in units of $1/L$, when periodic boundary conditions are applied to the supercell, are comprised in the following interval:

$$-\frac{\pi}{L} \leq \theta_i \leq \frac{\pi}{L} \quad \text{for } i \in (x, y, z) \tag{5.4}$$

The single-particle thermodynamic limit of an observable f defined on the crystal can be obtained by integrating over the whole set of allowed twist conditions. Therefore, the Brillouin zone summation defined in Eq. 4.4 for mean-field methods is equivalent to a twist integration as follows:

$$\frac{\Omega}{(2\pi)^3} \int_{\Omega_\mathrm{P}} d\mathbf{k} f(\mathbf{k}) \equiv \frac{1}{(2\pi)^3} \int_{-\pi/L}^{\pi/L} d\boldsymbol{\theta}\ f(\boldsymbol{\theta}), \tag{5.5}$$

where $f(\boldsymbol{\theta}) = \left\langle \Psi^{\boldsymbol{\theta}}(\mathbf{R}_\mathrm{el}) \right| f \left| \Psi^{\boldsymbol{\theta}}(\mathbf{R}_\mathrm{el}) \right\rangle$ and Ω_P is the volume of the BZ corresponding to the primitive cell.

The twist averaging technique for reducing single-particle FS effects was first introduced in the study of lattice models [159, 160, 161]. Its most important implementation within QMC is the standard twist-averaged boundary conditions (TABC) approach introduced in Ref. [162], which is nowadays the method of choice for most published works using QMC techniques.

In TABC the integral over twists in Eq. 5.5 is substituted by a summation over a finite number of twist conditions $\boldsymbol{\theta}$. It is obvious that a larger number of twists corresponds to a lower integration error. For choosing the quadrature, the same methods employed for $\mathbf{k}$-points sampling in DFT (see Sec. 4.2.5) can be used. It has been demonstrated [151] that a uniform Monkhorst-Pack [142] mesh suitably offsett from the origin ensures the most rapid convergence to the thermodynamic limit. A set of random twists is also often employed [163].

The most important feature of TABC is that averaging over twist conditions allows an effective reduction of the statistical noise on the computed quantities. In fact, if N_t twist conditions are used, the statistical error in TABC is given by:

$$\sigma_{\mathrm{TABC}} = \frac{1}{\sqrt{N_t}}\sigma[\boldsymbol{\theta}], \tag{5.6}$$

where $\sigma[\boldsymbol{\theta}]$ is the mean statistical error on a single twist. Notice that at a fixed amount of computer time, the attained $\sigma[\boldsymbol{\theta}]$ is roughly the same for all twists except the ones which can use a real-valued wavefunction which has a lower computational cost. Eq. 5.6 holds if one neglects all equilibration costs needed to thermalize the random walk when the twist condition is varied. This is an additional amount of time which decreases the TABC efficiency since no averages can be computed during this period. While in VMC the equilibration time is negligible, within DMC it is usually sizable and it can take a large part of the total simulation time. Therefore, for DMC, it is more efficient to use only a few twist conditions or rely on the alternative special twist method introduced in Sec. 5.4.

When using standard TABC there is, however, an important caveat: the number of particles is kept fixed for all twist conditions included in the average. TABC thus works in the canonical ensemble. However, as explained in Sec. 4.2.5, the correct independent particle thermodynamic limit is obtained by sampling the Fermi surface within the grand canonical ensemble with variable number of particles. If this is not verified, at each twist only the N single-particle orbitals with lowest energy are considered; therefore, even in the limit of an infinite number of twists, the standard TABC method does not sample the right Fermi surface in the independent-particle limit. This introduces a systematic error in the evaluation of kinetic energy; this bias slowly goes to zero with a power law $1/N^{\nu}$ [162, 151] when approaching the infinite supercell size limit.

The correct independent particle limit of the kinetic energy and momentum distribution can be restored by allowing the number of particles to vary among twist conditions [161, 162, 164, 152]. This method goes under the name of grand canonical TABC (GTABC). The number of particles associated with each twist can be simply determined analytically in the case of an isotropic Fermi surface, such as in the non-interacting homogeneous electron gas. For realistic metallic systems, the filling of single-particle states at each twist must be determined at a mean-field level using HF or DFT calculations. Notice that in doing so, it is assumed that the Fermi surface does not appreciably vary among the reference mean-field method and QMC. A similar assumption is made also when applying the special twist method detailed in Sec. 5.4 to correlated calculations. Although the validity of this hypothesis cannot be verified, one can indirectly probe it, for the systems under consideration, with the simple method detailed in Sec. 5.6.3.

The other side of the coin is that the variable number of particles reintroduces strong kinetic energy fluctuations [151]. The width of these fluctuations is typically larger than the kinetic energy bias introduced by the standard TABC. Furthermore, the many-body wavefunction needs to be optimized from scratch *separately* for each boundary condition. This strongly affects the overall efficiency of the method.

All the results obtained with twist averaging methods presented in this Chapter employ the standard TABC in the canonical ensemble.

5.4 Special twist methods

Special twist (ST) techniques provide an alternative approach for reducing FS effects due to shell filling of single-particle orbitals and, at the same time, maintaining the correct independent particle limit of the total energy. Here we introduce the theoretical foundations of ST methods within correlated QMC calculations and we propose a simple numerical procedure for finding special twist values within the supercell BZ which reproduce the mean-field thermodynamic limit with an adjustable numerical accuracy.

5.4.1 Theoretical foundations

As explained in the previous Chapter, within a mean-field framework as DFT the total energy of the infinite system is computed as a BZ integral:

$$\mathcal{E}_\infty = \frac{\Omega}{(2\pi)^3} \int_{\Omega_{\rm BZ}} d\mathbf{k}\, \mathcal{E}(\mathbf{k}), \tag{5.7}$$

where the energy $\mathcal{E}(\mathbf{k})$ has been already summed over the electronic bands. Here our focus is on the energy since it is the most basic quantity to evaluate in QMC. The mean-value theorem for definite integrals ensures the existence of a special point $\mathbf{k}^*$, the so-called mean-value point, for which the integrand in Eq. 5.7 equals the integral:

$$\mathcal{E}(\mathbf{k}^*) = \mathcal{E}_\infty. \tag{5.8}$$

The above relation is an *exact* alternative to the full BZ integration in Eq. 5.7. The validity of above theorem is restricted to continuous integrands. This is always the case for an insulator, wherease discontinuities in $\mathcal{E}(\mathbf{k})$ can arise in metals making Eq. 5.8 an approximation only.

Eq. 5.8 is valid within a single-particle theory, but it can be effectively exploited in correlated calculations for reducing one-body FS effects. Here we describe the theoretical foundations of the ST method for QMC. The numerical implementation we propose in

this thesis will be explained in the last part of this Section and is generally applicable to any correlated many-body method. As usual, we consider a supercell with N_S electrons described by the direct lattice vectors $\mathbf{L}_S$. This supercell is taken as an integer multiple of the primitive cell of the solid containing $N_P \leq N_S$ electrons. The basic idea behind the ST method, inspired by the aforementioned concept of a mean-value point, is to find a twist condition $\boldsymbol{\theta}_s$ which, in the single-particle limit of the many-body wavefunction, reproduces the exact mean-field infinite-size energy. The practical implementation of this idea consists in finding a solution to Eq. 5.8 within the given supercell. Let us consider a metallic system described by a single-particle wavefunction such as the single determinant Ψ_{SD} introduced in Eq. 2.11. In the previous Chapter (Sec. 4.2), we saw that the energy per particle within the Kohn-Sham approach can be written as:

$$\begin{aligned} \mathcal{E}_\infty[\rho_\infty] &= \frac{1}{N_p N_k} \sum_{i=1}^{N_k} \sum_{n\in\mathrm{occ}} \left\langle \Psi_{\mathrm{SD}}^{n,\mathbf{k}_i} \right| \mathcal{H}^{\mathrm{DFT}}[\rho_\infty] \left| \Psi_{\mathrm{SD}}^{n,\mathbf{k}_i} \right\rangle = \\ &= \frac{1}{N_p} \sum_{i=1}^{N_k} w_i \sum_{n\in\mathrm{occ}} F_\beta(\mathcal{E}_n^{\mathbf{k}_i}[\rho_\infty] - \mu_\infty)\, \mathcal{E}_n^{\mathbf{k}_i}[\rho_\infty], \end{aligned} \tag{5.9}$$

where we used the same notation as Eq. 4.16 omitting the spin index for the sake of simplicity. In Eq. 5.9, the energy is computed within a primitive cell and the number of particles can vary for each boundary condition through the smearing function F_β; thus $\mathcal{E}_\infty$ is a grand canonical quantity.

We would like now to reproduce the infinite size energy in Eq. 5.9 with a single special twist value $\boldsymbol{\theta}_s$. The DFT energy per particle computed for this special twist now reads:

$$\begin{aligned} \mathcal{E}_{\boldsymbol{\theta}_s}[\rho_{\boldsymbol{\theta}_s}] &= \frac{1}{N_s} \left\langle \Psi_{\mathrm{SD}}^{\boldsymbol{\theta}_s} \right| \mathcal{H}^{\mathrm{DFT}}[\rho_{\boldsymbol{\theta}_s}] \left| \Psi_{\mathrm{SD}}^{\boldsymbol{\theta}_s} \right\rangle = \\ &= \frac{1}{N_s} \sum_n H(\mathcal{E}_n^{\boldsymbol{\theta}_s} - \mu_{\boldsymbol{\theta}_s}) \mathcal{E}_n^{\boldsymbol{\theta}_s}[\rho_{\boldsymbol{\theta}_s}], \end{aligned} \tag{5.10}$$

where H is the Heaviside step function, $\rho_{\boldsymbol{\theta}_s}$ is the electronic density calculated for the special twist value only, and the chemical potential $\mu_{\boldsymbol{\theta}_s}$ is such that the effective number of electrons is equal to the true number of particles in the system. At variance with $\mathcal{E}_\infty[\rho_\infty]$, this energy is now computed within the QMC supercell and, furthermore, no fractional electron occupations are allowed. $\mathcal{E}_{\boldsymbol{\theta}_s}[\rho_{\boldsymbol{\theta}_s}]$ is thus computed in the *canonical ensemble* with a fixed number of particles.

Given this formalism, finding the special twist solution to Eq. 5.8 is equivalent to satisfying the following equality:

$$\mathcal{E}_{\boldsymbol{\theta}_s}[\rho_{\boldsymbol{\theta}_s}] = \mathcal{E}_\infty[\rho_\infty]. \tag{5.11}$$

Eq. 5.11 can be interpreted as a way to find the energy in the canonical ensemble which best approximates the correct thermodynamic limit evaluated within the grand canonical

ensemble. Once Eq. 5.11 is solved, it is possible to build a many-body wavefunction with the chosen twist value which automatically fulfills the correct (at the DFT level) independent-electron limit, at variance with the standard TABC method. As we will show in Sec. 5.6, this approach leads to a large reduction of the one-body FS errors in QMC calculations, whilst keeping at the same time an affordable computational cost.

In conclusion, it is important to remark that we use DFT(LDA) as a reference for single-particle theories.

5.4.2 The Baldereschi method

Eq. 5.11 can be solved approximately using the approach devised by Baldereschi in Ref. [165].

The Baldereschi method is based upon the fact that, within an independent electron framework, any intrinsic property of the crystal is periodic in the direct lattice identified by the vectors $\mathbf{L}_m$. This is valid for a primitive cell simulation box as well as for a supercell compatible with the primitive cell translations. The symmetry operations associated with the Bravais lattice of the crystal form a group referred as Γ_1. As usual, we consider the total energy as the reference observable. Belonging to the representation Γ_1 of the crystal point group imposes a set of constraints in the Fourier expansion of the energy which can be expressed as a linear combination of symmetrized plane waves as follows:

$$\mathcal{E}_{\boldsymbol{\theta}} = \sum_{m=0}^{\infty} a_m G_m^{(\Gamma_1)}(\boldsymbol{\theta}). \tag{5.12}$$

In Eq. 5.12 the index m spans the stars of equivalent lattice vectors within the Γ_1 representation, identified by $\Gamma_1(\mathbf{L}_m)$ and the plane waves are defined as $G_m^{(\Gamma_1)}(\boldsymbol{\theta}) = \exp\left[-i\boldsymbol{\theta}\cdot\Gamma_1(\mathbf{L}_m)\right]$ with $G_0^{(\Gamma_1)} = 1$. Notice that the functions $\{G_m^{(\Gamma_1)}\}$ only depends on the underlined crystal structure and not on the specific observable considered.

By inserting Eq. 5.12 in the BZ summation in Eq. 5.7, we notice that all terms of the summation except the first give a vanishing contribution, i.e. $\mathcal{E}_\infty = a_0$. This simply derives from the following plane waves property:

$$\int_{\Omega_{\mathrm{BZ}}} d\boldsymbol{\theta}\, e^{(-i\boldsymbol{\theta}\cdot\Gamma_1(\mathbf{L}_m))} = (2\pi)^3\delta(\Gamma_1(\mathbf{L}_m)),$$

where δ is the Dirac delta function.

Therefore, the special twist $\boldsymbol{\theta}_s$ is retrieved by requiring that all $G_m^{(\Gamma_1)}$ vanish except the first one. Baldereschi pointed out that this condition can be satisfied in an approximate way by requiring that only the first n Fourier components of Eq. 5.12 go to zero:

$$G_m^{(\Gamma_1)}(\boldsymbol{\theta}_s) = 0 \quad \text{for } m \in [1, n]. \tag{5.13}$$

n is generally chosen equal to 3 since we are seeking the 3 components of the special twist angle: $\boldsymbol{\theta}_s = (\theta_x^s, \theta_y^s, \theta_z^s)$. The Baldereschi method is a simple procedure for finding special twist values which best approximate the true mean-value point in either primitive or supercell calculations. However, the approximation in Eq. 5.13 is valid only if the observable has a smooth behavior in the reciprocal space. In this case, the higher Fourier components of the symmetrized plane wave expansion give a negligible contribution to the total sum. This is verified for insulators, whereas for metallic systems with a sharp Fermi surface, the approximation is not satisfactory and the Baldereschi method gives poor results when used for FS error correction. In the following we go beyond the Baldereschi approximation by proposing a procedure which can yield exact values of the special twist down to a desired numerical accuracy. We will see that the performance of our method is far superior to the simpler Baldereschi point for reducing one-body FS effects.

5.5 Exact special twist procedure

5.5.1 Detailed procedure

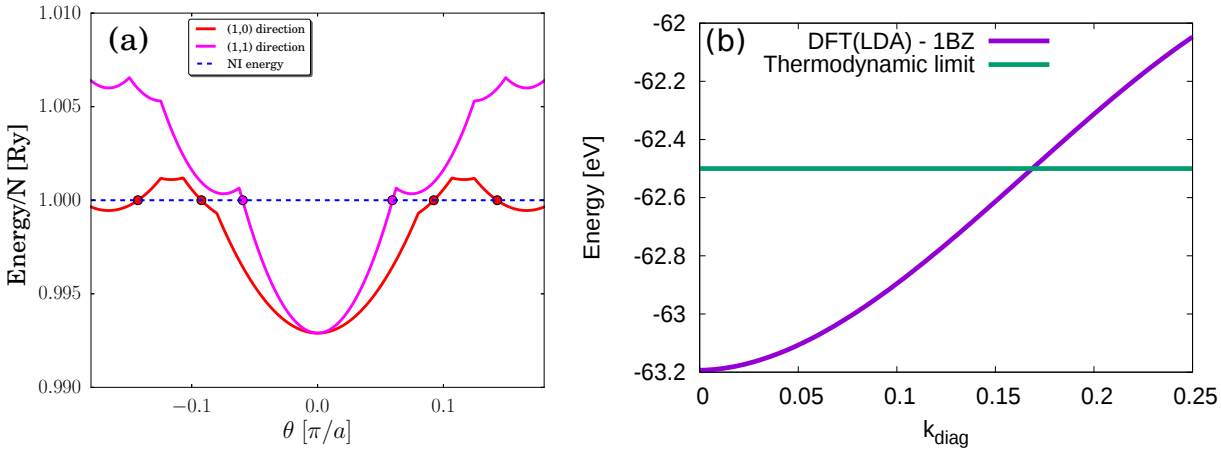
In this thesis we propose an alternative to the Baldereschi approximation which aims at finding an exact *numerical* solution of Eq. 5.11. The first and – to our knowledge – unique attempt to tackle this problem can be found in Refs. [144, 143]. Here the authors argued that a suitable $\boldsymbol{\theta}_s$ can be chosen in the set $\{\mathbf{G}_\mathrm{s}/2\}$, where $\mathbf{G}_\mathrm{s}$ are the supercell reciprocal lattice vectors. This choice ensures that the underlined $\mathbf{k}$-points mesh possesses inversion symmetry, thus allowing to employ a real-valued wavefunction and avoid complex arithmetics. The offset belonging to this set which provides the best thermodynamic limit is then determined via cheap DFT calculations in the LDA approximation at different supercell sizes.

At variance with the previous works, our procedure is not limited to real-valued wavefunctions thanks to the formalism developed in the previous Chapter. Our method, which we dub exact special twist (EST), allows to find the right special twist with an adjustable numerical accuracy. Furthermore, at variance with the Baldereschi method, it is system dependent and it can be used for any type of observable ranging from the total energy to structural parameters and correlation functions. The detailed steps of the EST are explained in the following:

1. At first we determine the thermodynamic converged energy $\mathcal{E}_\infty$ (Eq. 5.9) within an independent-particle approach. For the 3D-HEG calculations presented in this Section, the reference energy is the non-interacting (NI) energy: $\mathcal{E}_\infty^{\mathrm{HEG}} = \sum_{\mathbf{k}\in\mathrm{occ}} \mathbf{k}^2/2 \simeq \frac{2.21}{r_s^2}$ whose value is completely controlled by the Wigner-Seitz radius r_s, i.e. by

the electronic density. For realistic QMC runs (Sec. 5.6), we evaluate it at the DFT(LDA) level using a fully converged **k**-point mesh [142] and with the same basis set as QMC. All these calculations are carried out in the primitive cell with a negligible computational cost as compared to QMC.

Figure 5.2 – Panel (a): Energy landscape of the non-interacting (NI) 2D homogeneous electrons gas with 90 particles at a Wigner-Seize radius of $r_s = 1$ a.u. as a function of the twist angle θ in π/L units, where L is the cubic box parameter. Two representative directions in the Brillouin zone, (1,0) and (1,1), are shown. The exact value of the NI energy is represented by the dashed line. The energy surface presents some cusps which are due to discontinuous changes in the occupations of the electronic states. We notice that several twist conditions in both directions match the value of $\mathcal{E}_\infty^{\mathrm{HEG}}$. Panel (b): in analogy with the 2D homogeneous electron gas, for the DFT(LDA) case we scan a diagonal direction of the first Brillouin zone. In this plot we present a sample calculation on β-tin Si with a 16 atoms supercell. The diagonal direction is identified by a single parameter $\mathbf{k}_{\mathrm{diag}}$ again expressed in π/L units. The straight line indicates the value of the thermodynamically converged limit at the same level of theory.



2. The second step consists in solving numerically Eq. 5.11. In the case of the HEG, we select several directions in the BZ and we scan the reciprocal space along these directions in order to find the value of the twist $\boldsymbol{\theta}_s$ giving the exact non-interacting energy $\mathcal{E}_\infty^{\mathrm{HEG}}$. In Fig. 5.2a the simple case of the two dimensional electron gas with $r_s = 1$ a.u. is shown. The same procedure can be straightforwardly applied to the three dimensional electron gas that is the subject of the next Section. Similarly, in the case of realistic QMC calculations (Fig. 5.2b), we pick a direction in the first Brillouin zone (in general along a diagonal, thus characterized by just one parameter) and scan the DFT(LDA) band structure by computing energies at each twist on a uniform grid along the chosen direction. These runs must be performed within the same supercell used for QMC (Eq. 5.10) as explained at the beginning of this Section. We select the value which reproduces the thermodynamic limit within an energy range smaller than the accuracy required by QMC calculations.

In particular, for the metallic systems presented in Sec. 5.6, we choose an accuracy of the order of 0.001 eV/atom in determining the special twist, since we have a target precision of $\sim$ 0.01 eV/atom in the corresponding QMC runs. Notice that the selected twist will almost certainly require a complex-valued determinantal part in the QMC variational ansatz. In contrast with Refs. [144, 143], this is the case for all calculations presented in Sec. 5.6.

We notice from Fig. 5.2 that several twist values might satisfy the special twist condition in Eq. 5.11. In fact, the EST forms a surface in the reciprocal space; since the Fermi surface of the single-particle reference method is likely to be different with respect to the correlated QMC one, it is in principle possible that different choices of EST would give different total energies, thus posing a serious issue on the validity of our method. In Tab. 5.1 we demonstrate that, in practical calculations, this is not verified. In fact, it is apparent that for both 3D-HEG and realistic calculations the outcome at the VMC level is independent on the BZ direction chosen to compute the EST value. For all considered points on the EST surface, the VMC energies displays at most a difference of 3σ, thus they can be considered in statistical agreement.

Furthermore, we verified the dependence of the special twist on the reference single-particle framework. In order to do so, we apply our numerical procedure to β-tin Si using both DFT(LDA) and the simpler Hartree method (obtained imposing $V_{\mathrm{XC}} = 0$ instead of the full DFT exchange-correlation functional). The results are presented in Tab. 5.2, where we also present the energy differences between thermodynamically converged and EST calculations for both methods. Interestingly, we found that for all supercell size the EST value varies only very slightly when passing from one method to the other. This finding suggests that the single-particle reference framework does not represent a crucial issue for the EST determination; a logic explanation is that the EST is not too sensitive to small changes in the Fermi surface, as confirmed also from the results in Tab. 5.1. Since a simple Hartree calculation is extremely cheap and does not require any self-consistency, one can simply construct an automatic procedure for finding the exact special twist.

3. Once the special twist value $\boldsymbol{\theta}_s$ is found, we perform a final DFT(LDA) supercell calculation with the selected twist. The resulting wavefunction is used as the determinantal part of the total QMC ansatz.

4. Given the JSD wavefunction built in the previous steps, we carry out the Jastrow optimization using the real flavor of the stochastic reconfiguration method discussed in Chap. 2. Both the linear coefficients and the Gaussian exponents of the Jastrow

Table 5.1 – We report the VMC energies of several values of the exact special twist chosen following different directions in the reciprocal space. The EST values $(\theta_x, \theta_y, \theta_z)$ are in crystal coordinates with respect to the orthorhombic supercell used in the calculations. Calculations are presented for the 3D-HEG with 54 particles (top) as well as for bcc-Li structure and β-tin Si (bottom), two realistic metals which will be thoroughly investigated in Sec. 5.6. We used a 16 atoms supercell for bcc-Li and a supercell containing 8 atoms for β-tin Si. Energies are expressed in Ry for the 3D-HEG results and in eV/atom for the realistic calculations. It is apparent that for all considered values on the EST surface, the resulting energies at the VMC level are in optimal statistical agreement.

3D-HEG

BZ direction	r_s [a.u.]	$\mathcal{E}^{\mathrm{EST}}_{\mathrm{JSD}}$
(1,1,1)	10	-0.106446(12)
(1,1,0)	10	-0.106418(13)
(1,1,1)	1	1.15536(18)
(1,1,0)	1	1.15513(20)

DFT(LDA)

$(\theta_x, \theta_y, \theta_z)$	System	$\mathcal{E}^{\mathrm{EST}}_{\mathrm{VMC}}$
(0.2638, 0.2638, 0.2638)	Li	-7.1667(7)
(0.2122, 0.5000, 0.2122)	Li	-7.1684(7)
(0.1785, 0.2500, 0.5000)	Li	-7.1663(12)
(0.1362, 0.3181, 0.5000)	Li	-7.1675(8)
(0.2118, 0.2118, 0.2118)	Si	-106.5213(58)
(0.2500, 0.2500, 0.1961)	Si	-106.5225(58)
(0.4167, 0.0833, 0.2500)	Si	-106.5277(61)
(0.2500, 0.2118, 0.2118)	Si	-106.5211(54)

are optimized. In the case of the bcc Li structure we test the effectiveness of the special twist method in predicting structural properties by performing full QMC crystal cell relaxations. In this case, the QMC determinant is also fully optimized using the twist-independent method introduced in Sec. 4.5.2. Molecular dynamics benchmark calculations on liquid hydrogen are also performed for the same purposes.

5. The final QMC energy is evaluated with the VMC and LRDMC methods using the JSD ansatz. The procedure is repeated at different supercell sizes in order to perform an extrapolation to the infinite size limit. For realistic calculations we extrapolate the results with a linear polynomial in $1/N$, which should include the correct leading order of our curves. For the HEG case we use a quadratic polynomial

Table 5.2 – Comparison of the special twist values extracted with the procedure explained in the text using DFT(LDA) or simple Hartree theory as independent particle reference methods. In the first two rows we show the energy difference – in eV/atom – between the fully converged Hartree and DFT(LDA) energies and the EST energy at the same level of theory. The difference is always kept of the order of 0.001 eV/atom. In the last two rows we report the values of the special twists. We choose them along a diagonal direction in the BZ for the both methods. DFT(LDA) and simple Hartree theory values of EST show a very good agreement for all supercell sizes.

		# of Si atoms			
		8	16	64	96
energy difference	Hartree	-0.0012	-0.0023	0.0016	-0.0009
	DFT(LDA)	-0.0031	-0.0021	0.0011	0.0018
$\boldsymbol{\theta}_s$ value	Hartree	0.2096	0.1862	0.2618	0.2644
	DFT(LDA)	0.2123	0.1835	0.2612	0.2625

in $1/N$ instead.

In conclusion, our procedure for finding special twist values is cheap and its accuracy can be adjusted depending on the considered physical problem. The way we select the special twist ensures the correct independent-particle limit for the many-body QMC wavefunction without relying on the grand canonical ensemble formalism as required by the TABC technique. Besides these features, our approach also possesses the advantages of employing only a single twist: the computational cost of an EST calculation is thus about 2.5 times the cost of a simple gamma point simulation.

5.5.2 Benchmark calculations on 3D homogeneous electron gas

The homogeneous electron gas in 3 dimensions (3D-HEG) is certainly the most studied benchmark model for correlated metallic systems. Its importance is not limited to the model level; it also constitutes the basis for building the local density approximation routinely employed in density functional theory[16].

In this work, we simulated the 3D-HEG at an electronic density corresponding to a Wigner-Seitz radius of $r_s = 10$ a.u. This value has been used in several published works [162, 153] carried out with the TABC method, and it is therefore convenient for the sake of comparison. In Fig. 5.3 we present the FS size extrapolation of the HEG total energy per electron as a function of the inverse number of particles. We analyze the performance of the EST method by comparing it with simple PBC [158] and with TABC calculations [162] both carried out with the same Slater-Jastrow trial wavefunction as ours. The arrow in Fig. 5.3 indicates the infinite size limit as presented in the

original manuscript in Ref. [162]. We did not apply any many-body FS corrections to our results. We can immediately notice that both the EST and TABC methods are effective in suppressing shell fluctuations as the number of particles grows. Both methods yield a very smooth curve and we can easily extrapolate our EST results to the infinite size limit using a quadratic polynomial fit. In principle the next leading correction is decreasing as $N^{-4/3}$ in 3D. However, being this exponent quite close to the leading one, it is difficult to see it in the considered range of supercells. We have found much better fit using N^{-2} as next leading-order correction. Such choice is justified in the case that the corresponding coefficient is rather small, namely the curvature in the corresponding fits is almost negligible. In fact, in this range of sizes the $\mathcal{O}(N^{-2})$ contribution dominates over the $\mathcal{O}(N^{-4/3})$ and $\mathcal{O}(N^{-5/3})$ ones, which become eventually more relevant only for much larger number of particles. Furthermore, we can directly match the thermodynamic limits of the variational energy yielded by the two methods. The TABC and EST results are in agreement within the statistical error bar $\sigma \sim 2 \times 10^{-5}$ Ry, the extrapolated EST energy being $-0.10558(2)$ Ry/N, with $-0.10561(5)$ Ry/N the corresponding TABC energy.

At variance with the other methods, the shell fluctuations in the PBC energies are too large to perform any extrapolation, as expected. The difficulty is apparent also if one considers the infinite-size estimate of $-0.10549(2)$ Ry/N reported in Ref. [158]. We notice that it still displays a discrepancy of the order of $\sim 5\sigma$ with respect to the EST and TABC infinite size limits. This disagreement is likely due to residual one-body FS errors dependence which cannot be suppressed despite the very large number of particles employed for this simulation.

In Fig. 5.4 we present an additional comparison of our EST method with a calculation performed using the GTABC approach [153]. At variance with the TABC results presented in Fig. 5.3, a backflow correlated variational wavefunction is used for the GTABC results. This ansatz gives a better energy and produces the rigid shift in the thermodynamic energy estimate present in Fig. 5.4.

In Fig. 5.4 we also report the EST results (blue line) corrected with the aforementioned procedure based on the small wavelengths RPA structure factor [153] for curing many-body errors in both the kinetic energy and the interaction parts of the Hamiltonian. The comparison is interesting since it proves the effectiveness of combining many-body corrections with our approach. We notice that both EST and GTABC corrections drastically improve the estimation of the energy at small numbers of particles, leading to a convergence below 1 mRy when including more than 54 electrons in the simulation.

Figure 5.3 – Comparison of different method for alleviating FS effects in the 3D electron gas. Energies are computed with the VMC method. The orange arrow indicates the infinite size limit as presented in Ref. [162]. The solid line through the EST results is obtained with a simple quadratic polynomial fit.

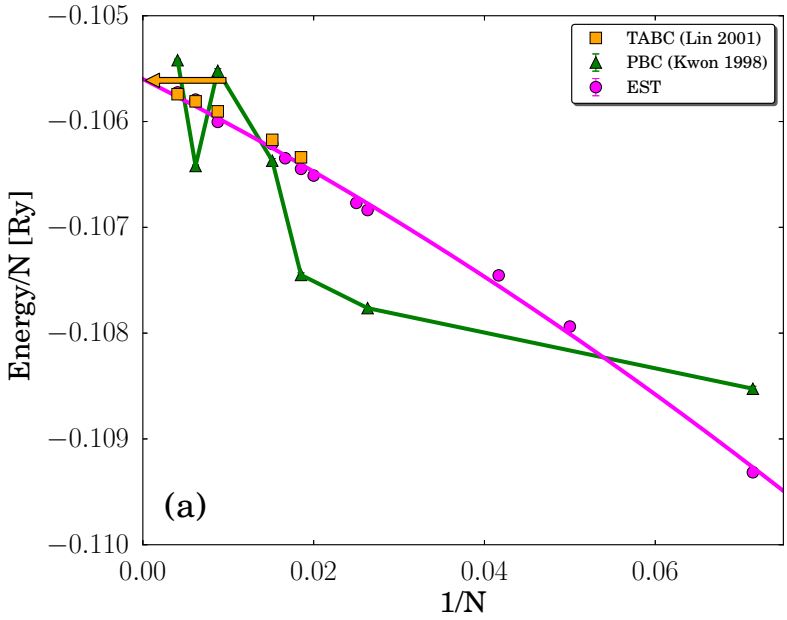


Figure 5.4 – Comparison of the EST method with the GTABC results taken from Ref. [153]. The blue line shows the EST points corrected for many-body errors [153]. The effectiveness of many-body error corrections combined with our EST method is shown. Notice that there is a rigid shift between the EST and GTABC thermodynamic limits since the GTABC points are obtained with a backflow correlated wavefunction which ensures a lower variational energy.

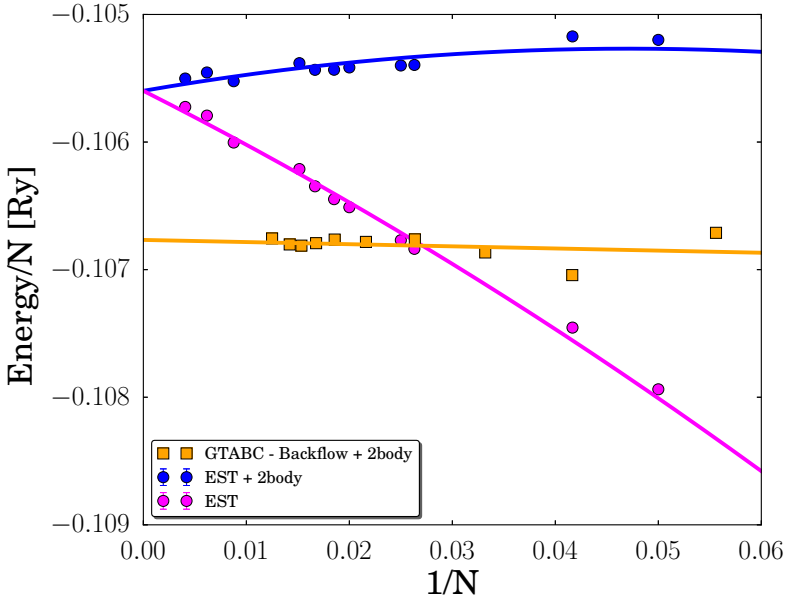
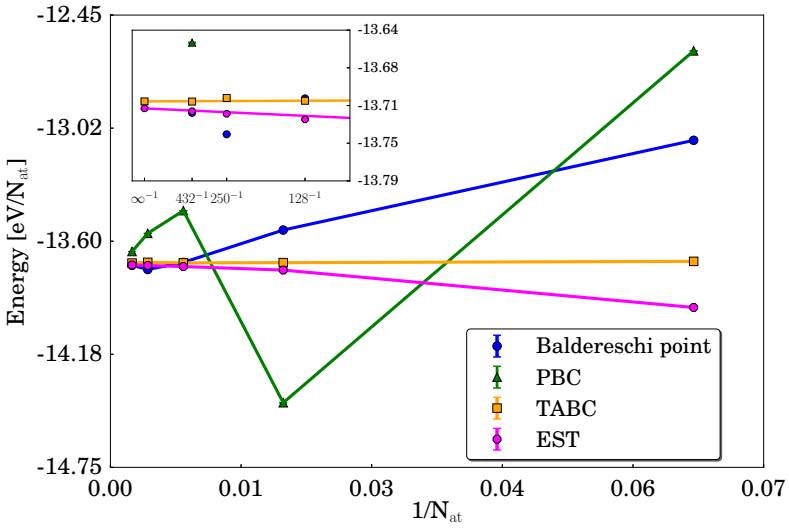


Figure 5.5 – VMC extrapolation to the infinite-size limit for solid bcc structure of hydrogen at a Wigner-Seitz radius $r_s = 1.32$. Results are corrected for two-body errors with the KZK method [**?**]. The x axis reports the inverse number of atoms in the supercell. In the inset we zoom into the results for the largest supercells in order to appreciate the thermodynamic limit convergence. The x axis reports the inverse number of atoms. For EST and TABC values we also show in the inset a linear fitting curve – as well as its energy extrapolation at $N = \infty$ (leftmost point labelled as ∞^{-1}) – carried out by excluding the smallest supercell with 16 atoms.



5.6 Realistic systems with the EST method

This Section contains an analysis of FS effects for several realistic systems by means of the EST method previously introduced. We test it against the use of simple periodic boundary conditions (PBC), the standard twist-averaged boundary conditions (TABC) technique [162] explained in Sec. 5.3, and a different special twist determined with the Baldereschi method [165] (see Sec. 5.4), which we dub Baldereschi point.

At first, we analyze the effectiveness of the EST method in extrapolating the total energy to the infinite size limit. The last part of Sec. 5.6 analyzes FS effects on correlation functions using the EST method. In particular, we present results on the bcc lithium lattice constant evaluated with a zero temperature structural relaxation of QMC ionic forces. Finally, we report some benchmark calculations on hydrogen-hydrogen pair distribution function extracted from QMC based molecular dynamics simulations at high temperature [166, 139, 8].

5.6.1 Many-body errors correction

If not otherwise specified, the energies are corrected for many-body FS effects using the KZK energy functional introduced in Ref. [154]. The DFT(KZK) calculations are per-

formed with the built-in DFT code of the *TurboRVB* package. This allows to consistently use the same simulation setup for all calculations needed by the EST method.

The two-body FS corrections [136] we apply to the total QMC energy read:

$$\epsilon_{2b}^{\mathrm{KZK}} = \mathcal{E}_{\mathrm{LDA}}^{N_s} - \mathcal{E}_{\mathrm{KZK}}^{N_s}, \tag{5.14}$$

where both LDA and KZK energies are computed within the same N_s electrons supercell and the same twist(s) condition(s) as the corresponding QMC calculation. We verified for some sample cases that the application of the Chiesa corrections [153] leads to very similar results.

The effectiveness of the KZK method for reducing many-body FS effects is tested against error estimations directly extracted from the QMC calculations, as we will see in detail in Sec. 5.6.3.

5.6.2 Total energy

In this Section we analyze the FS effects on the energetics of three paradigmatic metallic systems of increasing degrees of complexity. We believe these systems constitute an exhaustive testing ground for the EST method and they pave the way for applying EST to more complex compounds. For TABC calculations we choose a uniform Monkhorst-Pack[142] mesh offsett from the Γ point of the supercell Brillouin zone. The number of independent twist conditions is reduced using the point group symmetry operations of the supercell lattice. In order to ensure convergence, the mesh size is varied at each supercell size such that the corresponding number of inequivalent atoms times the number of twists is kept constant and appropriately large. As a premise, it is important to remark that we carried out linear extrapolations excluding from the fitting procedure the smallest supercells for all considered systems. In fact, for these sizes, the residual shell effects is still too high to carry out the linear fit. Our choice is also motivated by the fact that QMC production runs will unlikely use such small supercell sizes for extrapolation purposes, although they are useful in our benchmark calculations for determining the overall behavior of the size convergence. The points used for the fitting curve are specified for each system we address.

The first metallic system we address is solid hydrogen in the bcc structure. Despite being the simplest element in the periodic table, hydrogen displays very intriguing properties and its phase diagram under pressure is far from being completely understood. In particular, in the region up to $\sim$ 300 GPa, solid hydrogen undergoes numerous phase transitions displaying exotic quantum properties which are not well characterized yet either experimentally or theoretically [167, 168, 169]. It is well established that FS effects

Figure 5.6 – Energy extrapolation on Li in the high temperature bcc phase. We show VMC results (panel a) and LRDMC energies (panel b). The energies are compared with different techniques. Two-body corrections are applied with the KZK functional approach. In the inset, a zoom of the results for the largest supercells is shown. The x axis reports the inverse number of atoms. For EST and TABC values we also show in the inset a linear fitting curve – as well as its energy extrapolation at $N = \infty$ (leftmost point labelled as ∞^{-1}) – carried out by excluding the smallest supercell with 16 atoms.

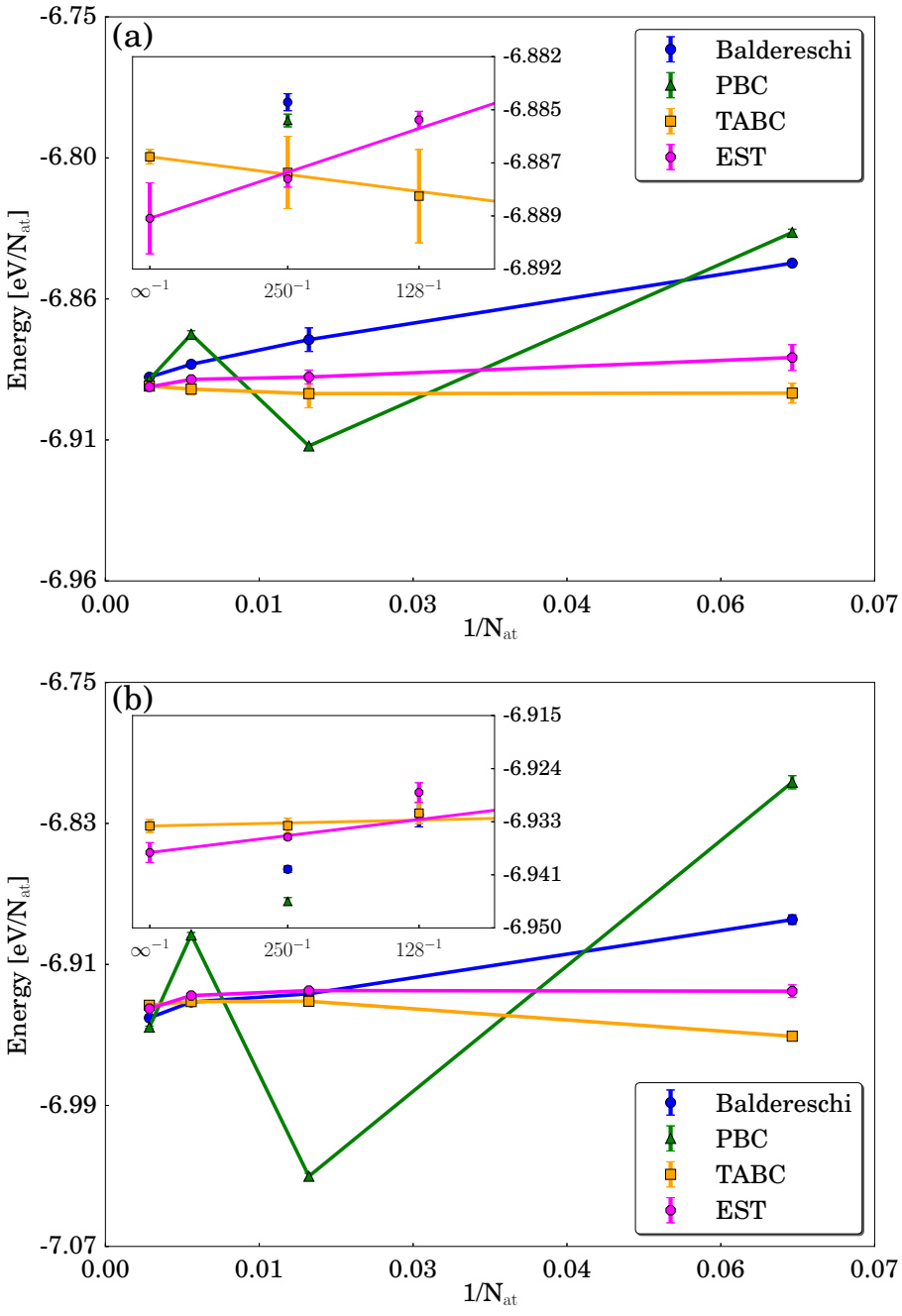
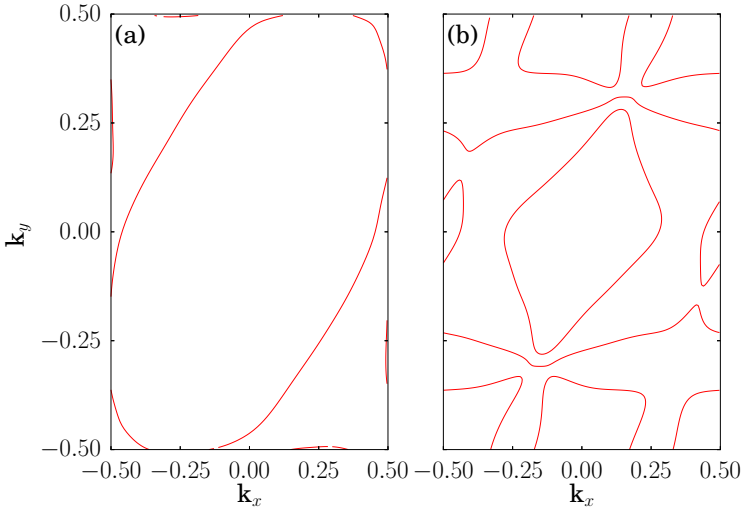


Figure 5.7 – Fermi surface contour plots along the $k_z = 0$ plane for bcc-Li (left panel) and β-tin Si (right panel). The contour has been taken at $k_z = 0$. The calculations have been performed with the software package Wannier90[171] based on DFT(LDA) results obtained with the Quantum ESPRESSO[18] program.



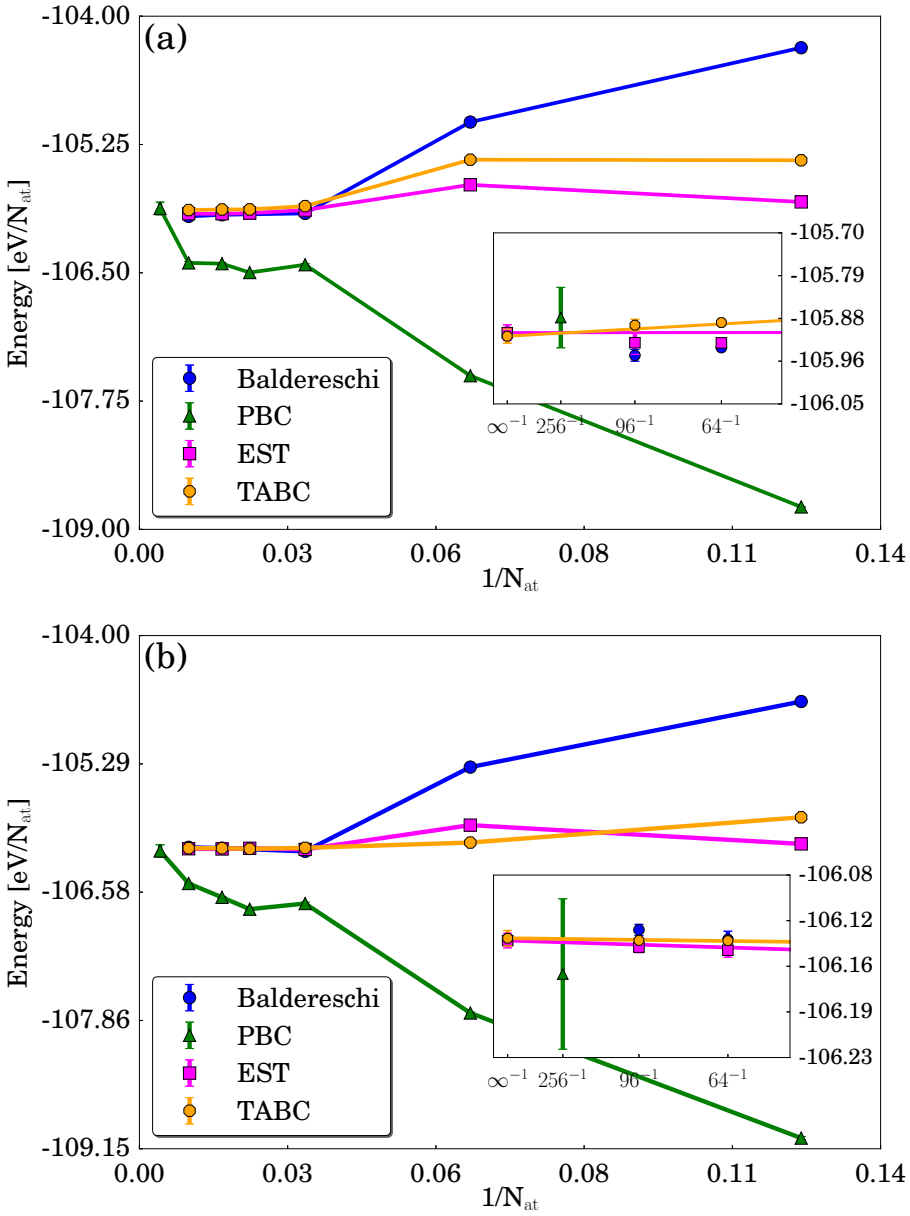
represent an important source of error in many-body simulations and the size of the simulation supercell is crucial for obtaining accurate correlation functions in molecular dynamics simulations of liquid hydrogen, as we will show in Sec. 5.6.4. Here we study the bcc structure of solid hydrogen which has not been observed in nature yet, but is one of the candidate structures for the high-pressure atomic phase due to its dense packing of atoms [170].

We use a primitive Gaussian basis set of $[J](2s)[D](2s)$, where J refers to the Jastrow and D to the determinantal part. The exponents of the determinant are taken from a previous fully optimized calculation [139]. Despite its small size, this basis has been proven accurate in describing both energetics and the most important correlation functions [166, 139] of hydrogen. The FS extrapolation at a the VMC level is presented in Fig. 5.5. As apparent, both TABC and EST results show a very smooth convergence toward the thermodynamic limit, indicating that most of the shell fluctuations have been eliminated.

By performing extrapolation to infinite size, we obtain a total energy of -13.7117(38) eV/atom for TABC and -13.7197(43) eV/atom using our EST method. These two FS correction methods are therefore in excellent agreement for a simple, but relevant system such as hydrogen. If the Baldereschi point [165] is used to offset the twist grid, the energy fluctuations are mostly suppressed, but the extrapolation procedure does not yield a satisfactory result (see the inset of Fig. 5.5).

We turn now our attention to metallic bcc lithium. Bulk Li has been the subject of

Figure 5.8 – Finite-size extrapolation with the VMC (panel a) and LRDMC (panel b) methods on Si in the high pressure β-tin phase. Again a comparison is shown among various FS correction methods and the inset reports a magnified view close to the thermodynamic limit. Two-body errors are corrected with the KZK method. The x axis reports the inverse number of atoms. For EST and TABC values we show in the inset a linear fitting curve – as well as its energy extrapolation at $N = \infty$ (leftmost point labelled as ∞^{-1}) – obtained by excluding the smallest (8 and 16 atoms) supercells.



intense studies due to the emergence of exotic quantum states, including superconductivity [172, 173] in its phase diagram under pressure and also to its extensive application in battery development. Previous QMC investigations [174, 175] provided very accurate results and, at the same time, established the important role of FS effects in determining the converged ground state energy [175, 163]. For treating this system the localized basis set used in this work is $[J](2s2p)[D](4s4p)$ and the $1s$ core electron is replaced with a BFD pseudopotential [52]. A comparison among several FS correction methods is presented in Fig. 5.6a for VMC and Fig. 5.6b for LRDMC energies. Both the TABC and EST methods, in combination with KZK corrections, ensure an almost complete suppression of energy fluctuations and provide a well converged result already for the 54-atom supercell, at variance with the Baldereschi point which displays a much slower convergence with supercell size. In order to fully appreciate the convergence to the thermodynamic limit, a zoom into the results for the largest systems is reported in the inset. The final extrapolated results obtained using TABC and EST are in agreement to 0.002 eV/atom for both VMC and LRDMC. This value is of the order of the attained statistical error. These results demonstrate that the TABC and EST methods offer similar performance in controlling FS effects in this system, although the former displays a slightly flatter curve.

The last system we address for benchmarking our method is the high pressure β-tin structure of silicon. Upon application of a pressure of around 12 GPa, Si displays a structural phase transition from the semiconductor diamond phase to a β-tin metallic phase. The transition develops on a very narrow energy scale [176] and standard DFT techniques yield unsatisfactory and functional-dependent results. Due to its sensitivity, this phenomenon is a perfect ground for benchmarking advanced first principles methods such as QMC and it has been extensively studied [177, 178, 136, 179, 180]. The tiny energy scale ($\sim$ 0.05 eV/atom) to be probed in order to spot the correct transition pressure requires a very accurate control of finite-size effects [136]. Metallic Si offers a perfect playground for testing the reliability of the EST method when tackling systems with complex and discontinuous Fermi surfaces. To be more explicit, in Fig. 5.7 we show a comparison of the LDA Fermi surfaces of bcc-Li and β-tin Si, where the contour has been taken along the $k_z = 0$ plane. We notice that the Li Fermi surface (Fig. 5.7a) displays practically no features except for a large electron pocket centered at the Γ point. β-tin Si (Fig. 5.7b) is considerably more challenging, in particular as a result of the small electron pockets present at the Brillouin zone borders. As already presented in the Li case, we report FS extrapolation in Si with KZK corrections for curing two-body FS effects. The final results obtained with the VMC and LRDMC methods are shown in Fig. 5.8a and Fig. 5.8b respectively. The EST method gives excellent results, comparable to the more expensive TABC in eliminating shell filling effects. In this particular case, the energy

curve obtained with the EST method is flatter than the TABC one. However, we notice that the energy of the 16-atom supercell (second point from the right) is shifted towards higher energies. We believe that this issue is related to the particularly poor sampling of the Fermi surface for that set of **k**-points. This fact together with the relatively small number of atoms in the supercell – which corresponds to a coarse **k**-points mesh – may give rise to the shift observed in EST results. We verified, at the DFT level, that this shift can be partially recovered by using a rotated supercell with lower symmetry. In Fig. 5.8 we do not include this point in the extrapolation procedure.

The infinite-size extrapolations of the TABC and EST total energies are in agreement up to 0.005 eV/atom using the VMC method, a value below the attained statistical error. A larger difference is apparent when using the LRDMC; beyond-leading-order extrapolation might be more appropriate in this case. However, when the largest supercell (96 atoms for EST, Baldereschi and TABC; 256 atoms at Γ point) is used, all results, independently from the method used, are converged up to 0.01 eV/atom for all techniques, an accuracy small enough to obtain a correct transition pressure [136]. We did not estimate this transition pressure since this has already been done with the same method as ours in Ref. [136] and, moreover, our main concern here is testing the EST method.

5.6.3 Comparison of errors in the EST method

In this Section we present a more quantitative discussion on the impact of FS effects in the special twist approach.

By definition (Eq. 5.11), the EST method cancels out all FS errors derived from the one-body contribution at mean-field level. However, if one switches electron correlation on in QMC, the Fermi surface can vary from the single-particle estimation, thus reintroducing some one-body finite-size effects to the EST-corrected results. Their size can be estimated via the TABC technique, which ensures a denser sampling of the Fermi surface. Thus, we provide an estimation of this residual contribution to the one-body FS errors directly within QMC. For this estimate VMC is the method of choice as it likely provides an outcome similar to LRDMC concerning FS effects and, since it is much cheaper, it can be used in production runs for correcting the LRDMC energy results. The residual one-body error at the VMC level reads:

$$\epsilon_{1\mathrm{b}}^{\mathrm{VMC}} = \mathcal{E}_{\mathrm{VMC}}^{\mathrm{TABC},N_s} - \mathcal{E}_{\mathrm{VMC}}^{N_s}, \tag{5.15}$$

where $\mathcal{E}_{\mathrm{VMC}}^{N_s}$ is the EST energy in a N_s-atom supercell, whereas $\mathcal{E}_{\mathrm{VMC}}^{\mathrm{TABC},N_s}$ is the correspondent fully converged TABC result in the same supercell. Results for Li and Si are presented in Tab. 5.3 (4th column); they constitute an indirect probe of the changes in

	ϵ_{2b}^{KZK}	ϵ_{2b}^{VMC}	ϵ_{1b}^{VMC}
# of atoms	bcc-Li		
16	0.1958	0.198(7)	-0.013(6)
54	0.0378	0.061(5)	-0.006(5)
128	0.0245	0.026(5)	-0.003(2)
250	0.0125	0.013(5)	0.0002(17)
	β-tin Si		
16	0.3545	0.185(6)	0.026(2)
48	0.1181	0.096(7)	0.036(7)
64	0.0886	0.067(6)	0.042(6)
96	0.0574	0.042(8)	0.035(2)

Table 5.3 – Comparison of many-body FS effects estimated with different methods for both bcc-Li and β-tin Si . All presented results are in eV per atom. Many-body errors are compared between the KZK and VMC corrections, both detailed in the text. We notice an overall good agreement between the two estimations for both Li and Si, except the case of the 16 atoms Si supercell which displayed similar issues also in the energy extrapolation (see Fig. 5.8).

the Fermi surface when going from the DFT to the QMC level. In the bcc-Li case, one can see that the one-body residual corrections are very small (one order of magnitude smaller than the many-body effects) and decrease fast when the size is increased. This implies that, in this case, the unknown exact value of the EST is supposedly very close to the one obtained at the DFT(LDA) level. This is further confirmed by the fact that the EST value changes only slightly from simple Hartree to DFT(LDA) mean-field estimates (see Tab. 5.2). In other words, the EST is rather insensitive to the underlying theory used to determine it, which makes the EST evaluation quite robust.

On the other hand, we notice that for β-tin Si the behavior of ϵ_{1b}^{VMC} is less systematic as a function of the system size. This could be related to the limitation of the TABC approach used to estimate the one-body corrections at the VMC level. Indeed, as already mentioned, TABC works in the canonical ensemble and can introduce a bias in the energy values due to the wrong **k**-point occupations, that can be particularly severe in the case of β-tin Si where the Fermi surface is much more complex than in the Li case (see Fig. 5.7).

Similarly to the one-body part, the residual many-body contribution to FS errors can also be evaluated at the VMC level. In Tab. 5.3 we compare the many-body error estimations obtained with the standard KZK method in Eq. 5.14 (2nd column) and directly

within VMC (3rd column) using the relation:

$$\epsilon_{2b}^{\mathrm{VMC}} = \mathcal{E}_{\mathrm{VMC}}^{\mathrm{extr}} - \mathcal{E}_{\mathrm{VMC}}^{\mathrm{TABC,N_s}} \tag{5.16}$$

where $\mathcal{E}_{\mathrm{VMC}}^{\mathrm{extr}}$ is the VMC energy extrapolated to the infinite-size limit. Notice that this extrapolation is obtained using KZK corrected values; however, the infinite size limit must obviously be the same.

Tab. 5.3 demonstrate that KZK and direct VMC estimations are in good agreement concerning many-body FS errors for both Li and Si, thus supporting the use of the cheap the KZK approach for production QMC runs. The only relevant discrepancy is for the 16-atom Si supercell which is a particularly delicate case both for both the EST and TABC methods, as previously mentioned.

In conclusion, it is important to remark that the special twist values used for Li and Si calculations have been determined using DFT(LDA) energies, and could be slightly different from the values obtained directly with the DFT(KZK) functional. This discrepancy could lead to some spurious contributions to the KZK estimate of the many-body errors. However, we verified that the variation of the EST between the two functionals is negligible, in line with what said before; thus, $\epsilon_{2b}^{\mathrm{KZK}}$ is considered as purely many-body and directly comparable with $\epsilon_{2b}^{\mathrm{VMC}}$.

5.6.4 Energy derivatives

In Sec. 5.6.2 we demonstrated the reliability of the EST method in extrapolating QMC energies to the infinite-size limit. However, as already pointed out in Ref. [162], sampling the Fermi surface with a single point might not be sufficient to account for more sensitive properties of the system such as the potential energy or correlation functions. In this Section we focus on testing the EST method with a particularly important type of correlation function: the ionic forces. Their evaluation within QMC has been the subject of intense research activity due to the intrinsic difficulty of finding an efficient and finite variance algorithm for computing many-body energy derivatives. As already mentioned, the AAD technique [72] introduced in Chap. 2 offers a solution to this issue. With our approach it is thus possible to perform both zero-temperature structural relaxation and molecular dynamics simulations based on QMC forces[166, 139, 7].

In Tab. 5.4 we present the optimization of the cell parameter of bcc Li carried out with full QMC force minimization. The EST results are obtained with the same value of special twist previously used for energy extrapolation.

Thanks to its denser sampling of the Fermi surface, the TABC method performs better than single-twist methods, with the cell parameter already converged for the 54 atoms

Table 5.4 – Lithium cell parameters obtained from VMC structural relaxation at different supercell sizes. The EST results are compared with standard PBC calculations and with the most accurate TABC method. We report also values obtained from fully converged DFT calculations in the LDA approximation, performed with the Quantum ESPRESSO[18] program using a $15 \times 15 \times 15$ **k**-point mesh and norm conserving pseudopotentials. Experimental cell parameters are also shown.

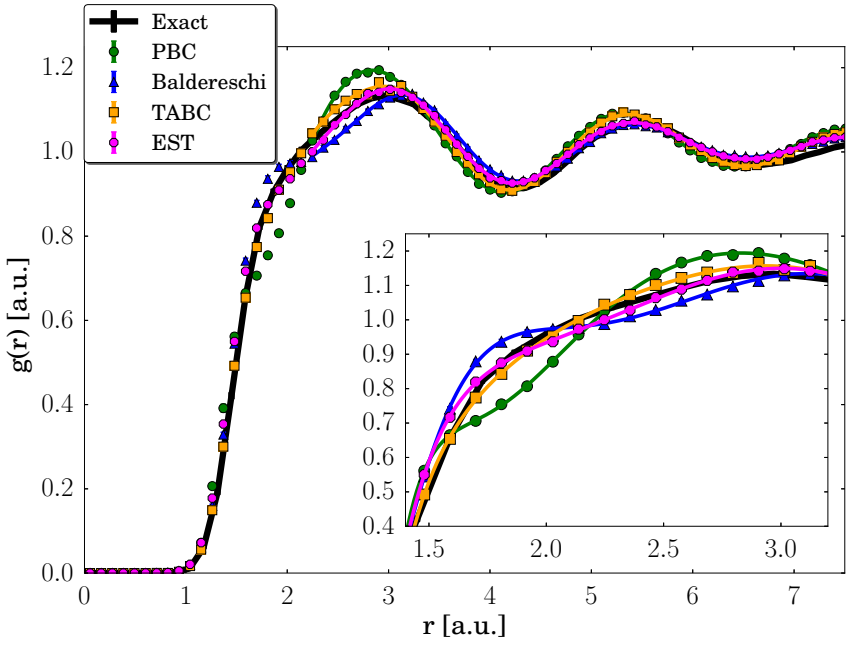
# of atoms	Cell parameter [Å]		
	PBC	EST	TABC
16	3.497(6)	3.457(4)	3.454(4)
54	3.469(3)	3.476(3)	3.496(3)
128	3.521(3)	3.505(3)	3.502(2)
250	3.510(5)	3.506(2)	3.499(3)
DFT(LDA)	3.3537		
EXP[181]	3.482		

supercell. The EST method is slightly slower to converge towards the infinite-size limit than the TABC method, but it displays a much smoother behavior with respect to simple PBC calculations, thus allowing an easy extrapolation to infinite size. By performing a linear extrapolation we obtain 3.508(3) Å for TABC and 3.504(5) Å for EST, which are in statistical agreement. The residual discrepancy with the experimental value [181] ($\sim$ 0.025 Å) could be due to temperature effects which are not taken into account by our calculations. However, the QMC results are already a substantial improvement with respect to DFT(LDA) calculations.

The lower computational cost of the EST method makes this approach the appropriate choice in the case of structural relaxation of more complex crystal cells, requiring the use of large supercells or the simultaneous optimization of several structural parameters, which would be computationally infeasible using the TABC technique, unless one uses the twist-independent parameterization introduced in Sec. 4.5 of this thesis. The latter technique allows us, with certain limitations, to optimize the wavefunction by collecting the statistics of all twist conditions used in TABC for an efficient computation of the energy derivatives.

The last part of this Section is devoted to benchmark calculations on the radial pair distribution function ($g(r)$) of liquid hydrogen. The $g(r)$ is extracted from QMC-based molecular dynamics (MD) simulations at a temperature of 1800 K and an estimated pressure of $\simeq$ 260 GPa. Forces are computed with the AAD technique and the MD is carried out with the methods introduced in Refs. [166, 139]. At these conditions, in our simulations, liquid hydrogen is in the atomic phase [166]. This phase is metallic, hence

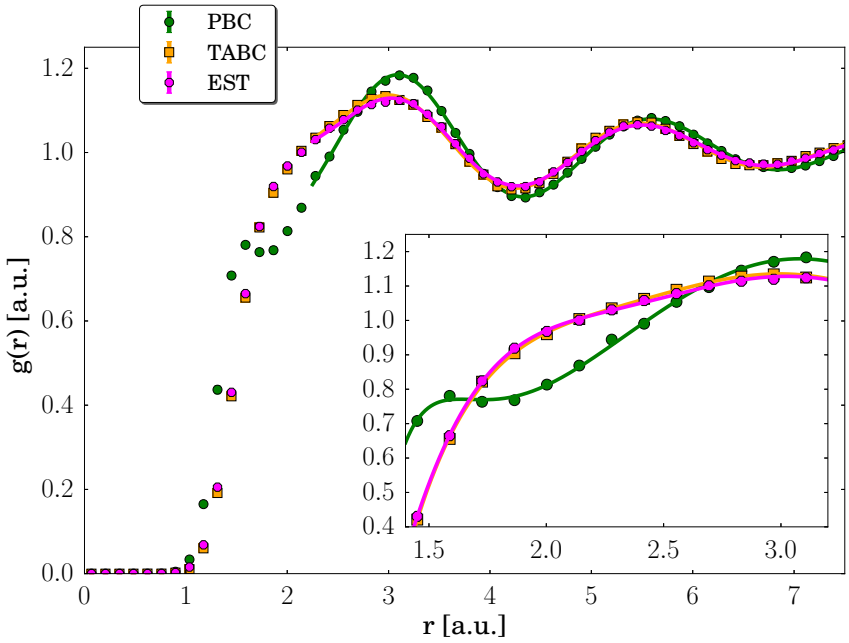
Figure 5.9 – Hydrogen radial pair distribution function extracted from a molecular dynamics simulation of liquid hydrogen at a density given by $r_s = 1.36$ a.u. For this plot we use a 64-atom supercell. We compare our EST method with the standard PBC, with the Baldereschi point and with TABC results performed with a $4 \times 4 \times 4$ uniform mesh (64 twists). For the sake of comparison, we also added a simulation with TABC using a 128-atom supercell. We consider this result as our best estimate and we labelled it as "exact" (with respect to size effects). The solid lines are obtained with a polynomial interpolation as guide for the eye. Despite small discrepancies in the $g(r)$ around the peaks, TABC and EST display an overall good agreement each other as well as with the more accurate result obtained with the 128-atom supercell. We notice instead some spurious features in the results obtained with the Baldereschi point, in particular in the region zoomed in the inset ($r \in [1.5, 3.2]$).



FS effects are likely to be important and need to be reduced in order to obtain a reliable description of the system.

The evaluation of the special twist for liquid hydrogen MD is more challenging. Due to its disordered nature, the system can be considered spherically symmetric and thus one can assume that the Baldereschi point for cubic systems $(1/4, 1/4, 1/4)$ would provide a good approximation for the special twist, since the disorder favors the lowest Fourier components of Brillouin zone integrals. In order to investigate this issue better, we extract several configurations from a previous molecular dynamics simulation carried out with the same conditions and apply the EST procedure. The special twist is obtained by averaging among all the special values found. In the case of the 64 atoms supercell, we found that the special twist has no relevant fluctuations among different configurations and that it tends to a value of $(1/4, 1/4, 0)$, in contrast with the initial assumption; we conclude that within this relatively small supercell the system tends to break the spherical (cubic) symmetry

Figure 5.10 – Same as Fig. 5.9 but performed with a larger supercell containing 128 hydrogen atoms. In this case the disorder of the system prevails and the system can be considered spherically symmetric; for this reason the special twist coincides with the Baldereschi point. We notice that at this supercell size the EST and TABC curves are practically indistinguishable, while PBC is still far from the other two methods, especially in the description of the first peak around $r = 3.2$ a.u.



given by its disordered nature. On the contrary, for the 128-atom supercell we found a special twist very close to the cubic Baldereschi point, as expected. This demonstrates that this supercell is sufficiently large to correctly account for the disordered nature of liquid hydrogen.

In Fig. 5.9 the $g(r)$ is reported for the 64-atom supercell. The overall behavior of the EST curve is in good agreement with the TABC curve with only small discrepancies around the peaks. This is not the case for simulations carried out with the Baldereschi point, where we notice an anomalous feature close to the first peak at $r \sim 2.8$ a.u. An excellent agreement between the two methods is instead obtained with a 128 atoms supercell, as shown in Fig. 5.10 where TABC and EST (Baldereschi) curves perfectly superpose. At variance with the EST method, even for this larger supercell, the $g(r)$ extracted from PBC calculations qualitatively differs from the more accurate TABC result.

The EST method remains reliable for the evaluation of ionic forces, in both simple structural relaxation and pair distribution functions extracted from MD simulations. However, accordance with the more expensive TABC technique is achieved only when relatively large supercells are employed. Hence, if one has to compute correlations functions using the EST method, a careful assessment of the impact of supercell size is needed before

starting production runs.

5.7 Flavor twist method

In this Section we test the *flavor-twisted* boundary condition (FTBC) method. This technique represents an alternative approach to the problem of FS errors in correlated calculations; in fact, the sole common thread with the previous methods is the use of twisted boundary conditions in the many-body wavefunction. At variance with the other techniques presented in this Chapter, FTBC is constructed to provide an effective remedy not only to the kinetic energy fluctuations, by means of an appropriate twist-averaging, but also to a part of the many-body FS effects, thanks to an approximate mapping from a large and in principle infinite simulation cell to a smaller and tractable system. Despite being a promising approach in this sense, the application of FTBC to condensed matter physics has been very limited and it appears – to our knowledge – only in Ref. [182].

In fact, the general idea of FTBC has been mostly employed in the domain of high-energy physics [183, 184]. In this context, FTBC enable the study of quantum chromodynamics (QCD) lattice models with allowed quantized momenta different from the ones imposed by the standard use of plain PBC. The purpose is again to reduce volume effects and study the behavior of the quantum fields at different momenta.

Following Ref. [182], let us consider a simulation setup consisting of a cubic supercell of total size L subjected to PBC. Electrons in this simulation cell are described by the Hamiltonian denoted with $\hat{H}_L$. The basic FTBC idea can be better understood by expressing the many-body electronic Hamiltonian (Eq. 2.2) in Fourier space. The standard second quantization expression for this operator reads:

$$\begin{aligned}\hat{H}_L = &\sum_{\mathbf{G},\sigma} \epsilon^{\sigma}_{\mathbf{G}} C^{\dagger}_{\mathbf{G},\sigma} C_{\mathbf{G},\sigma} + \\ &+ \frac{1}{L} \sum_{\mathbf{Q}} \left\{ \sum_{\mathbf{G},\sigma} \sum_{\mathbf{G}',\sigma'} U^{\sigma,\sigma'}_{\mathbf{G},\mathbf{G}',\mathbf{Q}} C^{\dagger}_{\mathbf{G}+\mathbf{Q},\sigma} C^{\dagger}_{\mathbf{G}'-\mathbf{Q},\sigma'} C_{\mathbf{G}',\sigma'} C_{\mathbf{G},\sigma} \right\},\end{aligned} \tag{5.17}$$

where the operator $C_{\mathbf{G},\sigma}(C^{\dagger}_{\mathbf{G},\sigma})$ annihilates (creates) a particle with momentum $\mathbf{G}$ and spin σ; $\{\epsilon^{\sigma}_{\mathbf{G}}\}$ and $\{U^{\sigma,\sigma'}_{\mathbf{G},\mathbf{G}',\mathbf{Q}}\}$ are the bare energy dispersion (one-body part of the Hamiltonian) and the interaction function (two-body part of the Hamiltonian) respectively. The vectors $\{\mathbf{G}\}$ and $\{\mathbf{G}'\}$ belong to the reciprocal lattice of the large supercell of size L.

The wavevector $\mathbf{Q}$ represents the momentum transferred between two electrons when they interact through the particle-particle Coulomb potential and it also belongs to the reciprocal lattice of the larger system.

Starting from the large system with size L, one can then select a smaller subsystem of

size $l < L$. Without loss of generality, we can assume that this subsystem is contained within the initial supercell and therefore the ratio L/l is a positive integer.

Within this small subsystem, the reciprocal lattice vectors $\mathbf{g}, \mathbf{g}'$ are defined on a coarser grid with respect to the larger system of size L; furthermore, since we assumed the smaller system to be contained in the larger one, we have that: $\{\mathbf{g}\} \subset \{\mathbf{G}\}$.

From these definitions, it follows that any vectors in reciprocal lattice of the larger system can be rewritten as a function of a corresponding vector in the smaller subsystem in the following way:

$$\begin{aligned} \mathbf{G} &= \bar{\mathbf{\Theta}}_1^{\ \sigma} + \mathbf{g} \\ \mathbf{G}' &= \bar{\mathbf{\Theta}}_2^{\ \sigma'} + \mathbf{g}' \\ \mathbf{Q} &= \bar{\mathbf{Q}} + \mathbf{q} \end{aligned} \tag{5.18}$$

$\{\bar{\mathbf{\Theta}}_1^{\ \sigma}\}, \{\bar{\mathbf{\Theta}}_2^{\ \sigma'}\}$ and $\{\bar{\mathbf{Q}}\}$ are L/l offset vectors applied to the set of reciprocal lattice of the large supercell and which fill the interstitial region between the finer grids $\{\mathbf{G}\}, \{\mathbf{G}'\}, \{\mathbf{Q}\}$ and the coarser grids $\{\mathbf{g}\}, \{\mathbf{g}'\}, \{\mathbf{q}\}$. These wavevectors belong to the first BZ of the large supercell and play the role of a twist condition for the many-body wavefunction. We define them as dependent from the particle spin. Therefore, according to the implementation of FTBC discussed here, the role of the flavor is represented by the spin of the particles.

The FTBC method is based on two approximations. The first one consists in restricting the transferred momentum $\mathbf{Q}$ only to the coarser mesh defined by the smaller subsystem, namely:

$$\mathbf{Q} = \bar{\mathbf{Q}} + \mathbf{q} \rightarrow \mathbf{Q} = \mathbf{q}. \tag{5.19}$$

This approximation limits the range of momenta to which a particle can be scattered to the coarse grid formed by the l points $\mathbf{q}$ in the reciprocal lattice of the small system.

Within this approximation, it can be shown that the Hamiltonian $\hat{H}_L$ of the larger system can be rewritten as follows:

$$\hat{H}_L \simeq \frac{L}{l} \sum_{\bar{\mathbf{\Theta}}_1, \sigma_1} \sum_{\bar{\mathbf{\Theta}}_2, \sigma_2} \hat{H}_l(\bar{\mathbf{\Theta}}_1^{\sigma_2}, \bar{\mathbf{\Theta}}_2^{\sigma_2}, \mathbf{q}), \tag{5.20}$$

where we explicitly write the dependence of the smaller system Hamiltonian on the interstitial wavevectors defined in Eq. 5.18.

In Eq. 5.20, the pairs $(\bar{\mathbf{\Theta}}_1^{\sigma_1}, \bar{\mathbf{\Theta}}_2^{\sigma_2})$ are all linked, thus the cost of this sum would still be $\mathcal{O}(L(L/l)^2) \sim \mathcal{O}(L^3)$, the same complexity as the full problem within the large system, even after the application of the aforementioned approximation. The FTBC method is obtained by enforcing another approximation in Eq. 5.20, which consists in neglecting all the links among the twist pairs $(\bar{\mathbf{\Theta}}_1^{\sigma_1}, \bar{\mathbf{\Theta}}_2^{\sigma_1})$ and integrating them as if they were independent. Such an approximation restricts the final complexity of Eq. 5.20 to $\mathcal{O}((L/l)^2) \ll \mathcal{O}(L^2)$.

The FTBC method consists therefore in an approximate expression of the Hamiltonian of the large system with dimension L as a function of the Hamiltonian of a smaller system with dimension $l \ll L$, suitably integrated over a set of spin-dependent twist conditions.

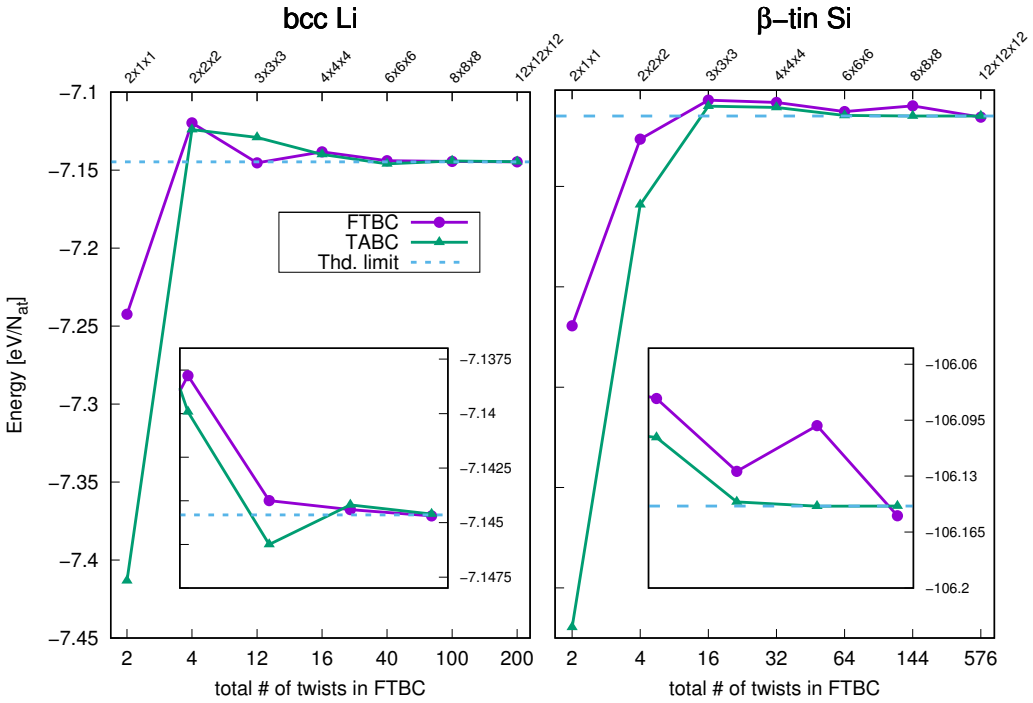


Figure 5.11 – DFT(LDA) calculations with the FTBC method on the Li bcc structure (left panel) and Si β-tin structure (right panel). In the bottom x axis we report the total number of twists used in the FTBC method, whereas the twist mesh used for standard TABC is shown in the top x axis.

The ground state eigenvector of the Hamiltonian $\hat{H}_l(\bar{\mathbf{\Theta}}_1^{\sigma_1}, \bar{\mathbf{\Theta}}_2^{\sigma_2})$ satisfies the generalized many-body Bloch condition as follows:

$$\Psi(\mathbf{r}_1, \ldots, \mathbf{r}_i^{\uparrow} + \mathbf{l}, \ldots, \mathbf{r}_j^{\downarrow} + \mathbf{l}, \ldots \mathbf{r}_N) = e^{i(\bar{\mathbf{\Theta}}_1^{\uparrow} + \bar{\mathbf{\Theta}}_2^{\downarrow})\cdot \mathbf{l}} \Psi(\mathbf{r}_1, \cdots, \mathbf{r}_N) \quad \forall\, i, j \in [1, N/2] \quad (5.21)$$

Thanks to the formalism developed in the previous Chapter, we are able to assign an arbitrary twist to each flavor (spin) and thus to fulfill the gauge condition on the many-body wavefunction required in Eq. 5.21. This allows the application of FTBC to realistic QMC simulations of metallic systems for the first time. It is worth remarking that the standard TABC method is retrieved from FTBC by setting $\bar{\mathbf{\Theta}}_1^{\uparrow} = \bar{\mathbf{\Theta}}_2^{\downarrow}$ and again performing the integration considering independent twists.

In the practical QMC implementation of FTBC within the *TurboRVB* code, we define two uniform meshes of size N_t in the supercell BZ with an offset among them of 0.5 the BZ length along the three Cartesian directions. From the first mesh we select the twists for $\uparrow$ particles, whereas from the second mesh we extract the corresponding twists

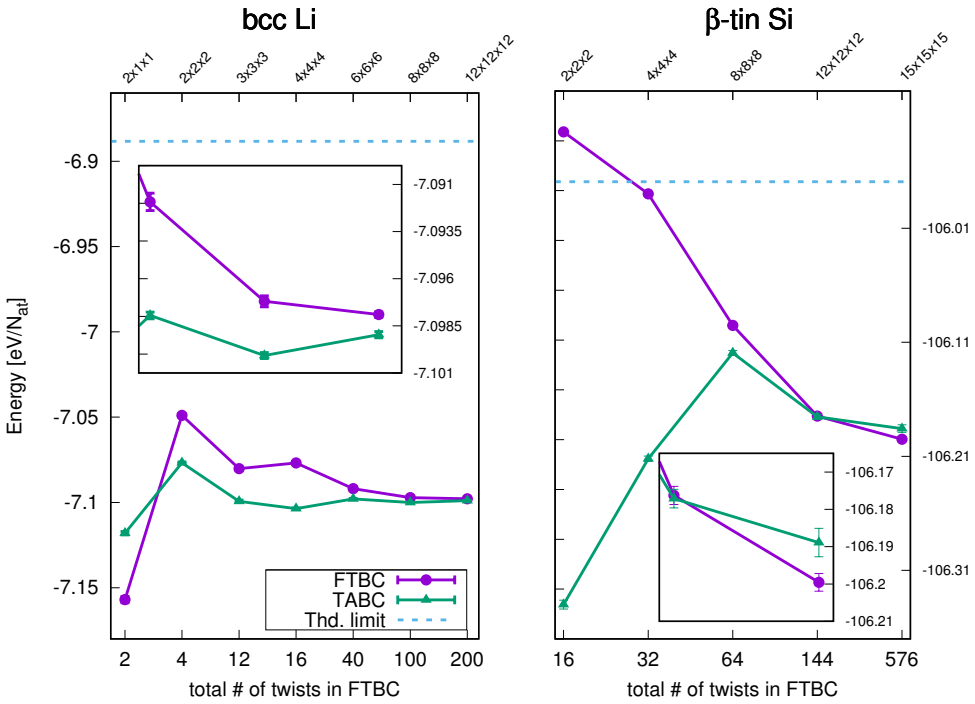


Figure 5.12 – VMC calculations with the FTBC method on the Li bcc structure (left panel) and Si β-tin structure (right panel). In the bottom x axis we report the total number of twists used in the FTBC method, whereas the twist mesh used for standard TABC is shown in the top x axis. The straight lines indicate the infinite-size limit coming from the extrapolation obtained in the previous Section without many-body error corrections.

for the down spin sector. In this way we create the $N_t \times N_t$ matrix in the reciprocal space required to perform the integration of the small system Hamiltonian in Eq. 5.20. We apply the FTBC method to two metallic systems studied in Sec. 5.6. Simulations are carried out in small supercells without the application of any corrections as we are interested in determining the performance of this technique for reducing both one-body and many-body effects when increasing the number of twist conditions. For the bcc Li structure we use a 16-atom supercell, whereas for the Si β-tin system a 8-atom supercell is employed.

We first assess this approach at the DFT(LDA) level by comparing it with the standard TABC approach with a uniform **k**-points grid [142]. In this case, no explicit electron-electron interaction is present; thus, FTBC should give similar results to the standard TABC calculation in the convergence to the thermodynamic limit. This is confirmed by the results presented in Fig. 5.11 where the DFT(LDA) convergence with the number of twists is shown for bcc Li and β-tin Si. We report also the thermodynamically converged DFT(LDA) value; it is apparent that both approaches reach this limit with the same trend as expected. The residual difference between FTBC and TABC results at this level of

theory has a pure one-body origin. We verified this assumption by performing benchmark TABC calculations where we assigned to the spin down sector a **k**-point grid shifted with respect to the spin up sector, as done for FTBC. In this way we recovered exactly the same DFT(LDA) energy for the two methods.

A correlated many-body framework such as VMC is a more meaningful approach for testing the reliability of the FTBC approach. In Fig. 5.12 we present the VMC results again for bcc Li and β-tin Si. In this case, the straight line in the plot corresponds to the infinite-size limit obtained with the extrapolation procedure presented in the previous Section. If the mapping constructed in Eq. 5.20 is a valid approximation for the large system, we expect this method to converge more rapidly to the correct thermodynamic limit indicated by the straight line in the plots with respect to the standard TABC, when using a small supercell. However, our results show that both in the lithium (left panel) and silicon (right panel) cases, FTBC displays a very similar behavior to the standard twist-averaging technique and many-body errors remain sizable in both situations. More than to the independent-twist approximation, the poor performance of FTBC in curing many-body size effects is likely linked to a choice of a too coarse grid for the transferred momenta **q** (Eq. 5.19). The integration error arising from this substantial coarsening of the **q** grid is too large to ensure a suppression of the leading $1/N$ behavior in the extrapolation to infinite size. This assumption should be validated by performing a similar calculation on a larger supercell for both systems studied here. This is left for future investigations.

5.8 Conclusions

In this Chapter we introduced and discussed the most important source of errors in QMC calculations of solids, namely the systematic effect arising from the finite size of the simulation cell. We focus our attention on the contribution to FS errors coming from the kinetic energy fluctuations due to shell filling effects in the one-particle orbitals. In order to mitigate these effects we presented here two methods representing an alternative to the mainstream TABC technique used in most part of modern QMC investigations.

The first approach is based on a novel procedure, which we dubbed EST, to find special twist values in the Brillouin zone corresponding to the simulation cell which reproduce the mean-field infinite-size energy up to an arbitrarily high numerical accuracy. .

This procedure has several advantages with respect to TABC technique. From a computational point of view, it is a single-twist technique and it is therefore more affordable, especially within diffusion Monte Carlo calculations characterized by a significant equilibration time. It is thus not only possible to accurately determine thermodynamic converged total energies, but also to perform structural relaxation of complex supercells

or even large molecular dynamics simulations within the QMC framework. On the other hand, the EST method is constructed in order to maintain the exact mean-field thermodynamic limit of the many-body variational energy. This feature allows to avoid any bias in the kinetic energy evaluation and provides, in principle, a more reliable description of the Fermi surface when a large supercell is used.

The EST method ensures a reduction of shell fluctuations comparable to the TABC technique on energy extrapolation to infinite size. We demonstrate this at first on the 3D electron gas model, and then on total energy extrapolations of three realistic metallic systems of increasing degree of complexity: solid bcc hydrogen, bcc Li and the high-pressure β-tin phase of silicon.

The calculation of correlation functions such as ionic forces is more delicate. We show for both zero temperature structural relaxation and molecular dynamics simulations that EST performs better than any other single-twist method. However, the TABC technique still shows a better performance, thanks to its denser sampling of the Fermi surface. However, for reasonably large supercell sizes, EST and TABC techniques are in perfect agreement. Therefore, a careful study of the supercell size dependence is necessary before applying the EST method in QMC production runs for the calculation of correlation functions. We believe that EST procedure here introduced can be the method of choice for reducing FS effects in many practical situations, particularly when the complexity of the system or the required supercell size make the more demanding TABC calculations infeasible.

The second method we presented is the flavor twist boundary conditions approach and it is an evolution of TABC which introduces different twist conditions for up and down spin sectors. It is applicable within our framework thanks to the generic twist conditions on the many-body wavefunction introduced in the previous Chapter. The basic idea of this approach is to perform an approximate mapping of the Hamiltonian of a large supercell onto the one of a much smaller system thanks to a suitable integration over a set of spin-dependent twist conditions. Whereas FTBC should in principle cure both one-body and many-body finite size errors, by applying the method to bcc Li and β-tin silicon we show that its performance is very similar to the TABC technique: it successfully cures one-body errors, but it does not lead to an appreciable reduction of many-body errors.

The ideas and methods discussed in this Chapter are of paramount importance to obtain reliable QMC simulations of solid systems and they will be thus extensively used in our QMC study of the high-temperature superconductor FeSe presented in the last Chapter of this thesis.

Chapter 6

Applications to iron-based superconductors: the case of iron selenide

6.1 High-temperature superconductivity from first principles

6.1.1 Introduction

High-temperature (HTc) superconductivity is one of the major open problems in modern physics. The discovery of the cuprate HTc family in 1986 [29] triggered an intense research effort that nowadays counts more than 15000 published works. These newly discovered materials immediately gave rise to several issues from both theoretical and applied points of view. Indeed, superconductivity in cuprates is not explained by the Bardeen-Cooper-Schrieffer (BCS) theory since the electron-phonon coupling constant is too small for a BCS-like electron pairing mechanism; therefore cuprates are often called *unconventional superconductors*. A comprehensive and predictive microscopic theory of unconventional superconductivity is still far from reach despite the decades old efforts.

In 2006, unconventional and HTc superconductivity were found in the iron-based layered compound LaOFeP [185]. This discovery led to a new impetus in HTc research and over the past years a plethora of novel iron-based superconductors (FeSCs) with different properties has been produced. At variance with cuprates, whose parent compounds are generally strongly correlated Mott insulators, FeSC parent compounds display a metallic or semi-metallic conductive behavior and are found, in general, to be weakly correlated [186]; furthermore, these materials show a quasi-2D layered crystal structure,

which is relatively simple with respect to the cuprates. These features combined with good mechanical properties make FeSCs more appealing than copper-oxides for potential commercial applications. Nevertheless, their discovery added several questions to the puzzle and broke the conventional wisdom that Fe should not play a direct role in superconductivity.

Despite their difference in structure and parent compounds, cuprates and FeSCs show a striking similarity in electronic structure properties [187], which leads to a quite universal phase diagram under doping or pressure application. The common thread which might explain these analogies is the proximity and often coexistence of the superconducting regime with an antiferromagnetic (AFM) phase [188]. This finding has been supported by many experimental works [189] and it has led to the hypothesis that the unconventional pairing mechanism may be generally mediated by (bosonic) AFM spin fluctuations in both families of superconductors [190].

However, a theoretical verification of this claim is still far from reach. Indeed, in order to provide a reliable description of HTc, theoretical tools need to cope with several complications such as the multi-band nature of these compounds and the diverse building blocks of their physics: weak Van der Waals forces, which bind layers together, magnetic fluctuations, and sometimes also strong electron correlation [187, 191]. These features give rise to a highly non-trivial behavior upon doping and also, differently from cuprates, pressure application, and to a rich phase diagram which is very sensitive to tiny variations of the crystal structure and internal parameters [192, 193]. Nevertheless, the similarities between copper oxides and FeSCs suggest that a single accurate theoretical framework may give an exhaustive description of high-temperature superconductivity.

6.1.2 QMC: an appropriate first-principle framework

Ab initio simulations of HTc and FeSCs in particular were mostly carried out with DFT-based methods until a few years ago. However, it is well established that the aforementioned building blocks driving their behavior are poorly reproduced by DFT. A great deal of research effort has been spent to construct density functional theories specifically designed to tackle the HTc problem and to account for the role of spin fluctuations in these materials. These techniques are based on a generalization of the Kohn-Sham formalism, which enforces an explicit functional dependence of the DFT Hamiltonian on the superconducting order parameter [194, 195]. Alternative ways have also been proposed, such as mixed frameworks combining DFT for the initial band structure and correlated GW or dynamical mean-field theory (DMFT) methods for calculating spin susceptibilities and other observables [191, 196]. Besides the dependence on the chosen energy functional,

there is another intrinsic issue behind these attempts. In fact, the electron condensation energy for the unconventional electron pairing is much smaller than in BCS-like superconductivity and a first principles approach would need a precision $\ll 0.01$ eV/atom (subchemical accuracy) in order to correctly account for it. As discussed in Chap. 3 for open systems, this condition is hard to meet when using effectively independent electron methods such as DFT.

Continuum QMC represents a very promising alternative since, as we have stressed throughout this thesis, it is in principle able to cope with all the ingredients required to describe FeSCs physics (dynamical spin fluctuations, tiny structural variations and strong electron correlation) within a unified and truly first principles approach. The *ab initio* framework proposed in this thesis can provide a direct comparison with experimental results without requiring any adjustable input parameters or functionals, as are required in the case of DFT. Moreover, at variance with DMFT and its cluster variants, does not require the "downfolding" of the continuum Hamiltonian to lattice model Hamiltonians [197].

However, the high computational cost of QMC for solids prevented the extensive use of this technique in the field of unconventional superconductivity until recent years. Nowadays, this trend has started to change and FeSCs are accessible by QMC thanks to the computational power of modern supercomputers. Previous published works focused on the properties of copper-oxides materials [198, 131] and demonstrated the capabilities of QMC in tackling the ground state magnetic properties of these systems. A limited number of applications of continuum QMC to FeSCs has appeared so far, for instance Ref. [149]. Our work on FeSe published in *Physical Review B* [135], issued from part of the results presented in this Chapter and is one of the first attempts in this direction. Notice that, among all QMC flavors, our implementation detailed in Chap. 2 and 4 is particularly well suited for HTCs. In fact, the JAGP wavefunction offers a direct probe of the electron pairing function and allows one to accurately recover its symmetries and to make an estimate of the pair condensation energy and hence of the critical temperature, as shown in Ref. [149].

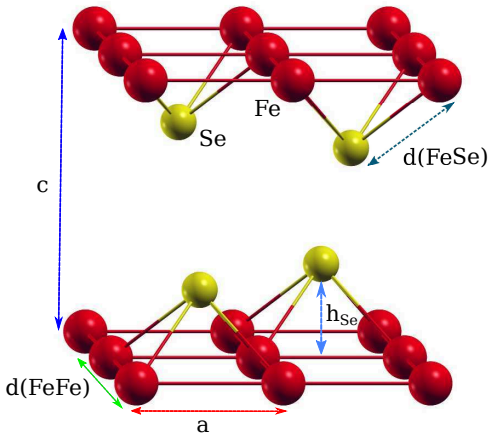
6.1.3 The case of FeSe: simplicity does not exclude complexity

In the context of HTc, a reliable first principles approach must first of all provide a reliable description of the electronic properties of its normal state. In this Chapter we present a thorough investigation of a representative member of the FeSCs family of chalcogenides (S,Se,Te) [199]: the iron selenide (FeSe) compound.

FeSe is certainly one the most intriguing and best studied of the FeSCs. This is due,

on one hand, to its very simple crystal structure (only 2 atoms in the unit cell) composed of a stack of iron planes without any intercalating charge reservoir as in cuprates or other FeSCs. The room temperature tetragonal P_4/nmm structure of FeSe is shown in Fig. 6.1. The atomic and axis notations reported in this Figure are used throughout this Chapter.

Figure 6.1 – Layered structure of the FeSe high-temperature superconductors. At room conditions, this compound belongs to the P_4/nmm tetragonal point group with a unit cell containing 2 atoms. We report here the labelling we used in Sec. 6.3 to identify the FeSe structural parameters. The only internal degree of freedom is represented by the Se height above the iron planes which we denote as h_{Se}. Around 90 K FeSe undergoes a tetragonal-to-orthorhombic transition (passing from P_4/nmm to $Cmme$ point group) with a small distorsion yielding a b/a ratio of ~ 1.005.



However, simplicity does not exclude physical complexity; in this respect, the case of FeSe is emblematic. This can be easily understood by looking at one of the first published FeSe pressure-temperature phase diagrams in Fig. 6.2a [200]. This experimental work was carried out with the neutron scattering technique and has been supported by several successive investigations. First of all, we notice that numerous structural phases appears as a function of both pressure and temperature. The tetragonal-to-orthorhombic transition revealed around 90 K is particularly important; this phase transition is usually called the "nematic" transition and it has triggered a great interest since it is a universal feature of FeSCs and, its origin, either magnetic or orbital driven, has not been elucidated yet [201]. We will not discuss this aspect of FeSe physics in this Chapter.

Moreover, the FeSe superconducting properties undergo significant changes when pressure is applied, as the critical temperature raises from 8 K at room conditions to $\sim$ 36 K at around $7 - 9$ GPa. Not only the change in structural parameters caused by the application of pressure, but also the thickness of the sample strikingly affects the superconducting properties. Indeed, it has been surprisingly found that if a single-layer of FeSe

is deposited on a $SrTiO_3$ substrate, the critical temperature rises to the record value for FeSCs of $\sim$ 100 K [202, 203]. Recently, a similar increase in the critical temperature has been spotted also on different substrates such as MgO. It is likely driven by an intrinsic doping of FeSe single-layer when interacting with the substrate [204].

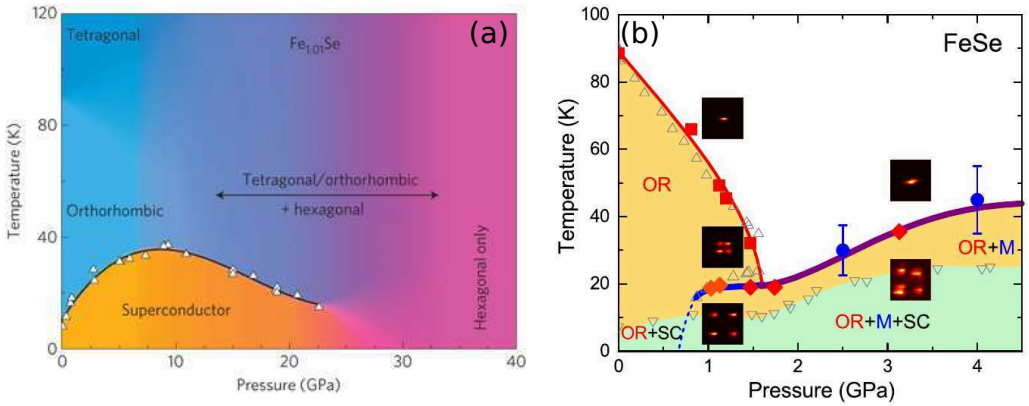


Figure 6.2 – Two phase diagrams of FeSe as a function of pressure. Panel (*a*) is adapted from Ref. [200] and dates back to 2009, right after the discovery of FeSe. It shows numerous structural phase transitions and no long range magnetic ordering. Panel (*b*) reports a very recent phase diagram adapted from Ref. [205] in which a new antiferromagnetic phase appears around 0.8 GPa. In this panel, the label OR indicates the orthorhombic ("nematic") phase, SC is the superconducting region and M identifies the newly discovered antiferromagnetic long-range ordered phase. We notice that at low temperature there is a coexistence of the magnetic and SC phases.

Whereas the superconducting properties are reasonably well-established, the nature of the electronic ground state of FeSe is still under intense debate. Until a few months ago, it was conventionally accepted that this compound did not have any long-range magnetic ordering over the whole phase diagram, as reported in Ref. [200]. A possible explanation for this exceptional behavior, different from any other FeSCs, lies in the detection of strong antiferromagnetic (AFM) spin fluctuations [200, 206], which could suppress the static ordering of the spins. However, this picture has been very recently questioned. Using a highly sensitive muon spin rotation technique [207] and Mössbauer spectroscopy [205], a new phase diagram of FeSe has been found which is presented in Fig. 6.2b. We notice that a new phase appears above 0.8 GPa close to the onset of superconductivity. This phase has been clearly identified with the long-range AFM phase discovered previously [208]. This novel scenario adds several puzzling questions to the fundamental physics of FeSe, especially because of the uncommon coexistence [209] in this compound of magnetism and superconductivity above 0.8 GPa.

It is evident that extracting reliable experimental results from bulk FeSe is challenging. This is mainly due to the intrinsic difficulty in obtaining crystals with the correct stoi-

chiometry. In such situation, *ab initio* calculations are of paramount importance to guide experimental results.

However, the precise calculation of the properties of this material remains challenging for mean-field methods such as DFT. In the particular case of FeSe, the main reason for this failure is the strong electron correlation which has been clearly found with DMFT calculations [191]. For example, the DFT(PBE) band structure is in poor agreement with experiments, which report a considerably narrower bandwidth [210]. Furthermore the FeSe lattice constants display an average error of ~ 0.1 Å independent of the exchange correlation functional employed (see for instance Ref. [211] and Tab. 6.3).

6.1.4 Organization of the Chapter

In this Chapter we present the first extensive investigation of the properties of the normal state of FeSe with the advanced QMC techniques discussed in this thesis work. All details and validation of the QMC calculations are reported in Sec. 6.2. In Sec. 6.3 we focus on crystal structure predictions and show that our method provides the best agreement with experimental results among the available *ab initio* approaches. A significant part of the discussion is devoted to the internal parameter of the compound, the Se height above the iron planes. We show that its evaluation is a challenging task, but that it can be successfully used as a parameter for tuning the properties FeSe and understanding where DFT-based techniques fail. In Sec. 6.4 we investigate several possible candidate magnetic patterns for the ground state of FeSe and study their behavior upon the application of pressure. We found that QMC predicts the collinear $(0, \pi)$ magnetic state as ground state, in disagreement with DFT. Furthermore, a clear distinction between collinear and non-collinear magnetic orderings under pressure is shown. We also comment on our results in the paramagnetic FeSe phase. The very accurate picture of structure and energetics yielded by QMC sets the stage for elucidating several fundamental connections between geometry, magnetism and charge degrees of freedom, which we present in Sec. 6.5. A simple explanation based on the Hund's coupling rule allows us to reconcile most part of our results on this compound. Finally, in Sec. 6.6 we draw the conclusions to this Chapter.

6.2 Technical details of the QMC calculations

6.2.1 Two complementary QMC flavors

In this work, carried out in collaboration with the group of Lucas K. Wagner at the University of Illinois, we thoroughly compare two distinct QMC flavors. This allows us to validate and complement the description of FeSe by QMC.

We dub the first method QMC(opt), where QMC stands for VMC or DMC depending on the chosen method. QMC(opt) results have been fully carried out during this thesis and are obtained by optimizing the single-particle orbitals directly at the QMC level, using the techniques explained in Chaps. 2 and 4. The diffusion Monte Carlo calculations are obtained with the lattice regularized variant (Sec. 2.4.2) and structural relaxation is performed by direct ionic forces minimization. QMC(opt) calculations are performed with the *TurboRVB* package [1] and they have been all carried out during this thesis work.

The second QMC flavor uses a different strategy. The single-particle orbitals are frozen to the DFT values obtained with the hybrid density functional PBE0, with a hybrid mixing parameter of 0.25. We dub this technique QMC(PBE0), where again QMC stands for VMC or DMC. DFT calculations for generating the orbitals are carried out with the Crystal [25] program whereas QMC runs are performed with the open source software *QWalk* [212]. Diffusion Monte Carlo results are obtained with the standard Green function DMC algorithm with non-local moves [213] for dealing with the BFD pseudopotentials. Furthermore, in this case the relaxation of the internal parameter is carried out with the standard Morse fitting procedure of the total energy curve. The QMC(PBE0) calculations have been mainly performed by Brian Busemeyer, graduate student in the group of Lucas K. Wagner.

For both methods we employ energy consistent [53] pseudopotentials for Fe (16 electrons) as well as for Se (6 electrons). We verified with all-electron DFT calculations that the pseudopotential error is below our target accuracy of 0.01 eV/Fe.

As we will see throughout this Chapter, QMC(opt) and QMC(PBE0) are complementary and offer a coherent overall picture of FeSe physics. We believe that their comparison strongly supports the validity of the QMC approach for tackling FeSCs, independently of the particular implementation of the method. Furthermore, the reliability of the less expensive QMC(PBE0) method is also confirmed.

Since the QMC(opt) calculations have been all carried out during this thesis, the technical details presented in this Section are related to this method only. We refer the reader to our work in Physical Review B [135] for further information on convergence and validation of the QMC(PBE0) flavor.

6.2.2 Simulation setup: supercells and magnetic phases

We perform calculations on several FeSe supercells subjected to PBC. The smaller one is composed by $2 \times 2 \times 1$ primitive cells with 4 Fe atoms and 4 Se atoms. The largest one is a $4 \times 4 \times 1$ simulation cell with 16 Fe and 16 Se atoms. All QMC(PBE0) calculations have been instead carried out with a rotated supercell containing 8 formula units (8 Fe and 8 Se atoms). Some test calculations have also been performed on a $2 \times 2 \times 2$ supercell with 16 atoms which explicitly includes an additional Fe plane. For all results presented, we use the tetragonal P_4/nmm structure. As we already mentioned, the low temperature ground state structure (see Fig. 6.2) belongs to the $Cmme$ orthorhombic point group. However, the orthorhombic distorsion is of the order of 0.5% [214, 215], hence the use of the simpler tetragonal structure should not affect the overall electronic structure properties.

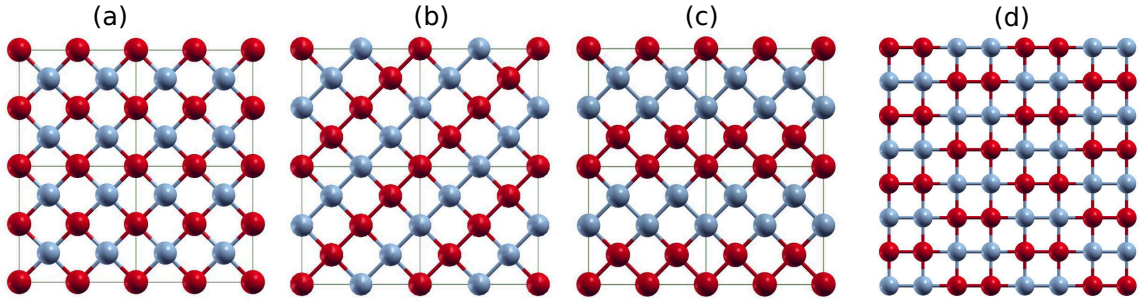
Concerning the magnetic ordering, we investigate several possible candidates for the possible and still unknown magnetic ground state of FeSe (see Fig. 6.2b). These magnetic patterns are enforced directly at the preliminary DFT stage of the calculations by imposing the determinant orbitals to possess the right spin pattern. This is then kept during the following QMC simulation, even in the case of orbital optimization. Notice that, due to the relatively small size of our supercells, long-range spin fluctuations which could destroy these orderings cannot be described. The main magnetic patterns studied in this Chapter are presented in Fig. 6.3. They include the simple AFM or checkerboard pattern (Fig. 6.3a) with antiparallel nearest neighbor spins, the collinear configuration with chains of aligned spins with alternate signs (Fig. 6.3b). Some QMC(PBE0) calculations are also performed with the bicollinear pattern, which is built with double chains of aligned spins (Fig. 6.3c), and with the staggered dimer configuration in Fig. 6.3d. Along with magnetic phases, several results have been taken within the paramagnetic configuration.

6.2.3 Wavefunction and basis set

We use the Jastrow-single determinant (JSD) ansatz introduced in Eq. 2.41 for all FeSe calculations presented in this Chapter.

For the Slater determinant, we choose a primitive Gaussian basis set tailored for tackling periodic solids (Eqs. 2.14 and 4.9). The number of Gaussians in the primitive basis set is $(7s6p6d)$ for Fe and $(5s4p4d)$ for Se. To reduce the number of variational parameters we found it more convenient in this situation to use standard atomic basis contractions. The optimal contractions can be obtained restricting the procedure outlined in Chap. 3 to a single angular momentum channel. A contracted atomic orbital is then defined as a linear

Figure 6.3 – Different magnetic patterns investigated in the FeSe compound. The two colors represent atoms with majority of ↑ and ↓ spin electrons. In panel (a) is shown the simple antiferromagnetic or checkerboard structure. In panels (b) and (c) are reported two flavors of spin "chain" configurations: the first one is the collinear pattern with a single chain of aligned spin and the second one is the bicollinear configuration with a rotated double spin chain. In panel (d) we show the staggered dimer configuration, which is the DFT ground state for FeSe [216].



combination of primitive basis orbitals sharing the same angular momentum channel:

$$\tilde{\chi}^a_{l,m} = \sum_{n=1}^{n_l} a_n \chi^a_{l,m,n} \tag{6.1}$$

where we use the same notation as Eq. 2.14. The coefficients $\{a_i\}$ are now taken as variational parameters in the wavefunction optimization procedure. If $N_{\text{contr}} \ll N_{\text{bas}}$ contracted orbitals are chosen, the total number of parameters P_{contr} in the full wavefunction reduces to: $P_{\text{contr}} \propto N_{\text{contr}} \times N_{\text{tot}} \ll N_{\text{tot}} \times N_{\text{tot}}$. The number of contracted orbitals employed for each angular momentum channel is chosen to attain the best compromise between size and quality of the contraction. We determine the quality of the contracted basis by performing DFT(LDA) calculations and checking both the resulting total energy and the overlap $S = \langle \Psi_{\text{full}} | \, \Psi_{\text{contr}} \rangle$ between the wavefunction expanded over the primitive basis and the contracted basis sets respectively. In Tab. 6.1 we report the results of this test for some of the contracted basis sets we tried. We notice that the contracted basis composed by $[3s2p2d]$ for Fe and $[2s2p1d]$ for Se offers the best compromise both in terms of overlap and total energy difference with the uncontracted wavefunction. This is used for all QMC production runs.

The Jastrow factor is much cheaper to evaluate with respect to the determinantal part. Therefore we employ a primitive basis set of size $(3s2p1d)$ for Fe and $(2s2p)$ for Se without any contraction. We use the complete charge and spin Jastrow form introduced in Eq. 2.35. This choice has been supported by test calculations against a simple charge Jastrow obtained by setting to zero the matrix elements $S^{a,b}_{\mu,\nu}$ in Eq. 2.35. In Tab. 6.2 we report the energies per Fe atom obtained using the two aforementioned Jastrow flavors

Table 6.1 – We report a detailed study on the contracted basis set convergence for FeSe. Several contractions of the primitive basis Fe(7*s*6*p*6*d*)Se(5*s*4*p*4*d*) are tested. The final atomic contraction is chosen as the best compromise between the overlap S defined in the text (second column) and the total energy difference ΔE between the contracted and uncontracted wavefunctions at the DFT(LDA) level (third column). In the fourth column we also show the corresponding number of optimizable variational parameters in the QMC determinant.

Basis set	S	ΔE [Ha]	# of parameters
Uncontracted basis	1.0000	0.0000	5060
Fe[2s2p2d]Se[2s2p1d]	0.9979	19.8863	641
Fe[3s2p2d]Se[2s2p1d]	0.9987	1.9108	692
Fe[3s3p3d]Se[2s2p2d]	0.9993	1.8909	1117

at different magnetic orderings.

Table 6.2 – FeSe energies with different forms of the Jastrow factor. The total energies are in Ha per Fe atom, while energy gains from simple charge to the full Jastrow factor are reported in eV per Fe atom. The tetragonal state is used within the $4 \times 4 \times 1$ supercell and experimental geometry [214]. Calculations are performed at the Gamma point. We verified that the energy gain is independent of the chosen twist condition. For all magnetic configurations the gain is above 0.04 eV/atom, a value which allows us to distinguish between different magnetic phases. We notice that for the VMC energies, which are much more sensitive to the quality of the variational wavefunction, this gain can be up to 10 times larger than in the LRDMC results.

Method	Paramagnetic	Checkerboard	Collinear
VMC - charge Jastrow	-133.1851(4)	-133.1868(4)	-133.1987(5)
VMC - charge+spin Jastrow	-133.2106(2)	-133.2228(3)	-133.2396(5)
En. gain VMC	-0.694(12)	-0.980(14)	-1.113(19)
LRDMC - charge Jastrow	-133.2861(8)	-133.3136(6)	-133.3092(6)
LRDMC - charge+spin Jastrow	-133.2876(7)	-133.3197(7)	-133.3135(5)
En. gain LRDMC	-0.041(29)	-0.166(25)	-0.117(21)

It is apparent from these results that the inclusion of the spin sector in the Jastrow term allows a considerable reduction ($\geq$ 0.04 eV/atom) of the variational energy for all considered magnetic configurations. We notice that the gain obtained from using the full Jastrow factor at a VMC level is around one order of magnitude larger than the corresponding gain in LRDMC calculations. Since VMC is much more dependent on the quality of the trial wavefunction, this highlights the significant improvement in the variational ansatz obtained by including dynamical spin fluctuations within the Jastrow factor. The full Jastrow factor is used for all QMC(opt) calculations presented in this Chapter.

It is important to remark that from the results in Tab. 6.2 one can infer that the simple AFM checkerboard state is lower in energy than the collinear configuration, in contrast with what was mentioned in the introduction and presented in Sec. 6.4. This apparent discrepancy is due to the lack of treatment of finite-size errors as the calculations presented here are performed at the Gamma point only. We address this problem in the next Section.

6.2.4 Finite-size effects and lattice extrapolation

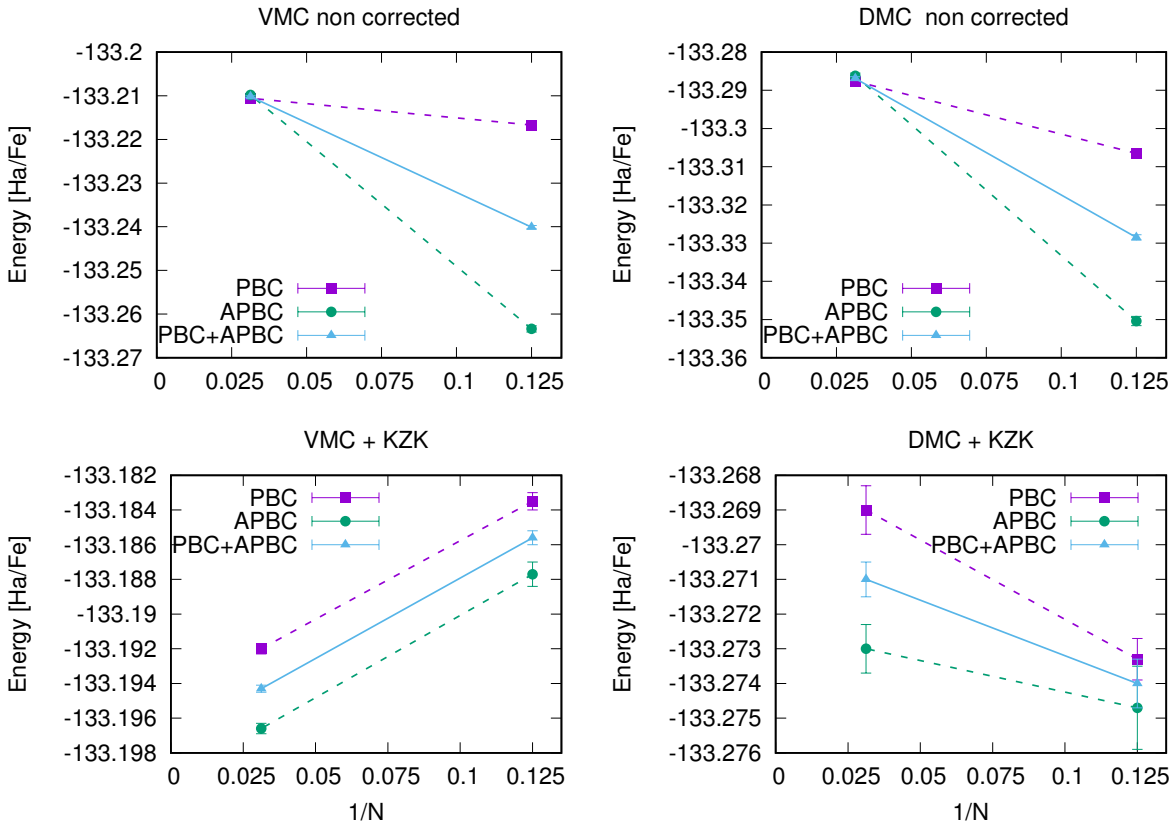


Figure 6.4 – Size extrapolation of the VMC and LRDMC energies in the FeSe paramagnetic phase at different twist conditions. The top panels report the energies without any two-body corrections, whereas the bottom panels show the KZK [154] corrected values. We notice that these corrections give a great benefit to the convergence rate with the system size. The final finite size error in VMC and LRDMC energies is below the target statistical error of 0.01 eV/atom.

Finite-size effects have a large impact on the quality of FeSe calculations both in the energetics and in crystal structure predictions, as we will see in Sec. 6.3. Within the QMC(opt) method, the one-body finite-size effects (see Chap. 5) are cured by averaging over two boundary conditions: the simple PBC with no twist and the fully antiperiodic

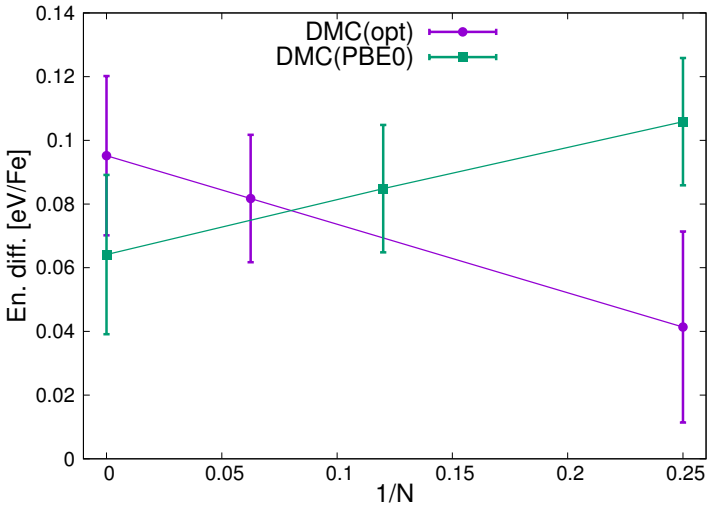


Figure 6.5 – Finite-size extrapolation at experimental ambient conditions of the energy differences between collinear and checkerboard magnetic ordering for DMC(opt) (thick outline) an DMC(PBE0) (thin outline). DMC(opt) data points are averaged over periodic and fully antiperiodic boundary conditions with two-body error corrections, while DMC(PBE0) are averaged over 8 twists. The extrapolations are in agreement within statistical error.

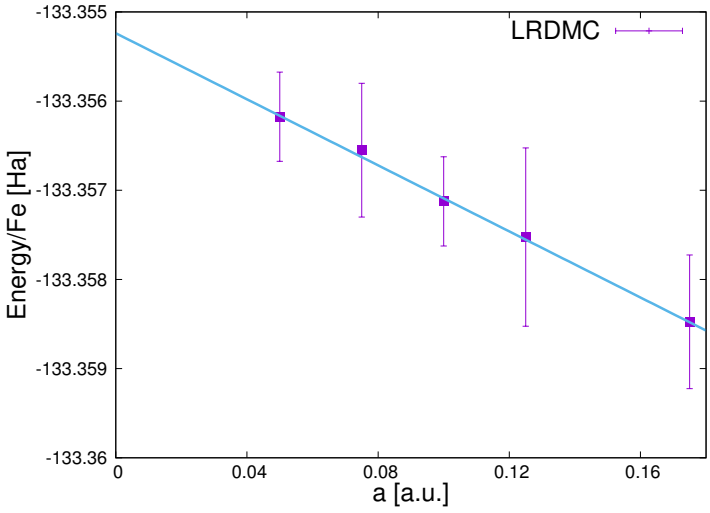


Figure 6.6 – Lattice space extrapolation to $a \to 0$ of the LRDMC energies in the $2 \times 2 \times 1$ supercell of the paramagnetic FeSe. The continuous curve is obtained from a linear fit. The final extrapolated energy at $a = 0$ is $-533.4214(6)$ Ha. A lattice space of $a = 0.125$ a.u. gives an error of ~ 0.01 eV per Fe atom which is our target accuracy. Therefore, for all final results presented in this Chapter this lattice space is used.

boundary conditions (APBC) equivalent to a twist condition of $(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$ in reciprocal lattice units. This approach is less sophisticated than the EST method presented in Chap. 5 since these calculations were performed when the implementation of complex-valued wavefunctions within the *TurboRVB* package was still under development.

The two-body errors are treated by means of the KZK functional method elucidated in Sec. 5.6 of the previous Chapter. We report the supercell-size extrapolations in Fig. 6.4 for both the VMC and DMC methods. We notice the effectiveness of the KZK correction in smoothing the impact of two-body errors by observing the different energy scale between top and bottom panels. After applying the two-body corrections, the residual finite-size effects in the LRDMC energies are of the order of ~ 0.005 eV/atom, which is below our target accuracy. The VMC error is slightly higher (~ 0.0085 eV/atom), but still below the target precision of 0.01 eV/atom.

In QMC(PBE0) calculations the average is instead performed over 8 twist conditions and no two-body corrections are applied since they likely cancel out when taking the energy difference. We verified the consistency of the two QMC flavors concerning size effects by checking the energy differences between collinear and checkerboard magnetic configurations with the typical setting used for production runs: 8 twists for QMC(PBE0) and 2 twists with KZK corrections for DMC(opt). The results are presented in Fig. 6.5. We find that the extrapolated values as well as the largest supercell sizes we used are in perfect statistical agreement.

We conclude this Section by assessing the additional error source represented by the finite lattice space used in LRDMC calculations. We select the best lattice space by performing test calculations on the paramagnetic $2 \times 2 \times 1$ supercell of FeSe. Results are shown in Fig. 6.6. We choose for productions runs a lattice space of $a = 0.125$ a.u. which gives an error of ~ 0.01 eV/atom with respect to the extrapolated value.

6.3 Structural properties

Obtaining the correct crystal structure for FeSe is a major challenge. This is mainly due to the non-bonded interaction among iron layers and to the strong link between crystal parameters and magnetic degrees of freedom, as we will see later on. The c lattice parameter in particular is affected by Van der Waals (VdW) interactions whereas electron correlation plays a key role in determining the in-plane physics. The behavior of FeSe's superconducting properties under pressure (see Fig. 6.2) gives another clue to the importance of structural variations in its description. The prediction of its lattice parameters is thus an important test for assessing the accuracy of first principles methods. Since structural relaxations at a QMC level are computationally costly, we limited our study to the tetragonal phase of FeSe since we can safely assume that the small orthorhombic distorsion does not affect the final results presented in this Section. On the other hand, the b/a ratio could be parametrically varied in order to investigate the tetragonal-to-orthorhombic transition in FeSe. This is left for future work.

Source	Magnetic Ord.	a	c	$\overline{\text{FeFe}}$	h_{Se}
DFT(PBE)	paramagnetic	3.6802	6.1663	2.6023	1.3862
DFT(PBE)	collinear	3.8007	6.2363	2.6966	1.4568
DFT(DF2)	paramagnetic	3.7181	5.2778	2.6323	1.3980
VMC ($2 \times 2 \times 1$)	paramagnetic	3.59(1)	5.296(8)	2.536(8)	1.40(1)
VMC ($2 \times 2 \times 2$)	paramagnetic	3.81(1)	5.35(1)	2.69(1)	1.38(1)
VMC ($2 \times 2 \times 2$)	collinear	4.10(1)	5.13(1)	2.902(9)	1.331(15)
VMC ($4 \times 4 \times 1$)	paramagnetic	3.71(1)	5.49(1)	2.62(1)	1.43(1)
VMC ($4 \times 4 \times 1$)	collinear	3.72(1)	5.68(1)	2.63(1)	1.56(1)
experimental [217] - T 7 K		3.7646(1)	5.47920(9)		1.4622
experimental [215] - T 8 K		3.7685(1)	5.5194(9)	2.6647(3)	1.5879
experimental [214] - T 300 K		3.7724(1)	5.5217(1)		1.4759

Table 6.3 – FeSe optimal structural parameters with different computational methods. DFT calculations have been performed with the software Quantum ESPRESSO [18] using a $10 \times 10 \times 10$ **k**-points mesh, an energy cutoff of 75 Ry and norm conserving pseudopotentials for both Fe and Se. The PBE and Van der Waals corrected DF2 functionals are compared. All VMC results presented in this Table are obtained with the QMC(opt) method, since with the simpler QMC(PBE0) the complete structural relaxation cannot be achieved. VMC calculations are carried out at the Gamma point only for all supercell sizes and magnetic configurations presented.

The equilibrium lattice parameters of FeSe for various magnetic orderings and cell sizes are presented in Table 6.3. As mentioned in the previous sections, these results are obtained with a direct optimization of FeSe cell parameters at the VMC(opt) level of theory.

If we consider our best result in the $4 \times 4 \times 1$ supercell, the in-plane FeSe properties should be well captured by QMC since the a lattice parameter is in close agreement with experimental results (within $\sim 4\ \sigma$) independently of the chosen magnetic configuration. Results obtained using both collinear and paramagnetic wavefunctions shows also a general improvement with respect to DFT(PBE) concerning the c lattice parameter. This provides evidence of the accuracy in treating VdW interactions with the QMC wavefunction, mainly achieved with the long-range part of our Jastrow factor. Notice that the importance of weak inter-planes interactions is also demonstrated by the outcome of the DFT(DF2) calculations [218], which explicitly treat VdW long-range dispersion within the Kohn-Sham Hamiltonian. This approach improves plain DFT(PBE) results in the paramagnetic phase for all considered crystal parameters.

It is important to remark that the geometry showing the best agreement with experiments is given by the paramagnetic phase; as we shall see in the next Section, this outcome is in contradiction with the energetics, where the paramagnetic displays a much larger energy value with respect to all other magnetic configurations.

The evaluation of the inter-plane c distance might be affected by the dispersion along the z-axis. In order to address this issue, we relaxed the structure of a $2 \times 2 \times 2$ supercell which explicitly contains an additional Fe plane. We notice that in the paramagnetic phase the differences with respect to the single plane $2 \times 2 \times 1$ supercell are negligible whereas both are quite far from our best result. Also the c parameter shows a slight improvement with respect to the single plane supercell. The collinear configuration displays completely wrong crystal parameters with respect to both experiment and its $4 \times 4 \times 1$ counterpart. Such a finding is certainly due to the poor description of the in-plane magnetism caused by the too tiny size of the in-plane simulation cell; this affects the overall description of the FeSe crystal structure as modeled by this supercell.

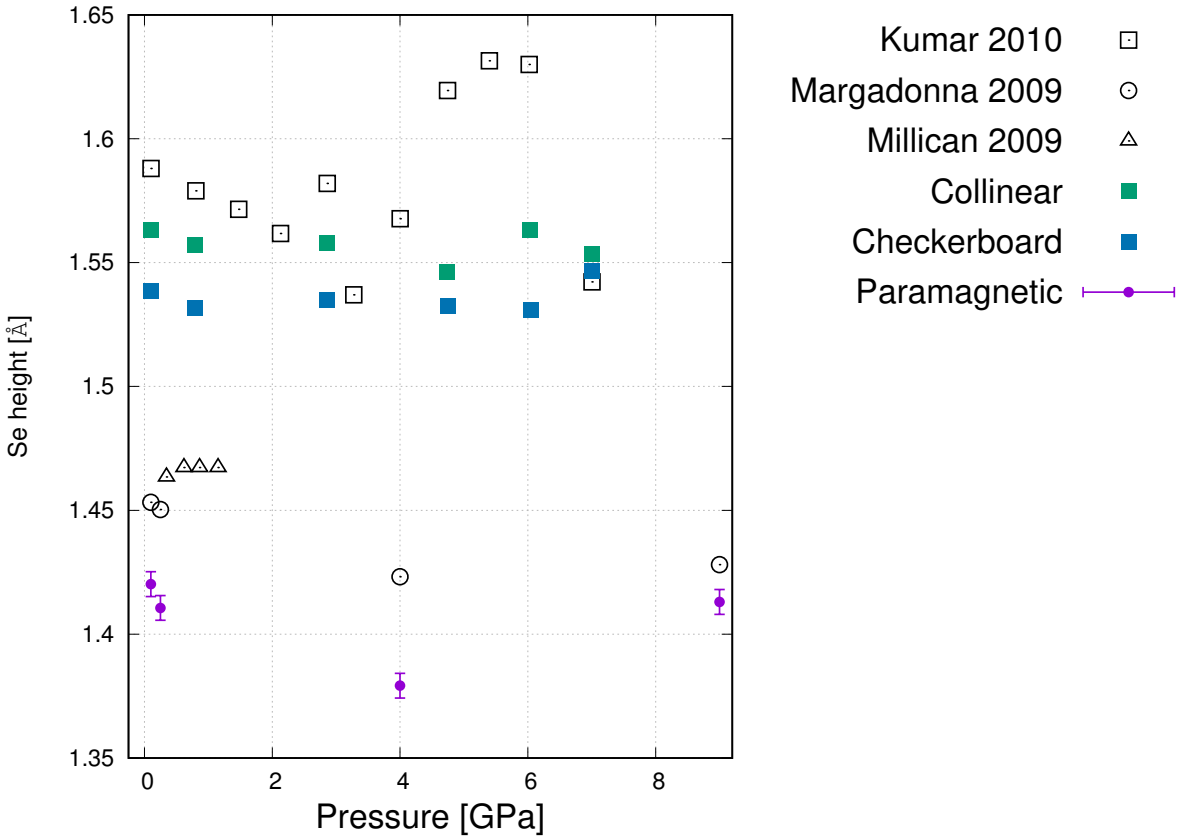


Figure 6.7 – Behavior of the Se height as a function of pressure. Here we compare several experimental results with the QMC outcome at different magnetic configurations. The paramagnetic results are obtained with the VMC(opt) technique by direct VMC force minimization, whereas collinear and checkerboard results are evaluated by fitting total DMC(PBE0) energy curves with a Morse function. A strong dependence of the Se height both on the QMC method (variational or diffusion MC) and on the chosen magnetic ordering is present. We notice also that the experimental results are quite scattered, highlighting the intrinsic difficulty to accurately measuring h_{Se} both theoretically and experimentally.

Let us now focus on the internal parameter h_{Se}, namely the height of the Se anion

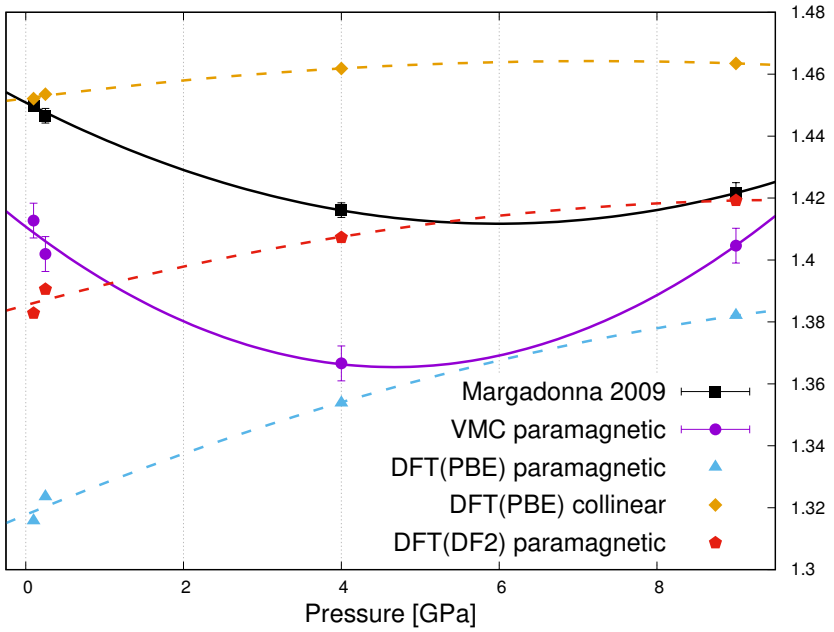


Figure 6.8 – Comparison of VMC(opt) and DFT curve of the Se height as a function of pressure. DFT calculations are performed with the software QUANTUM ESPRESSO [18] using a $10 \times 10 \times 10$ **k**-points mesh, an energy cutoff of 75 Ry and norm conserving pseudopotentials for both Fe and Se. Due to the complication in h_{Se} evaluation, we focus here on the paramagnetic phase only and on the experimental outcome provided by Ref. [214], and we use the Se height as a parameter for tuning FeSe properties. Van der Waals corrected DFT(DF2) shows a monotonic increasing curve, in contrast with VMC(opt) and experiments. h_{Se} behavior is linked to both in-plane and out of plane properties of FeSe, therefore this discrepancy suggests a relevant different between the VMC(opt) and DFT description of the physics of this compound. We provide a simple explanation to this issue in Fig. 6.9.

above the plane. It has been experimentally demonstrated [192] that there is a universal correlation, common to a large number of FeSCs, between variations of this parameter and the superconducting properties, the critical temperature in particular. We collect all our calculations of h_{Se}, as well as some experimental results in Fig. 6.7. From this Figure it is evident that a reliable measurement of h_{Se} is very hard to obtain from both experimental and theoretical points of view.

In particular, within QMC we find this parameter to be strongly affected by the magnetic configuration chosen (see also Tab. 6.3). Furthermore, the Se height is very sensitive to finite-size effects, either out-of-plane or in-plane. Due to this sensitivity, we decided to validate our QMC methods by performing DMC(PBE0) and DMC(opt) calculations with exactly the same setup. For this purpose, we used a small $2 \times 2 \times 1$ supercell in the collinear configuration at the Γ point. We computed the optimal Se height by fitting with a Morse function the total energy curve with both methods and we found 1.46(2) Å for the DMC(opt) method and 1.40(5) Å for the DMC(PBE0) method. The two values lie within

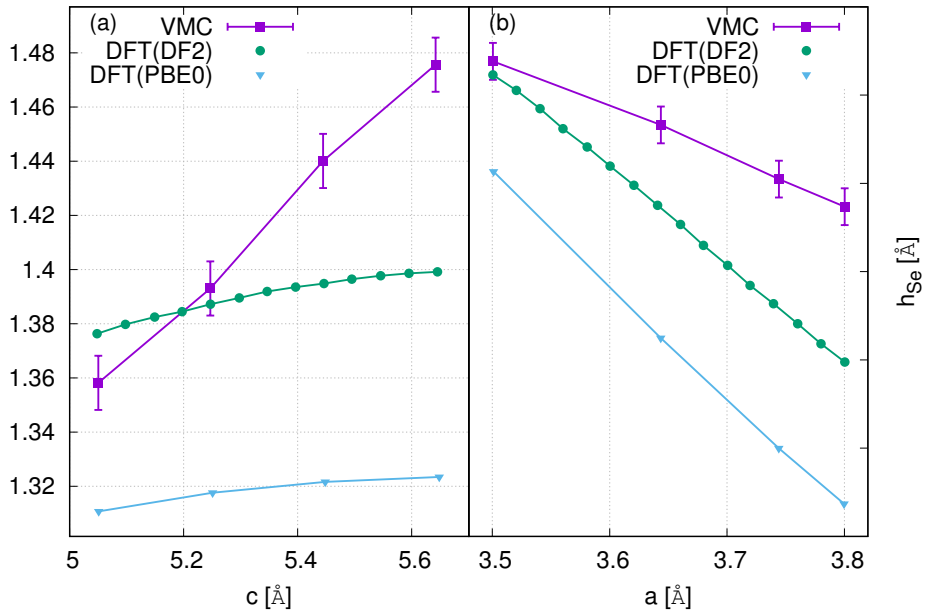


Figure 6.9 – We present in this graphs how the Se height behavior is affected by changing the parameters a and c of the FeSe crystal structure at different levels of theory. DFT(DF2) calculations are performed with the software QUANTUM ESPRESSO [18] using a $10 \times 10 \times 10$ **k**-points mesh, an energy cutoff of 75 Ry and norm conserving pseudopotentials for both Fe and Se. DFT(PBE0) calculations are instead performed using the software VASP [19] with a $5 \times 5 \times 5$ **k**-points mesh, an energy cutoff of 100 Ry and norm conserving pseudopotentials for both Fe and Se. We notice that the DFT outcome shows a different slope as a function of both a (underestimated) and c (overestimated). As detailed in the text, this behavior can be elucidated by a simple physical explanation. This accounts also for the monotonic behavior of h_{Se} under pressure seen at a DFT level (Fig. 6.8).

one standard deviation of each other, thus we consider them in statistical agreement.

Given the supercells that we studied, we found a variation in h_{Se} of approximately 0.05 Å. When using the experimental lattice parameters at ambient conditions, our best estimate of the Se height in the collinear phase is 1.56(1) Å, in agreement with the experimental results of Ref. [215], whereas the paramagnetic phase yields an estimation of 1.43(1) Å, close to the experimental values presented by several works [214, 192, 217].

The determination of h_{Se} is thus very challenging, however we can shed lights on the reliability of first principles methods by considering it as a parameter only and by looking at how it varies with the the other FeSe cell parameters. With this aim, we perform a full set of QMC(opt) calculations within the paramagnetic phase and with the $4 \times 4 \times 1$ supercell. In Fig. 6.8 we present the Se height curve under pressure compared between VMC(opt) and DFT predictions. Due to the importance of weak VdW interactions, we include also results obtained with the DF2 functional. While in the paramagnetic configuration DFT(PBE) is considerably far from all experimental results, a better agreement

with Ref. [214] is displayed when the collinear configuration is simulated or by including VdW dispersion. However, in both cases, DFT predicts a monotonic increasing behavior of the internal parameter, in contrast with QMC and experiments [192, 214].

This discrepancy can be linked to a difference in the underlying physical description offered by QMC and DFT methods. One can argue that the h_{Se} trend is mainly driven by two competing effects which arise upon pressure application. From one side, the shrink of the inter-plane distance induces a lowering of the Se height due to the stronger VdW repulsion. On the other hand, a contraction in the a parameter causes an increase in h_{Se} driven by the repulsion of the nearest Fe nuclei. In the experimental pressure range considered [214] up to 4 GPa, the variation of c is much larger than the correspondent a one. Therefore the lowering of h_{Se} induced by c prevails over the a effect until ~ 5.5 GPa, where the minimum is located. For pressure larger than this value the effect of a dominates and h_{Se} starts to increase again. Following this reasoning, we perform Se height relaxation in the tetragonal FeSe at several values of a and c by keeping the volume constant. The results are presented in Fig. 6.9. With respect to the QMC results, DFT strongly underestimates the slope as a function of c (Fig. 6.9a) whereas an overestimation is evident in Fig. 6.9b for the h_{Se} versus a curve. The results are independent of the functional employed and are not changed by a direct inclusion of VdW dispersion into the functional. Therefore, the effect due to the shrink of a parameter is always dominant, thus leading to the monotonic increasing behavior we observe in Fig. 6.8. This finding provides a simple explanation of why DFT, at difference with our approach, cannot reproduce the characteristic minimum of the h_{Se} under pressure shown in several experiments [214, 192].

The last structural parameter we compute is the bulk modulus, extracted by fitting the DMC(PBE0) energies with an equation of state previously used by Anton *et al.* [219]. The collection of ambient-pressure bulk-moduli results is reported in Tab. 6.4, in units of GPa. For all these calculations, experimental lattice constants [215] have been used. As for the internal parameter, we notice a very strong dependence on the magnetic ordering enforced in the QMC determinant.

The DMC(PBE0) calculations demonstrate excellent agreement with all three experiments reported if the collinear magnetic ordering is imposed, but there is less agreement if other patterns are used instead. The hybrid functional PBE0 calculations are also in somewhat good agreement with DMC(PBE0). On the other hand, PBE bulk moduli are significantly lower than both experiment and the other calculations, generally predicting bulk moduli between 7 and 10 GPa, depending only slightly on the magnetic ordering.

In this Section we presented the first full QMC prediction of FeSe structural parameters. Our results display a much better agreement with experiments when compared with DFT(PBE). This poses a strong constraint on the validity of DFT(PBE)-based investi-

Table 6.4 – Comparison of ambient-pressure bulk moduli from experiments and various calculations with various magnetic orderings. All calculations were performed using the room temperature experimental structure. We notice a strong dependence of the bulk modulus on the chose magnetic orderings. At a QMC level, the collinear structure displays the best agreement with experiment.

Source	Magnetic Ord.	Bulk Mod. (GPa)
Margadonna [214]	–	30.7(1.1)
Millican [220]	–	31
Kumar [215]	–	30.9(3)
DMC(PBE0)	bicollinear	26.4(8)
DMC(PBE0)	checkerboard	43.1(8)
DMC(PBE0)	collinear	31.2(7)
PBE0	bicollinear	27.6(1)
PBE0	checkerboard	40(1)
PBE0	collinear	29.7(2)
PBE	bicollinear	7.03(3)
PBE	checkerboard	8.9(2)
PBE	collinear	7.3(1)

gations on the FeSe structure. Furthermore, we demonstrated a strong inter-dependence between structure and magnetism which is evident both in the internal parameter predictions and in the bulk moduli. This effect is somewhat present also at a DFT(PBE) level for the internal parameter predictions; however, due to the discrepancy of DFT(PBE) structure with all available experimental results, the application of QMC is unavoidable to elucidate this connection, as we shall see in the next Section.

6.4 Magnetic properties

6.4.1 FeSe energetics under pressure

As discussed in the introduction, the true ground state magnetic configuration of FeSe is still under intense debate and it is not clear yet whether static long-range ordering is present or not in this compound. However, magnetic QMC calculations are performed within periodic boundary conditions on a relatively small simulation cell and thus they cannot describe sufficiently long-range spin fluctuations which can eventually suppress magnetic order in the crystal, as suggested by several works [200, 221]. On the other hand, the paramagnetic QMC state which we included in our calculations is found to be more than 0.5 eV/Fe higher in energy than any magnetically ordered wavefunction. More

Table 6.5 – Fe magnetic moments computed with DMC(PBE0) and DMC(opt) at different magnetic orderings. For sake of comparison we also included PBE0 and PBE results. All calculations were performed using the room temperature experimental structure.

Source	Magnetic Ord.	Mag. Mom. (μ_B)
DMC(opt)	checkerboard	3.134(6)
DMC(opt)	collinear	3.014(8)
DMC(PBE0)	bicollinear	3.518(5)
DMC(PBE0)	checkerboard	3.500(7)
DMC(PBE0)	collinear	3.443(6)
PBE0	bicollinear	3.492
PBE0	checkerboard	3.473
PBE0	collinear	3.433
PBE	bicollinear	2.653
PBE	checkerboard	2.332
PBE	collinear	2.561

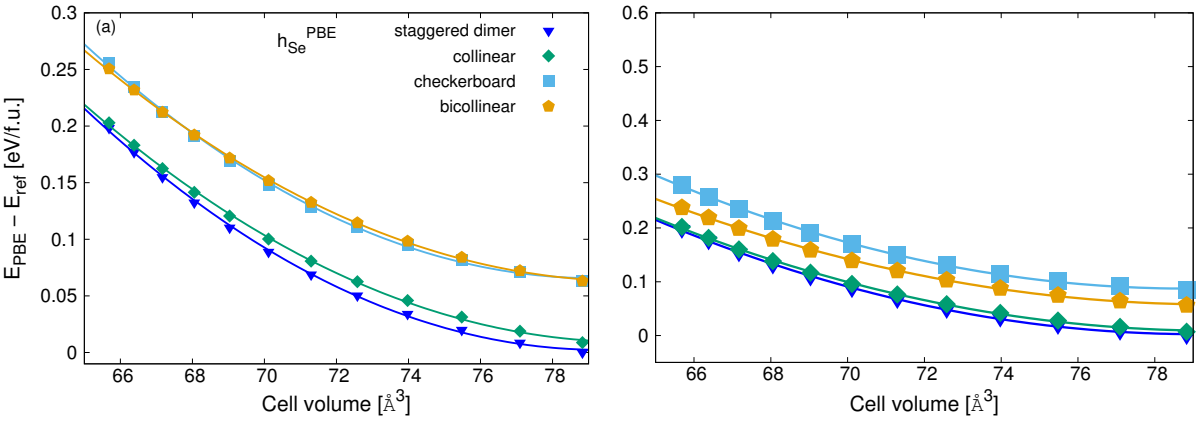
details about the paramagnetic calculations are presented later on in this Section. The ground state thus seems to require large local moments on the Fe atoms.

For the experimental crystal structure, the collinear magnetic ordering is the lowest in energy in our calculations and is observed to be the dominant short-range order experimentally [206]. If a static magnetic configuration is present in FeSe, it is also likely to be of collinear ordering [208, 207]. The energetic cost of introducing a "defect" (spin flip) into the magnetic order is quite small as we will see later on. Both the DMC(opt) and DMC(PBE0) flavors result in a rather large magnetic moment on the Fe atom. For the collinear magnetic ordering we obtain a value of $\sim 3.4\ \mu_B$ for DMC(PBE0), and a slightly lower $\sim 3.1\mu_B$ for the fully optimized calculations. In both cases the magnetic moment is close to the atomic limit. Since the exchange interaction plays a major role in determining the local magnetic moments in FeSe [222], it is not surprising the accordance between QMC and DFT(PBE0) predictions shown in Tab. 6.5. The local treatment of the exchange interaction in DFT(PBE) leads to an underestimation of the magnetic moment with respect to the other methods. All these results are contained in Tab. 6.5 and they are obtained at ambient conditions.

For the two QMC approaches we use, the energy difference between different magnetic orderings is in agreement within statistical errors (see also Fig. 6.5), so there is good reason to believe that the simpler and cheaper DMC(PBE0) technique is accurate.

We first focus on the DFT(PBE) calculations which are the most common in literature.

Figure 6.10 – FeSe energies in eV/f.u. as a function of the cell volume at the DFT(PBE) level of theory. The calculations have been performed with the software QUANTUM ESPRESSO [18] using a $10\times10\times10$ **k**-points mesh, an energy cutoff of 75 Ry and norm conserving pseudopotentials for both Fe and Se for all considered magnetic orderings. The reference is always taken as the energy of the "staggered dimer" configuration which is the DFT(PBE) ground state over the whole considered pressure range. In panel (a) the calculations are performed with the relaxed values of h_{Se} whereas on panel (b) we use experimental values [215] of the internal parameter.



In Fig. 6.10 we report the DFT(PBE) energy differences as a function of cell volume for different magnetic orderings. It is apparent that the relative energies are quite different from the DMC ones reported in Figs. 6.11a and 6.11b. The DFT(PBE) energy landscape is considerably more flat (see also Fig. 6.11d) and, moreover, even the lowest magnetic phase is different among the two methods. DFT(PBE) predicts as ground state the "staggered dimer" configuration with an energy gain of ~ 0.006 eV/atom with respect to the collinear phase over the whole considered pressure range. This is confirmed by several recent calculations [216, 223, 224]. On the other hand the DMC ground state turns out to be the collinear configuration independently from the value of the Se height.

It appears that hybrid DFT calculations in the PBE0 approximation (Figs. 6.11e and 6.11f) obtain reasonably good magnetic energy differences in comparison to DMC; thus this functional may be capturing some of the correct physics for the magnetic properties of this material. However, the PBE0 functional predicts an insulating gap [225] for FeSe independently from the chosen magnetic phase in contrast to DMC and experiment, as quantitatively demonstrated later on in this Section.

The validation of the DMC(PBE0) energy differences with the more accurate DMC(opt) allowed us to apply this method for an extensive investigation of the FeSe energy landscape as a function of pressure at different magnetic orderings. The idea is to find a common trends which could better elucidate how the magnetic ground state of FeSe evolves when the crystal structure is modified. A particular care is taken of the internal parameter

h_{Se} due to its importance for superconductivity [192]. Therefore energetics has been calculated with method relaxed internal parameter (z_{Se}^{theory}) as well as with experimental values of the Se height (z_{Se}^{exp}), in order to assess its impact.

Figure 6.11 – FeSe relative energies as a function of pressure with fixed values of h_{Se} for different levels of theory. For each calculation (QMC, PBE, and PBE0): (Right panel) Total energies for 8 f.u. cell for various magnetic orderings, as a function of the cell volume, choosing experimental [214] values of h_{Se}. (Left panel) Same as right, but choosing optimized values of h_{Se}. For the top QMC plots, energies are referenced to the collinear energy at around 77 Å^3 which corresponds to ambient pressure. The DFT calculations are referenced to the h_{Se} minimum energy for that type of calculation. DFT(PBE) calculations are performed with the Quantum ESPRESSO [18] package with fully converged **k**-point mesh and norm conserving pseudopotentials. The DFT(PBE0) results are obtained from fully converged calculations performed with the Crystal [25] code using the BFD pseudopotentials as QMC. The DMC(opt) paramagnetic energies are ~ 0.85 eV/f.u. higher than the reference collinear energy (see Tab. 6.6) and they thus not reported here.

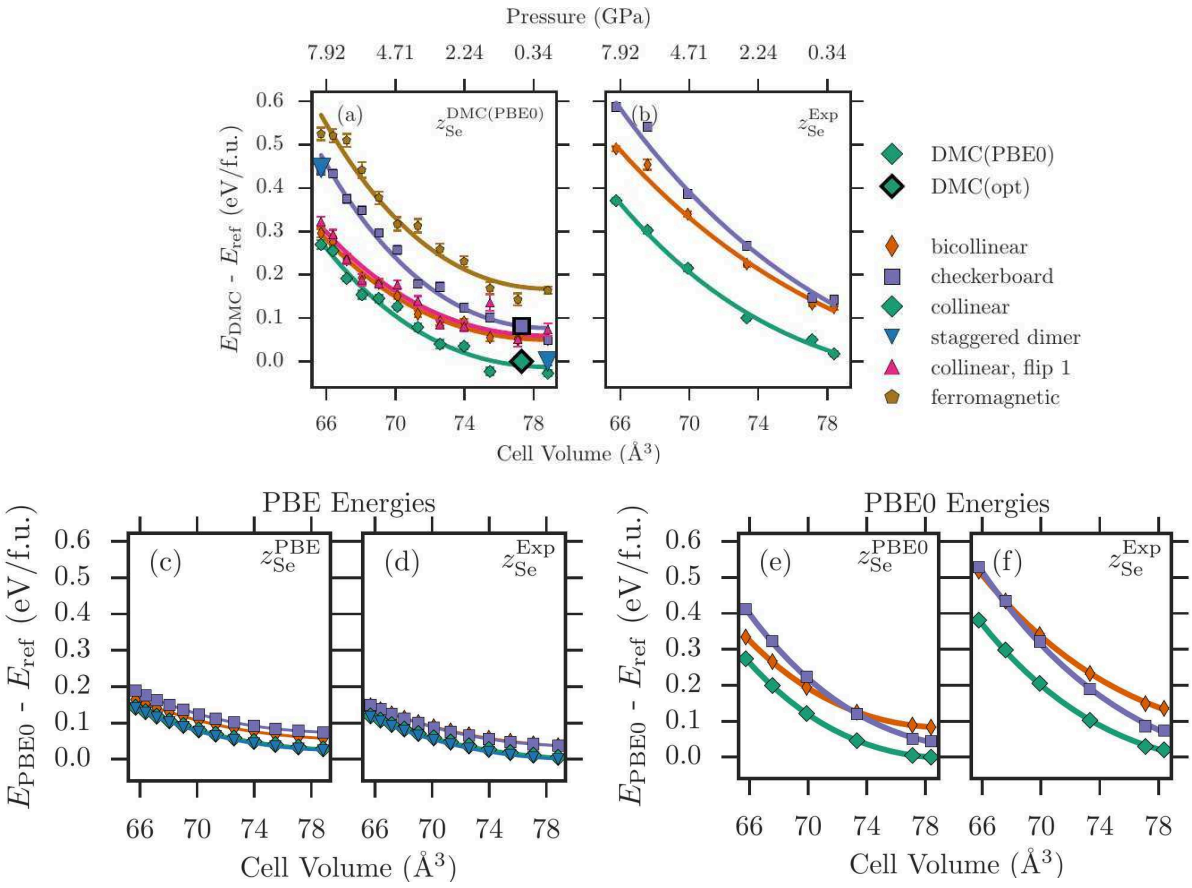


Fig. 6.11 summarizes all our energy results with different computational methods. Finite-size effects and other source of errors in DMC(opt) and DMC(PBE0) are addressed with the techniques explained in Sec. 6.2.4. In our investigation we include also the ferromagnetic phase with all spin aligned and the configuration labelled by "collinear, flip

1" which is obtained by introducing one defect, i.e. spin flip, in the collinear ordering.

Under pressure, the checkerboard, ferromagnetic, and staggered dimer magnetic orderings rise in energy compared to the lowest energy collinear ordering (Fig. 6.11a and Fig. 6.11b). On the other hand, the stripe-like orderings, including the bicollinear and collinear orderings with defects converge with applied pressure.

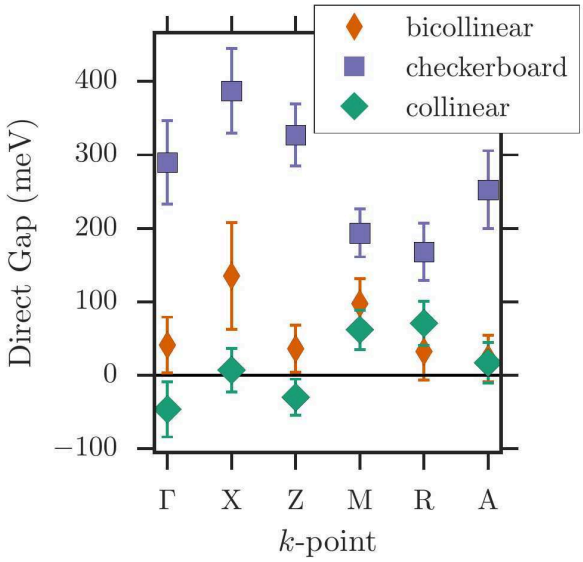
From Fig. 6.11c and Fig. 6.11d the failure of PBE in capturing this trend is apparent. PBE0 magnetic energies are reasonably close to the DMC results (Fig. 6.11e and Fig. 6.11f). However, as mentioned before, they cannot be considered reliable due to the wrong estimation of conducting properties of FeSe [225]. This is not the case for our ground state QMC configuration, as demonstrated by the gap calculations discussed in the following.

We calculate the direct optical gap at the DMC(PBE0) level by promoting the highest energy orbital in the Slater determinant part of the trial wavefunction to the next excited state orbital at the same twist condition. This constructs a wave function *ansatz* for an electron-hole excitation and the direct gap is obtained by taking the energy difference between the ground state and excited wavefunctions. The results are shown in Fig. 6.12. Interestingly, the DMC(PBE0) gap is 0 within statistical uncertainties for collinear orderings despite the fact PBE0 estimates a rather large gap for all orderings [225] and the determinant orbitals are frozen to the PBE0 ones. The closing of the gap is thus solely derived from correlation effects introduced with the Jastrow factor. Experimentally [226], the gap is no more than 80 meV at any point in the reciprocal space, which is consistent with our results for the bicollinear and collinear magnetic ordering. Only the checkerboard state is gapped according to QMC predictions.

Given all this data available to us, we can determine some properties that are robust to the finite-size errors and to the uncertainty in h_{Se} present in our calculations. The first one is that the relative energetics of magnetic orders changes strongly as a function h_{Se} and pressure. In DMC and PBE0, which would *a priori* be expected to be more accurate, the collinear and bicollinear orders become closer in energy as with increasing pressure. According to DMC, this effect is robust against h_{Se} variations, depending mainly on the change in the relative magnetic energies as a function of pressure. Furthermore, the energetic cost of reversing a single spin in the collinear ordered state, follows the bicollinear energy quite closely. Because this cost decreases with pressure, we can surmise that magnetic fluctuations become more energetically available as pressure is increased. In order to be confirmed, this result requires further experimental investigations.

It is evident from the energetics and gap calculations presented in this Section that within FeSe there is a strong interplay between structural variations and spin degrees of freedom. On the other hand, the direct gap calculations shows that the latter are also

Figure 6.12 – Direct optical gap calculated by promoting an electron to the next excited orbital with the DMC(PBE0) method at different twist conditions. Both collinear and bicollinear configurations displays a vanishing gap in agreement with experimental outcome. This introduces a clear distinction with the simple AFM checkerboard state which is has a gap of ~ 0.2 eV.



tightly coupled with the charge mobility in the system. This complex interplay between several effects occurring in FeSe is discussed and elucidated in Sec. 6.5 of this Chapter.

6.4.2 Comments on the paramagnetic phase

At ambient conditions, no experimental evidence of a static magnetic ordering has been found in FeSe. Several theoretical investigations tried to understand the origin of this paramagnetic phase. DFT and lattice model calculations [216, 227] surmise that the paramagnetism in this compound originates from the frustration derived due to the competition of several spin fluctuations at different wavevectors. The claim of an emergence of static long-range order at ~ 0.8 GPa [208, 207, 205] calls for further investigations on this scenario. However, the coexistence of magnetism and superconductivity above 0.8 GPa is unconventional [209] and sorely needs further investigations. The paramagnetic phase of FeSe remains interesting since it is the ground state at ambient conditions (< 0.8 GPa) and may give important insights on the emergence of the magnetic phase.

We simulate with QMC the paramagnetic phase of FeSe by constraining the determinant orbitals to be the same for $\uparrow$ and $\downarrow$ spin channels both at the preliminary DFT level and at optimization stage. However, as previously mentioned, this configuration shows a much higher energy with respect to magnetic configurations, independently from the chosen supercell size and the QMC method used. We present a more quantitative proof

of this outcome in Tab. 6.6, where we show the energy differences of the paramagnetic phase with the AFM checkerboard and collinear configurations compared between different computational methods. QMC(opt) calculations are performed with the standard setup. The inclusion of dispersion along the z-axis using a $2 \times 2 \times 2$ supercell does not modify the results.

The QMC energies show an energy difference between paramagnetic and magnetic configurations which is much larger ($\sim$ one order of magnitude) with respect to simple DFT with local functionals (LDA and PBE). A possible explanation of the large discrepancy can be attributed to the over-delocalization of valence electrons, a well known deficiency of local density functionals (such as LDA and GGA). This is one of the main problems in the simulation of strongly correlated (and therefore localized) quantum systems with DFT. Within QMC, this tendency is partially cured thanks to a more accurate treatment of the many-body effects involved in the exchange mechanism. Therefore exchange-driven magnetic orders are considerably more stable, as apparent from our results in Tab. 6.6. Hybrid functionals mix the local DFT exchange energy with the non-local exact Hartree-Fock exchange. Exchange is therefore strongly enhanced within this technique. With the PBE0 hybrid functional [225], the energy of FeSe magnetic configurations are lower than those of the paramagnetic configuration by about a factor 2 with respect to QMC outcomes for both checkerboard and collinear orders. As expected, we find QMC results to lie in between results obtained using an over-delocalized framework such as PBE, and results obtained using a theory which includes exact exchange such as PBE0.

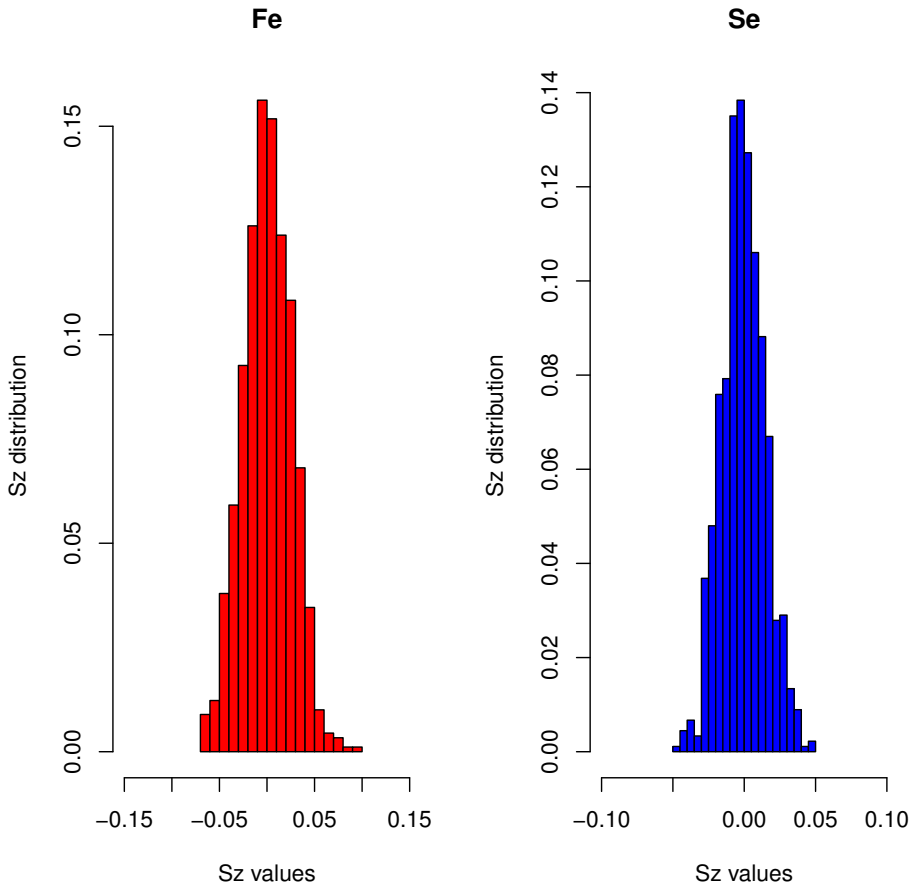
In order to validate the aforementioned frustration scenario it is important to understand whether our paramagnetic wavefunction is able to simulate the emergence of sizable local magnetic moments, which would imply the presence of spin fluctuations in the compound.

Table 6.6 – VMC(opt) and DMC(opt) energy differences between magnetic states (E_{chb} for the simple AFM and E_{col} for the collinear state) and the paramagnetic configuration (E_{para}). The tetragonal state is assumed using the experimental geometry [214] at ambient conditions. DFT(PBE) and DFT(LDA) calculations have been performed with the Quantum ESPRESSO [18] software.

Source	E_{chb}-E_{para} [eV/Fe]	E_{col}-E_{para} [eV/Fe]
DFT(PBE)	-0.192	-0.244
DFT(LDA)	-0.099	-0.136
DFT(PBE0) [225]	-1.425	-1.45
VMC(opt)	-0.269(10)	-0.5851(9)
DMC(opt)	-0.765(20)	-0.843(20)

The simplest estimator for probing the presence of local magnetic moments is the

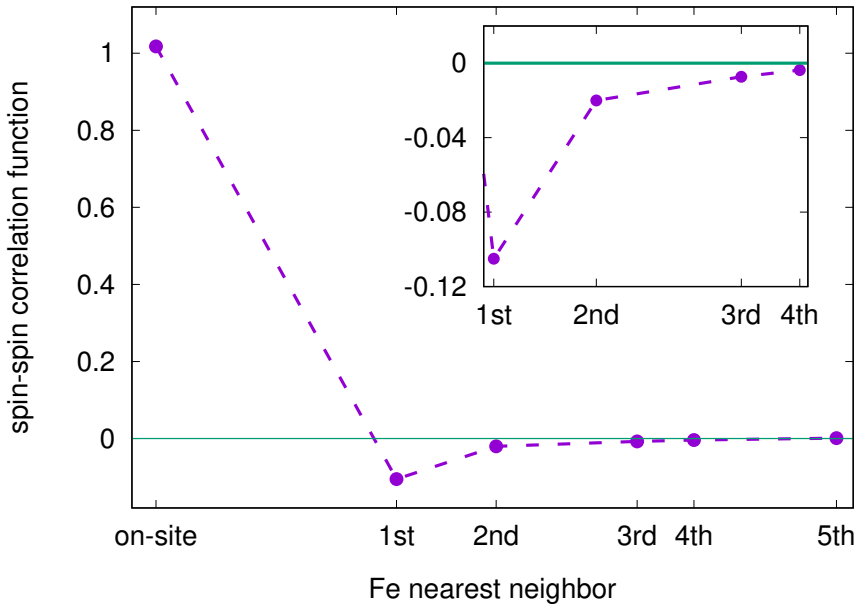
Figure 6.13 – Histogram of the atomic spin distribution function evaluated on the electronic configurations sampled during the QMC simulation. The mixed estimator formalism has been employed. Both functions are strongly peaked around 0, thus proving that no local magnetic moments are forming within paramagnetic FeSe as described by the QMC method.



atomic spin distribution function. Notice that we consider the spin conventionally aligned along the z-axis. This function is evaluated using a post-processing tool which analyzes the ensemble of electronic configurations sampled during the QMC simulation and the final results are obtained using the mixed estimator formalism outlined in Sec. 2.4.2. In Fig. 6.13 we present the spin distribution function obtained with our best paramagnetic variational ansatz projected onto Fe (left panel) and Se (right panel) atoms. We notice that this function is strongly peaked around 0 for both atomic species. If local magnetic moments were forming within our paramagnetic state, we would have instead expected to have a bimodal distribution with two clear peaks around the values of $\pm M$ with M the magnitude of the local magnetic moment. Given the distribution in Fig. 6.13, we can conclude that our paramagnetic wavefunction is not sufficient to support the emergence of local magnetic moments.

In order to have more insights into this behavior, we compute the spin-spin correlation

Figure 6.14 – Spin-spin correlation function $G_{\mathrm{Sz}}(\mathbf{R}_i, \mathbf{R}_j)$ projected onto Fe sites in the FeSe paramagnetic phase; this observable has been evaluated by assuming the translational invariance of the atomic lattice and with the mixed estimator formalism. The decay of $G_{\mathrm{Sz}}(\mathbf{R}_i, \mathbf{R}_j)$ is very sharp and points out that the spin fluctuations simulated by our approach are very short-range.

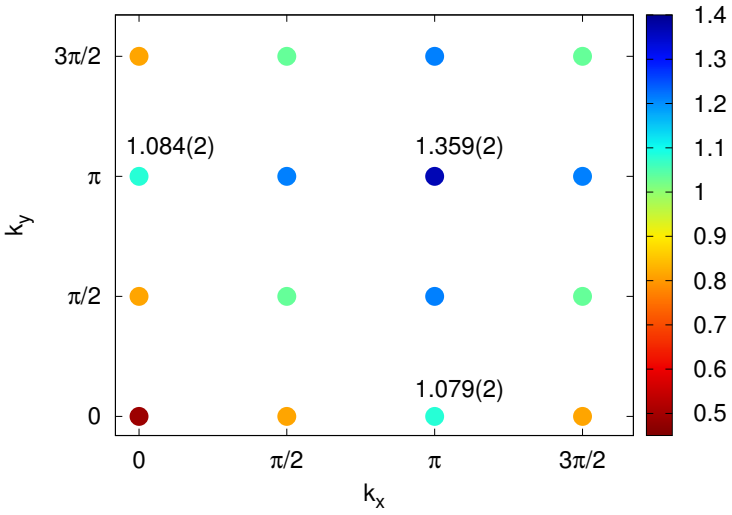


function $G^{\mathrm{Fe}}_{\mathrm{Sz}}(\mathbf{R}_i, \mathbf{R}_j) = \langle S_z(\mathbf{R}_i) S_z(\mathbf{R}_j) \rangle$ among Fe sites, where $S_z(\mathbf{R}_i)$ is the spin operator along the z axis associated with the i-th Fe site located at $\mathbf{R}_i$. The evaluation of the spin-spin correlation function requires the average spin $\langle S_z(\mathbf{R}_i) \rangle$ associated to the i-th site; we compute this quantity as the difference between the number of $\uparrow$-spin electrons and $\downarrow$-spin electrons falling within a sphere of radius r_c centered on the ion during the QMC random walk. r_c is chosen to be close to the atomic radius of Fe. We verify that this choice gives the correct atomic charge for both Fe and Se atoms. The final values of these observables and an estimation of their error bars are obtained by averaging over the Monte Carlo generations and using the mixed estimator formalism (see Chap. 2).

We find that $G^{\mathrm{Fe}}_{\mathrm{Sz}}(\mathbf{R}_i, \mathbf{R}_j)$ has a very rapid decay, vanishing already at the 2nd neighboring atom. The results are shown in Fig. 6.14 for our best paramagnetic wavefunction. This behavior is likely due to the fact that our variational ansatz takes into account spin fluctuations only within the Jastrow factor whereas the determinant can only describe a static long-range order. Multiple single determinant wavefunctions such as the Pfaffian [39] have been devised in order to accommodate distinct spin configurations within the same function. In this way, spin fluctuations can be consistently incorporated within the full QMC variational ansatz. It is probable that such an ansatz could enhance the role played by magnetic fluctuations in our simulations of FeSe. This would disfavor the

long-range magnetic orderings with respect to the paramagnetic state and may lead to a change in the ground state configuration at ambient conditions, producing results in agreement with experiments.

Figure 6.15 – Spin structure factor of the $4 \times 4 \times 1$ paramagnetic supercell at ambient conditions restricted to the reciprocal space of the Fe sublattice. The values in a.u. of the structure factor for the most important wavevectors ($(0, \pi)$, $(\pi, 0)$ and (π, π)) are reported in the graph.



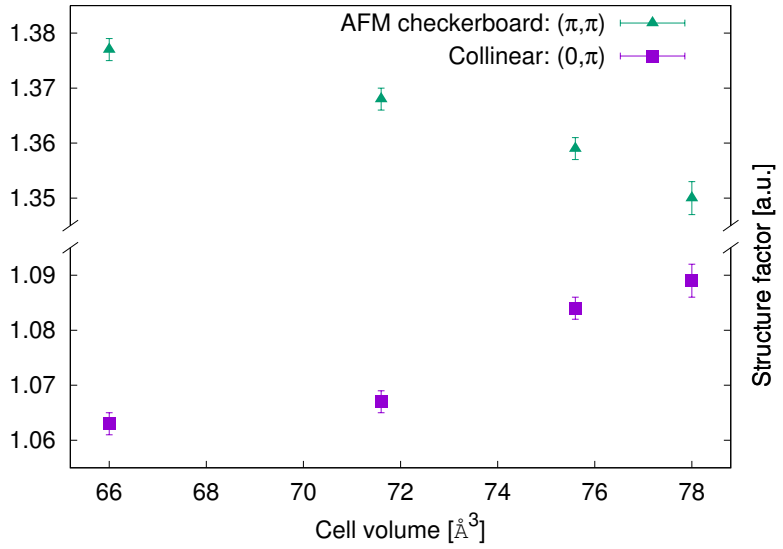
Despite the absence of localized magnetic moments, our paramagnetic state displays a value of the on-site spin-spin correlation function $G^{\mathrm{Fe}}_{\mathrm{Sz}}(0) = 2\sqrt{\langle S^2_z(0)\rangle}$ of 2.016(1) μ_{B}. We demonstrated that the localization of like-spin particles is not captured by our ansatz; hence this non-negligible value of the on-site spin fluctuations solely derives from the superexchange mechanism. In order to understand how magnetic fluctuations at different wavevectors contribute to this mechanism, we compute the Fe spin structure factor. This is defined as the Fourier transform of the spin-spin correlation function restricted to the Fe sublattice:

$$F^{\mathrm{Fe}}(\mathbf{q}) = \frac{1}{N_{\mathrm{Fe}}} \sum_{\mathbf{R}} e^{i\mathbf{q}\cdot\mathbf{R}} \langle S_z(\mathbf{R}_{\mathrm{ref}}) S_z(|\mathbf{R} - \mathbf{R}_{\mathrm{ref}}|)\rangle_{\mathrm{Fe}} \tag{6.2}$$

where we used the translational invariance of the crystal lattice by taking the atom located at $\mathbf{R}_{\mathrm{ref}}$ as reference. In Fig. 6.15 we show this quantity for the $4 \times 4 \times 1$ FeSe supercell at ambient conditions. The wavevector (π, π) corresponds to the simple AFM checkerboard pattern, whereas the wavevectors $(0, \pi)$ and $(\pi, 0)$ are associated with collinear spin fluctuations along the x and the y directions respectively; the values of $F^{\mathrm{Fe}}(\mathbf{q})$ for these wavevectors are reported on the plot. Taking into account the aforementioned limitations of our ansatz, it is however apparent from Fig. 6.15 that there is a strong competition between several magnetic fluctuations; interestingly, at ambient conditions the fluctuations

along the π direction in Fourier space display the largest value of the spin structure factor – in particular the (π, π) wavevector represents the dominant contribution – whereas the collinear fluctuations show a lower value. We can elucidate this behavior by looking at

Figure 6.16 – Spin structure factor of Fe sublattice as a function of the cell volume for the $(\pi, 0)$ and (π, π) spin fluctuations corresponding to collinear and checkerboard orderings respectively.



the values of $F^{\mathrm{Fe}}(\mathbf{q})$ for collinear and checkerboard orderings as a function of the cell volume which are shown in Fig. 6.16. We notice that the competition between these two competing orderings increases as the cell volume increases; the maximum level of frustration is reached at a cell volume which corresponds to ambient conditions. This is in agreement with the relative energetics presented in Fig. 6.11a, where the checkerboard state becomes closer in energy to the collinear configuration at larger cell volumes. With our paramagnetic solution we are therefore able to capture the competition among different orderings within the compound based only on the superexchange mechanism. These findings support the idea that magnetic frustration might be the cause of the disappearance of static long-range ordering below 1.2 GPa, as confirmed by all experimental results available nowadays. Despite the interest of these results, the aforementioned limitations of our paramagnetic ansatz are evident in the fact that we are not able to predict the dominance of the collinear magnetic fluctuation, as clearly seen in neutron scattering experiments [206].

6.5 Interaction between structure and magnetism

6.5.1 Relevance of stripe-like magnetic orderings

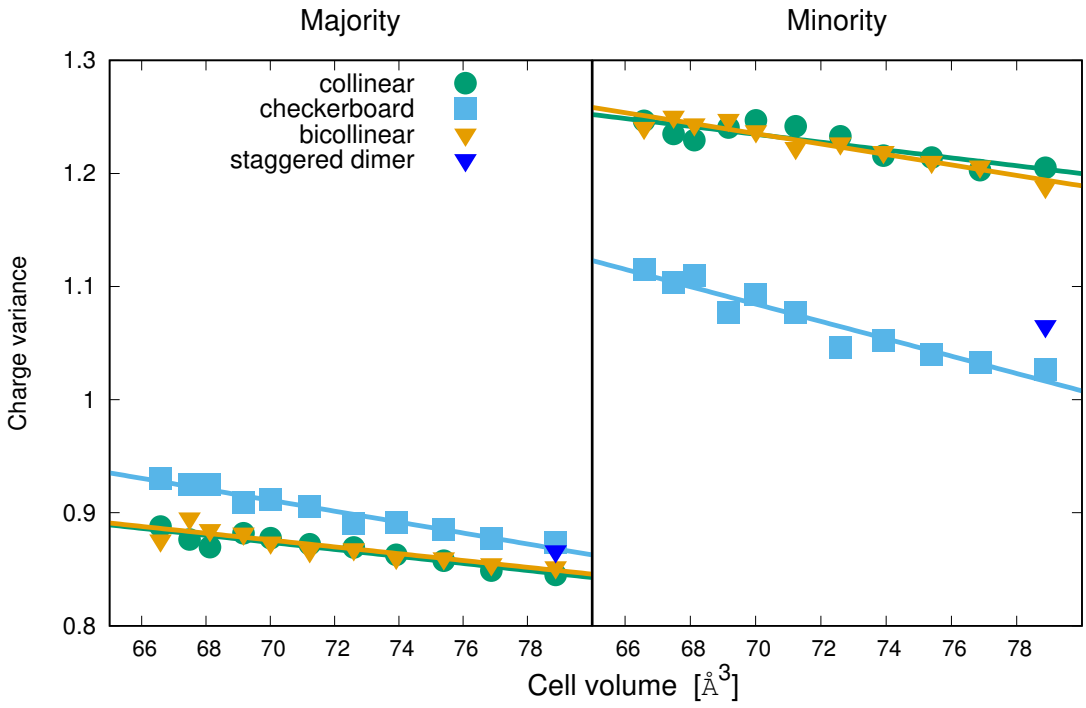
From the energetic properties explained in the previous Section, we note two classes of magnetic order in FeSe: ones which are stripe-like, and ones which are not stripe-like. The stripe-like orderings (collinear and bicollinear mainly) converge in energy with pressure, while the checkerboard and staggered dimer pattern increases in energy relative to those orderings. Similarly, the gap calculated in DMC(PBE0) distinguishes between different orderings, with metallic character in the stripe-like ordering.

There is hence an evident interplay between magnetic configurations and structural changes due to pressure application. However this effect strongly depends on the magnetic ordering imposed on the compound. In this Section we relate this different behavior to the electronic properties of FeSe; this allows us to provide a simple explanation to this behavior based on the Hund's coupling rule. In order to characterize the differences between the spin channels, we evaluate the local charge compressibility of the Fe sites: $\langle(n_{i,\sigma} - \langle n_{i,\sigma}\rangle)^2\rangle$, where n_{i_σ} is the number of electrons within a Voronoi polyhedron around the ith Fe site of spin σ. Larger values of the compressibility indicate electrons are more likely to hop between atoms. For a Fe atom with net $\uparrow$ spin, the $\uparrow$ electrons are labeled majority electrons and the $\downarrow$ minority, and vice versa for Fe atoms with net $\downarrow$ spin.

In Fig. 6.17, these results are presented as a function of pressure for the QMC(PBE0) method using mixed estimators. For all magnetic orders, the majority spin is very similar and shows a low local charge compressibility, while the minority spin is different between different magnetic orders, and its local charge compressibility is larger than the majority channel by around $0.3n_e^2$ (electron number squared). This suggests the minority electrons are more mobile, however, their ability to hop is affected by the local magnetic order. For the stripe-like orders, the minority electrons are least constrained, and their minority channel compressibility is about $0.1n_e^2$ more than the checkerboard and staggered dimer state. This measure of mobility seems to be correlated with the optical gap calculations, which predict that the checkerboard pattern induces a gap, in contrast to the other magnetic orders.

We turn our attention now to the single-particle orbital occupations. The standard approach for computing this quantity is through the diagonalization of the one-body reduced density matrix computed with the QMC variational wavefunction. This method is used with the QMC(PBE0) approach. For the QMC(opt) technique we implemented instead a different methodology which is extensively described in Appendix A. In Fig. 6.18, we present the occupations of the d orbitals in the minority spin channel for different magnetic states. Notice that for the paramagnetic state the two channels are obviously equiv-

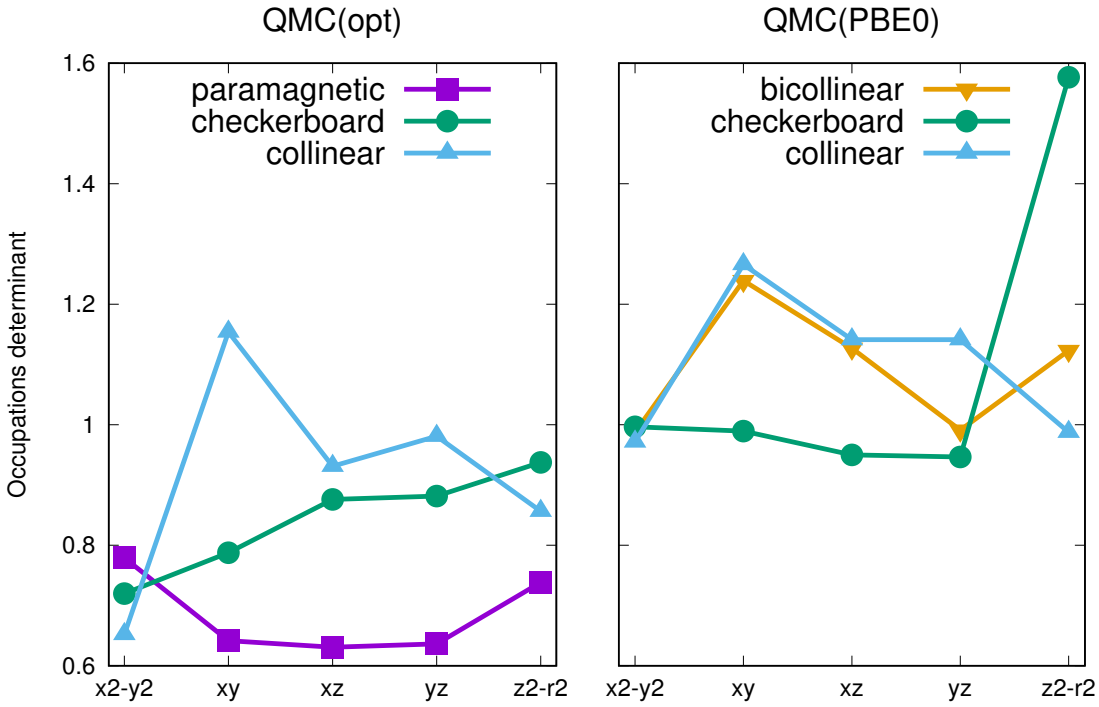
Figure 6.17 – Charge variance – in units of number of electrons squared – for different magnetic orderings in the majority and minority spin channels as a function of cell volume. This graph illustrates that the minority spin channel is more mobile whereas the checkerboard ordering's electrons are more constrained to a given iron site.



alent. QMC(opt) predictions are Fig. 6.18a while the QMC(PBE0) ones are in Fig. 6.18b. The occupations of the checkerboard state from QMC(PBE0) are strongly peaked on the $3z^2 - r^2$ orbital whose highest probability regions are placed outside the Fe plane along the z axis. From Fig. 6.18b is evident that the determinant optimization has an impact on the electron occupations. Whilst the collinear case is in good accordance with the simpler QMC(PBE0) approach, the checkerboard case shows a considerably smoother behavior. However, the qualitatively trend is maintained with the out-of-plane $3z^2 - r^2$ orbital being the most occupied at variance with all other considered magnetic configurations.

This finding gives a simple explanation for the differences in the local charge compressibility: the checkerboard pattern leads to a less hybridized electronic arrangement, as apparent from Fig. 6.19c. This idea can be confirmed by checking the off-diagonal one-body density matrix elements between Fe atoms with parallel and antiparallel net spins, as shown in Fig 6.19a for the QMC(PBE0) method. The atomic orbitals are more hybridized between parallel spin Fe atoms for the stripe-like orders. The charge degrees

Figure 6.18 – *d* orbitals occupation extracted from QMC simulations with the mixed estimators. In the left panel we show the QMC(opt) occupations computed with the procedure explained in Appendix A while in the right panel we present the QMC(PBE0) occupations.

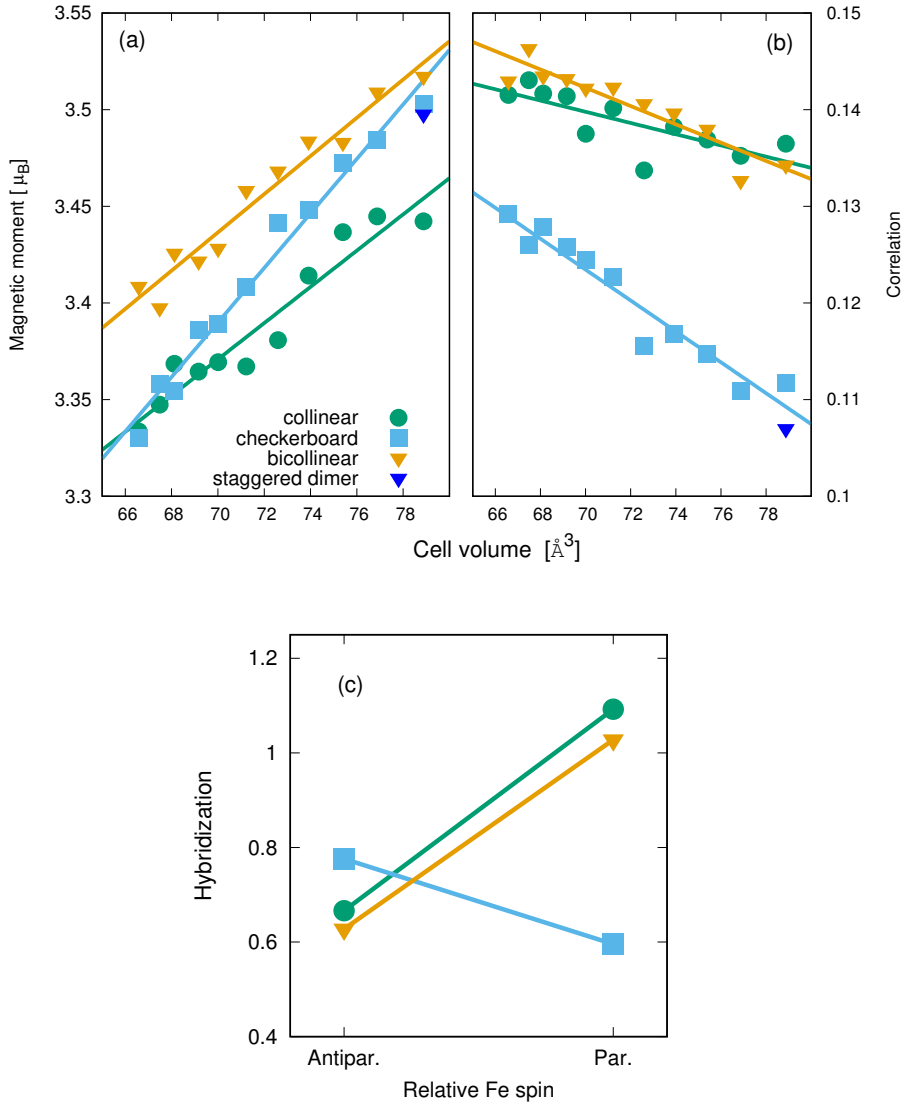


of freedom, which are mainly the minority spins from the Fe, interact strongly with the magnetic ordering. This effect also interacts with the net magnetic moment and on-site correlations as shown in Fig. 6.19b.

6.5.2 Connection with Hund's rule coupling

The importance of Hund's coupling in tuning correlation effects of multiband materials has been extensively demonstrated by means of dynamical mean field theory (DMFT) calculations [228, 229, 230]. These results highlighted its role in determining the bad metallic behavior in iron-based superconductors [231], which are therefore sometimes referred as Hund's metals. DMFT studies of $BaFe_2As_2$ have predicted that kinetic energy should be lower in the paramagnetic state, although the total energy is lower in the spin-polarized states due to Hund's coupling [232]. Correspondingly, we find that comparing the paramagnetic state and collinear state, the kinetic energy is 10.7(4) eV/f.u. larger in collinear, while the total energy is ~ 0.85 eV/f.u. lower for collinear (see Tab. 6.6). Also due to Hund's coupling, DMFT studies have predicted that the high-spin state should

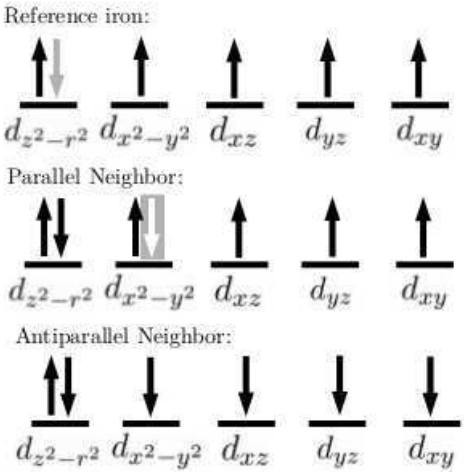
Figure 6.19 – **Panel (a)**: Magnetic moments inside each iron's Voronoi polyhedron. **Panel (b)**: Magnitude of on-site correlations, measured by $|\mathrm{Cov}(n^{\uparrow}, n^{\downarrow})|/[\mathrm{Var}(n^{\uparrow})\mathrm{Var}(n^{\downarrow})]^{1/2}$ where n^{σ} is the number of electrons of spin σ within a reference iron's Voronoi polyhedron. **Panel (c)**: Hybridization of different orderings, as measured by an average of the off-diagonal elements of the one-body density matrix, broken down by interactions between antiparallel (Antipar.) and parallel (Par.) aligned irons.



be the only highly probable state [233]. Correspondingly, we find that the iron magnetic moment fluctuates around 3.4 μ_B, with a standard deviation of 1.5 μ_B.

A simple picture based on Hund's coupling can explain the energetics and other properties presented in the results section. Hund's rules dictate that for an atom with a partially filled shell, we expect the electrons to have total spin S that maximizes the multiplicity $2S + 1$. This is consistent with our computed magnetic moment, which finds that the

Figure 6.20 – Diagrams depicting the occupations of the d-orbitals of a reference iron, one of its neighbors with parallel net magnetic moment, and another neighbor with antiparallel magnetic moment. The minority channel is spin down for the top two, and is spin up for the last. The minority electron on the reference atom is most likely to hop to a neighbor, for example, the greyed out down electron on $d_{z^2-r^2}$. It may easily hop to its parallel neighbor, which may fill its $d_{x^2-y^2}$ orbital as suggested by the grey box. It may not hop to any of the orbitals of the antiparallel neighbor, since the down spin channel is filled. Any hopping from the reference iron to its antiparallel neighbor must occur in the spin-up channel, which consequently violates Hund's rule for the reference iron.



majority channel is mostly filled, bringing the moment to around 3.1-3.4 μ_B. The spin occupation of the d-states in a reference iron is diagrammatically shown in the top row of Fig. 6.20. Also due to Hund's coupling, the electron that is most likely to hop to nearby iron atoms would be the electron in the minority channel, to keep a large S. As illustrated in Fig. 6.20, this minority channel is already filled for neighboring irons that are antiparallel, so only majority spin electrons can hop to those atoms. Conversely, minority electrons can hop to neighboring parallel irons, since that spin channel is not filled. Thus, irons with parallel spins allow the minority electrons to more easily hop about the aligned iron sites, therefore decrease the kinetic energy. As seen in Fig. 6.18 the magnetic ordering affects the occupation of the d states, hence affects the labeling of the states in Fig. 6.18, but the basic idea is unchanged.

While the minority spins require at least some parallel iron magnetic moments, the large localized magnetic moments also interact antiferromagnetically, leading to a competition between these two mechanisms. As a compromise, antiferromagnetic configurations with ferromagnetic chains emerge as the lowest energy configurations.

This picture unifies many of the observations from our calculations. The checkerboard state is distinguished from the other states by its lack of parallel nearest neighbors, similar to how the ferromagnetic state is distinguished by its lack of antiparallel neighbors. These

two extremes are higher in energy, and are disfavored as pressure increases the importance of Fe-Fe interactions. Because the checkerboard has no parallel nearest neighbor, its iron d-electrons are more trapped at a single site, leading to a low charge variance, and states that primarily occupy the $d_{z^2-r^2}$ orbital. All stripe-like states have a combination of antiparallel and parallel nearest neighbors, and allow the electrons to delocalize along the irons chains, leading to higher correlations, higher variance, and more Fe-Fe hybridization. Although the staggered dimer ordering is energetically competitive at low pressures, its energy, charge variance, and magnetic moment are similar to checkerboard, and at high pressures, becomes energetically unfavorable just as the checkerboard ordering does. Although the staggered dimer does allow some delocalization between the dimered parallel spins, the itinerant spins are still trapped on the dimers, and therefore this state's energetics follows the checkerboard behavior at higher pressures. This competition of interactions sets up a fine balance between many qualitatively different magnetic configurations. Parameters in the structure can tilt this balance one way or another, leading to a strong magneto-structural coupling. This is evident both from the strong magnetic dependence of the bulk modulus seen in Tab. 6.4 and from the dependence of the Se height on magnetic ordering.

6.6 Conclusions

In this Chapter we presented an extensive QMC investigation of FeSe. This material is one of the most studied compounds among FeSCs, however its ground state properties and its highly non-trivial behavior upon pressure application have not been completely elucidated yet. Our focus is on the normal state of this material, a fundamental step towards the description of the superconducting phase. We demonstrate the robustness and reliability of our approach by combining the results of two different yet complementary QMC flavors which have been carefully compared in this work. All main source of errors, such as finite-size effects and time step/lattice space errors in diffusion Monte Carlo have been addressed in great detail and shown to be small enough to distinguish among different magnetic configurations of FeSe.

We performed, for the first time, a fully QMC structural relaxation of this compound. With a single comprehensive *ab initio* framework, we are able to describe the main building blocks of FeSe physics, yielding an overall prediction of structural parameters as well as bulk modulus which is considerably improved with respect to previous DFT based investigations. We showed that the sole internal parameter, namely the Se height above Fe planes, is very sensitive to finite-size effects, but it can be successfully used as a tunable parameter to understand which ingredients – mainly Van der Waals forces and in-plane

electron correlation – are missing from the DFT description to yield a correct prediction.

In the second part of this Chapter we investigate several long-range magnetic patterns as a function of pressure which are possible candidates for the ground state of FeSe. At variance with DFT(PBE), both QMC flavors predict the collinear state as the lowest energy configuration over the whole pressure range; not only the ground state, but also the overall energy trend is qualitatively different at QMC level. By focusing on the energy differences relative to the collinear state, we identified two distinct magnetic classes: the "striped" patterns characterized by chains of like spin particles which are converging under pressure, and non-striped phases (such as simple AFM checkerboard pattern and the ferromagnetic phase) whose energy difference with respect to the collinear state increases when pressure is applied. Even though our wavefunction cannot support the formation of localized magnetic moments, the paramagnetic spin structure factor displays a strong competition among collinear and non collinear paramagnetic spin fluctuations. We clearly find that the contribution to spin fluctuations of the simple AFM checkerboard phase increases under pressure with respect to the collinear, similarly to what has been found for the relative energetics. This scenario points to the role of spin fluctuations in disfavoring the long-range magnetic ordering in FeSe at ambient conditions.

These findings, and the fact that only striped patterns are gapless in according to experiments, demonstrate a strong link between magnetic, structural and charge degrees of freedom in the material. This scenario can be elucidated by looking at the charge mobility (or charge variance) in the system; our results show that, for the striped patterns, the minority electrons on the high-spin Fe state are considerably freer to move with respect to configurations without aligned spin chains. This delocalization effect is strong enough to change the occupation of atomic orbitals depending on the magnetic ordering, therefore it is larger than the crystal field splitting of the orbitals. This behavior is in agreement with the prescriptions imposed by the Hund's coupling rule for like and unlike spin particles. We find therefore that a simple scenario based on the Hund's coupling can reconcile most part of our results, providing a strong link with previous studies carried out with DMFT.

The work presented in this Chapter demonstrated that the QMC approach is mature enough to attempt an *ab initio* description of unconventional superconductivity within the prototypical FeSe compound.

Chapter 7

Conclusions and outlook

In this thesis we developed and applied several original improvements to the *ab initio* quantum Monte Carlo (QMC) approach. We aimed at increasing the performance and capabilities of these stochastic first principles methods towards the description of more realistic quantum systems whose simulation by QMC will be possible using the next generation exascale supercomputing machines.

The new techniques presented here have been implemented and tested within the *TurboRVB* code [1], a comprehensive and self-contained software package capable of performing QMC calculations on a wide range of systems. This manuscript has been organized following the chronological order of the work performed during the thesis, starting from molecular cluster simulations and evolving towards a more reliable and cheaper QMC description of crystalline solids.

Part 1: QMC methods for open systems

In the first part of this manuscript we focused on finite systems. In this field, the accuracy of QMC makes this method a direct competitor of advanced quantum chemistry approaches such as coupled cluster techniques, but QMC has the great advantages of a better scaling with the number of particles and almost no parallelization overhead thanks to the stochastic nature of its principal algorithms. The main ingredient of any *ab initio* QMC calculation is the variational ansatz. In Chap. 2 we describe the powerful Jastrow antisymmetrized geminal power (JAGP) functional form. The JAGP wavefunction is the *ab initio* realization of the resonating valence bond theory, an alternative to the well-known molecular orbital theory for the description of chemical bonds. Being the two-electron generalization of the Slater determinant, the JAGP evaluation cost displays the same N^3 scaling and, furthermore, it contains the Hartree-Fock determinant as a special case. These features combined with the use of a Gaussian localized single-particle basis

set enable a fine tuning of the level of correlation in the system by keeping a flexible and very compact representation.

Flexibility is also crucial in order to obtain a well optimized wavefunction. Indeed, a wavefunction as close as possible to the true ground state is unavoidable since the main QMC algorithms used in this work – variational Monte Carlo (VMC) and lattice regularized diffusion Monte Carlo (LRDMC) – obey the well-known zero-variance principle. To this purpose, we need two ingredients. At first, a good starting guess which, in our implementation, is the DFT(LDA) single determinant wavefunction; this is obtained with a built-in DFT package which uses exactly the same setup (basis set, pseudopotential) as the QMC part. The second ingredient is an efficient wavefunction optimization method. We used an energy minimization scheme based on the stochastic reconfiguration (SR) technique which allows us to optimize up to 10^4 variational parameter *at the same* time using Monte Carlo integration. The SR method requires us to compute the derivatives of a few quantities, such as the wavefunction and the local energy, with respect to the whole set of parameters, thus the number of inputs (variational parameters) is several order of magnitude larger than the number of outputs (wavefunction and local energy). We attempt to solve this problem by the adjoint algorithmic differentiation (AAD), a technique which allows the computations of any number of derivatives of a function implemented as a computer program, with the *same efficiency* as the function itself. AAD also adds modularity and flexibility to the QMC code. In fact, the implementation of a new variational ansatz only requires one to devise an efficient algorithm to evaluate it, whereas its derivatives are automatically computed with practically no coding effort. This aspect enables an efficient and quick implementation of several important wavefunction forms which can considerably expand the capabilities of the *TurboRVB* code. They include the Pfaffian pairing form [39], the generalization of JAGP to multiple orthogonal geminals, known as the antisymmetric geminal product [234], or even multideterminantal CI-like wavefunctions [9].

Despite its efficiency, QMC remains considerably more expensive with than effective one-electron methods such as DFT, in particular when tackling large molecular clusters. A crucial step to cure this issue is the reduction of the basis set size. In Chap. 3 we presented an original scheme for effectively decreasing the number of basis-set parameters with a negligible loss of accuracy. Our procedure allows us to generate optimal hybrid orbitals defined as eigenvectors of a suitably projected density matrix. We showed that the geminal embedded orbitals (GEOs) display better performance with respect to the standard natural orbitals and apply them to the paradigmatic case of the H_2O molecule. With ~ 2 times fewer parameters in the basis, we are able to reach a total energy which is only ~ 1 mHa higher than our best ansatz as well as very close to the best results

published in the literature.

Another useful feature of GEOs is that they are fully portable to larger systems. We transfer them to the protonated water dimer $H_5O_2^+$, a simple but extensively studied model of proton transfer (PT) reactions in liquid water. Due to its very sensitive thermal behavior – responsible for the abnormally high proton diffusion along the hydrogen bond network – PT requires an accuracy in the energetics lower than the chemical accuracy of 1 Kcal/mol and is thus very difficult to describe with single-particle theories such as DFT. We demonstrated that our framework is able to reach this accuracy and to reproduce, within the statistical error, energetics and structural properties as predicted by the quantum chemistry "gold standard" CCSD(T). However, as shown by a test on a larger model of PT, QMC is able to simulate much larger clusters than CCSD(T), which suffers from a scaling of $\sim N^7$ with the number of electrons.

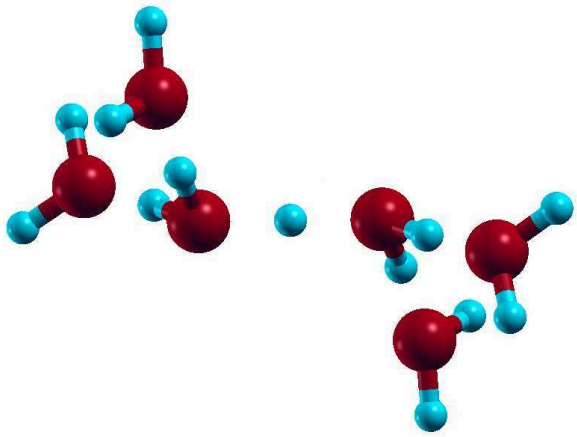


Figure 7.1 – Larger proton transfer cluster composed by 6 H_2O molecules and one excess proton. This model fills the first solvation shell of the Zundel ion.

These very encouraging results open the way to several future directions in the research on PT, whose exact microscopic mechanism still represents an open problem in chemistry with important applications, for instance, in the design of fuel cells:

- The most direct continuation of the work presented in Chap. 3 is the study of solvation effects on the zero-temperature energy landscape and static proton transfer barriers of $H_5O_2^+$. By employing the same setup, we can simulate larger molecular clusters such as the one shown in Fig. 7.1, which completes the first solvation shell of the Zundel ion. The investigation of larger clusters is unavoidable in order to understand the best simulation model accounting for all necessary physical ingredients needed by PT.

- Once the aforementioned model has been identified, the next step will beto apply

the recently developed *ab initio* molecular dynamics (AIMD) technique [8] based on Born-Oppenheimer QMC forces to carry out finite-temperature calculations at ambient conditions. In this way, QMC predictions of experimentally available quantities – for instance the oxygen-oxygen and oxygen-hydrogen radial distribution functions or the PT diffusion coefficient – would be possible. If the accuracy shown by QMC at zero-temperature is confirmed within finite temperature simulations, our technique will likely be able to achieve an agreement with experimental results on PT which is still missing nowadays. Moreover, the simulation of more than 128 H_2O molecules by QMC-AIMD will be feasible in the near future. This would enable us to directly access concerted proton transfer reactions within a sufficiently large portion of the hydrogen bond network. Hence we believe that an accurate simulation of the microscopic PT mechanism based on quantum Monte Carlo is within reach.

Part 2: QMC methods for periodic systems

The second part of this thesis work was devoted to the implementation of long-established as well as original techniques for an accurate description of extended periodic systems. To this purpose, QMC is one of the few highly correlated *ab initio* techniques effectively capable of predicting physical properties of realistic solids. In Chap. 4, which possesses the same structure as Chap. 2, we explained the most important modifications needed by the QMC algorithms for working with extended systems. On one hand, they come from the fact that the single-particle basis orbitals for periodic systems need to satisfy the well-known Bloch theorem, thus requiring complex-valued functions. Switching from real to complex arithmetic implied a considerable coding effort which concerned every step of a typical QMC run: the DFT(LDA) initial calculations, the wavefunction optimization as well as the diffusion Monte Carlo algorithm. On the other hand, solids must be simulated within appropriate supercells. This leads to the generalization of the concept of a **k**-point to its many-body equivalent: the twist condition. With a smart choice of the twist conditions for $\uparrow$ and $\downarrow$ spin particles we have been able to write the determinantal part of our ansatz in a twist independent way, thus enabling the full optimization of multiple twist conditions at the cost of a single point.

The use of many twists is indeed crucial to reduce finite size (FS) effects on the energy and other observables coming from the one-body part of the quantum Hamiltonian. In fact, twist-averaged boundary conditions (TABC) [162], the most important method devised to this purpose, is based on averaging over several twists. In Chap. 5 we presented two alternative approaches to this technique, all based on the concept of twisted boundary conditions. The first one is the exact special twist (EST) method devised to find values of the twist condition reproducing the mean-field infinite-size limit of the energy with an

arbitrary numerical accuracy. By studying several paradigmatic metallic systems (bcc hydrogen, bcc lithium and β-tin silicon), we showed that the EST method, with a single twist, performs similarly to TABC for total energy extrapolation, whereas it converges slightly slower to the thermodynamic limit for the more delicate case of ionic forces. The second method, the flavor twisted boundary conditions (FTBC) approach, is an evolution of TABC based on an approximate mapping of the Hamiltonian defined on a large supercell into the Hamiltonian of a smaller cell. In principle, this should go beyond one-body FS corrections; however, the quality of the approximation made by FTBC is not sufficient to suppress the diverging $1/q$ term in the Coulomb potential. Hence, the FTBC method produces results similar to those obtaine using the standard TABC method in curing FS errors.

The evolution of the *TurboRVB* code to treat solids with non-trivial twist conditions paves the way to a vast amount of possible applications, ranging from the large domain of low-energy phase transitions requiring an accurate treatment of the Brillouin zone sampling – for instance high-temperature superconductivity – to molecular dynamics simulations of disordered liquids – such as hydrogen, water and biological systems – where FS effects are extremely relevant. As a direct continuation of Chap. 5, it is important to test the EST method with more extensive molecular dynamics simulations to understand whether we can generally assume that the EST value varies only a little from one sampled configuration to another. On the other hand, since we noticed that the EST values do not appreciably depend on the reference mean-field method (see Tab. 5.2), we can conceive an automatic determination of the EST by using the simple and extremely cheap Hartree method as reference. In this way EST can be computed "on the fly" at each configuration in the dynamics.

The developments of Chaps. 4 and 5 will be likely used in the continuation of the work presented in Chap. 6. This is devoted to the application of continuum QMC for solids to high-temperature superconductivity. In collaboration with the group of Lucas K. Wagner at the University of Illinois, Urbana-Champaign, we studied the structural and magnetic properties of the normal state of iron selenide (FeSe) by means of two complementary QMC flavors. We show that QMC is able to considerably improve structural predictions with respect to previous DFT-based studies. The overall energetics of different magnetic orderings as functions of pressure are also found to be qualitatively different. In this respect, we demonstrated a strong coupling between the structure and the magnetic ordering in the system; in particular, we distinguished between two magnetic configurations: the collinear patterns whose energies become degenerate as pressure is increased, and the non-collinear orderings which become more different in energy for larger pressures. We elucidated our results on energetics by studying the electron mobility and noticed that it

is mainly driven by minority spin electrons, which are more mobile in the collinear configurations. The minority electrons simply obey the Hund's coupling rule, a straightforward application of which allows us to reconcile most of our findings.

Our successful investigation on the normal state of FeSe opens several interesting research directions. Furthermore, the developments carried out in Chaps. 4 and 5 – implemented in parallel to the work on FeSe presented in this manuscript – will improve the quality of our QMC description of FeSe and high-temperature superconductors in general. Among the possible continuations to our work on FeSe, the most relevant are the following:

- The investigation of the tetragonal-to-orthorhombic transition in FeSe. In particular, the GEOs introduced in Chap. 3 are able to probe the crystal field splitting of the orbitals [103]. This can give new insights on the debated orbital origin of this transition.

- As already mentioned, the implementation of a Pfaffian wavefunction would allow us to include within our QMC approach both static and dynamic spin-spin correlation. This improvement may be essential in order to understand the origin of the paramagnetism of FeSe at ambient conditions and to provide an *ab initio* confirmation of the paramagnetic-to-AFM phase transition identified in recent experimental work [205, 207, 208]. Moreover, it would allow us to unravel the possible magnetic origin of the tetragonal-to-orthorhombic transition at ambient pressure, thus giving a definite *ab initio* answer to this long-standing problem.

- The investigation of the normal state of the parent compound of the iron chalcogenide family: iron telluride (FeTe). Since this material shows collinear ordering, but no superconductivity over the whole phase diagram, comparing its magnetic properties with FeSe can shed new lights on the connection of magnetism and superconductivity within FeSCs.

- The transition from bulk to few-layer and eventually to single-layer FeSe changes the electronic and superconducting properties [202, 203]. Our JAGP variational ansatz is tailored for investigating superconductivity through a direct determination of the electron pairing function within an intrinsically unconventional context since no electron-phonon coupling is included in our description. Additionally, the developments carried out in Chaps. 4 and 5 allow us to resolve the QMC pairing function in reciprocal space. The comparison between the pairing functions obtained at different thicknesses of the FeSe crystal (bulk, few layers, single layers) may help

us understand the sharp increase of the critical temperature as the thickness is reduced.

The perspectives outlined in this Section represent a small part of the possibilities opened up by the algorithmic and programming developments presented in this thesis, that have already been employed in several to-be-published works.

We showed in this work that *ab initio* QMC methods are now feasible on a wide range of small to medium size systems with an amount of computer time achievable using most supercomputing machines available nowadays. Subchemical accuracy (< 0.01 eV/atom $= 1$ Kcal/mol) can be reached both on molecular clusters and crystalline solids, thus ensuring a reliable description of materials at the quantum level. Furthermore, the information accessible through the many-body wavefunction, fully avaliable within the QMC approach, provides an invaluable tool for investigating the role of electron correlation at the microscopic level.

Very large-scale simulations containing several thousands of atoms will likely not be possible using QMC techniques even in the era of exascale computing. However, for the aforementioned reasons supported by this thesis work, we strongly believe that QMC will soon become the standard method for studying the wide range of microscopic phenomena and fundamental problems which require less than 1000 atoms to be simulated.

Bibliography

[1] Sandro Sorella. TurboRVB, quantum Monte Carlo software for electronic structure calculations. http://people.sissa.it/ sorella/web/.

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

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

[4] D. M. Arnow, M. H. Kalos, Michael A. Lee, and K. E. Schmidt. Green's function Monte Carlo for few fermion problems. *J. Chem. Phys.*, 77, 1982.

[5] Michele Casula, Claudia Filippi, and Sandro Sorella. Diffusion Monte Carlo method with lattice regularization. *Phys. Rev. Lett.*, 95:100201, 2005.

[6] D. M. Ceperley and B. J. Alder. Quantum Monte Carlo for molecules: Green's function and nodal release. *J. Chem. Phys.*, 81, 1984.

[7] Andrea Zen, Ye Luo, Guglielmo Mazzola, Leonardo Guidoni, and Sandro Sorella. *Ab initio* molecular dynamics simulation of liquid water by quantum Monte Carlo. *J. Chem. Phys.*, 142:144111, 2015.

[8] S. Sorella and G. Mazzola. Accelerated ab-initio molecular dynamics: probing the weak dispersive forces in dense liquid hydrogen. *arXiv:1605.08423*[cond-mat.str-el], 2016.

[9] 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. *J. Chem. Phys.*, 131, 2009.

[10] Shiwei Zhang and Henry Krakauer. Quantum Monte Carlo method using phase-free random walks with Slater determinants. *Phys. Rev. Lett.*, 90:136401, 2003.

[11] Jeffrey C. Grossman. Benchmark quantum Monte Carlo calculations. *J. Chem. Phys.*, 117, 2002.

[12] N. A. Benedek, I. K. Snook, M. D. Towler, and R. J. Needs. Quantum Monte Carlo calculations of the dissociation energy of the water dimer. *J. Chem. Phys.*, 125:104302, 2006.

[13] P. Hohenberg and W. Kohn. Inhomogeneous electron gas. *Phys. Rev.*, 136:B864–B871, 1964.

[14] W. Kohn and L. J. Sham. Self-consistent equations including exchange and correlation effects. *Phys. Rev.*, 140:A1133–A1138, 1965.

[15] Efthimios Kaxiras. *Atomic and Electronic Structure of Solids*. Cambridge University Press, 2003. Cambridge Books Online.

[16] D. M. Ceperley and B. J. Alder. Ground state of the electron gas by a stochastic method. *Phys. Rev. Lett.*, 45:566–569, 1980.

[17] J. P. Perdew and Alex Zunger. Self-interaction correction to density-functional approximations for many-electron systems. *Phys. Rev. B*, 23:5048–5079, 1981.

[18] 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.: Condens. Matter*, 21(39):395502, 2009.

[19] G. Kresse and J. Furthm˙Efficiency of *ab initio* total energy calculations for metals and semiconductors using a plane-wave basis set. *Comput. Mater. Sci.*, 6:15 – 50, 1996.

[20] D. Alfè and M. J. Gillan. Efficient localized basis set for quantum Monte Carlo calculations on condensed matter. *Phys. Rev. B*, 70:161101, 2004.

[21] William D. Parker, C.J. Umrigar, Dario Alfè, F.R. Petruzielo, Richard G. Hennig, and John W. Wilkins. Comparison of polynomial approximations to speed up planewave-based quantum Monte Carlo calculations. *J. Comp. Phys.*, 287:77 – 87, 2015.

[22] F. R. Petruzielo, Julien Toulouse, and C. J. Umrigar. Compact and flexible basis functions for quantum Monte Carlo calculations. *J. Chem. Phys.*, 132:094109, 2010.

[23] M. Burkatzki, Claudia Filippi, and M. Dolg. Energy-consistent pseudopotential database. Available at http://www.burkatzki.com/pseudos/index.2.html.

[24] Available at https://bse.pnl.gov/bse/portal.

[25] A. Erba C. M. Zicovich-Wilson B. Civalleri S. Casassa L. Maschio M. Ferrabone M. De La Pierre P. D'Arco Y. Noel M. Causa M. Rerat B. Kirtman R. Dovesi, R. Orlando. *Int. J. Quantum Chem.*, 114:1287, 2014.

[26] S. Azadi, C. Cavazzoni, and S. Sorella. Systematically convergent method for accurate total energy calculations with localized atomic orbitals. *Phys. Rev. B*, 82:125112, 2010.

[27] Dolly Vijay, Hidehiro Sakurai, and G. Narahari Sastry. The impact of basis set superposition error on the structure of $\pi - \pi$ dimers. *Int. J. Quantum Chem.*

[28] E. Anderson, Z. Bai, C. Bischof, S. Blackford, J. Demmel, J. Dongarra, J. Du Croz, A. Greenbaum, S. Hammarling, A. McKenney, and D. Sorensen. *LAPACK Users' Guide*. Society for Industrial and Applied Mathematics, Philadelphia, PA, third edition, 1999.

[29] J. G. Bednorz and K. A. Müller. Possible high-T_c superconductivity in the BaLaCuO system. *Z. Phys. B: Condensed Matter*, 64:189–193, 1986.

[30] Linus Pauling. *The Nature of the Chemical Bond and the Structure of Molecules and Crystals: An Introduction to Modern Structural Chemistry*. Cornell University Press, Ithaca, New York, 1939.

[31] L. Pauling. A resonating-valence-bond theory of metals and intermetallic compounds. *Proc. R. Soc. A*, 196:343–362, 1949.

[32] P.W. Anderson. Resonating valence bonds: A new kind of insulator? *Mater. Res. Bull.*, 8:153 – 160, 1973.

[33] P. W. Anderson, G. Baskaran, Z. Zou, and T. Hsu. Resonating valence-bond theory of phase transitions and superconductivity in la_2cuo_4-based compounds. *Phys. Rev. Lett.*, 58:2790–2793, 1987.

[34] J. Bardeen, L. N. Cooper, and J. R. Schrieffer. Theory of superconductivity. *Phys. Rev.*, 108:1175–1204, 1957.

[35] Michele Casula and Sandro Sorella. Geminal wave functions with Jastrow correlation: A first application to atoms. *J. Chem. Phys.*, 119, 2003.

[36] Michele Casula. *New QMC aproaches for the simulation of electronic systems: a first application to aromatic molecules and transition metal compounds*. PhD thesis, International School for Advanced Studies, Trieste (Italy), 2005.

[37] Nicolas Dupuy, Samira Bouaouli, Francesco Mauri, Sandro Sorella, and Michele Casula. Vertical and adiabatic excitations in anthracene from quantum Monte Carlo: Constrained energy minimization for structural and electronic excited-state properties in the JAGP ansatz. *J. Chem. Phys.*, 142:214109, 2015.

[38] C. Lhuillier J.P. Bouchaud, A. Georges. Pair wave functions for strongly correlated fermions and their determinantal representation. *J. Phys. (Paris)*, 49, 1987.

[39] M. Bajdich, L. Mitas, L. K. Wagner, and K. E. Schmidt. Pfaffian pairing and backflow wavefunctions for electronic structure quantum Monte Carlo methods. *Phys. Rev. B*, 77:115112, 2008.

[40] 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, 2007.

[41] Tosio Kato. On the eigenfunctions of many-particle systems in quantum mechanics. *Comm. Pure App. Math.*, 10:151–177, 1957.

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

[43] Nicholas Metropolis, Arianna W. Rosenbluth, Marshall N. Rosenbluth, Augusta H. Teller, and Edward Teller. Equation of state calculations by fast computing machines. *J. Chem. Phys.*, 21, 1953.

[44] H. F. Trotter. On the product of semi-groups of operators. *Proc. Amer. Math. Soc.*, 10:545–551, 1959.

[45] Masuo Suzuki. Decomposition formulas of exponential operators and Lie exponentials with some applications to quantum mechanics and statistical physics. *J. Math. Phys.*, 26, 1985.

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

[47] Julien Toulouse, Roland Assaraf, and Cyrus J. Umrigar. Chapter fifteen - introduction to the variational and diffusion monte carlo methods. In Philip E. Hoggan and Telhat Ozdogan, editors, *Electron Correlation in Molecules – ab initio Beyond Gaussian Quantum Chemistry*, volume 73 of *Advances in Quantum Chemistry*, pages 285 – 314. Academic Press, 2016.

[48] R.N Barnett, P.J Reynolds, and W.A Lester. Monte Carlo algorithms for expectation values of coordinate operators. *J. Comp. Phys.*, 96:258 – 276, 1991.

[49] 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, 1999.

[50] D. M. Ceperley. The statistical error of Green's function Monte Carlo. *J. Stat. Phys.*, 43:815–826, 1986.

[51] A. Ma, N. D. Drummond, M. D. Towler, and R. J. Needs. All-electron quantum Monte Carlo calculations for the noble gas atoms He to Xe. *Phys. Rev. E*, 71:066704, 2005.

[52] M. Burkatzki, C. Filippi, and M. Dolg. Energy-consistent pseudopotentials for quantum Monte Carlo calculations. *J. Chem. Phys.*, 126:234105, 2007.

[53] M. Burkatzki, Claudia Filippi, and M. Dolg. Energy-consistent small-core pseudopotentials for 3d-transition metals adapted to quantum Monte Carlo calculations. *J. Chem. Phys.*, 129:164115, 2008.

[54] William W. Tipton, Neil D. Drummond, and Richard G. Hennig. Importance of high-angular-momentum channels in pseudopotentials for quantum Monte Carlo. *Phys. Rev. B*, 90:125110, Sep 2014.

[55] 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, 1990.

[56] Luboš Mitáš, Eric L. Shirley, and David M. Ceperley. Nonlocal pseudopotentials and diffusion Monte Carlo. *J. Chem. Phys.*, 95:3467–3475, 1991.

[57] 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.*, 132:154113, 2010.

[58] Sandro Sorella and Luca Capriotti. Green function Monte Carlo with stochastic reconfiguration: An effective remedy for the sign problem. *Phys. Rev. B*, 61:2599–2612, 2000.

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

[60] Claudia Filippi and Stephen Fahy. Optimal orbitals from energy fluctuations in correlated wave functions. *J. Chem. Phys.*, 112, 2000.

[61] Friedemann Schautz and Claudia Filippi. Optimized Jastrow–Slater wave functions for ground and excited states: Application to the lowest states of ethene. *J. Chem. Phys.*, 120, 2004.

[62] Sandro Sorella. Generalized Lanczos algorithm for variational quantum Monte Carlo. *Phys. Rev. B*, 64:024512, 2001.

[63] Sandro Sorella. Wave function optimization in the variational Monte Carlo method. *Phys. Rev. B*, 71:241103, 2005.

[64] J. H. Bartlett, J. J. Gibbons, and C. G. Dunn. The normal helium atom. *Phys. Rev.*, 47:679–680, 1935.

[65] 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, 1988.

[66] 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.*, 36:217–230, 1989.

[67] M. Meierovich, A. Mushinski, and M. P. Nightingale. Van der Waals clusters in the ultraquantum limit: A Monte Carlo study. *J. Chem. Phys.*, 105, 1996.

[68] Jeongnim Kim, Kenneth P. Esler, Jeremy McMinis, Miguel A. Morales, Bryan K. Clark, Luke Shulenburger, and David M. Ceperley. QMCPACK. http://qmcpack.cmscc.org/.

[69] Michele Casula, Claudio Attaccalite, and Sandro Sorella. Correlated geminal wave function for molecules:an efficient resonating valence bond approach. *J. Chem. Phys.*, 121:7110–7126, 2004.

[70] Ye Luo. Ab initio *molecular dynamics of water by quantum Monte Carlo*. PhD thesis, International School for Advanced Studies, Trieste (Italy), 2015.

[71] 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 Comp.*, 9:4332–4350, 2013.

[72] Sandro Sorella and Luca Capriotti. Algorithmic differentiation and the calculation of forces by quantum Monte Carlo. *J. Chem. Phys*, 133, 2010.

[73] Roland Assaraf and Michel Caffarel. Zero-variance zero-bias principle for observables in quantum Monte Carlo: Application to forces. *J. Chem. Phys.*, 119:10536–10552, 2003.

[74] Simone Chiesa, D. M. Ceperley, and Shiwei Zhang. Accurate, efficient, and simple forces computed with quantum Monte Carlo methods. *Phys. Rev. Lett.*, 94:036404, 2005.

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

[76] Andreas Griewank. *Evaluating Derivatives: Principles and Techniques of Algorithmic Differentiation*. Frontiers Appl. Math., Philadelphia, 2000.

[77] Luca Capriotti. Fast greeks by algorithmic differentiation. *J. Comput. Finance*, 14:3, 2011.

[78] Institut National de Recherche en Informatique et Automatique Tropics team. TAPENADE, au automatic differentiation engine.

[79] Pier Luigi Silvestrelli and Michele Parrinello. Structural, electronic, and bonding properties of liquid water from first principles. *J. Chem. Phys.*, 111:3572–3580, 1999.

[80] 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.*, 122:014515, 2005.

[81] 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 Ih with the Perdew–Burke–Ernzerhof and Becke–Lee–Yang–Parr functionals. *J. Chem. Phys.*, 130:221102, 2009.

[82] 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.*, 126:164501, 2007.

[83] Dominik Marx. Proton transfer 200 years after von grotthuss: Insights from ab initio simulations. *ChemPhysChem*, 7:1848–1870, 2006.

[84] Johann Deisenhofer, Otto Epp, Irmarg Sinning, and Hartmut Michel. Crystallographic refinement at 2.3 å resolution and refined model of the photosynthetic reaction centre from rhodopseudomonas viridis. *J. Mol. Biol.*, 246:429–457, 1995.

[85] M. Eigen. Proton transfer, acid-base catalysis, and enzymatic hydrolysis. part i: Elementary processes. *Angew. Chem. Int. Ed.*, 3:1–19, 1964.

[86] Omar F. Mohammed, Dina Pines, Ehud Pines, and Erik T.J. Nibbering. Aqueous bimolecular proton transfer in acid–base neutralization. *Chem. Phys.*, 341:240–257, 2007.

[87] Udo W. Schmitt and Gregory A. Voth. The computer simulation of proton transport in water. *J. Chem. Phys.*, 111, 1999.

[88] Hadas Lapid, Noam Agmon, Matt K. Petersen, and Gregory A. Voth. A bond-order analysis of the mechanism for hydrated proton mobility in liquid water. *J. Chem. Phys.*, 122, 2005.

[89] Z. Luz and S. Meiboom. The activation energies of proton transfer reactions in water. *J. Am. Chem. Soc.*, 86:4768–4769, 1964.

[90] Agmon Noam. Hydrogen bonds, water rotation and proton mobility. *J. Chim. Phys.*, 93:1714–1736, 1996.

[91] Dominik Marx, Mark E. Tuckerman, Jurg Hutter, and Michele Parrinello. The nature of the hydrated excess proton in water. *Nature*.

[92] Dominik Marx, Amalendu Chandra, and Mark E. Tuckerman. Aqueous basic solutions: Hydroxide solvation, structural diffusion, and comparison to the hydrated proton. *Chem. Rev.*, 110:2174–2216, 2010. PMID: 20170203.

[93] Ali Hassanali, Federico Giberti, Jérôme Cuny, Thomas D. Kühne, and Michele Parrinello. Proton transfer through the water gossamer. *Proc. Natl. Acad. Sci. U.S.A.*, 110:13723–13728, 2013.

[94] E. Wicke, M. Eigen, and H. Ackermann. *Z. Phys. Chem.*, 1:340–364, 1954.

[95] G. Zundel and H. Metzger. *Z. Phys. Chem.*, 58:225–245, 1968.

[96] Yaoming Xie, Richard B. Remington, and Henry F. Schaefer. The protonated water dimer: Extensive theoretical studies of $H_5O_2^+$. *J. Chem. Phys.*, 101:4878–4884, 1994.

[97] Edward F. Valeev and Henry F. Schaefer. The protonated water dimer: Brueckner methods remove the spurious C1 symmetry minimum. *J. Chem. Phys.*, 108:7197–7201, 1998.

[98] Jeffrey M. Headrick, Eric G. Diken, Richard S. Walters, Nathan I. Hammer, Richard A. Christie, Jun Cui, Evgeniy M. Myshakin, Michael A. Duncan, Mark A. Johnson, and Kenneth D. Jordan. Spectral signatures of hydrated proton vibrations in water clusters. *Science*, 308:1765–1769, 2005.

[99] Mark E. Tuckerman, Dominik Marx, Michael L. Klein, and Michele Parrinello. On the quantum nature of the shared proton in hydrogen bonds. *Science*, 275:817–820, 1997.

[100] Xinchuan Huang, Bastiaan J. Braams, and Joel M. Bowman. *Ab initio* potential energy and dipole moment surfaces for H_5O_+ . *J. Chem. Phys.*, 122:044308, 2005.

[101] Sorella Sorella, Nicolas Devaux, Mario Dagrada, Guglielmo Mazzola, and Michele Casula. Geminal embedding scheme for optimal atomic basis set construction in correlated calculations. *J. Chem. Phys.*, 143:244112, 2015.

[102] Karen A. Hallberg. New trends in density matrix renormalization. *Adv. Phys.*, 55:477–526, 2006.

[103] N. Devaux, M. Casula, F. Decremps, and S. Sorella. Electronic origin of the volume collapse in cerium. *Phys. Rev. B*, 91:081101, 2015.

[104] Per-Olov Löwdin and Harrison Shull. Natural orbitals in the quantum theory of two-electron systems. *Phys. Rev.*, 101:1730–1739, 1956.

[105] I. Mayer. Non-orthogonal localized orbitals and orthogonal atomic hybrids derived from Mulliken's population analysis. *Chem. Phys. Lett.*, 242:499 – 506, 1995.

[106] Frank Neese, Frank Wennmohs, and Andreas Hansen. Efficient and accurate local approximations to coupled-electron pair approaches: An attempt to revive the pair natural orbital method. *J. Chem. Phys.*, 130:114108, 2009.

[107] Alan E. Reed, Robert B. Weinstock, and Frank Weinhold. Natural population analysis. *J. Chem. Phys.*, 83:735–746, 1985.

[108] Ernest R. Davidson and David Feller. Molecular properties of water. *Chem. Phys. Lett.*, 104:54 – 58, 1984.

[109] Trygve Helgaker, Wim Klopper, Henrik Koch, and Jozef Noga. Basis-set convergence of correlated calculations on water. *J. Chem. Phys.*, 106, 1997.

[110] Mariapia Marchi, Sam Azadi, Michele Casula, and Sandro Sorella. Resonating valence bond wave function with molecular orbitals: Application to first-row molecules. *J. Chem. Phys.*, 131:154116, 2009.

[111] W. S. Benedict, N. Gailar, and Earle K. Plyler. Rotationvibration spectra of deuterated water vapor. *J. Chem. Phys.*, 24:1139–1165, 1956.

[112] Arne Lüchow and Reinhold F. Fink. On the systematic improvement of fixed-node diffusion quantum Monte Carlo energies using pair natural orbital CI guide functions. *J. Chem. Phys.*, 113:8457–8463, 2000.

[113] David Feller, Caroline M. Boyle, and Ernest R. Davidson. Oneelectron properties of several small molecules using near Hartree–Fock limit basis sets. *J. Chem. Phys.*, 86:3424–3440, 1987.

[114] Mario Dagrada, Michele Casula, Antonino M. Saitta, Sandro Sorella, and Francesco Mauri. Quantum Monte Carlo study of the protonated water dimer. *J. Chem. Theory Comp.*, 10:1980–1993, 2014.

[115] David J. Wales. Rearrangements and tunneling splittings of protonated water dimer. *J. Chem. Phys.*, 110:10403–10409, 1999.

[116] Alexander A. Auer, Trygve Helgaker, and Wim Klopper. Accurate molecular geometries of the protonated water dimer. *Phys. Chem. Chem. Phys.*, 2:2235–2238, 2000.

[117] Uwe Bergmann, Andrea Di Cicco, Philippe Werner, Emiliano Principi, Pieter Glatzel, and Anders Nilsson. Nearest-neighbor oxygen distances in liquid water and ice observed by x-ray Raman based extended x-ray absorption fine structure. *J. Chem. Phys.*, 127:174504, 2007.

[118] Dongqing Wei and D. R. Salahub. Hydrated proton clusters and solvent effects on the proton transfer barrier: A density functional study. *J. Chem. Phys.*, 101:7633–7642, 1994.

[119] Shaumo Sadhukhan, David Muñoz, Carlo Adamo, and Gustavo E. Scuseria. Predicting proton transfer barriers with density functional methods. *Chem. Phys. Lett.*, 306:83–87, 1999.

[120] Eric Neuscamman. Size consistency error in the antisymmetric geminal power wave function can be completely removed. *Phys. Rev. Lett.*, 109:203001, 2012.

[121] Michael W. Schmidt, Kim K. Baldridge, Jerry A. Boatz, Steven T. Elbert, Mark S. Gordon, Jan H. Jensen, Shiro Koseki, Nikita Matsunaga, Kiet A. Nguyen, Shujun Su, Theresa L. Windus, Michel Dupuis, and John A. Montgomery. General atomic and molecular electronic structure system. *J. Comput. Chem.*, 14:1347–1363, 1993.

[122] John A. Pople, Martin HeadGordon, and Krishnan Raghavachari. Quadratic configuration interaction. a general technique for determining electron correlation energies. *J. Chem. Phys.*, 87, 1987.

[123] So Hirata, Rafał Podeszwa, Motoi Tobita, and Rodney J. Bartlett. Coupled-cluster singles and doubles for extended systems. *J. Chem. Phys.*, 120, 2004.

[124] Michaël Deleuze, Joseph Delhalle, Barry T. Pickup, and Jean-Louis Calais. Size-consistency aspects and physical interpretation of many-body Green's-function calculations on extended chains. *Phys. Rev. B*, 46:15668–15682, 1992.

[125] Jun-Qiang Sun and Rodney J. Bartlett. Convergence of many-body perturbation methods with lattice summations in extended systems. *J. Chem. Phys.*, 106, 1997.

[126] James J. Shepherd, Thomas M. Henderson, and Gustavo E. Scuseria. Range-separated brueckner coupled cluster doubles theory. *Phys. Rev. Lett.*, 112:133002, 2014.

[127] James McClain, Johannes Lischner, Thomas Watson, Devin A. Matthews, Enrico Ronca, Steven G. Louie, Timothy C. Berkelbach, and Garnet Kin-Lic Chan. Spectral functions of the uniform electron gas via coupled-cluster theory and comparison to the *GW* and related approximations. *Phys. Rev. B*, 93:235139, 2016.

[128] G. H. Booth, A. Gruneis, G. Kresse, and A. Alavi. Towards an exact description of electronic wavefunctions in real solids. *Nature*, 493:365–370, 2013.

[129] Chandrima Mitra, Jaron T. Krogel, Juan A. Santana, and Fernando A. Reboredo. Many-body ab initio diffusion quantum Monte Carlo applied to the strongly correlated oxide NiO. *J. Chem. Phys.*, 143:164710, 2015.

[130] Juan A. Santana, Jaron T. Krogel, Jeongnim Kim, Paul R. C. Kent, and Fernando A. Reboredo. Structural stability and defect energetics of ZnO from diffusion quantum Monte Carlo. *J. Chem. Phys.*, 142:164705, 2015.

[131] Kateryna Foyevtsova, Jaron T. Krogel, Jeongnim Kim, P. R. C. Kent, Elbio Dagotto, and Fernando A. Reboredo. *Ab initio* quantum Monte Carlo calculations of spin superexchange in cuprates: The benchmarking case of Ca_2CuO_3 . *Phys. Rev. X*, 4:031003, 2014.

[132] Stefan Grimme. Accurate description of van der waals complexes by density functional theory including empirical corrections. *J. Comp. Chem.*, 25:1463–1473, 2004.

[133] N. D. Drummond and R. J. Needs. Quantum Monte Carlo, density functional theory, and pair potential studies of solid neon. *Phys. Rev. B*, 73:024107, 2006.

[134] Henri Hay, Guillaume Ferlat, Michele Casula, Ari Paavo Seitsonen, and Francesco Mauri. Dispersion effects in SiO_2 polymorphs: an *ab initio* study. *Phys. Rev. B*, 92:144111, 2015.

[135] Brian Busemeyer, Mario Dagrada, Sandro Sorella, Michele Casula, and Lucas K. Wagner. Competing collinear magnetic structures in superconducting FeSe by first-principles quantum Monte Carlo calculations. *Phys. Rev. B*, 94:035108, 2016.

[136] Sandro Sorella, Michele Casula, Leonardo Spanu, and Andrea Dal Corso. *Ab initio* calculations for the β-tin diamond transition in silicon: Comparing theories with experiments. *Phys. Rev. B*, 83:075119, 2011.

[137] C. R. Hsing, C. M. Wei, and M. Y. Chou. Quantum Monte Carlo investigations of adsorption energetics on graphene. *J. Phys.: Cond. Mat.*, 24:395002, 2012.

[138] Ozan Karalti, Dario Alfe, Michael J. Gillan, and Kenneth D. Jordan. Adsorption of a water molecule on the MgO(100) surface as described by cluster and slab models. *Phys. Chem. Chem. Phys.*, 14:7846–7853, 2012.

[139] Claudio Attaccalite and Sandro Sorella. Stable liquid hydrogen at high pressure by a novel *Ab Initio* molecular-dynamics calculation. *Phys. Rev. Lett.*, 100:114501, 2008.

[140] Claudio Attaccalite. *RVB phase of hydrogen at high pressure: towards the first* ab initio *Molecular Dynamics by Quantum Monte Carlo.* PhD thesis, International School for Advanced Studies, Trieste (Italy), 2005.

[141] Gregory H. Wannier. Dynamics of band electrons in electric and magnetic fields. *Rev. Mod. Phys.*, 34:645–655, 1962.

[142] Hendrik J. Monkhorst and James D. Pack. Special points for Brillouin-zone integrations. *Phys. Rev. B*, 13:5188–5192, 1976.

[143] G. Rajagopal, R. J. Needs, A. James, S. D. Kenny, and W. M. C. Foulkes. Variational and diffusion quantum Monte Carlo calculations at nonzero wave vectors: Theory and application to diamond-structure germanium. *Phys. Rev. B*, 51:10591–10600, 1995.

[144] G. Rajagopal, R. J. Needs, S. Kenny, W. M. C. Foulkes, and A. James. Quantum Monte Carlo calculations for solids using special k points methods. *Phys. Rev. Lett.*, 73:1959–1962, 1994.

[145] G. Ortiz, D. M. Ceperley, and R. M. Martin. New stochastic method for systems with broken time-reversal symmetry: 2D fermions in a magnetic field. *Phys. Rev. Lett.*, 71:2777–2780, 1993.

[146] F. Bolton. Fixed-phase quantum Monte Carlo method applied to interacting electrons in a quantum dot. *Phys. Rev. B*, 54:4780–4793, 1996.

[147] Matthew D. Jones, Gerardo Ortiz, and David M. Ceperley. Released-phase quantum Monte Carlo method. *Phys. Rev. E*, 55:6202–6210, 1997.

[148] Steffen Bücheler, Dirk Engel, Jörg Main, and Günter Wunner. Quantum Monte Carlo studies of the ground states of heavy atoms in neutron-star magnetic fields. *Phys. Rev. A*, 76:032501, 2007.

[149] M. Casula and S. Sorella. Improper s-wave symmetry of the electronic pairing in iron-based superconductors by first-principles calculations. *Phys. Rev. B*, 88:155125, 2013.

[150] Jun Yang, Weifeng Hu, Denis Usvyat, Devin Matthews, Martin Schütz, and Garnet Kin-Lic Chan. Ab initio determination of the crystalline benzene lattice energy to sub-kilojoule/mole accuracy. *Science*, 345:640–643, 2014.

[151] N. D. Drummond, R. J. Needs, A. Sorouri, and W. M. C. Foulkes. Finite-size errors in continuum quantum Monte Carlo calculations. *Phys. Rev. B*, 78:125106, 2008.

[152] M. Holzmann, R. C. Clay, M. A. Morales, N. M. Tubman, D. M. Ceperley, and C. Pierleoni. Theory of finite size effects for electronic quantum Monte Carlo calculations of liquids and solids. *arXiv:1603.03957*[physics.comp-ph], 2016.

[153] Simone Chiesa, David M. Ceperley, Richard M. Martin, and Markus Holzmann. Finite-size error in many-body simulations with long-range interactions. *Phys. Rev. Lett.*, 97:076404, 2006.

[154] Hendra Kwee, Shiwei Zhang, and Henry Krakauer. Finite-size correction in many-body electronic structure calculations. *Phys. Rev. Lett.*, 100:126404, 2008.

[155] Louisa M. Fraser, W. M. C. Foulkes, G. Rajagopal, R. J. Needs, S. D. Kenny, and A. J. Williamson. Finite-size effects and Coulomb interactions in quantum Monte Carlo calculations for homogeneous systems with periodic boundary conditions. *Phys. Rev. B*, 53:1814–1832, 1996.

[156] P. R. C. Kent, Randolph Q. Hood, A. J. Williamson, R. J. Needs, W. M. C. Foulkes, and G. Rajagopal. Finite-size errors in quantum many-body simulations of extended systems. *Phys. Rev. B*, 59:1917–1929, 1999.

[157] P. P. Ewald. Die berechnung optischer und elektrostatischer gitterpotentiale. *Ann. Phys. (Berlin)*, 369:253–287, 1921.

[158] Yongkyung Kwon, D. M. Ceperley, and Richard M. Martin. Effects of backflow correlation in the three-dimensional electron gas: Quantum Monte Carlo study. *Phys. Rev. B*, 58:6800–6806, 1998.

[159] Claudius Gros. The boundary condition integration technique: results for the Hubbard model in 1D and 2D. *Z. Phys. B: Condensed Matter*, 86:359–365, 1992.

[160] J. Tinka Gammel, D.K. Campbell, and E.Y. Loh Jr. Extracting infinite system properties from finite size clusters: "phase randomization/boundary condition averaging". *Synth. Met.*, 57:4437 – 4442, 1993.

[161] Claudius Gros. Control of the finite-size corrections in exact diagonalization studies. *Phys. Rev. B*, 53:6865–6868, 1996.

[162] C. Lin, F. H. Zong, and D. M. Ceperley. Twist-averaged boundary conditions in continuum quantum Monte Carlo algorithms. *Phys. Rev. E*, 64:016702, 2001.

[163] Sam Azadi and W. M. C. Foulkes. Systematic study of finite-size effects in quantum Monte Carlo calculations of real metallic systems. *J. Chem. Phys*, 143, 2015.

[164] Markus Holzmann, Bernard Bernu, and David M Ceperley. Finite-size analysis of the Fermi liquid properties of the homogeneous electron gas. *J. Phys.: Conf. Ser.*, 321:012020, 2011.

[165] A. Baldereschi. Mean-value point in the Brillouin zone. *Phys. Rev. B*, 7:5212–5215, 1973.

[166] Guglielmo Mazzola, Seiji Yunoki, and Sandro Sorella. Unexpectedly high pressure for molecular dissociation in liquid hydrogen by electronic simulation. *Nat. Commun.*, 5, 2014.

[167] Jeffrey M. McMahon, Miguel A. Morales, Carlo Pierleoni, and David M. Ceperley. The properties of hydrogen and helium under extreme conditions. *Rev. Mod. Phys.*, 84:1607–1653, 2012.

[168] N. D. Drummond, Bartomeu Monserrat, Jonathan H. Lloyd-Williams, P. Lopez Rios, Chris J. Pickard, and R. J. Needs. Quantum Monte Carlo study of the phase diagram of solid molecular hydrogen at extreme pressures. *Nat. Commun.*, 6, 2015.

[169] Sam Azadi, Bartomeu Monserrat, W. M. C. Foulkes, and R. J. Needs. Dissociation of high-pressure solid molecular hydrogen: A quantum Monte Carlo and anharmonic vibrational study. *Phys. Rev. Lett.*, 112:165501, 2014.

[170] J. R. Chelikowsky and S. G. Louie. *Quantum theory of real materials*. Springer, 1996.

[171] Arash A. Mostofi, Jonathan R. Yates, Young-Su Lee, Ivo Souza, David Vanderbilt, and Nicola Marzari. wannier90: A tool for obtaining maximally-localised Wannier functions. *Comp. Phys. Comm.*, 178:685 – 699, 2008.

[172] Shanti Deemyad and James S. Schilling. Superconducting phase diagram of Li metal in nearly hydrostatic pressures up to 67 GPa. *Phys. Rev. Lett.*, 91:167001, 2003.

[173] Anne Marie Schaeffer, Scott R. Temple, Jasmine K. Bishop, and Shanti Deemyad. High-pressure superconducting phase diagram of ^{6}Li: Isotope effects in dense lithium. *Proc. Natl. Acad. Sci. U.S.A.*, 112:60–64, 2015.

[174] D. Nissenbaum, L. Spanu, C. Attaccalite, B. Barbiellini, and A. Bansil. Resonating-valence-bond ground state of lithium nanoclusters. *Phys. Rev. B*, 79:035416, 2009.

[175] K. M. Rasch and L. Mitas. Fixed-node diffusion Monte Carlo method for lithium systems. *Phys. Rev. B*, 92:045122, 2015.

[176] A. Mujica, Angel Rubio, A. Muñoz, and R. J. Needs. High-pressure phases of group-IV, III–V, and II–VI compounds . *Rev. Mod. Phys.*, 75:863–912, 2003.

[177] D. Alfè, M. J. Gillan, M. D. Towler, and R. J. Needs. Diamond and β-tin structures of Si studied with quantum Monte Carlo calculations. *Phys. Rev. B*, 70:214102, 2004.

[178] Wirawan Purwanto, Henry Krakauer, and Shiwei Zhang. Pressure-induced diamond to β-tin transition in bulk silicon: A quantum Monte Carlo study. *Phys. Rev. B*, 80:214116, 2009.

[179] Ryo Maezono, N. D. Drummond, A. Ma, and R. J. Needs. Diamond to β-tin phase transition in Si within diffusion quantum Monte Carlo. *Phys. Rev. B*, 82:184108, 2010.

[180] R. G. Hennig, A. Wadehra, K. P. Driver, W. D. Parker, C. J. Umrigar, and J. W. Wilkins. Phase transformation in Si from semiconducting diamond to metallic β-Sn phase in QMC and DFT under hydrostatic and anisotropic stress. *Phys. Rev. B*, 82:014101, 2010.

[181] S.N. Vaidya, I.C. Getting, and G.C. Kennedy. The compression of the alkali metals to 45 Kbar. *J. Phys. Chem. Solids*, 32:2545 – 2556, 1971.

[182] Wei-Guo Yin and Wei Ku. Flavor-twisted boundary condition for simulations of quantum many-body systems. *Phys. Rev. B*, 80:180402, 2009.

[183] Paulo F. Bedaque. Aharonov–Bohm effect and nucleon–nucleon phase shifts on the lattice. *Physics Letters B*, 593:82 – 88, 2004.

[184] C.T. Sachrajda and G. Villadoro. Twisted boundary conditions in lattice simulations. *Phys. Lett. B*, 609:73 – 85, 2005.

[185] Yoichi Kamihara, Hidenori Hiramatsu, Masahiro Hirano, Ryuto Kawamura, Hiroshi Yanagi, Toshio Kamiya, and Hideo Hosono. Iron-based layered superconductor: LaOFeP. *J. Am. Chem. Soc.*, 128:10012–10013, 2006.

[186] W. L. Yang, A. P. Sorini, C-C. Chen, B. Moritz, W.-S. Lee, F. Vernay, P. Olalde-Velasco, J. D. Denlinger, B. Delley, J.-H. Chu, J. G. Analytis, I. R. Fisher, Z. A. Ren, J. Yang, W. Lu, Z. X. Zhao, J. van den Brink, Z. Hussain, Z.-X. Shen, and T. P. Devereaux. Evidence for weak electronic correlations in iron pnictides. *Phys. Rev. B*, 80:014508, 2009.

[187] Luca de' Medici, Gianluca Giovannetti, and Massimo Capone. Selective Mott physics as a key to iron superconductors. *Phys. Rev. Lett.*, 112:177001, 2014.

[188] Pengcheng Dai. Antiferromagnetic order and spin dynamics in iron-based superconductors. *Rev. Mod. Phys.*, 87:855–896, 2015.

[189] Shen Yao Pan Bingying Hao Yiqing Ma Mingwei Zhou Fang Steffens P. Schmalzl K. Forrest T. R. Abdel-Hafiez M. Chen Xiaojia Chareev D. A. Vasiliev A. N. Bourges P. Sidis Y. Cao Huibo Wang, Qisi and Jun Zhao. Strong interplay between stripe spin fluctuations, nematicity and superconductivity in FeSe. *Nat. Mater.*, 15:159–163, 2015.

[190] D. J. Scalapino. A common thread: the pairing interaction for unconventional superconductors. *Rev. Mod. Phys.*, 84:1383–1417, 2012.

[191] Markus Aichhorn, Silke Biermann, Takashi Miyake, Antoine Georges, and Masatoshi Imada. Theoretical evidence for strong correlations and incoherent metallic state in FeSe. *Phys. Rev. B*, 82:064504, 2010.

[192] H. Okabe, N. Takeshita, K. Horigane, T. Muranaka, and J. Akimitsu. Pressure-induced high-T_c superconducting phase in FeSe: Correlation between anion height and T_c. *Phys. Rev. B (R)*, 81:205119, 2010.

[193] T. Mousavi, C. R. M. Grovenor, and S. C. Speller. Structural parameters affecting superconductivity in iron chalcogenides: a review. *Materials Science and Technology*, 30:1929–1943, 2014.

[194] M. Lüders, M. A. L. Marques, N. N. Lathiotakis, A. Floris, G. Profeta, L. Fast, A. Continenza, S. Massidda, and E. K. U. Gross. *Ab initio* theory of superconductivity. I. Density functional formalism and approximate functionals. *Phys. Rev. B*, 72:024545, 2005.

[195] P. Cudazzo, G. Profeta, A. Sanna, A. Floris, A. Continenza, S. Massidda, and E. K. U. Gross. *Ab Initio* description of high-temperature superconductivity in dense molecular hydrogen. *Phys. Rev. Lett.*, 100:257001, 2008.

[196] Johannes Lischner, Timur Bazhirov, Allan H. MacDonald, Marvin L. Cohen, and Steven G. Louie. First-principles theory of electron-spin fluctuation coupling and superconducting instabilities in iron selenide. *Phys. Rev. B*, 91:020502, 2015.

[197] Markus Aichhorn, Leonid Pourovskii, Veronica Vildosola, Michel Ferrero, Olivier Parcollet, Takashi Miyake, Antoine Georges, and Silke Biermann. Dynamical mean-field theory within an augmented plane-wave framework: Assessing electronic correlations in the iron pnictide lafeaso. *Phys. Rev. B*, 80:085101, Aug 2009.

[198] Lucas K. Wagner and Peter Abbamonte. Effect of electron correlation on the electronic structure and spin-lattice coupling of high-T_c cuprates: quantum Monte Carlo calculations. *Phys. Rev. B*, 90:125129, 2014.

[199] S. Qimiao, Y. Rong, and A. Elihu. High-temperature superconductivity in iron pnictides and chalcogenides. *Nat. Rev. Mater.*, 1:16017, 2016.

[200] S. Medvedev, T. M. McQueen, I. A. Troyan, T. Palasyuk, M. I. Eremets, R. J. Cava, S. Naghavi, F. Casper, V. Ksenofontov, G. Wortmann, and C. Felser. Electronic and magnetic phase diagram of bold β-$Fe_{1.01}Se$ with superconductivity at 36.7 K under pressure. *Nature Materials*, 8:630–633, 2009.

[201] R. M. Fernandes, A. V. Chubukov, and J. Schmalian. What drives nematic order in iron-based superconductors? *Nat. Phys.*, 10:97–104, 2014.

[202] Defa Liu, Wenhao Zhang, Daixiang Mou, Junfeng He, Yun-Bo Ou, Qing-Yan Wang, Zhi Li, Lin Wang, Lili adn Zhao, Shaolong He, Yingying Peng, Xu Liu, Chaoyu Chen, Li Yu, Guodong Liu, Xiaoli Dong, Jun Zhang, Chuangtian Chen, Zuyan Xu, Jiangping Hu, Xi Chen, Xucun Ma, Qikun Xue, and X.J. Zhou. Electronic origin of high-temperature superconductivity in single-layer FeSe superconductor. *Nat. Commun.*, 3:931, 2012.

[203] Jian-Feng Ge, Zhi-Long Liu, Canhua Liu, Chun-Lei Gao, Dong Qian, Qi-Kun Xue, Ying Liu, and Jin-Feng Jia. Superconductivity above 100 K in single-layer FeSe films on doped $SrTiO_3$. *Nat. Mater.*, 14:285–289, 2015.

[204] Ito Y. Mitsuhashi T. Nojima T. Shiogai, J. and A. Tsukazaki. Electric-field-induced superconductivity in electrochemically etched ultrathin FeSe films on $SrTiO_3$ and MgO. *Nat. Phys.*, 12:42–46, 2016.

[205] K. Kothapalli, A. E. Böhmer, W. T. Jayasekara, B. G. Ueland, P. Das, A. Sapkota, V. Taufour, Y. Xiao, E. E. Alp, S. L. Bud'ko, P. C. Canfield, A. Kreyssig, and A. I. Goldman. Strong cooperative coupling of pressure-induced magnetic order and nematicity in FeSe. *arXiv:1603.04135*[cond-mat.supr-con], 2016.

[206] M. C. Rahn, R. A. Ewings, S. J. Sedlmaier, S. J. Clarke, and A. T. Boothroyd. Strong $(\pi, 0)$ spin fluctuations in $\beta - \mathrm{FeSe}$ observed by neutron spectroscopy. *Phys. Rev. B*, 91:180501, 2015.

[207] Taichi Terashima, Naoki Kikugawa, Shigeru Kasahara, Tatsuya Watashige, Takasada Shibauchi, Yuji Matsuda, Thomas Wolf, Anna E. Böhmer, Frédéric Hardy,

Christoph Meingast, Hilbert v. Löhneysen, and Shinya Uji. Pressure-induced antiferromagnetic transition and phase diagram in FeSe. *J. Phys. Soc. Jpn*, 84:063701, 2015.

[208] M. Bendele, A. Amato, K. Conder, M. Elender, H. Keller, H.-H. Klauss, H. Luetkens, E. Pomjakushina, A. Raselli, and R. Khasanov. Pressure induced static magnetic order in superconducting $FeSe_{1-x}$. *Phys. Rev. Lett.*, 104:087003, 2010.

[209] Dai Pengcheng, Jiangping Hu, and Elbio Dagotto. Magnetism and its microscopic origin in iron-based high-temperature superconductors. *Nat. Phys.*, 8:709–718, 2012.

[210] Alain Audouard, Fabienne Duc, Loïc Drigo, Pierre Toulemonde, Sandra Karlsson, Pierre Strobel, and André Sulpice. Quantum oscillations and upper critical magnetic field of the iron-based superconductor FeSe. *EPL*, 109:27003, 2015.

[211] F. Caglieris, F. Ricci, G. Lamura, A. Martinelli, A. Palenzona, I. Pallecchi, A. Sala, G. Profeta, and M. Putti. Theoretical and experimental investigation of magneto-transport in iron chalcogenides. *Sci. Tech. Adv. Mater.*, 13:054402, 2012.

[212] Lucas K. Wagner, Michal Bajdich, and Lubos Mitas. QWalk: A quantum Monte Carlo program for electronic structure. *J. Comp. Phys.*, 228:3390 – 3404, 2009.

[213] Michele Casula. Beyond the locality approximation in the standard diffusion Monte Carlo method. *Phys. Rev. B*, 74:161102, 2006.

[214] S. Margadonna, Y. Takabayashi, Y. Ohishi, Y. Mizuguchi, Y. Takano, T. Kagayama, T. Nakagawa, M. Takata, and K. Prassides. Pressure evolution of the low-temperature crystal structure and bonding of the superconductor FeSe ($T_c = 37$ K). *Phys. Rev. B*, 80:064506, 2009.

[215] Ravhi S. Kumar, Yi Zhang, Stanislav Sinogeikin, Yuming Xiao, Sathish Kumar, Paul Chow, Andrew L. Cornelius, and Changfeng Chen. Crystal and electronic structure of FeSe at high pressure and low temperature. *J. Phys. Chem. B*, 114:12597–12606, 2010.

[216] J. K. Glasbrenner, I. I. Mazin, Harald O. Jeschke, P. J. Hirschfeld, R. M. Fernandes, and Roser Valentí. Effect of magnetic frustration on nematicity and superconductivity in iron chalcogenides. *Nat. Phys.*, 1:953–958, 2015.

[217] Despina Louca, K. Horigane, A. Llobet, R. Arita, S. Ji, N. Katayama, S. Konbu, K. Nakamura, T.-Y. Koo, P. Tong, and K. Yamada. Local atomic structure of superconducting $FeSe_xTe_{1-x}$. *Phys. Rev. B*, 81:134524, 2010.

[218] Kyuho Lee, Éamonn D. Murray, Lingzhu Kong, Bengt I. Lundqvist, and David C. Langreth. Higher-accuracy van der Waals density functional. *Phys. Rev. B*, 82:081101, 2010.

[219] H. Anton and P. C. Schmidt. Theoretical investigations of the elastic constants in Laves phases. *Intermetallics*, 5:449–465, 1997.

[220] Jasmine N. Millican, Daniel Phelan, Evan L. Thomas, Juscelino B. Leão, and Elisabeth Carpenter. Pressure-induced effects on the structure of the FeSe superconductor. *Solid State Commun.*, 149:707–710, 2009.

[221] F. Essenberger, P. Buczek, A. Ernst, L. Sandratskii, and E. K. U. Gross. Paramagnons in FeSe close to a magnetic quantum phase transition: *ab initio* study. *Phys. Rev. B*, 86:060412, 2012.

[222] Weicheng Lv, Frank Krüger, and Philip Phillips. Orbital ordering and unfrustrated $(\pi, 0)$ magnetism from degenerate double exchange in the iron pnictides. *Phys. Rev. B*, 82:045125, 2010.

[223] Hai-Yuan Cao, Shiyou Chen, Hongjun Xiang, and Xin-Gao Gong. Antiferromagnetic ground state with pair-checkerboard order in FeSe. *Phys. Rev. B*, 91:020504, 2015.

[224] C. Tresca, F. Ricci, and G. Profeta. Strain effects in monolayer iron-chalcogenide superconductors. *2D Mater.*, 2:015001, 2015.

[225] Wei Wu. Modelling the electronic structure and magnetic properties of LiFeAs and FeSe using hybrid-exchange density functional theory. *Solid State Commun.*, 161:23–28, 2013.

[226] T. Shimojima, Y. Suzuki, T. Sonobe, A. Nakamura, M. Sakano, J. Omachi, K. Yoshioka, M. Kuwata-Gonokami, K. Ono, H. Kumigashira, A. E. Böhmer, F. Hardy, T. Wolf, C. Meingast, H. V. Löhneysen, H. Ikeda, and K. Ishizaka. Lifting of xz/yz orbital degeneracy at the structural transition in detwinned FeSe. *Phys. Rev. B*, 90:121111, 2014.

[227] F. Wang, S. A. Kivelson, and D. Lee. Nematicity and quantum paramagnetism in FeSe. *Nat. Phys.*, 11:959–963, 2015.

[228] Luca de' Medici, Jernej Mravlje, and Antoine Georges. Janus-faced influence of Hund's rule coupling in strongly correlated materials. *Phys. Rev. Lett.*, 107:256401, 2011.

[229] Luca de' Medici. Hund's coupling and its key role in tuning multiorbital correlations. *Phys. Rev. B*, 83:205112, 2011.

[230] Antoine Georges, Luca de' Medici, and Jernej Mravlje. Strong correlations from Hund's coupling. 4:137–178, 2013.

[231] T. Tzen Ong and Piers Coleman. Local quantum criticality of an iron-pnictide tetrahedron. *Phys. Rev. Lett.*, 108:107201, 2012.

[232] Z. P. Yin, K. Haule, and G. Kotliar. Magnetism and charge dynamics in iron pnictides. *Nat. Phys.*, 7:294–297, 2011.

[233] Z. P. Yin, K. Haule, and G. Kotliar. Kinetic frustration and the nature of the magnetic and paramagnetic states in iron pnictides and iron chalcogenides. *Nat. Mater.*, 10:932–935, 2011.

[234] Werner Kutzelnigg. Direct determination of natural orbitals and natural expansion coefficients of manyelectron wavefunctions. I. natural orbitals in the geminal product approximation. *J. Chem. Phys.*, 40, 1964.

Appendix A

Computation of electron occupations within QMC

This Appendix reports a detailed discussion on the procedure to compute orbital occupations within a quantum Monte Carlo simulation. The single-particle occupations are given by the diagonal elements of the one-body density matrix. If we consider a N particles many-body wavefunction $\Psi(\mathbf{r}_1, \cdots, \mathbf{r}_N)$ expandend over a basis set $\{\chi_i(\mathbf{r})\}$, the one-body density matrix can be conventionally written as:

$$\rho(\mathbf{r}, \mathbf{r}') = \sum_{i,j} \rho_{ij} \chi_i(\mathbf{r}) \chi_j(\mathbf{r}') \tag{A.1}$$

where the sum goes over the whole basis set dimension.

In the context of geminal wavefunction formalism within QMC introduced in Chaps. 2 and 4, the one-body density operator (1DM) has the following definition:

$$\hat{\rho} = \sum_{\alpha,\beta}^{N_{\text{at}}} \sum_{i,j}^{N_{\text{bas}}} |\chi_i^\alpha\rangle \, \lambda_{ij}^{\alpha\beta} \left\langle \chi_j^\beta \right| \tag{A.2}$$

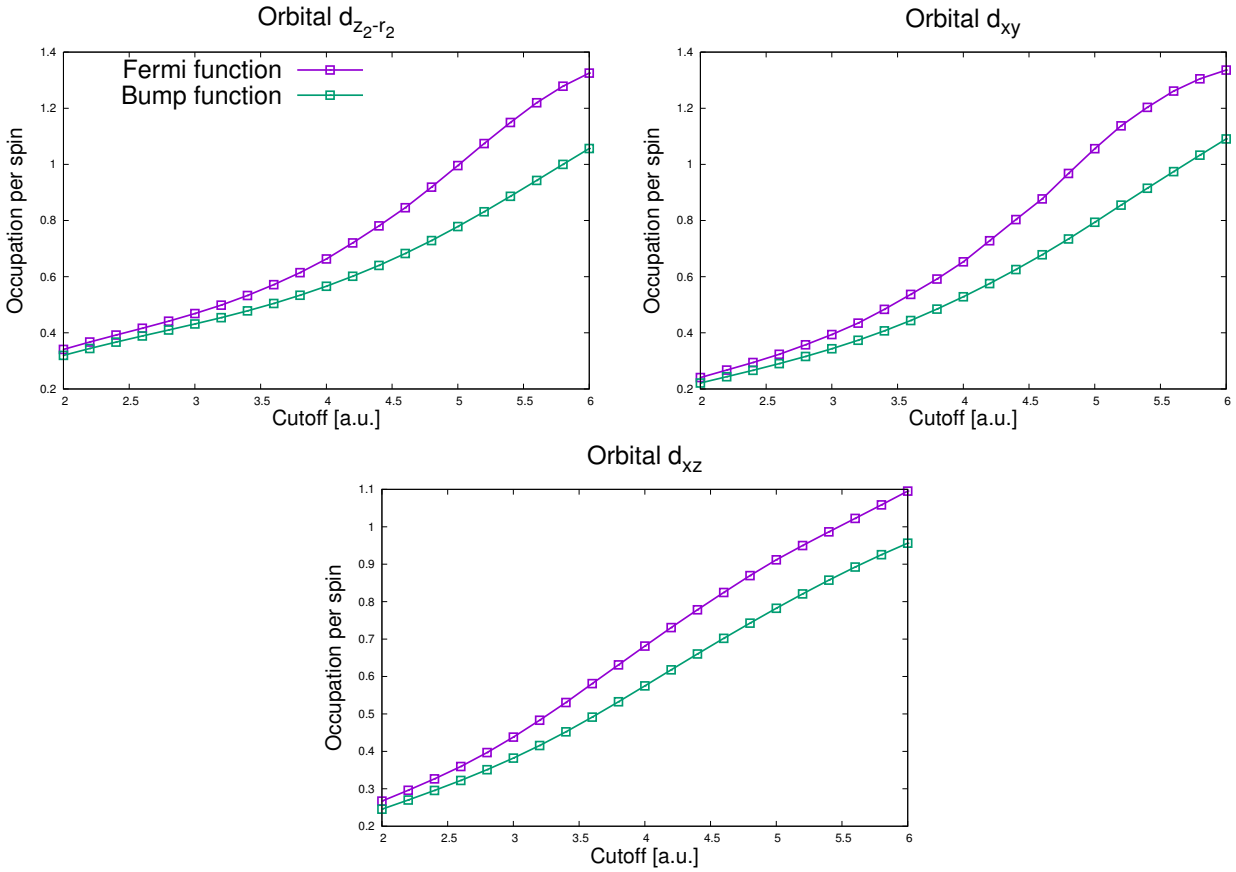
which is analogous to Eq. A.1. We used the same notation as the main text. The total normalization of 1DM is defined such that: $\text{Tr}(\hat{\rho}) = \text{Tr}(S\lambda) = N$ where S is the basis overlap matrix.

In order to filter out the occupations of individual atoms and/or specific angular momentum channels, we apply a generalized atomic projector $\mathcal{P}$ to Eq. A.1. We consider a generic non-orthogonal basis set $\{\tilde{\chi}_i^\alpha\}$ which can be defined for both open and extended systems. The projector operator is defined as:

$$\mathcal{P}_{\text{at}} = \sum_{\alpha,\beta} \sum_{i,j} |\tilde{\chi}_i^\alpha\rangle \, \tilde{S}_{ij}^{-1} \left\langle \tilde{\chi}_j^\beta \right| \tag{A.3}$$

where the indices (i, j) goes over the whole projection basis set dimension (which can be different with respect to the wavefunction basis set) and $\tilde{S}$ is the overlap matrix. The

Figure A.1 – Electron occupations of Fe atoms computed with mixed estimators for three sample orbitals of the antiferromagnetic (checkerboard) phase of FeSe. All curves show an inflexion point which determines the border between atomic and interstitial regiones, but the change in concavity is more pronounced in the case of Fermi orbitals.



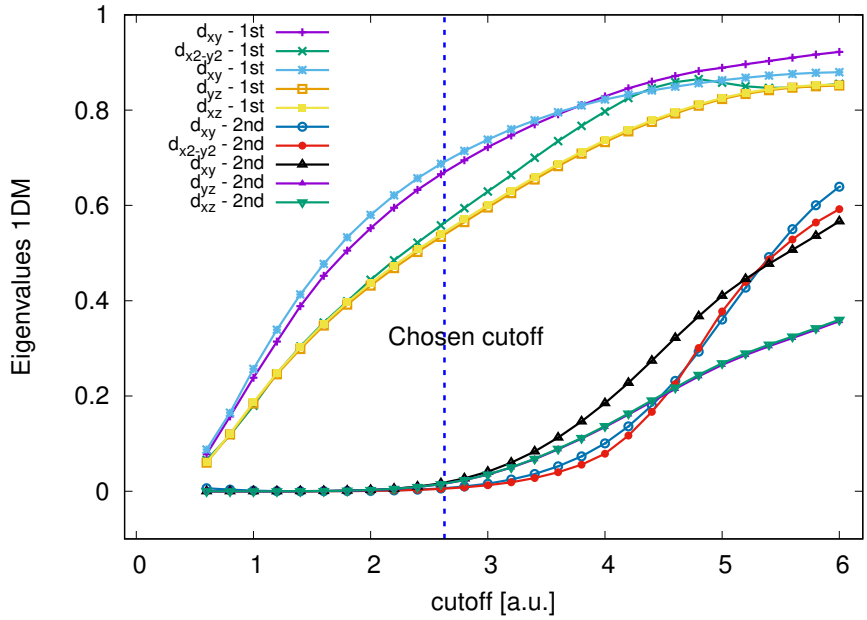
choice of the basis set for the projection is arbitrary. In our implementation we tested two possible functional forms, both based on a modification of the localized atomic orbitals employed for expanding the QMC wavefunction. For the first choice, which we denote as as "bump" orbitals, we define these orbitals as a stepwise function which reads:

$$\chi^P(r) = \begin{cases} \chi^a_\mu(r) \exp\left[-\ln\left(1-\frac{\zeta r^2}{\gamma}\right)\right]^2 & \text{if} \quad \zeta r^2 > \gamma \\ \\ 0 \quad \text{if} \quad \zeta r^2 \geq \gamma \end{cases} \tag{A.4}$$

where $r = \mathbf{r} - \mathbf{R}$. The continuity of this function at the cutoff condition $\zeta \mathbf{r}^2 = \gamma$ is easily demonstrated. γ is a cutoff parameter which is tuned to isolate the contribution of the considered atom in the simulation supercell from the others.

We tested a second choice based for the projection which smoothly decreases when going away from the cutoff condition. For this purpose we use a Fermi-Dirac function and

Figure A.2 – First and second eigenvalues of the one-body density matrix projected on the d angular momentum channel of a reference Fe atom. The calculations have been performed using the mixed estimator formalism in the $4 \times 4 \times 1$ supercell of paramagnetic FeSe.



we denote the projection basis set as:

$$\chi^P(r) = \chi^a_\mu(r) \frac{1}{\exp\left(\frac{r-\gamma}{\Delta}\right) + 1} \tag{A.5}$$

where Δ is a parameter which determines the steepness of the Fermi-Dirac distribution. We use a value of $\Delta = 0.05$ for our calculations. γ is equivalent to the cutoff parameter defined before and it must be tuned to filter out the occupations of the considered atom. We denote this choice as "Fermi" orbitals.

Once the projection basis set is defined, we can compute electronic occupations by projecting the 1DM in Eq. A.2 in the following way:

$$\mathrm{Tr}\left[M\right] = \mathrm{Tr}\left[\mathcal{P}_{\mathrm{at}}\, \hat{\rho}\, \mathcal{P}_{\mathrm{at}}\right] \tag{A.6}$$

Mixed estimator formalism can be used for obtaining a more reliable results in the case of DMC calculations. Notice that the computation of the matrix M involves the evaluation of the basis set overlap between the normal and projected basis sets. Occupations of a particular atom, angular momentum channel or orbital can be extracted by tracing over the 1DM suitably restricted to the desired sector.

The choice of the cutoff parameter γ is very delicate since it can lead to very different results. In Fig. A.1 we present a study of the occupations as a function of γ for three representative d orbitals of a particular Fe atom in the simple antiferromagnetic checkerboard

configuration of FeSe (see Chap. 6 for more details). We notice that the occupation curves as a function of the cutoff present a characteristic point where the concavity changes. We can choose the value of this inflexion point to determine the right value of the cutoff for the considered atomic species, since it is a good indicator of the border between atomic and interstitial regions. It is evident in Fig. A.1 that the inflexion points are more pronounced in the case of the Fermi orbitals. This facilitates the definition of the cutoff parameter and therefore we choose them for the final implementation.

An alternative procedure to determine the cutoff is to consider the eigenvalues of 1DM projected onto an angular momentum channel for a reference Fe atom. 1st and 2nd eigenvalues (in descending order) of 1DM projected onto d-orbital in paramagnetic FeSe are plotted in Fig. A.2. We choose the value of the cutoff radius such as the magnitude of the 2nd eigenvalue is below a certain thresold – which we set at 10^{-3} – for each d-orbital. In this way, we avoid to include any virtual state when computing the occupations. The chosen cutoff is indicated by the straight vertical line in Fig. A.2.

The two aforementioned procedures yield a very similar value of the radial cutoff. However the second one is more robust and it can be easily automatized for the implementation of methods requiring orbital occupations such as the DFT+U framework.

Appendix B

Parallel scheme for Brillouin zone integration

Table B.1 – Basic instructions for splitting the MPI communicators to create an appropriate scheme for BZ integration in DFT. The same scheme is used in QMC calculations when averaging over multiple twist conditions.

```
! number of processes in rows and columns
columns_nproc = nk
rows_nproc    = (world_nproc)/nk
! create the grid and split the
    communicators
irow = (world_rank)/(rows_nproc)
jcol = mod(world_rank,rows_nproc)
call MPI_comm_split(world_comm,irow,jcol,
    rows_comm,ierr)
call MPI_comm_split(world_comm,jcol,irow,
    columns_comm,ierr)
! rows/columns process IDs
call MPI_comm_rank(rows_comm,rows_rank,
    ierr)
call MPI_comm_rank(columns_comm,
    columns_rank,ierr)
! synchronize
call MPI_barrier(world_comm,ierr)
```

In this short Appendix we present a code snippet of the algorithm used for **k**-points

sampling with the DFT code described in the text. For interfacing purposes, the same algorithm is used also for QMC when using multiple twist conditions.

The two subsequent calls to the instruction `MPI_comm_split` allows to split the considered MPI communicator (denoted as `world_comm` in two independent groups which we denote as rows (communicator `rows_comm` and columns (communicator `columns_comm`). Each column accommodates all **k**-points included in the calculation, whereas each rows is composed by a pool of processors devoted to a single boundary condition.

In this way, in DFT calculations the pool of processors of each row deals with the parallelization over the real space grid for each boundary independently. When an average over **k**-points is required, it is sufficient to perform collective operations within the columns communicator. Within QMC, we can straightforwardly apply the same scheme by assigning to each processor in the row pool a single walker instead of a slice of the integration mesh as in DFT.

It is worth remarking that this scheme offers a very simple way for interfacing DFT and QMC codes, but it has a sort of drawback as the number of processors must be multiple of the number of **k**-points/twist conditions. However, within modern supercomputers equipped with a large amount of CPUs this does not constitute a problem.

Appendix C

Automatic algorithmic differentiation with complex variables

In this Appendix we present how to apply the automatic algorithmic differentiation (AAD, see Sec. 2.5.3 in the main text for more details) of two algebraic operations for complex variables ubiquitous in the QMC code developed during this thesis: the dot product between two vectors and the matrix-matrix multiplication. Let us consider the dot product between two generic complex vectors $\mathbf{a} = \mathbf{a}^R + \mathbf{a}^I$ and $\mathbf{b} = \mathbf{b}^R + \mathbf{b}^I$ of any dimension. The product reads:

$$c^R + ic^I = \sum_j (a_j^R + ia_j^I)(b_j^R + ib_j^I) = \sum_j (a_j^R b_j^R - a_j^I b_j^I) + i(a_j^R b_j^I + a_j^I b_j^R) \tag{C.1}$$

where i is the imaginary unit and the superscripts R/I identify the real and imaginary parts of the vector components. Let us focus for sake of simplicity on a single component. The adjoint of real and imaginary parts of the vector components are defined as an arbitrary linear combination of the derivatives of these quantities with respect to the outputs c^R/c^I of the direct algorithm. By setting all coefficients of the linear expansion of the Jacobian to 1, the adjoints read:

$$\bar{a}_j^R = \bar{a}_j^R + \frac{\partial c^R}{\partial a_j^R}\bar{c}^R + \frac{\partial c^I}{\partial a_j^R}\bar{c}^I = \bar{a}_j^R + b_j^R\bar{c}^R + b_j^I\bar{c}^I \tag{C.2}$$

$$\bar{a}_j^I = \bar{a}_j^I + \frac{\partial c^R}{\partial a_j^I}\bar{c}^R + \frac{\partial c^I}{\partial a_j^I}\bar{c}^I = \bar{a}_j^R - b_j^I\bar{c}^R + b_j^R\bar{c}^I \tag{C.3}$$

$$\bar{b}_j^R = \bar{b}_j^R + \frac{\partial c^R}{\partial b_j^R}\bar{c}^R + \frac{\partial c^I}{\partial b_j^R}\bar{c}^I = \bar{a}_j^R + a_j^R\bar{c}^R + a_j^I\bar{c}^I \tag{C.4}$$

$$\bar{b}_j^I = \bar{b}_j^I + \frac{\partial c^R}{\partial b_j^I}\bar{c}^R + \frac{\partial c^I}{\partial b_j^I}\bar{c}^I = \bar{a}_j^R - a_j^I\bar{c}^R + a_j^R\bar{c}^I \tag{C.5}$$

The final version of the adjoint dot product can be written in a compact form using vector notation as follows:

$$\bar{\mathbf{a}} = \bar{\mathbf{a}} + \mathbf{b}^* \bar{\mathbf{c}} \tag{C.6}$$

$$\bar{\mathbf{b}} = \bar{\mathbf{b}} + \mathbf{a}^* \bar{\mathbf{c}} \tag{C.7}$$

The above expression works also in the case that one of the vectors, for instance $\mathbf{a}$, is real by imposing that also the corresponding adjoint $\bar{\mathbf{a}}$ is real simply by disregarding its fictious imaginary part.

We turn now our attention to the matrix-matrix multiplication, which is certainly the most important and widely used algebraic operation in the *TurboRVB* code due to its very efficient, machine optimized implementation in standard linear algebra packages such as LAPACK [28]. We consider two generic complex matrices A and B and a standard case of matrix-matrix multiplication present in the LAPACK library which reads:

$$C = \beta C + \alpha AB = (\beta_R + i\beta_I)(C_R + iC_I) + (\alpha_R + i\alpha_I)(A_R + iA_I)(B_R + iB_I) \tag{C.8}$$

where α and β are generic complex numbers and the subscripts R/I denote as usual real and imaginary parts. By following a recipe analogous to the one used for the dot product and using the linearity property of the adjoints, we find the following set of adjoints of real and imaginary parts of the input matrices:

$$\bar{A}_R = \bar{A}_R + \alpha_R \bar{C}_R B_R^T + \alpha_I \bar{C}_I B_I^T \tag{C.9}$$

$$\bar{A}_I = \bar{A}_I - \alpha_I \bar{C}_R B_I^T + \alpha_R \alpha_R \bar{C}_I B_R^T \tag{C.10}$$

$$\bar{B}_R = \bar{B}_R + \alpha_R \bar{C}_R A_R^T + \alpha_I \bar{C}_I A_I^T \tag{C.11}$$

$$\bar{B}_I = \bar{B}_R - \alpha_I \bar{C}_R A_I^T + \alpha_R \bar{C}_I A_R^T \tag{C.12}$$

$$\bar{C}_R = \beta_R \bar{C}_R \tag{C.13}$$

$$\bar{C}_I = -\beta_I \bar{C}_I \tag{C.14}$$

Notice that the adjoint of the output of the direct algorithm C must not be updated. In compact matrix notation, we can write the final adjont for the operation in Eq. C.8 as:

$$\bar{A} = \bar{A} + \alpha^* \bar{C} B^\dagger \tag{C.15}$$

$$\bar{B} = \bar{B} + \alpha^* \bar{C} A^\dagger \tag{C.16}$$

$$\bar{C} = \beta^* \bar{C} \tag{C.17}$$

It is straightforward to generalize the above set of equations to the other cases included in the LAPACK libraries.