



THEORY AND NUMERICAL SIMULATION OF CONDENSED MATTER

---

CLAUDIO GENOVESE

GEMINAL POWER IN QMC

---

PHD THESIS

---

SUPERVISOR:  
PROF. SANDRO SORELLA

---

ACADEMIC YEAR 2019/2020

*to Sara,  
because this thesis, like everything else,  
would not have value without her.*

# Contents

<b>Introduction</b>	<b>3</b>
<b>1 Introduction to Quantum Monte Carlo</b>	<b>13</b>
1.1 Variational Monte Carlo . . . . .	14
1.2 Optimization . . . . .	16
1.2.1 Derivatives Calculation . . . . .	17
1.2.2 Optimization Strategy . . . . .	20
1.3 Diffusion Monte Carlo . . . . .	23
<b>2 Wave Functions</b>	<b>31</b>
2.1 Basis Set . . . . .	35
2.2 Fermionic Wave Functions . . . . .	39
2.2.1 Slater Determinant . . . . .	39
2.2.2 Pairing Function . . . . .	40
2.2.3 AGP . . . . .	42
2.2.4 Pfaffian Wave Function . . . . .	44
2.2.5 AGP and Pf with a constrained number of molecular orbital (AGPn and Pfn) . . . . .	50
2.2.6 Conversion of the Wave Functions . . . . .	55
2.2.7 Diagonalization of a antisymmetric generally complex matrix $\lambda$ . . . . .	57
2.3 Jastrow Factor . . . . .	63

2.3.1	The $S^2$ operator . . . . .	66
<b>3</b>	<b>Hydrogen Chain</b>	<b>72</b>
3.1	Insulating phase . . . . .	73
3.1.1	Antiferromagnetism. . . . .	73
3.1.2	Dimerization . . . . .	75
3.2	Metallic phase . . . . .	78
3.2.1	Insulator-to-metal transition . . . . .	78
3.2.2	Origin of the MIT and properties of the metallic phase . . . . .	79
<b>4</b>	<b>The <math>H_4</math> model system</b>	<b>82</b>
4.1	Wave functions comparison: the failure of the Slater Determinant	84
4.2	The role of the Jastrow Factor . . . . .	88
<b>5</b>	<b>Atoms, Dimers and Molecules</b>	<b>91</b>
5.1	Magnetic Dimers . . . . .	94
5.1.1	Carbon . . . . .	94
5.1.2	Nitrogen . . . . .	102
5.1.3	Oxygen . . . . .	105
5.2	Benzene . . . . .	108
5.2.1	The benzene dimer . . . . .	110
	<b>Conclusions</b>	<b>113</b>
	<b>Acknowledgements</b>	<b>116</b>
	<b>Bibliography</b>	<b>118</b>

# Introduction

The Schrödinger equation[1] together with its relativistic counterpart, the Dirac equation[2], can probably be considered the biggest deceit in the history of knowledge. In our hands we hold the keyhole for the full understanding of the microscopic world but there is no key available. We can exactly describe the problem that we want to solve and obtain its constraint to reality but we cannot solve it, we cannot access this knowledge. It is for this reason that every PhD thesis dealing with quantum chemistry or condensed matter needs to have in its introduction (if not even before!) this sentence from Paul A. M. Dirac:

*"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. It therefore becomes desirable that approximate practical methods of applying quantum mechanics should be developed, which can lead to an explanation of the main features of complex atomic systems without too much computation."*[3]

This sentence is indeed the key for the understanding of a large part of the modern research in quantum mechanics and, consequently, the key to understand the purpose of this thesis. The problem is well known and eminent scientists have written papers and books on this topic, so I will try to cut the

story short on what I can assume is common background knowledge.

From the perspective of atomistic simulations, when we look at the Schrödinger equation for atomic systems

$$i\hbar\frac{\partial\Psi}{\partial t} = \left[\hat{T} + \hat{V}_{ii} + \hat{V}_{ei} + \hat{V}_{ee}\right] \Psi, \quad (1)$$

where  $\hat{T}$  is the kinetic energy operator,  $\hat{V}_{ii}$  is the ion-ion potential,  $\hat{V}_{ei}$  the electron-ion potential and  $\hat{V}_{ee}$  the electron-electron potential. We can recognize in this last term the origin of the main challenges in this research field. It introduces the correlation between the electrons and makes impossible to find the exact solution for almost every interesting system. Indeed, if the electron-electron potential is absent electrons do not "see" each other and they can be treated as independent particles, with the only requirement of satisfying the Pauli exclusion principle. A many-electron problem is thus reduced to the antisymmetrization of the single electron solution of the Schrödinger equation and for  $N$  electrons the Hilbert space associated with this problem is basically equivalent to  $N$  times the Hilbert space of a single particle problem. When we introduce the electron-electron potential, the electrons are influenced by the surrounding ones and we cannot decompose the problem in  $N$  single-particle ones. The Hilbert space dimension grows exponentially with the number of electrons and the exact ground state may become impossible to calculate. In principle, a solution is always possible, but it has a cost that grows exponentially with the size of the system, and so for interesting problems it becomes immediately incompatible with the time necessary for a Phd, a lifetime or the universe existence. Thus, it is important to find approximations and techniques to obtain an approximate solution for the Schrödinger equation, yielding the required degree of accuracy.

Many problems can indeed be addressed without even going to the quantum world or even without maintaining an atomic resolution, while others require to push the available approximations and calculation facilities to their limits.

Thus, it is necessary to find a balance between accuracy and computational cost, targeting the cheapest solution allowing a meaningful description of the system under study. This has stimulated an impressive effort of the scientific community to find, tune, and optimize these techniques in the constant attempt to push always a bit further the limit of what was possible and what wasn't. Many techniques have been developed ranging from the Molecular Dynamics to Hartree-Fock, to Density Functional Theory, Density Matrix Renormalization Group, Coupled Cluster, Dynamical Mean Field Theory, Quantum Monte Carlo and many other mores, each one valid and affordable in a certain range.

In this thesis I will deal with the Quantum Monte Carlo (QMC) framework: it consists of a series of techniques meant to address the electronic problems with a high level of accuracy. The name Monte Carlo is a huge collector of methods that are fundamentally based on the concept of stochasticity. They are applied in many different tasks requiring numerical calculation, even in some cases that we perceive as very far from science. Monte Carlo based methods appear in finance, in risk management, in project planning, in quality control, in machine learning and many other fields. Just to give an idea, during these four years, with other three colleagues (Lucas Kohn, Matteo Seclì and Davide Tisi), we successfully applied Monte Carlo to process monitoring and task optimization problems in business, in the context of a collaboration between SISSA and Esteco (a software company based in Trieste). I personally find extremely gratifying the idea of distilling information from random processes. The concept behind QMC is exactly to apply stochastic processes to obtain information on the quantum world. In particular, in this thesis I will present the technical advances and the related "experiments" with two different QMC techniques: the variational Monte Carlo (VMC)[4, 5, 6] and the diffusion Monte Carlo (DMC)[5, 6, 7, 8]. I will describe the technical details of these methods in the next chapter but here let me just point out their

basic working mechanism and why they are, in my opinion, very promising techniques for the near future. Both VMC and DMC are wave function (WF) based methods. This means that their workflow depends on a WF that is the central object of the calculation. The two pillars of VMC and DMC are respectively the variational principle and the projection technique. These two foundations give a very robust and reliable background to the theories and the techniques developed. In particular, in the last years many developments have been made in the variational search thanks also to the increasing interest of the scientific community in machine learning optimization problems. In the machine learning nomenclature it is straightforward to identify VMC as an unsupervised learning problem. The variational search is performed attempting to minimize the total energy of the system that, in this case, plays the role of a cost function. DMC, instead, is a projection algorithm performed statistically using the information on the sign contained in the given WF, dubbed here as guiding function. In this way, we can considerably improve the description of the ground state (GS), projecting onto the lowest possible energy WF with the same signs of the guiding WF. In the ideal case of a guiding function that, for each configuration, has the same sign of the GS, the above described DMC algorithm provides the exact solution[5, 6]. The core DMC and VMC algorithms are performed using operation on dense matrices and for this reason they allow us to exploit the peak performance of the architectures used for the calculation. This feature will allow a relatively easy transition to all the architectures that will be developed in the future that will be, most likely, designed to achieve the best performances on matrix operations. Moreover, VMC and DMC algorithms have a moderate scaling with the number of the electrons in the system,  $O(N^3)$  and  $O(N^4)$  respectively with  $N$  being the number of electrons, leading to a good scaling with the size of the systems and so they are particularly indicated for large systems. For the time being, the most important limits are the QMC prefactors of the computational cost, constrain-

ing the maximum system size to a few hundreds of electrons. However, this problem is going to disappear in the future thanks to the introduction of new technologies for the calculation facilities and software advances. In this way these prefactors will become affordable while leaving intact the good scaling properties of QMC, making it a good candidate for very accurate calculations on large systems in the next future.

The aim of this thesis and of the work I have done in these years is to introduce new technical advances and to try to understand the limits and strengths of the different possibilities available for the calculations to unveil physical and chemical properties of systems that are not accessible with the more standard approaches. Since we do not have control over the performances of the devices used for the calculations, the main contribution that we can give to the growth and development of QMC is to improve the quality of the results that can be obtained at a given cost. Introducing new techniques and protocols for simulations can indeed increase the quality of the results of the calculations and also the control that the scientists can have with a given method. As I already mentioned, VMC and DMC are two WF based methods: the workflow of the computations and the quality of the result are strongly dependent on the WF chosen. Hence, it is critical to find the best balance possible between cost and accuracy. As I will explain later, a successful DMC calculation relies on a reasonably accurate WF and so it is crucial to tackle correctly the variational search of the WF. As in every variational problem, the first and more important decision that we have to face when we approach a new system with VMC is the *ansatz* choice. This choice may look like a bare mathematical problem, but, instead, it is determinant from a physical point of view. For example, an *ansatz* may not be able to describe a particular physical feature fundamental for the system in exam. By choosing the *ansatz* we often make some physical hypotheses and constrain the variational search. Except for some very limited cases, it is impossible to know in advance which will be the best *ansatz* for

the system in exam, but narrowing the choice is possible and fundamental to obtain meaningful results and to save computational resources.

In the QMC framework it is very common to factorize the WF in two different elements: a fermionic mean-field that defines the WF behavior and implements the Pauli exclusion principle [9] and a Jastrow factor (JF) correlator that gives an exponential modulation to the former factor to take into account explicitly the electronic correlation. The simplest choice for the mean-field part is the Slater Determinant (SD) [10]; in two words it is the antisymmetrized product of a set of orthonormal orbitals that take the name of molecular orbitals (MO). In terms of computational cost, the SD is the cheapest fermionic WF. It can be obtained directly from mean field calculations like Hartree-Fock (HF) or Density Functional Theory (DFT) and then further optimized if necessary. It is possible in some cases to obtain very high accuracies already with the SD and it is even more likely when it is combined with a JF (JSD) [11, 12]. Unfortunately, there are many systems where the JSD cannot capture properly the correlation of the system and the description of the chemical properties is very poor [13, 14, 15, 16]. In order to overcome this problem, different approaches have been elaborated and many of them involve the use of more advanced WFs. One of the most popular choices in this sense is the use of multi-determinant WFs: they are calculated as the linear combination of several SDs, each one with a different set of MOs. This approach is in some cases very effective and allows to reach extremely high accuracies [16, 17, 18, 19, 20]. It is systematically improvable and with a large enough number of SDs it is possible in principle to find the exact ground state WF. Unfortunately, this method comes with some drawbacks, indeed, the number of SDs that have to be considered, and thus the computational cost, scales exponentially with the size of the system [21]. Therefore, even if it represents a very powerful technique for small systems, it easily becomes too expensive when increasing the number of electrons. Thus, it is fundamental to find techniques to improve the

accuracy of the SD while remaining with a similar computational cost.

The theory and the results presented in this work move exactly in this direction, by using *ansatze* that have a cost similar to the SD, but able to describe much more accurate solutions. In particular, in this thesis, I will present and analyze the Jastrow correlated Antisymmetrized Geminal Power *ansatze*, by introducing the most general one and trying to understand the effectiveness and the limitations of the different possibilities. The use of the pairing function replaces the single-particle description of the SD with a richer one in terms of electron pairs. The corresponding WF is a natural extension of the SD *ansatz* and represents a direct and efficient implementation of the Anderson's resonating valence bond (RVB) theory of many-electron WFs. In particular, it provides a direct description of the singlet and triplet correlations that are absent in the SD. Even if the pairing functions cannot be improved systematically, these WFs have a much larger variational freedom than the SD with a similar computational cost. In the following chapters I will describe the technical details required for the calculation, but let me sketch here the available possibilities. The pairing function describes a pair of electrons given their spins and positions and has the following explicit form

$$\begin{aligned}
 g(\vec{r}_1\sigma_1, \vec{r}_2\sigma_2) &= \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle)g_+(\vec{r}_1, \vec{r}_2) \\
 &+ \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle)g_-(\vec{r}_1, \vec{r}_2) \\
 &+ |\uparrow\uparrow\rangle g_\uparrow(\vec{r}_1, \vec{r}_2) + |\downarrow\downarrow\rangle g_\downarrow(\vec{r}_1, \vec{r}_2), \tag{2}
 \end{aligned}$$

here some symmetry constraints have to be applied to ensure the proper fermionic behavior, but this will be discussed in the next chapters. Depending on the definition of the pairing function, qualitatively distinct WFs can be obtained. The differences between these *ansatze* depend on the terms kept in Eq.(2) leading to peculiar magnetic properties of the electron pair. There are three important cases: i) when no triplet correlations are allowed, we have

a perfect singlet and we denote it by AGP[22, 23, 24]; ii) when only the parallel spin term of the triplet component are omitted (namely the last line in Eq.(2)), the WF can break the spin symmetry but the magnetic order parameter can be directed only in the z-quantization axis, and in this case we will refer to AGP<sub>u</sub>[25]; iii) the most important case is the most general one that contains all triplet contributions and it is known as Pfaffian WF (Pf)[26, 27, 28]. The Pfaffian WF literally realizes the most general Antisymmetrized Geminal Power and its use together with a new general spin and charge dependent JF is the main technical contribution to the TURBORVB[29] code and to the QMC framework presented in this thesis. However, for the pairing functions, the optimization of a large number of non-linear variational parameters can become a serious limitation if not handled efficiently. Indeed, in order to exploit the full potential of these *ansatze* it has been fundamental to use the most recent techniques for the calculation of the derivatives and optimization strategies. The AGP *ansatz* has already been successfully used in the past and it has represented the key concept at the basis of the TURBORVB package. Also the Pf WF is not a novelty in literature [26, 27], but in the previous attempts it has been used by exploiting only a very small fraction of the large variational freedom of the *ansatz*. The results were not encouraging and the energies obtained with this *ansatz* did not improve the ones of the AGP that, in turn, has a lower computational cost. Despite the Pfaffian was no longer used in the electronic systems, the experience with lattice models has shown that the Pf WF can improve considerably the description of magnetic and correlated systems[30]. Moreover, the introduction of a powerful JF and the recent results obtained in combination with the AGP [31, 13, 32] encouraged us to look for the unexpressed potential of the Pf WF. Finally, I will introduce the Pfaffian WF with a constrained number of MOs (Pfn). This is a variant of the Pf WF that is useful to reduce the number of variational parameters still maintaining many of the properties of the original Pf. Even if it has not

been deeply studied and carefully benchmarked it has been implemented in TURBORVB and represents a promising alternative to explore in the future. Note that, in general, all the above mentioned WFs take the prefix J when they are combined with a JF correlator.

The content of this thesis is based on the research papers published in these years[12, 28, 33, 13, 29] and on the one in preparation[34]. For the sake of completeness and for a possible reader not expert in the technicalities of QMC, in the first chapter I will discuss the VMC and DMC algorithms, the calculation of the derivatives using the Adjoint Algorithmic Differentiation[35] (known in many other communities as Backpropagation) as well as the optimization techniques used. In the second chapter I will describe the WFs mentioned above and the transformations required to convert the WFs from an *ansatz* to another one and to fix the possible numerical instabilities that can occur, finally describing an efficient procedure to calculate the expectation value of the  $S^2$  operator.

Starting from the third chapter I will deal with the physical systems that have been investigated using the techniques described in the first two chapters. Here I will discuss the problem of the hydrogen chain[12], a one dimensional lattice of equispaced atoms, a study conducted in collaboration with the Simons Foundation. This system lies at the boundary between models and realistic systems, being the simplest realization of an atomic lattice and still resembling the Hubbard model. Despite its simplicity, this model shows a rich phase diagram and, surprisingly, a metal-insulator transition due to a self-doping mechanism. For this study we used a JSD *ansatz* that, despite the presence of more powerful alternatives, provides excellent results, also in comparison with the other methods used in the study[11, 12].

In the fourth chapter, instead, I will discuss an emblematic failure of the JSD *ansatz*, the  $H_4$  model system[13], that has been also extremely useful to demonstrate that the AGP WF, when not combined with a JF, can lead to

misleading results and also that the JF is fundamental to obtain the correct nodes in the AGP optimization.

In the fifth chapter I will show the benchmark of the JPf, JAGP and JSD WFs on the atoms and diatomic molecules of the first row of the periodic table and on the RVB prototypical case of the Benzene molecule[28, 33, 29]. Among the different dimers, I will particularly focus on the carbon, nitrogen, and oxygen ones, three systems where the magnetic interaction plays a fundamental role. In these cases only the JPf or multi-determinant WFs can reach a high accuracy while the JSD and JAGP have quantitatively and qualitatively wrong results. Here I will also show that for non-magnetic systems, like the Benzene, the use of the JAGP is equivalent to the JPf. Finally, we will see that also for non-covalent bonds, like for the benzene dimer, the use of optimized pairing functions like the JAGP leads to significantly more accurate results compared to the JSD also at DMC level.

Finally, I will try to summarize what we have discussed in the different sections to reach a better understanding of the potential limitations of the JAGP and the JPf, to highlight the possible future developments and systems that the JPf could help to investigate. But for the moment let us go to the most technical part to get some insight on how these methods work and how these *ansatze* are defined.

# Chapter 1

## Introduction to Quantum Monte Carlo

QMC is a widespread acronym used to indicate many different techniques for the numerical simulation of electronic systems, exploiting different principles and approximations, and all sharing the use of stochastic methods of calculation.

The statistical foundation of the Monte Carlo (MC) technique is standard and is out of the scope of this thesis. The interested reader is encouraged to visit the related Wikipedia web page[36] that will provide a decent introduction. For the technical details and a quantitative understanding of MC working principles I sincerely advise you to read the book "Quantum Monte Carlo Approaches for Correlated Systems" written by Federico Becca and Sandro Sorella[5]. The first chapter of the book describes the MC in general, while the later chapters deal with many QMC approaches for lattice models and the last chapter deals with QMC for realistic system simulations. For what concerns us, however, it is enough to say that if we have a quantity  $\mathbf{x}$  distributed according to a certain probability distribution  $p(\mathbf{x})$ , a MC algorithm can provide an arbitrarily large set of samples  $\{\mathbf{x}_i\}$  distributed according to the probability distribution  $p(\mathbf{x})$ .

There are different ways to build an algorithm to study a quantum system

using a MC method, but, as I have already anticipated in this thesis I will use two different techniques that are suitable to investigate atomistic realistic systems: the variational Monte Carlo (VMC) and the diffusion Monte Carlo (DMC). Within these methods, the quantity  $\mathbf{x}$  to be sampled for a system of  $N$  electrons is the real space configuration of the  $3N$  electronic coordinates and the  $N$  spin values  $\mathbf{x} = (\mathbf{r}_1\sigma_1, \mathbf{r}_2\sigma_2, \dots, \mathbf{r}_N\sigma_N)$ . We can notice that despite the VMC and DMC working mechanisms are built around the concept of WF, they can be described without specifying the particular *ansatz* chosen. The *ansatze* will be the subject of the next chapter. In the following sections we will analyze these two different methods and the optimization techniques implemented in TURBORVB, and for this reason I will follow the review paper that we recently published summarizing its features[29].

## 1.1 Variational Monte Carlo

As we can easily guess from its name, the VMC method working principle relies on the variational principle for the calculation and optimization of the WF  $\Psi$  in order to find the best possible approximate GS within a given *ansatz*. The expectation value of the energy  $\langle E \rangle$  can be written as:

$$\langle E \rangle = \frac{\int d\mathbf{x} \Psi^2(\mathbf{x}) \cdot \hat{\mathcal{H}}\Psi(\mathbf{x}) / \Psi(\mathbf{x})}{\int d\mathbf{x} \Psi^2(\mathbf{x})}, \quad (1.1)$$

where  $\mathbf{x} = (\mathbf{r}_1\sigma_1, \mathbf{r}_2\sigma_2, \dots, \mathbf{r}_N\sigma_N)$  here and henceforth is a shorthand notation for the  $N$  electron coordinates and their spins. If we define

$$e_L(\mathbf{x}) = \frac{\hat{\mathcal{H}}\Psi(\mathbf{x})}{\Psi(\mathbf{x})}, \text{ and } \pi(\mathbf{x}) = \frac{\Psi^2(\mathbf{x})}{\int d\mathbf{x}' \Psi^2(\mathbf{x}')}, \quad (1.2)$$

the so-called local energy and the probability of the configuration  $\mathbf{x}$ , respectively, we can recast Eq. (1.1) as

$$\langle E \rangle = \int d\mathbf{x} e_L(\mathbf{x}) \pi(\mathbf{x}), \quad (1.3)$$

This multidimensional integration can be evaluated stochastically by generating a set of configurations  $\{\mathbf{x}_i\}$  according to the distribution  $\pi(\mathbf{x})$  using the Markov chain MC such as the (accelerated [37, 38]) Metropolis method and by averaging the obtained local energies  $e_L(\mathbf{x}_i)$ :

$$E_{\text{VMC}} = \langle e_L(\mathbf{x}) \rangle_{\pi(\mathbf{x})} \approx \frac{1}{M} \sum_{i=1}^M e_L(\mathbf{x}_i), \quad (1.4)$$

which has an associated statistical error of  $\sqrt{\text{Var}[e_L(\mathbf{x}_i)]/\tilde{M}}$ , where  $\text{Var}[e_L(\mathbf{x}_i)]$  is the variance of the sampled local energies, and  $\tilde{M}$  is the sampling size  $M$  divided by the autocorrelation time. This indicates that the precision of the VMC evaluation is inversely proportional to the square root of the number of samplings (*i.e.*, of the computational cost). It is worth to notice that, if  $\Psi(\mathbf{x})$  is an eigenfunction of  $\hat{\mathcal{H}}$ , say with eigenvalue  $E_0$ , than  $e_L(\mathbf{x}) = E_0$  for each  $\mathbf{x}$ , implying that the variance of the local energy is zero and  $E_{\text{VMC}} = E_0$  with no stochastic uncertainty. This feature is known as the zero-variance property.

This scheme for the calculation of the expectation value of the energy can be extended to all the other observables that can be evaluated in real space. Considering a generic operator  $\hat{O}$  we have that its expectation value  $\langle O \rangle$  is

$$\langle O \rangle = \frac{\int d\mathbf{x} \Psi^2(\mathbf{x}) \cdot \hat{O}\Psi(\mathbf{x}) / \Psi(\mathbf{x})}{\int d\mathbf{x} \Psi^2(\mathbf{x})} = \int d\mathbf{x} O_L(\mathbf{x}) \pi(\mathbf{x}), \quad (1.5)$$

where  $\pi(\mathbf{x})$  has the same definition of Eq. (1.2) and the local value of the operator  $O_L(\mathbf{x})$  is defined as

$$O_L(\mathbf{x}) \equiv \frac{\hat{O}\Psi(\mathbf{x})}{\Psi(\mathbf{x})}. \quad (1.6)$$

With the same argument used for the energy we can define the VMC expectation value of an operator  $O_{\text{VMC}}$  as

$$O_{\text{VMC}} = \langle O_L(\mathbf{x}) \rangle_{\pi(\mathbf{x})} \approx \frac{1}{M} \sum_{i=1}^M O_L(\mathbf{x}_i). \quad (1.7)$$

A similar procedure is used to average the derivatives of the energy with respect to the WF parameters. Given a set of parameters  $(\alpha_1, \alpha_2, \dots, \alpha_p)$  that define the WF  $\Psi(\mathbf{x}, \alpha)$  one can use the *variational theorem*

$$E_{\text{VMC}}(\alpha) = \int d\mathbf{x} e_L(\mathbf{x}, \alpha) \pi(\mathbf{x}, \alpha) \geq E_{\text{exact}}. \quad (1.8)$$

Indeed, in order to obtain an accurate description of the GS, the optimization of the WF is a crucial step for VMC and, as a consequence, also for DMC. The optimization techniques are the topic of the next section.

## 1.2 Optimization

The optimization of a many-body WF is a difficult challenge. Not only optimizing a cost function containing several parameters is a complex numerical task, due to the presence of several local minima in the energy landscape, but also this is further complicated by the presence of statistical errors in the QMC evaluation of any quantity.

Nevertheless, a great improvement in this field has been achieved when the QMC optimization techniques have made use of the explicit evaluation of energy derivatives with finite statistical errors. For the calculation of the generalized forces ( $f_i = -\frac{\partial E}{\partial \alpha_i}$ ) TURBORVB adopts the adjoint algorithmic differentiation (AAD)[35]. This technique exploits the chain rule for the calculation of the derivatives.

Unfortunately, the bare knowledge of the derivatives is not enough for an optimization problem that can involve thousands of parameters and a good strategy is also necessary. For VMC optimization it is particularly efficient to use the so-called "stochastic reconfiguration" (SR) [39, 40] combined sometimes with "the linear method" [41, 42, 43]. The following subsection will deal with the calculation of the derivatives through the AAD and the SR optimization strategy.

### 1.2.1 Derivatives Calculation

The derivatives of total energies with respect to variational parameters represent an essential ingredient for optimizing a many-body WF. Also forces (derivative with respect to atomic positions) are essential to perform structural optimization or molecular dynamics calculation. However, in a complex code, and especially in QMC, the evaluation of the functional derivatives, necessary for the WF optimization, is very complicated, mainly for the complexity of the algorithm that, in turn, may lead to a very inefficient implementation, though recent progress has been done [44]. For instance, a simple approach is to compute them with finite difference expressions, that leads to a too large computational time, because it is obviously proportional to the number of targeted derivatives.

The AAD is a method capable of solving all the above problems, essentially by a smart application of the differentiation chain rule. This method is particularly efficient when the number of input parameters is much larger than the corresponding output and allows the calculation of all possible derivatives in a computational time proportional (with a small prefactor, see *e.g.*, Fig. 1.1) to the one for computing the target function (*i.e.*, the energy or the WF value). AAD is therefore the ideal method for Quantum Monte Carlo when the variational WF contains several variational parameters.

But how does AAD work? I will not provide a rigorous and complete explanation, but it can be useful to gain a general intuition about this technique that is now widely used in different fields. We can define an algorithm as the function  $W$  mapping a set of inputs  $X_0$  into a set of output targets  $Y$  such that  $Y = W(X_0)$ . In particular, an algorithm goes through a set of  $l$  code assignment operations  $\{W_i\}$  that here I will assume to be all differentiable. In this way, the input  $X_0$  is transformed in  $X_1 = W_1(X_0)$ , and, in general,  $X_i = W_i(X_{i-1})$ , until the last operation  $Y = W_l(X_{l-1})$ . Obviously in many

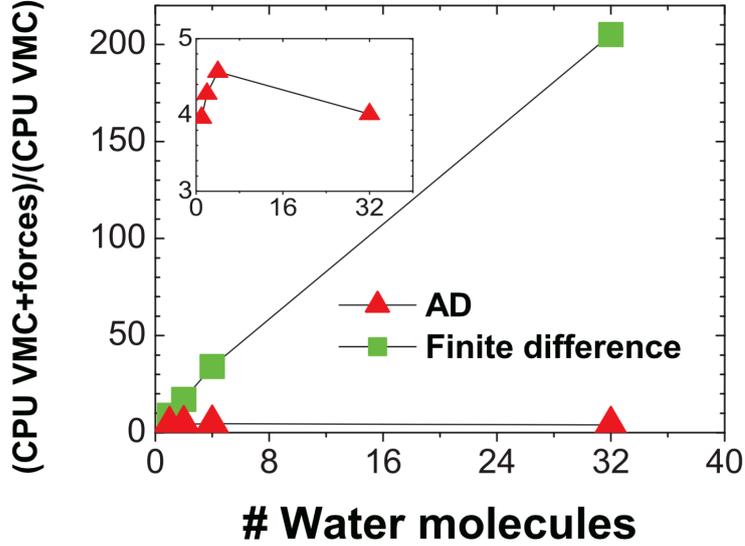


Figure 1.1: Ratio of CPU time required to compute energies and all force components referenced to the one required for the simple energy calculation within VMC. The calculations refer to 1, 2, 4, and 32 water molecules. The inset is an expansion of the lower part of the plot. From reference [35]

operations not all the inputs are modified and for simplicity we can consider as if an identity is applied to them. The algorithm  $W$ , that takes the name of direct algorithm, is decomposed in this way in a set of small easily differentiable operations. Then it can be described as

$$Y = W_l(W_{l-1}(W_{l-2}(\dots W_1(X_0))\dots)). \quad (1.9)$$

If we apply the chain rule for the derivation of a composed function, we obtain that the derivatives of the output  $Y^i$  with respect to an input  $X^j$  can be calculated as

$$\frac{\partial Y^i}{\partial X^j} = \frac{\partial W_l^i}{\partial X_{l-1}} \Big|_{X_{l-1}} \cdot \frac{\partial W_{l-1}}{\partial X_{l-2}} \Big|_{X_{l-2}} \dots \frac{\partial W_1}{\partial X_0} \Big|_{X_0}. \quad (1.10)$$

So, starting from the direct algorithm we can derive every single operation with respect to their input and then calculating the derivatives of the final outputs with respect to the initial inputs by following *backward* the composition order,

defining an adjoint algorithm. Using this technique it is possible to calculate the derivatives of each function that can be written in the form of an algorithm under the only hypothesis that all the operations involved are differentiable.

As already mentioned, all the derivatives in TURBORVB are calculated by applying AAD. In the following we will focus on the calculation of the derivatives of the energy with respect to the variational parameters. Even if it is also possible to calculate the interatomic forces, they have not been deeply used in the work presented here and thus their derivation is not presented.

The derivatives of the energy with respect to a given real variational parameter  $\alpha_k$  (one complex parameter can be thought to be composed by two real ones, its real and imaginary part, respectively) is represented as a generalized force:

$$f_k = -\frac{\partial E(\alpha)}{\partial \alpha_k} = -\frac{\partial}{\partial \alpha_k} \frac{\langle \Psi_\alpha | \hat{\mathcal{H}} | \Psi_\alpha \rangle}{\langle \Psi_\alpha | \Psi_\alpha \rangle}. \quad (1.11)$$

In variational Monte Carlo, the derivative can be evaluated using  $M$  configurations of electron coordinates [5]:

$$\begin{aligned} f_k &= -2 \operatorname{Re} \left[ \sum_{\mathbf{x}} \frac{e_L^*(\mathbf{x}) (O_k(\mathbf{x}) - \bar{O}_k) |\Psi_\alpha(\mathbf{x})|^2}{\sum_{\mathbf{x}} |\Psi_\alpha(\mathbf{x})|^2} \right] \\ &\approx -2 \operatorname{Re} \left[ \frac{1}{M} \sum_{i=1}^M e_L^*(\mathbf{x}_i) (O_k(\mathbf{x}_i) - \bar{O}_k) \right], \end{aligned} \quad (1.12)$$

where  $e_L(\mathbf{x})$  is the local energy,  $O_k(\mathbf{x})$  is the logarithmic derivative of the WF (*i.e.*,  $O_k(\mathbf{x}) = \frac{\partial \ln \Psi_\alpha(\mathbf{x})}{\partial \alpha_k}$ ), and  $\bar{O}_k$  is its average over  $M$  samples (*i.e.*,  $\bar{O}_k = \frac{1}{M} \sum_{i=1}^M O_k(\mathbf{x}_i)$ ). In TURBORVB, the logarithmic derivative ( $O_k(\mathbf{x})$ ) is computed very efficiently by using the algorithm defined with the AAD. Notice that the derivatives of the local energy are not needed here because the Hamiltonian does not depend on any variational parameter. Instead, these terms are necessary in order to calculate ionic forces (*i.e.*, derivatives of the total energies with respect to atomic positions). If the WF is an exact eigenstate of the hamiltonian, the generalized forces  $f_k$  exactly vanish without sta-

tistical errors because the local energy is no longer dependent on  $\mathbf{x}$ . In other words, the derivatives have the zero-variance property and represent therefore the fundamental ingredients for an efficient WF optimization.

### 1.2.2 Optimization Strategy

The bare knowledge of the derivatives, for such a complicated variational problem, is unfortunately only a first step for the optimization. A good optimization strategy is indeed necessary to deal with a large number of parameters, defining the WFs in a highly non-linear way. Once the energy derivatives are computed, the most straightforward strategy to optimize a WF is to employ the gradient descent method, where the WF parameters are iteratively updated as follows:

$$\alpha_k \rightarrow \alpha'_k = \alpha_k + \delta\alpha_k \quad (1.13)$$

$$\delta\alpha_k = \Delta \mathbf{f}_k, \quad (1.14)$$

where  $\Delta$  is a small constant and  $f_k \equiv -\frac{\partial E}{\partial \alpha_k}$  is the generalized force already defined in Eq. (1.11). However, this method does not work well when optimizing highly non-linear WF parameters because a small change of a given variational parameter may produce a very different WF, whereas another parameter change may weakly affect the WF. Of course, one can use more sophisticated methods such as the Newton-Raphson method, the conjugate gradient, the quasi-Newton method, but the straightforward implementation of these optimizations do not work efficiently within a stochastic approach like QMC. In order to overcome this difficulty, a more efficient change in the variational parameters has been defined by means of a positive-definite *preconditioning* matrix  $\mathbf{S}$  and the generalized force vector  $\mathbf{f}$ :

$$\alpha_k \rightarrow \alpha_k + \Delta \cdot (\mathbf{S}^{-1} \mathbf{f})_k, \quad (1.15)$$

where the matrix  $\mathcal{S}$  is stochastically evaluated by means of  $M$  configuration samples  $\mathbf{x} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_M\}$ :

$$\mathcal{S}_{k,k'} = \left[ \frac{1}{M} \sum_{i=1}^M (O_k(\mathbf{x}_i) - \bar{O}_k)^* (O_{k'}(\mathbf{x}_i) - \bar{O}_{k'}) \right], \quad (1.16)$$

where  $O_k(\mathbf{x}_i) = \frac{\partial \ln \Psi(\mathbf{x}_i)}{\partial \alpha_k}$  and  $\bar{O}_k = \frac{1}{M} \sum_{i=1}^M O_k(\mathbf{x}_i)$ . The resulting approach is the so-called stochastic reconfiguration (SR) method [39]. The matrix  $\mathcal{S}$  is essentially a *metric* for the parameter space, measuring the distance of two very close and normalized WFs[45]. Therefore, Eq. (1.15) is simply the steepest descent in this curved manifold. This observation connects the SR method with the so-called natural gradient method, widely used in the context of deep learning [46]. In this context, for each parameter  $\alpha$ ,  $(\hat{O}_k - \bar{O}_k)$  and  $\mathcal{S}_{k,k'}$  can be interpreted as the score function (*i.e.*, the gradient of the log-likelihood function) and the Fisher information matrix, respectively, while the WF square  $|\Psi(\mathbf{x})|^2$  plays the role of the likelihood function. In this sense, the stochastic reconfiguration method is essentially identical to the natural gradient optimization with the Fisher information matrix that has been intensively used in the machine-learning community.[46]

The straightforward implementation of the SR method is often not possible considering that both  $\mathcal{S}$  and  $\mathbf{f}$  are computed stochastically, that implies instabilities in the evaluation of  $\mathcal{S}'^{-1}\mathbf{f}$ . Therefore, in practice, the diagonal elements of the preconditioning matrix  $\mathcal{S}$  are shifted by a small positive parameter ( $\varepsilon$ ) as:

$$s'_{i,i} = s_{i,i}(1 + \varepsilon). \quad (1.17)$$

This modification improves the efficiency and stability of the optimization by several orders of magnitude, as shown in Fig. 1.2. Finally, the variational parameters are updated as:

$$\alpha_k \rightarrow \alpha_k + \Delta \cdot \left( \mathcal{S}'^{-1}\mathbf{f} \right)_k. \quad (1.18)$$

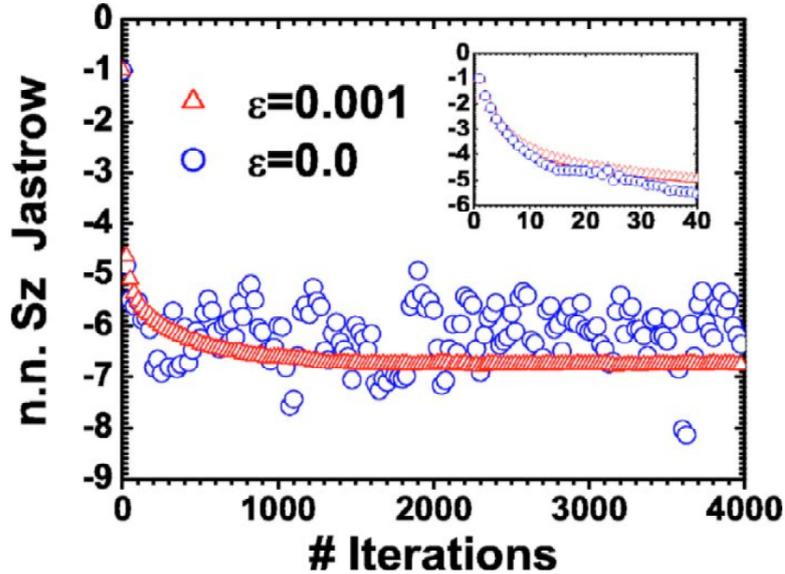


Figure 1.2: Optimization of the variational WF in the simple one-dimensional Heisenberg model  $\mathcal{H} = J \sum_i \vec{S}_i \cdot \vec{S}_{i+1}$  with the standard SR ( $\varepsilon = 0$ , open circles) and with the present regularization  $\varepsilon = 0.001$ , open triangles). In the figure, the evolution of the nearest neighbor spin-spin ( $S_z$ ) Jastrow parameter is plotted. This figure clearly shows that the SR method with the regularization adopted is several orders of magnitude more efficient than the standard SR for determining the variational parameter with a given statistical accuracy. The inset shows the first few iterations. From reference [47].

For the optimization of WFs with a large number of variational parameters, the inversion of the matrix  $\mathcal{S}'$  can easily become the bottleneck for the performance of the optimization procedure since it implies the inversion of a square matrix with leading dimension equal to the number of variational parameters. Recently a highly efficient parallel realization of the SR has been developed and it has been intensively used to obtain the results presented in this thesis. This alternative approach uses the conjugate gradient method to calculate  $\mathcal{S}'^{-1}\mathbf{f}$ . With the conjugate gradient the matrix  $\mathcal{S}'$  is used only to perform matrix vector multiplications. The  $M$  samples are distributed between the MPI tasks and the sample dimension is much smaller than the number of variables.

In this case, instead of calculating directly the matrix product  $\mathcal{S}'X$  where  $X$  is a generic vector, it is convenient to factorize the matrix  $\mathcal{S}'$  as

$$\mathcal{S} = \tilde{O}\tilde{O}^\dagger, \quad (1.19)$$

where  $\tilde{O}_{ij} = O_i(\mathbf{x}_j) - \bar{O}_i$ , and perform two separate matrix vector multiplications exploiting the fact that the matrix  $\tilde{O}$  is already distributed in the memory of all the processors used. This implementation makes it possible to use SR with an almost negligible cost even for several thousands of variational parameters.

### 1.3 Diffusion Monte Carlo

Until now I referred to the DMC technique without further specifying the algorithm. In TURBORVB a variant of the DMC that is known as Lattice regularized diffusion Monte Carlo (LRDMC) has been implemented. Initially proposed by M. Casula *et al.* [48], it is a projection technique that allows us to improve a variational *ansatz* systematically. This method is based on the lattice Green's function Monte Carlo (GFMC) [49, 50, 51], and filters out the ground state WF  $|\Upsilon_0\rangle$  from a given trial WF  $|\Psi_T\rangle$ . As explained in the following, the GFMC relies on the matrix projection technique for determining the largest eigenvalue and its corresponding eigenvector.

Since the eigenstates of the Hamiltonian have the completeness property, the trial WF can be expanded as:

$$|\Psi_T\rangle = \sum_n a_n |\Upsilon_n\rangle, \quad (1.20)$$

where  $a_n$  is the coefficient for the  $n$ -th eigenvectors ( $\Upsilon_n$ ) of  $\hat{\mathcal{H}}$ . Therefore, by

applying  $\mathcal{G}^M = (\Lambda - \hat{\mathcal{H}})^M$ , one obtains

$$\begin{aligned} |\Upsilon_0\rangle &\propto \lim_{M \rightarrow \infty} (\Lambda - \hat{\mathcal{H}})^M |\Psi_T\rangle \\ &= \lim_{M \rightarrow \infty} (\Lambda - E_0)^M \left[ a_0 |\Upsilon_0\rangle + \sum_{n \neq 0} \left( \frac{\Lambda - E_n}{\Lambda - E_0} \right)^M a_n |\Upsilon_n\rangle \right], \end{aligned} \quad (1.21)$$

where  $\Lambda$  is a diagonal matrix with  $\Lambda_{x,x'} = \lambda \delta_{x,x'}$ , and  $E_n$  is  $n$ -th eigenvalue of  $\hat{\mathcal{H}}$ . Since  $\frac{\Lambda - E_n}{\Lambda - E_0} < 1$  ( $\lambda$  should be sufficiently large in order to fulfill this condition for any  $n \neq 0$ ), the projection filters out the ground state WF  $\Upsilon_0$  from a given trial WF  $|\Psi_T\rangle$ , as long as the trial WF is not orthogonal to the true ground state (*i.e.*,  $a_0 \equiv \langle \Psi_T | \Upsilon_0 \rangle \neq 0$ ).

In order to apply the LRDMC for *ab initio* electron calculations, the original continuous Hamiltonian is regularized by allowing electron hopping with step size  $a$ , in order to mimic the electronic kinetic energy. The corresponding Hamiltonian  $\hat{\mathcal{H}}^a$  is then defined such that  $\hat{\mathcal{H}}^a \rightarrow \hat{\mathcal{H}}$  for  $a \rightarrow 0$ . Namely, the kinetic part is approximated by a finite difference form:

$$\begin{aligned} \Delta_i f(x_i, y_i, z_i) &\approx \Delta_i^a f(x_i, y_i, z_i) \\ &\equiv \frac{1}{a^2} \{ [f(x_i + a) - f(x_i)] + [f(x_i - a) - f(x_i)] \} \\ &\leftrightarrow y_i \leftrightarrow z_i, \end{aligned} \quad (1.22)$$

and the potential term is modified as:

$$V^a(\mathbf{x}) = V(\mathbf{x}) + \frac{1}{2} \left[ \frac{\sum_i (\Delta_i^a - \Delta_i)(\mathbf{x})}{\Psi_G(\mathbf{x})} \right]. \quad (1.23)$$

The corresponding Green's function matrix elements are:

$$\mathcal{G}_{x,x'} = (\Lambda - \mathcal{H}_{x',x}^a), \quad (1.24)$$

and the single LRDMC iteration step is given by the following equation:

$$\Psi_{n+1}(x') = \sum_x \mathcal{G}_{x',x} \Psi_n(x). \quad (1.25)$$

The sketch of the LRDMC algorithm, a Markov chain that evolves the many-body WF according to the Eq. (1.25), is as follows [48]: (STEP 1) Prepare

a walker with configuration  $x$  and weight  $w$  ( $w_0 = 1$ ). (STEP 2) A new configuration  $x'$  is generated by the transition probability:

$$p_{x,x'} = \mathcal{G}_{x,x'}/b_x, \quad (1.26)$$

where

$$b_x = \sum_{x'} \mathcal{G}_{x',x} \quad (1.27)$$

is a normalization factor. By applying the discretized Hamiltonian to a given configuration ( $\hat{\mathcal{H}}^a |x\rangle$ ),  $(6N + 1)$  configurations  $|x'\rangle$  are determined according to the probability  $p_{x,x'}$  in Eq. (1.26), where  $N$  is the number of electrons in the system [52]. This allows the evaluation of the normalization factor  $b_x$  in Eq. (1.27) even in a continuous model. Notice that  $6N$  comes from the diffusion of each electron in two directions ( $\pm a$ ) and the remaining 1 stands for the starting configuration before the possible hopping (all  $N$  electrons)  $x$  (*i.e.*,  $x' = x$ ). (STEP 3) Finally, update the weight with  $w_{n+1} = w_n b_x$ , and return to the STEP I. After a sufficiently large number of iterations (the Markov process is equilibrated), one can calculate the ground state energy  $E_0$ :

$$E_0 \approx \frac{\langle w_n e_L(x_n) \rangle}{\langle w_n \rangle}, \quad (1.28)$$

where  $\langle \dots \rangle$  denotes the statistical average over many independent samples generated by the Markov chains, and  $e_L(x)$  is called the (bare) local energy that reads:

$$e_L(x) = \sum_{x'} \mathcal{H}_{x',x} = \lambda - b_x. \quad (1.29)$$

Indeed, the ground state energy can be calculated after many independent  $n$ -step calculations. A more efficient computation can be realized by using the so-called "correcting factor" technique: after a single simulation that is much larger than the equilibration time, one can imagine starting a projection of length  $p$  from each  $(n - p)^{th}$  iteration. The accumulated weight for each

projection is:

$$\mathcal{G}_n^p = \prod_{j=1}^p b_{n-j}. \quad (1.30)$$

Then, the ground state energy  $E_0$  can be estimated by:

$$E_0 \approx \frac{\sum_n \mathcal{G}_n^p e_L(x_n)}{\sum_n \mathcal{G}_n^p}. \quad (1.31)$$

This straightforward implementation of the above simple method is not suitable for realistic simulations due to fluctuations of weights, large correlation times, the sign problems, and so on. TURBORVB implements the following state-of-art techniques for real electronic structure calculations.

If the potential term (Eq. (1.23)) is unbounded (it is the case in *ab initio* calculations), the bare weight  $b_x$  (Eq. (1.27)) and the local energy  $e_L(x)$  (Eq. (1.29)) significantly fluctuate, making the numerical simulation very unstable and inefficient. To overcome this difficulty, the code employs the importance sampling scheme [5], in which the original Green's function is modified using the so-called guiding function  $\Psi_G$  as:

$$\tilde{\mathcal{G}}_{x',x} = \mathcal{G}_{x',x} \frac{\Psi_G(x')}{\Psi_G(x)}, \quad (1.32)$$

and the projection is modified as:

$$\Psi_G(x') \Psi_{n+1}(x') = \sum_x \tilde{\mathcal{G}}_{x',x} \Psi_G(x) \Psi_n(x). \quad (1.33)$$

In practice, the guiding function is prepared by a VMC calculation. The modified Green's function for importance sampling  $\tilde{\mathcal{G}}_{x',x}$  has the same eigenvalues as the original one, and this transformation does not change the formalism of LRDMC. The weight is updated by:

$$\tilde{b}_x = \sum_{x'} \tilde{\mathcal{G}}_{x',x}, \quad (1.34)$$

and the local energy with importance sampling is:

$$\tilde{e}_L(x) = \frac{\langle \Psi_G | \hat{\mathcal{H}} | x \rangle}{\langle \Psi_G | x \rangle} = \sum_{x'} \mathcal{H}_{x,x'} \frac{\Psi_G(x')}{\Psi_G(x)}. \quad (1.35)$$

Eq. (1.35) implies that if the guiding function  $\Psi_G$  is an exact eigenstate of the Hamiltonian, there are no statistical fluctuations, implying the zero variance property, namely the computational efficiency to obtain a given statistical error on the energy improves with the quality of the variational WF. In this respect, it is also important to emphasize that a meaningful reduction of the statistical fluctuations is obtained by satisfying the so-called cusp conditions. As long as they are satisfied, the resulting local energy does not diverge at the coalescence points where two particles overlap, despite the singularity of the Coulomb potential term ( $V(x)$  in Eq. (1.23)) [52]. In addition, the importance sampling maintains the electrons in a region away from the nodal surface, since the guiding function vanishes there (*i.e.*, the right-hand side of Eq. (1.32)  $\rightarrow 0$ ). This clearly enhances the efficiency of the sampling because the local energy diverges at the nodal surface.

It is easy to identify in Eq. (1.26) the source of a crucial issue of this technique known as the sign problem: if  $\mathcal{G}_{x,x'}$  is negative Eq. (1.26) cannot have a simple formulation in terms of stochastic process. In general the Green's function cannot be made strictly positive for fermions; therefore, in order to avoid the sign problem, the fixed-node (FN) approximation has to be introduced [5]. Indeed, the Hamiltonian is modified using the spin-flip term

$$\mathcal{V}_{\text{sf}}(x) = \sum_{x': s_{x,x'} > 0} \mathcal{H}_{x,x'} \Psi_G(x') / \Psi_G(x):$$

$$\mathcal{H}_{x,x'}^{\text{FN},\gamma} = \begin{cases} \mathcal{H}_{x,x} + (1 + \gamma) \mathcal{V}_{\text{SF}}(x) & \text{for } x' = x, \\ \mathcal{H}_{x,x'} & \text{for } x' \neq x, s_{x,x'} < 0, \\ -\gamma \mathcal{H}_{x,x'} & \text{for } x' \neq x, s_{x,x'} > 0, \end{cases} \quad (1.36)$$

where  $s_{x,x'} = \Psi_G(x) \mathcal{H}_{x,x'} \Psi_G(x')$  and  $\gamma \geq 0$  is a real parameter. The use of the fixed-node Green's function:

$$\tilde{\mathcal{G}}_{x,x'}^{\text{FN}} = (\mathbf{\Lambda} - \mathcal{H}_{x',x}^{\text{FN}}) \frac{\Psi_G(x')}{\Psi_G(x)} \quad (1.37)$$

can prevent the crossing of regions where the configuration space yields a sign

flip of the Green's function; therefore, the walkers are constrained in the same nodal pockets and avoid the sign problem.

TURBORVB also implements the many-walker technique and the branching (denoted as reconfiguration [53] in TURBORVB) scheme for a more efficient computation [5]. The code performs the branching as follows: (1) Set the new weights equal to the average of the old ones:

$$w'_{\alpha,n} = \bar{w} \equiv \frac{1}{N_w} \sum_{\beta} w_{\beta,n}. \quad (1.38)$$

(2) Select the new walkers from the old ones with a probability that is proportional to the old walkers' weights:

$$p_{\alpha,n} = \frac{w_{\alpha,n}}{\sum_{\beta} w_{\beta,n}}, \quad (1.39)$$

which does not change the statistical average of weights, but suppresses the fluctuations by dropping walkers having small weights. The code performs branching (reconfiguration) after a projection time  $\tau$ , that can be chosen as a user input parameter. In practice, within the many-walker and branching schemes, the average weights are stored and are set to one for all walkers after each branching. The accumulated weights can be retrieved at the end of the simulation:

$$\mathcal{G}_n^p = \prod_{j=1}^p \bar{w}_{n-j}, \quad (1.40)$$

and one can calculate the ground state energy:

$$E_0 \approx \frac{\sum_n \mathcal{G}_n^p e_L(x_n)}{\sum_n \mathcal{G}_n^p}, \quad (1.41)$$

where  $e_L(x_n)$  is the mean local energy averaged over the walkers, which reads

$$e_L(x_n) = \frac{\sum_{\alpha} w_{\alpha,n} e_L(x_{\alpha,n})}{\sum_{\alpha} w_{\alpha,n}} \quad (1.42)$$

and is evaluated just before each reconfiguration. Notice that  $p$  is also an input parameter, that has to be carefully chosen by the user to allow energy convergence.

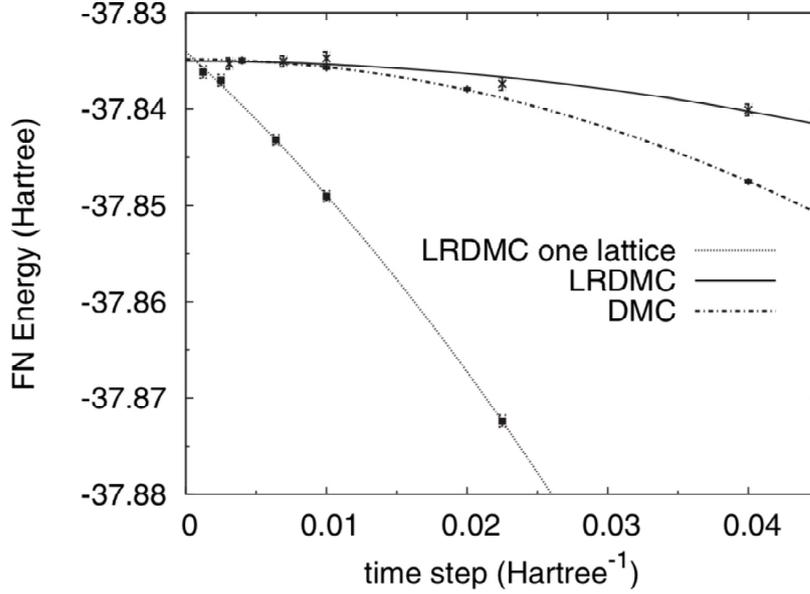


Figure 1.3: Fixed-node energies for the all-electron carbon atom computed within DMC, single-grid LRDMC (one lattice), and double-grid LRDMC. The lattice discretization parameter  $a$  is mapped to the time-step  $\tau$  as  $a = \sqrt{\tau}$ . From reference [48].

When  $\lambda$  is sufficiently large, the correlation time also becomes large because the diagonal terms of the Green's function become very close to one (*i.e.*, a walker remains in the same configuration), which causes a very large correlation time. This difficulty can be solved by considering in a different way the diagonal and non-diagonal moves. In a given interval of  $M$  iterations, the non-diagonal updates are efficiently calculated by sampling directly the probability to remain in the same configuration with the generation of a single random number. This technique can be generalized to the continuous-time limit, namely,  $M \rightarrow \infty$ , at  $\frac{M}{\Lambda} = \tau$  fixed. In the  $M \rightarrow \infty$  limit, the projection  $(\Lambda - \hat{\mathcal{H}})^M$  is equal to the imaginary time evolution  $\exp(-\tau\hat{\mathcal{H}})$ , apart for an irrelevant constant  $\Lambda^M$ . The branching (reconfiguration) is performed at the end of each time interval of length  $\tau$  within the many-walker and the branching implementation.

In practice, there are three important features in LRDMC. First, there is not a time-step error in LRDMC because the Suzuki-Trotter decomposition is not necessary, unlike the standard DMC algorithm [5]. Instead, there is a finite-size lattice size ( $a$ ) error due to the discretization of the Hamiltonian ( $a$ ). Therefore, in order to obtain an unbiased FN energy, it is important to extrapolate the LRDMC energy to the  $a \rightarrow 0$  limit by using several results corresponding to different lattice spaces [48]. This is then consistent with the standard DMC energy estimate (Fig. 1.3) obtained in the limit of an infinitely small time step. Probably one of the most important advantages of the LRDMC method is that the extrapolation to the  $a \rightarrow 0$  limit is very smooth and reliable, so that unbiased FN energies are easily obtained with low order polynomial fits. Secondly, LRDMC can straightforwardly handle different length scales of a WF by introducing different mesh sizes ( $a$  and  $a'$ ), so that electrons in the vicinity of the nuclei and those in the valence region can be appropriately diffused[48, 54], which defines the so-called double-grid LRDMC. This scheme saves a substantial computational cost in all-electron calculations, especially for a system including atoms with a large atomic number[54].

# Chapter 2

## Wave Functions

As already mentioned, the WF is the principal ingredient and the most crucial component for VMC and DMC calculations. In this chapter I will present the different WF *ansatze* that have been used for the studies presented in this thesis with a particular focus on the Pfaffian WF (Pf). Indeed, this last one, together with the related tools, can be considered the most important technical advance developed in TURBORVB during my PhD. For the description of the WFs I will mainly follow our review paper about TURBORVB[29] and our first paper published on the Pfaffian WF[28].

Both the accuracy and the computational efficiency of QMC approaches critically depend on the WF *ansatz*. The optimal *ansatz* is typically a tradeoff between accuracy and efficiency. On one hand, a very accurate *ansatz* can be involved and cumbersome, having many parameters and being expensive to evaluate. On the other hand, an efficient *ansatz* is described only by the most relevant parameters and can be quickly and easily evaluated. In particular, in the previous chapter, we have seen that QMC algorithms, both at the variational and fixed-node level, imply several calculations of the local energy  $e_L(\mathbf{x})$  and the ratio  $\Psi(\mathbf{x})/\Psi(\mathbf{x}')$  for different electronic configurations  $\mathbf{x}$  and  $\mathbf{x}'$ . The computational cost of these operations determines the overall efficiency of QMC and its scaling with systems size.

A common choice for QMC calculations is to employ a many-body WF

*ansatz*  $\Psi$  which can be written as the product of two terms:

$$\Psi = J(\mathbf{x}) \times \Phi_{\text{AS}}(\mathbf{x}), \quad (2.1)$$

where the term  $J$ , conventionally dubbed Jastrow factor, is symmetric under electron exchange, and the term  $\Phi_{\text{AS}}$ , also referred to as the determinant part of the WF, encodes the fermionic nature of the WF and is antisymmetric. The resulting WF  $\Psi$  is antisymmetric, thus fermionic.

In the majority of QMC applications, the chosen  $\Phi_{\text{AS}}$  is a SD  $\Phi_{\text{SD}}$ , *i.e.*, an antisymmetrized product of single-electron WFs. Clearly, SD alone does not include any correlation other than the exchange. However, when a Jastrow factor, explicitly depending on the inter-electronic distances, is applied to  $\Phi_{\text{SD}}$  the resulting *ansatz*, Jastrow correlated SD (JSD), often provides already at the variational level over 70% of the correlation energy. This last quantity is typically defined as the difference between the exact energy and the Hartree-Fock energy, which is the variational minimum for a SD *ansatz*. Thus, the Jastrow factor proves very effective in describing the correlation, employing only a relatively small number of parameters, and therefore providing a very efficient way to improve the *ansatz*. A JSD function yields a computational cost for QMC simulations – both VMC and FN – about  $\propto N^3$ , namely the same scaling of most DFT codes. Therefore, although QMC has a much larger prefactor, it represents an approach much cheaper than traditional quantum chemistry ones, at least for large enough systems.

While the JSD *ansatz* is quite satisfactory for several applications, there are situations where very high accuracy is required, and a more advanced *ansatz* is necessary. The route to improve JSD is not unique, and different approaches have been attempted within the QMC community. First, it should be mentioned that improving the Jastrow factors is not an effective approach to achieve higher accuracy at the FN level, as the Jastrow is positive and cannot change the nodal surface. A popular approach is through the employment of

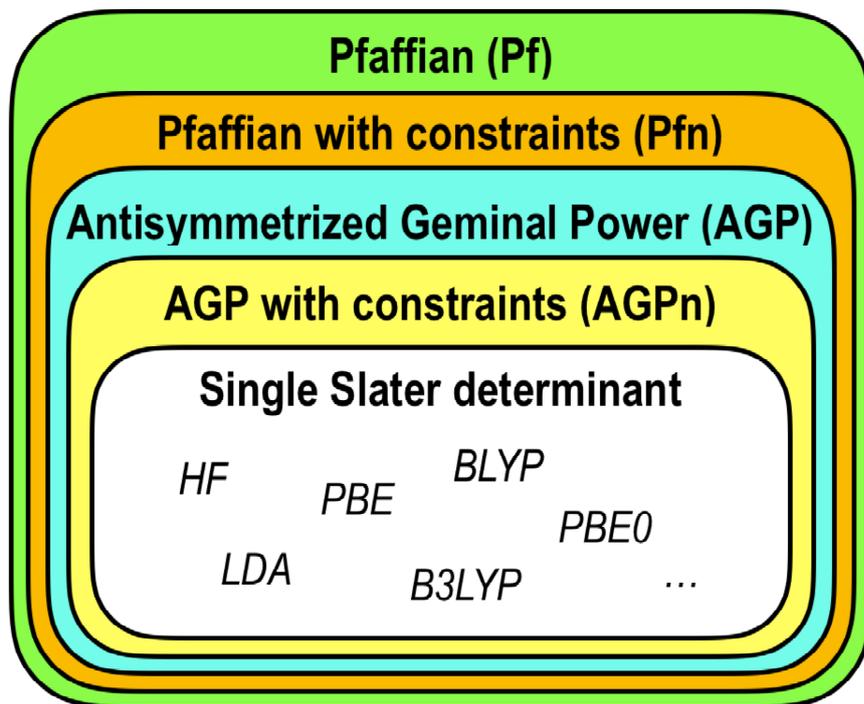


Figure 2.1: *Ansatz* hierarchy. The output of Hartree-Fock (HF) or DFT simulations with different exchange-correlation functionals are special instances of the SD *ansatz*. From reference [29].

backflow[55], which is a remapping of the electronic configurations that enters into  $\Phi_{AS}$  (SD as a special case) where each electron position is appropriately changed depending on nearby electrons and nuclei. Backflow is an effective way to recover correlation energy, both at the variational and FN level. However, it can be used at a price to increase significantly an already large computational cost. Indeed, with backflow each time an electron is moved, all the entries in  $\Phi_{AS}$  (or several of them, if cut-offs are used) have to be changed, resulting in a much more expensive algorithm. Another possibility is to improve  $\Psi_{SD}$  similarly to conventional quantum chemistry approaches, namely by considering  $\Phi_{AS}$  as a linear expansion of several Slater determinants. While this second approach can provide very high accuracy, it may be extremely expensive, as the number of determinants necessary to remain with a pre-defined accuracy

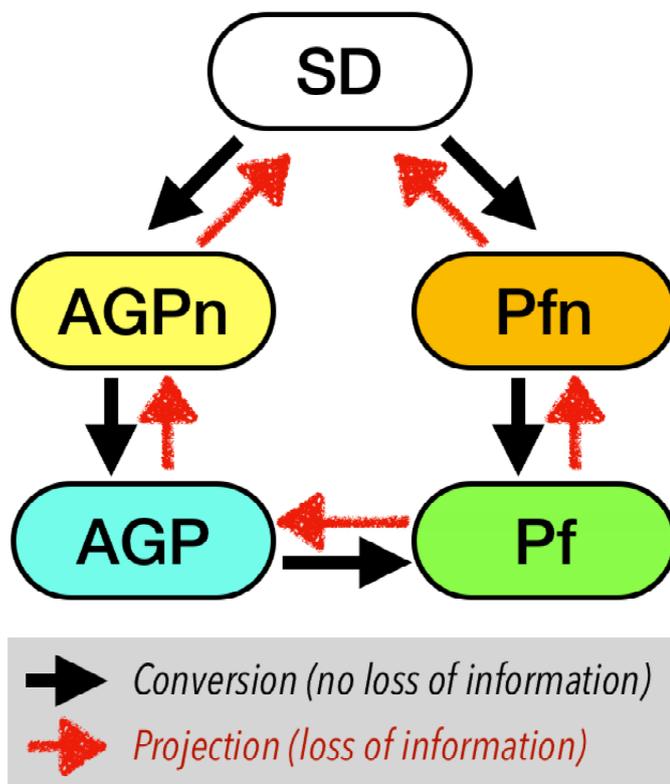


Figure 2.2: *Ansatz* conversion. From reference [29].

grows combinatorially with the system size.

The vision embraced in TURBORVB, that is also the inspiration of this work, is that the route toward an improved *ansatz* should not compromise the efficiency and good scaling of QMC. The main goal is instead to consider an *ansatz* that can be implicitly equivalent to a multi-determinant expansion but remains in practice as efficient as a single determinant.

In the following I will discuss five alternatives for the choice of  $\Phi_{AS}$ , which correspond to (i) the single Slater determinant, (ii) the Antisymmetrized Geminal Power (AGP), (iii) the Antisymmetrized Geminal Power with a constrained number of molecular orbitals (AGPn), (iv) the Pfaffian (Pf), (v) and the Pfaffian with constrained number of molecular orbitals (Pfn). In particular, the AGPn has not been used for the results presented in this thesis, but

it will be briefly introduced for the sake of completeness.

It is interesting to observe that all the other WFs can be obtained by introducing specific constraints on the most general Pf *ansatz*. The hierarchy of the five *ansatze* is represented in the Venn diagram of Fig. 2.1. Clearly, a more general *ansatz* is more accurate in the total energy but not necessarily in the energy differences. Moreover, it is described by more variational parameters, that could imply a more challenging optimization and a higher cost. TURBORVB includes several tools to go from one *ansatz* to another, as represented in Fig. 2.2. The ones concerning the Pf and the Pfn have been developed during the work for this thesis.

In this chapter I will first introduce the basis set used for the calculations, then I will deal with the different types of fermionic WFs: the SD, and, after an introduction to the pairing function, the AGP, the AGPn, the Pf, and the Pfn. In this section I will also discuss the possible conversions between the different *ansatze*, focusing on the operations necessary for the conversion of the Pf. After the fermionic part of the WF I will introduce the JF. I will finally present a technical advance strictly related to the JPf WF: a fast and efficient algorithm for the calculation of the spin squared expectation value.

## 2.1 Basis Set

In TURBORVB the electronic positions are described using localized basis sets. In particular, for the cases presented in this thesis, we expanded the *ansatze* in atom-centered basis sets of gaussian orbitals for the calculation of the JF, while often for the fermionic part of the WF a hybrid basis set has been employed in place of the simple uncontracted gaussian one. The CRYSTAL basis set has been used for systems with periodic boundary conditions (PBC).

The gaussian orbitals basis set is indicated as  $\{\phi_{I,\nu}(\mathbf{r})\}$ , with each element being the  $\nu$ -th orbital centred on the  $I$ -th atom at the position  $\mathbf{R}_I$ . The

elements in the basis set have the form

$$\phi_{I,\nu}(\mathbf{r}) \propto e^{-Z_\nu|\mathbf{r}-\mathbf{R}_I|^2}[Y_{l_\nu,m_\nu} \pm Y_{l_\nu,-m_\nu}], \quad (2.2)$$

where  $Z_\nu$  is a numerical coefficient that describes how diffuse the atomic orbital is around the atom, while  $Y_{l_\nu,m_\nu}$  is the spherical harmonic function with angular quantum numbers  $l_\nu$  and  $m_\nu$ . The sign of the combination  $[Y_{l_\nu,m_\nu} \pm Y_{l_\nu,-m_\nu}]$  is chosen to ensure the orbital type  $\nu$  to be real. This basis set has been used without further contractions for the description of the JF and the hydrogen atoms.

For all the other cases, for the fermionic part of our WF we have used hybrid atomic orbitals (HO)[24, 23] to expand the electron positions over a richer set of gaussian orbitals and, by means of the contractions, remaining with an affordable number of variational parameters. The HOs, indeed, are obtained as linear combinations of all the elements of the gaussian basis set used for a given atom, labeled by  $I$ :

$$\bar{\phi}_{I,\omega}(\mathbf{r}) = \sum_{\nu} \mu_{\omega,\nu} \phi_{I,\nu}(\mathbf{r}). \quad (2.3)$$

For the sake of compactness, we will indicate in the following all the basis elements as  $\{\phi_k(\mathbf{r})\}$  combining the indices  $\omega$  and  $I$ , and  $I$  and  $\nu$  in a single index  $k$  for a lighter notation. This basis set has been used for the description of the fermionic part of all the systems containing atoms with atomic number  $Z > 1$ .

The basis set presented above works without further modifications for all the systems in open boundary conditions (OBC), while for periodic systems modifications are necessary. The many-body WF of a system in PBC should satisfy the many-body Bloch condition [56, 57]:

$$\Psi_{k_s}(\mathbf{r}_1, \dots, \mathbf{r}_i + \mathbf{T}_s, \dots, \mathbf{r}_N) = e^{i\mathbf{k}_s \cdot \mathbf{T}_s} \Psi_{k_s}(\mathbf{r}_1, \dots, \mathbf{r}_i, \dots, \mathbf{r}_N), \quad (2.4)$$

which follows from the property that the many-body Hamiltonian is invariant under the translation of *any electron coordinate* by a simulation-cell vector  $\mathbf{T}_s$ ,

where  $\mathbf{T}_s = l\mathbf{a} + m\mathbf{b} + n\mathbf{c}$  is determined by arbitrary integers  $l, n, m$ , and the three vectors  $\mathbf{a}, \mathbf{b}$  and  $\mathbf{c}$  define the supercell.

The convention in TURBORVB is that a single-particle basis set satisfies the following conditions:

$$\phi_{I,\nu}^{\text{PBC}}(\mathbf{r} + \mathbf{T}_s; \zeta) = e^{-i\mathbf{k}_s \cdot \mathbf{T}_s} \phi_{I,\nu}^{\text{PBC}}(\mathbf{r}; \zeta) \quad (2.5)$$

where  $\mathbf{k}_s$  is a twist vector ( $\mathbf{k}_s = (k_s^x, k_s^y, k_s^z)$ ), and  $\mathbf{T}_s$  represents an arbitrary simulation cell vector. Notice that the use of a non-vanishing twist vector generally makes a many-body WF *complex*. To satisfy this condition TURBORVB implements the CRYSTAL periodic basis[58, 5, 59]:

$$\phi_{I,\nu}^{\text{PBC}}(\mathbf{r}; \zeta) = \sum_{\mathbf{T}_s} \phi_{I,\nu}(\mathbf{r} + \mathbf{T}_s; \zeta) e^{i\mathbf{k}_s \cdot \mathbf{T}_s} \quad (2.6)$$

where  $\phi_{I,\nu}$  is a non-periodic real atomic orbital, in our case a Gaussian orbital (Eq. 2.2). The use of Gaussian orbitals that rapidly decay far from nuclei guarantees that the above summation converges fast with a finite small number of  $\mathbf{T}_s$ . The same procedure is applied to the basis set for the Jastrow part, by using simple periodic boundary conditions[5], because the twists do not affect the Jastrow part of the WF, namely:

$$\chi_{l,m,I}^{\text{PBC}}(\mathbf{r}; \zeta) = \sum_{\mathbf{T}} \chi_{l,m,I}(\mathbf{r} + \mathbf{T}_s; \zeta), \quad (2.7)$$

which satisfies  $\chi_{l,m,I}^{\text{PBC}}(\mathbf{r} + \mathbf{T}_s; \zeta) = \chi_{l,m,I}^{\text{PBC}}(\mathbf{r}; \zeta)$ .

However, there are terms in the JF that are not defined in terms of the above periodic basis. These terms, that will be discussed in section 2.3 in detail, are the so-called one-body and two-body JF and have to be appropriately periodized. These terms are calculated using the electron-electron and electron-ion coordinate differences  $\mathbf{r}_d$ , expanded as:

$$\mathbf{r}_d = r_a\mathbf{a} + r_b\mathbf{b} + r_c\mathbf{c}, \quad (2.8)$$

where  $r_a, r_b$  and  $r_c$  are appropriate transformed coordinates, that are conveniently defined within a cube of unit length, because of the assumed periodicity of the supercell, namely  $|r_a|, |r_b|, |r_c| < 1/2$ . As a consequence, this mapping makes the physical electron-electron and electron-ion distance periodic by definition (*i.e.*, they refer to the minimum distance image of the supercell). However, there may be divergences or singularities at the boundaries of this unit cube. Therefore, before computing the distance corresponding to  $\mathbf{r}_d$ , these coordinates are transformed  $(r_a, r_b, r_c) \rightarrow (\bar{r}_a, \bar{r}_b, \bar{r}_c) = (p(r_a), p(r_b), p(r_c))$  by use of an appropriate function  $p(x)$ , with at least continuous first derivative for  $|x| < 1/2$ . This function is chosen to preserve the physical meaning at short distances, *i.e.*,  $p(x) = x$  in these cases, and being nonlinear elsewhere, in order to satisfy not only the periodicity but also the requirement of continuous first derivatives of the many-body WF  $\Psi_{k_s}$ . We have, therefore, defined  $p(x)$  as follows:

$$p(x) = \begin{cases} x & (-\frac{1}{4} < x < \frac{1}{4}) \\ -\frac{1}{8(1+2x)} & (-\frac{1}{2} \leq x \leq -\frac{1}{4}) \\ \frac{1}{8(1-2x)} & (\frac{1}{4} \leq x \leq \frac{1}{2}). \end{cases} \quad (2.9)$$

and, only for the case of the long range terms of the two-body and one-body terms of the JF, as

$$p(x) = \begin{cases} x & (-\frac{1}{6} < x < \frac{1}{6}) \\ -\frac{1}{54(1/2+x)^2} & (-\frac{1}{2} \leq x \leq -\frac{1}{6}) \\ \frac{1}{54(1/2-x)^2} & (\frac{1}{6} \leq x \leq \frac{1}{2}). \end{cases} \quad (2.10)$$

Though the modified relative distance diverges (*i.e.*,  $|\mathbf{r}_d| \rightarrow \infty$ ) at the edges of the Wigner-Seitz cell (*e.g.*,  $\mathbf{r} = \pm\frac{1}{2}\mathbf{a}, \pm\frac{1}{2}\mathbf{b}, \pm\frac{1}{2}\mathbf{c}$ ), the continuity and the periodicity of the whole JF and its derivatives is preserved.

Finally, we remark that the many-body WF also obeys the second Bloch condition [56, 57], namely:

$$\Psi_{k_p}(\{\mathbf{r}_i + \mathbf{T}_p\}) = e^{i\mathbf{k}_p \cdot \mathbf{T}_p} \Psi_{k_p}(\{\mathbf{r}_i\}), \quad (2.11)$$

where  $\mathbf{T}_p$  represents a *unit-cell* (not supercell) vector, and  $\mathbf{k}_p$  is the crystal momentum. This comes from the property that the many-body Hamiltonian is invariant under the *simultaneous translation* of all-electron coordinates by a unit-cell vector  $\mathbf{T}_p$ . Within TURBORVB, this condition can be employed by imposing the intra-unit cell translational symmetries on the Jastrow and the pairing function, as simple linear constraints in the variational parameters. However, this option is restricted to the case  $\mathbf{k}_p = 0$ .

It is possible to use different twists on each spin component, that has proven very effective for implementing the mentioned translation symmetries within pairing WFs[60].

## 2.2 Fermionic Wave Functions

The most relevant component of the WF is the fermionic element  $\Phi_{AS}(\mathbf{x})$ . It provides the spatial structure of the orbitals and contains the information about the sign and the nodal surface of the WF. In this section I will present the possible choices considered in this thesis: the SD and the *ansatze* based on the pairing function.

### 2.2.1 Slater Determinant

In the following we will describe the SD. This WF is particularly important for several reasons. It is the standard choice for the fermionic part of the WF  $\Phi_{AS}(\mathbf{x})$  and thus it represents not only a reasonably accurate *ansatz* but also a fundamental benchmark for other *ansatze*. Furthermore, it is necessary for the initialization of the pairing function.

From a theoretical and computational point of view, the SD is the simplest fermionic WF. It is built from the vacuum state by populating a number of orthogonal single-particle molecular orbitals (MO) equal to the number of electrons in the system. Henceforth, we omit the spin indices, by assuming

that to each spin component corresponds a different Slater determinant. In our basis the MOs are in the form

$$\Phi_{\alpha}^{mol}(\mathbf{r}) = \sum_{k=1} P_{\alpha,k} \phi_k(\mathbf{r}). \quad (2.12)$$

The MOs can be obtained directly from a DFT or HF calculation, but they can also be further optimized with VMC[61]. It is well known that the antisymmetric product of these MO leads to the determinant of the matrix in which every molecular orbital is evaluated for each electron position:

$$\Phi_{SD}(\mathbf{X}) = \det \begin{pmatrix} \Phi_1^{mol}(\mathbf{r}_1) & \Phi_1^{mol}(\mathbf{r}_2) & \cdots & \Phi_1^{mol}(\mathbf{r}_N) \\ \Phi_2^{mol}(\mathbf{r}_1) & \Phi_2^{mol}(\mathbf{r}_2) & \cdots & \Phi_2^{mol}(\mathbf{r}_N) \\ \vdots & \vdots & \ddots & \vdots \\ \Phi_N^{mol}(\mathbf{r}_1) & \Phi_N^{mol}(\mathbf{r}_2) & \cdots & \Phi_N^{mol}(\mathbf{r}_N) \end{pmatrix}. \quad (2.13)$$

For weakly correlated systems the JSD can often give reasonably good results with a reasonable computational cost and a limited number of variational parameters. It is also a common choice to use a linear combinations of SDs to improve the description of the WF, with *ansatze* that take different names depending on the type and number of SDs considered. In this thesis we will compare directly the results of our WFs to the ones obtained with some of the most successful multi-determinant WFs, the full valence complete active space (FVCAS) WF, the full configuration interaction (FCI), and the heat-bath configuration interaction (HCI).

### 2.2.2 Pairing Function

The use of the pairing function in correlated WFs allows an electronic description that goes beyond the single-particle picture of the SD, including also the pair correlation that was missing in the previous case. The building block of this WFs is named pairing (or geminal) function and has the following general form

$$g(\mathbf{r}_1\sigma_1, \mathbf{r}_2\sigma_2) = \sum_{k,l} \lambda_{k,l}^{\sigma_1\sigma_2} \phi_{k\sigma_1}(\mathbf{r}_1) \phi_{l\sigma_2}(\mathbf{r}_2), \quad (2.14)$$

where all the elements of the matrix  $\lambda$  represent most of the WF variational parameters. They depend on the orbitals considered and on the spin  $\sigma_1, \sigma_2$  of the so-called geminal function  $g$ . In principle when we break the spin symmetry the basis sets used for  $\uparrow$  and  $\downarrow$  electrons can be different, otherwise the basis chosen does not depend on the spin component. In order to set up a consistent many-body WF starting from the geminal several choices are possible depending on the criteria adopted for the definition of the geminal. To highlight the different possibilities we can recast Eq. (2.14) in a way in which the spin dependency is more explicit

$$\begin{aligned}
g(\mathbf{r}_1\sigma_1, \mathbf{r}_2\sigma_2) &= \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle)g_+(\mathbf{r}_1, \mathbf{r}_2) \\
&+ \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle)g_-(\mathbf{r}_1, \mathbf{r}_2) \\
&+ |\uparrow\uparrow\rangle g_\uparrow(\mathbf{r}_1, \mathbf{r}_2) + |\downarrow\downarrow\rangle g_\downarrow(\mathbf{r}_1, \mathbf{r}_2).
\end{aligned} \tag{2.15}$$

where

$$\begin{aligned}
g_\pm(\mathbf{r}_1, \mathbf{r}_2) &= g(\mathbf{r}_1 \uparrow, \mathbf{r}_2 \downarrow) \pm g(\mathbf{r}_1 \downarrow, \mathbf{r}_2 \uparrow), \\
g_\sigma(\mathbf{r}_1, \mathbf{r}_2) &= g(\mathbf{r}_1\sigma, \mathbf{r}_2\sigma) \text{ with } \sigma = \uparrow, \downarrow.
\end{aligned} \tag{2.16}$$

In order to satisfy the Pauli principle we have  $g_\pm(\mathbf{r}_1, \mathbf{r}_2) = \pm g_\pm(\mathbf{r}_2, \mathbf{r}_1)$  and  $g_\sigma(\mathbf{r}_1, \mathbf{r}_2) = -g_\sigma(\mathbf{r}_2, \mathbf{r}_1)$  for  $\sigma = \uparrow, \downarrow$ . Our WF is then obtained by antisymmetrizing the product over all the electron pairs considered that, by definition, occupy the *same* pairing function. For simplicity we will enumerate the spin up electrons from 1 to  $N_\uparrow$  and the spin down ones from  $N_\uparrow + 1$  to  $N$ .

As suggested by the name Antisymmetrized Geminal Power, our goal is to define a WF that is literally the antisymmetrized product of the geminals and the unpaired orbitals (if present), namely

$$\begin{aligned}
\Phi(\mathbf{X}) &= \sum_{\alpha} \text{Sgn}(\alpha) (g(\mathbf{r}_{1_\alpha}\sigma_{1_\alpha}, \mathbf{r}_{2_\alpha}\sigma_{2_\alpha})g(\mathbf{r}_{3_\alpha}\sigma_{3_\alpha}, \mathbf{r}_{4_\alpha}\sigma_{4_\alpha}) \cdots \\
&\quad g(\mathbf{r}_{p-1_\alpha}\sigma_{p-1_\alpha}, \mathbf{r}_{p_\alpha}\sigma_{p_\alpha})\Theta_1(\mathbf{r}_{p+1_\alpha}) \cdots \Theta_{N-p}(\mathbf{r}_{N_\alpha})),
\end{aligned} \tag{2.17}$$

where  $\alpha$  is one of the possible ways of distributing the  $N$  electrons between the  $p/2$  pairs and the  $N - p$  unpaired orbitals  $\Theta$  and  $\text{Sgn}(\alpha)$  is the sign of the corresponding permutation of the particles that is required to ensure the fermionic behavior. In particular, different choices of the pairing function, obtained by excluding one or more terms in the Eq. (2.16), lead to different ways to compute Eq. (2.17). These choices also impact quantitatively and qualitatively on the kind of physics that we can describe by means of this type of WF. Therefore, we will distinguish in the following among the possible distinct cases. If we consider only the singlet term in Eq. (2.16) we obtain the AGP. An extension of the AGP is the broken symmetry AGP (AGPu) where we break the spin symmetry including the  $S_z = 0$  triplet term, without considering the parallel spin triplet terms. The most general case, where we include all the triplet and the singlet terms, defines the Pf WF. In addition to these three cases, there is the possibility to define the AGP and the Pf WFs with a constrained number of MOs, respectively the AGPn and the Pfn.

### 2.2.3 AGP

Let me consider for the moment the unpolarized case  $N_\uparrow = N_\downarrow$ , the extension to the polarized cases will be straightforward and will be discussed later on. When no triplet correlations are allowed we build the WFs using only singlet pairs and the pairing function in Eq. (2.15) contains only the symmetric element  $g_+$

$$g(\mathbf{r}_1\sigma_1, \mathbf{r}_2\sigma_2) = \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle)g_+(\mathbf{r}_1, \mathbf{r}_2). \quad (2.18)$$

In this case we project a perfect singlet that we denote as AGP. The  $\lambda$  matrix elements in Eq. (2.18) are non zero only for  $\sigma_1 \neq \sigma_2$  and they are symmetric for spin exchange. In order to calculate the AGP we can write all the possible

combinations of pairs of opposite spin electrons in a matrix defined as

$$G = \begin{pmatrix} g(\mathbf{r}_1 \uparrow, \mathbf{r}_{N_\uparrow+1} \downarrow) & g(\mathbf{r}_1 \uparrow, \mathbf{r}_{N_\uparrow+2} \downarrow) & \cdots & g(\mathbf{r}_1 \uparrow, \mathbf{r}_N \downarrow) \\ g(\mathbf{r}_2 \uparrow, \mathbf{r}_{N_\uparrow+1} \downarrow) & g(\mathbf{r}_2 \uparrow, \mathbf{r}_{N_\uparrow+2} \downarrow) & \cdots & g(\mathbf{r}_2 \uparrow, \mathbf{r}_N \downarrow) \\ \vdots & \vdots & \ddots & \vdots \\ g(\mathbf{r}_{N_\uparrow} \uparrow, \mathbf{r}_{N_\uparrow+1} \downarrow) & g(\mathbf{r}_{N_\uparrow} \uparrow, \mathbf{r}_{N_\uparrow+2} \downarrow) & \cdots & g(\mathbf{r}_{N_\uparrow} \uparrow, \mathbf{r}_N \downarrow) \end{pmatrix}. \quad (2.19)$$

In this way to each row of the matrix corresponds an electron of spin  $\uparrow$ , and to each column an electron of spin  $\downarrow$ . The definition of the matrix  $G$  in this form is convenient because it allows the antisymmetrization requested by the Eq. (2.17) in a simple and efficient way. Indeed, it can be demonstrated[23] that the correct antisymmetrization of the pairs considered in this case is given by

$$\Phi_{AGP}(\mathbf{X}) = \det\{G\}. \quad (2.20)$$

This is somehow intuitive because we want to sum all the possible products of  $N/2$  matrix elements of  $G$ , where in all these factors a column element or a row element is present only once, exhausting all the possible configurations of the system considered with an appropriate  $\pm$  sign that, in this case, is just given by the one corresponding to the determinant of  $G$ .

When the system is polarized and  $N_\uparrow \neq N_\downarrow$ , we cannot build the solution using only the singlet terms, because the matrix  $G$  written as in Eq. (2.19) is a rectangular matrix and its determinant cannot be computed. Supposing for simplicity that  $N_\uparrow > N_\downarrow$ , in this case we have to add a number  $N_\uparrow - N_\downarrow$  of unpaired spin-up MOs  $\{\Theta_i(\mathbf{r})\}$  not only for fulfilling the polarization required but, most importantly, to turn the matrix  $G$  to a perfectly defined square matrix:

$$G = \begin{pmatrix} g(\mathbf{r}_1 \uparrow, \mathbf{r}_{N_\uparrow+1} \downarrow) & \cdots & \Theta_1(\mathbf{r}_1) & \cdots & \Theta_{N_\uparrow-N_\downarrow}(\mathbf{r}_1) \\ g(\mathbf{r}_2 \uparrow, \mathbf{r}_{N_\uparrow+1} \downarrow) & \cdots & \Theta_1(\mathbf{r}_2) & \cdots & \Theta_{N_\uparrow-N_\downarrow}(\mathbf{r}_2) \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ g(\mathbf{r}_{N_\uparrow} \uparrow, \mathbf{r}_{N_\uparrow+1} \downarrow) & \cdots & \Theta_1(\mathbf{r}_{N_\uparrow}) & \cdots & \Theta_{N_\uparrow-N_\downarrow}(\mathbf{r}_{N_\uparrow}). \end{pmatrix} \quad (2.21)$$

Also in this case a consistent antisymmetric wave function can be again

calculated as the determinant[23] of the matrix  $G$  exactly in the same way of the singlet pairing in Eq. (2.20).

### AGPu

The AGPu is an extension of the AGP where also the  $S_z = 0$  triplet component is included and only the parallel spin terms of the triplet are omitted. This means that the spin symmetry is broken and a magnetic order parameter can be directed only along the  $z$ -quantization axis. This WF is called broken symmetry AGP (AGPu) and the difference from the previous AGP is the presence of the antisymmetric  $g_-$  component in the definition of the pairing function in Eq. (2.15), that for this case is

$$\begin{aligned}
g(\mathbf{r}_1\sigma_1, \mathbf{r}_2\sigma_2) &= \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle)g_+(\mathbf{r}_1, \mathbf{r}_2) \\
&+ \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle)g_-(\mathbf{r}_1, \mathbf{r}_2). \quad (2.22)
\end{aligned}$$

In order to define this pairing function, we break the spin symmetry in the opposite electron spin case with  $\sigma_1 \neq \sigma_2$ , by keeping equal to zero the  $\sigma_1 = \sigma_2$  components of Eq. (2.14). With exactly the same procedure used in the case of the AGP, depending on the polarization, we can build the same matrix  $G$  of Eq. (2.19) or Eq. (2.21), that now is no longer symmetric. Even in this case the correct antisymmetrized sum of these pairs is given by the determinant[23]. Thus, analogously to Eq.(2.20) we obtain

$$\Phi_{AGPu}(\mathbf{X}) = \det\{G\}, \quad (2.23)$$

that implements the simplest broken symmetry *ansatz* based on the pairing function.

### 2.2.4 Pfaffian Wave Function

The Pfaffian WF is the most important among the pairing functions *ansatze* presented so far, being the most general one and encoding new variational

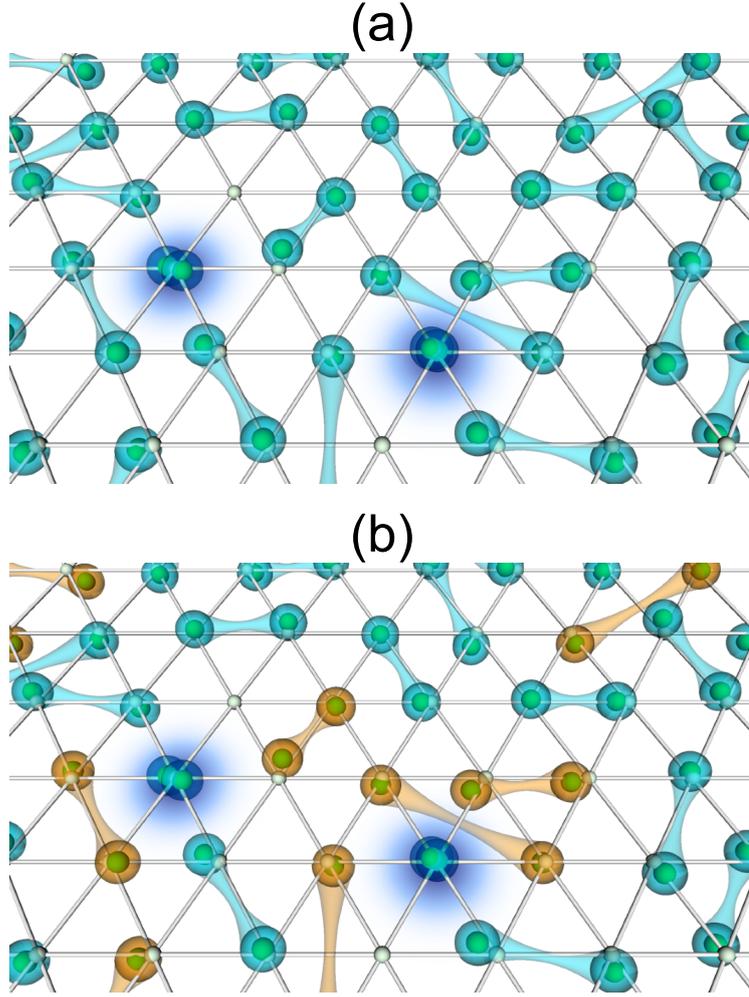


Figure 2.3: Cartoon picture of a typical valence bond generated by the AGP WFs (on the upper panel) and the Pf WF (on the lower panel). The balls indicate the electrons, singlet and triplet bonds are displayed with cyan and orange colors, respectively. From reference [33].

freedoms into the AGP and the AGPu. In next chapters it will become clear that it represents the most powerful description of the chemical bond within the paradigm developed in this work. This WF represents also the most general mean-field state, namely the GS of a mean-field Hamiltonian containing also BCS anomalous terms projected on a given number  $N$  of particles and total spin projection  $S_{tot}^z = \sum_{i=1} \sigma_i$  along the  $z$ -quantization axis. In this case the

definition of the pairing function is exactly the one in Eq. (2.15), containing all the terms including the parallel spin terms of the triplet as graphically pictured in Fig. 2.3. This means that now, when we build the Pf, we have to include in the WF also the parallel spin electron pairs. In this way the Pf can also describe a magnetic order parameter in any direction of the space, and thus it is also possible to rotate the spin component of the WF in any direction. This will allow us to break the symmetry along the spin quantization axis and then to rotate it. As we will explain later, this plays a crucial role when we use this WF in combination with our JF, since it allows us to preserve the total  $S_z$  of the molecules and include spin fluctuations.

Of course, we cannot create a WF using only pairs if the number of electrons in the system is odd, so, for the moment, let us assume  $N$  is even. The extension to the odd number of electrons is trivial and will be discussed immediately after. We will dub as  $W$  the matrix containing all the possible pairs

$$W = \begin{pmatrix} 0 & g(\mathbf{r}_1 \uparrow, \mathbf{r}_2 \uparrow) & \cdots & g(\mathbf{r}_1 \uparrow, \mathbf{r}_N \downarrow) \\ g(\mathbf{r}_2 \uparrow, \mathbf{r}_1 \uparrow) & 0 & \cdots & g(\mathbf{r}_2 \uparrow, \mathbf{r}_N \downarrow) \\ \vdots & \vdots & \ddots & \vdots \\ g(\mathbf{r}_{N-1} \downarrow, \mathbf{r}_1 \uparrow) & g(\mathbf{r}_{N-1} \downarrow, \mathbf{r}_2 \uparrow) & \cdots & g(\mathbf{r}_{N-1} \downarrow, \mathbf{r}_N \downarrow) \\ g(\mathbf{r}_N \downarrow, \mathbf{r}_1 \uparrow) & g(\mathbf{r}_N \downarrow, \mathbf{r}_2 \uparrow) & \cdots & 0 \end{pmatrix}, \quad (2.24)$$

where the matrix is antisymmetric for the fermionic commutation rules and thus the elements of the diagonal are set to zero. We can recast the  $W$  highlighting its different spin sectors as

$$W = \begin{pmatrix} W_{\uparrow\uparrow} & W_{\uparrow\downarrow} \\ W_{\downarrow\uparrow} & W_{\downarrow\downarrow} \end{pmatrix} \quad (2.25)$$

where  $W_{\uparrow\uparrow}$  and  $W_{\downarrow\downarrow}$  are respectively a  $N_{\uparrow} \times N_{\uparrow}$  and a  $N_{\downarrow} \times N_{\downarrow}$  antisymmetric matrices that take into account the parallel spin terms of the triplet, while  $W_{\uparrow\downarrow}$  is a  $N_{\uparrow} \times N_{\downarrow}$  matrix such that  $W_{\uparrow\downarrow} = -W_{\downarrow\uparrow}^T$  describing the remaining triplet and singlet contributions. In the case of AGP and AGPu we can also build a similar matrix where the matrices  $W_{\uparrow\uparrow}$  and  $W_{\downarrow\downarrow}$  are identically zero.

Analogously to the case of the AGP and AGPu, we have to identify a way to calculate the antisymmetric product of all the pairs considered. In this case we can identify the antisymmetrization procedure defined in Eq. (2.17) as the Pfaffian of the matrix  $W$ . After introducing a few algebraic definitions, the reason will be clear to the reader.

The Pfaffian is an algebraic operation acting on antisymmetric square matrices with an even number of rows and columns. Being  $N$  even, the matrix  $W$  satisfies these hypotheses. The usual definition of the Pfaffian, requires the introduction of the concept of partition of the matrix  $W$

$$A(\alpha) = \text{sign}(\alpha) \prod_{k=1}^{N/2} W_{i_k, j_k} \quad (2.26)$$

where all  $i_k$  and  $j_k$  are different,  $i_k < j_k$  for each  $k$  and  $i_1 < i_2 < \dots < i_N$ . The  $\text{sign}(\alpha)$  is given by the permutation that orders the vector of the indices  $\{i_1, j_1, i_2, j_2, \dots, i_M, j_M\}$ . In this way all the indices are considered only once. The Pfaffian is then defined as

$$\text{Pf}(W) = \sum_{\alpha} A(\alpha) \quad (2.27)$$

where the sum over  $\alpha$  is extended over all the possible partitions. However, an alternative definition[62] of the Pfaffian can better clarify the correspondence to the Eq. (2.17). It can indeed be defined alternatively as

$$\text{Pf}(W) = [(N/2)!2^{N/2}]^{-1} \sum_P \text{sign}(P) \prod_{k_P=1}^{N/2} W_{i_{k_P}, j_{k_P}} \quad (2.28)$$

where  $P$  now represents a generic permutation of the possible row and column indices of the matrix without any constraints and the  $\text{sign}(P)$  is the parity of the permutation. In this definition it is easy to recognize the antisymmetrized sum corresponding to the Eq. (2.17). Let us introduce now a further property of the Pfaffian that will be useful in the following. In the following we will

indicate with  $0$  a  $m \times m$  matrix containing only  $0$  and  $B$  a generic  $m \times m$  matrix, we have that

$$\text{Pf} \begin{bmatrix} 0 & B \\ -B^T & 0 \end{bmatrix} = (-1)^{m(m-1)/2} \det(B). \quad (2.29)$$

For an odd number of electrons it is necessary to use a spin-dependent unpaired orbital  $\Theta_\sigma(\mathbf{r})$  so that we can accommodate the remaining electron that is not considered by the product of the pairs. The unpaired orbital introduces a supplementary row and column to the matrix  $W$ . Being  $\Theta_\uparrow = (\Theta_\uparrow(\mathbf{r}_1), \Theta_\uparrow(\mathbf{r}_2), \dots, \Theta_\uparrow(\mathbf{r}_{N_\uparrow}))$  the vector containing the values of the unpaired orbital  $\Theta_\uparrow$  at the  $\uparrow$  electron positions and  $\Theta_\downarrow = (\Theta_\downarrow(\mathbf{r}_{N_\uparrow+1}), \Theta_\downarrow(\mathbf{r}_{N_\uparrow+2}), \dots, \Theta_\downarrow(\mathbf{r}_N))$  the one calculated for the  $\downarrow$  electron ones, we modify the matrix in Eq. (2.25) as

$$W = \begin{pmatrix} W_{\uparrow\uparrow} & W_{\uparrow\downarrow} & \Theta_\uparrow \\ W_{\downarrow\uparrow} & W_{\downarrow\downarrow} & \Theta_\downarrow \\ -\Theta_\uparrow^T & -\Theta_\downarrow^T & 0 \end{pmatrix}. \quad (2.30)$$

Also in this case the permutation sum implied by the Pfaffian leads to the correct antisymmetrization required from Eq. (2.17). The matrix  $W$  satisfies the hypothesis of the calculation having an even leading matrix dimension  $\bar{N} = N + 1$ . We can further notice that no assumption has been made on the polarization of the system and so no unpaired orbital is required except for a single one in case of odd  $N$ .

It is however possible in principle to introduce further pairs of unpaired orbitals, if, for example, we want to describe an AGP or AGPu with a full Pf WF. We define  $\Theta_{i\sigma}(\mathbf{r})$  as the set of the considered  $m$  unpaired orbitals and  $\Theta_{i\uparrow} = (\Theta_{i,\uparrow}(\mathbf{r}_1), \Theta_{i,\uparrow}(\mathbf{r}_2), \dots, \Theta_{i,\uparrow}(\mathbf{r}_{N_\uparrow}))$  the vector containing the values of the unpaired orbital  $\Theta_{i,\uparrow}$  for the  $\uparrow$  electron positions and  $\Theta_{i\downarrow} = (\Theta_{i,\downarrow}(\mathbf{r}_{N_\uparrow+1}), \Theta_{i,\downarrow}(\mathbf{r}_{N_\uparrow+2}), \dots, \Theta_{i,\downarrow}(\mathbf{r}_N))$  the one calculated for the  $\downarrow$  electron ones. We can modify the matrix in

Eq. (2.25) as

$$W = \begin{pmatrix} W_{\uparrow\uparrow} & W_{\uparrow\downarrow} & \Theta_{1\uparrow} & \cdots & \Theta_{m\uparrow} \\ W_{\downarrow\uparrow} & W_{\downarrow\downarrow} & \Theta_{1\downarrow} & \cdots & \Theta_{m\downarrow} \\ -\Theta_{1\uparrow}^T & -\Theta_{1\downarrow}^T & 0 & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -\Theta_{m\uparrow}^T & -\Theta_{m\downarrow}^T & 0 & \cdots & 0 \end{pmatrix}, \quad (2.31)$$

that is a  $\bar{N} \times \bar{N}$  matrix where  $\bar{N} = N + m$ . We can again antisymmetrize this product using the definition of the Pfaffian provided in Eq. (2.28). A careful reader could have noticed that, by applying the Pfaffian definition, we are antisymmetrizing not only over the electron indices but also over the orbital indices of the unpaired orbitals. This antisymmetrization, however, contains the one over the physical electrons and leads therefore to a physically allowed electronic wave function.

Moreover, we can notice that, by using the previous definition, we can identify the AGP and the AGPu as sub-cases of the general Pf. Indeed, by using the expressions of the pairing function and the unpaired orbitals of the AGP and AGPu we obtain  $W_{\uparrow\uparrow} = 0$ ,  $W_{\downarrow\downarrow} = 0$ ,  $\Theta_{i\downarrow} = 0$  and  $\bar{N} = 2N_{\uparrow}$ . By merging Eq. (2.21) and Eq. (2.25) we can define

$$W = \begin{pmatrix} 0 & G \\ -G^T & 0 \end{pmatrix}, \quad (2.32)$$

and this means that applying Eq. (2.29) we immediately obtain

$$\text{Pf}(W) = \pm \det(G), \quad (2.33)$$

where the sign only depends on the number of electrons and is constant, thus irrelevant. This shows in a straightforward way that the AGP and AGPu defined in the previous subsection are nothing but particular cases of the most general Pf.

## 2.2.5 AGP and Pf with a constrained number of molecular orbital (AGPn and Pfn)

A convenient way to impose constraints on the variational parameters defining the AGP or Pf WF is obtained by rewriting the expansion of the geminal in terms of molecular orbitals (MOs). As shown in Eq. (2.14), a geminal  $g(\mathbf{r}_1\sigma_1, \mathbf{r}_2\sigma_2)$  there are no restrictions on the nature of the orbitals  $\{\phi_{i\sigma}(\mathbf{r})\}$ . The sum can be extended over a set of atomic orbitals as well as over a set of orthogonal MOs. This last choice can be convenient to control the number of variational parameters while maintaining some of the AGP and Pf physical properties. Let us start considering Eq. (2.14) in terms of atomic orbitals. In order to simplify the notation here, let us merge the atomic orbital and spin indices in a unique one that is indicated with a greek symbol (*e.g.*,  $\mu \leftrightarrow (i, \sigma)$ ) running from 1 to the total dimension  $2L$  of the atomic orbitals used,  $L$  for each spin component. Using this shortened notation we rewrite Eq. (2.14) as

$$g(\mathbf{i}, \mathbf{j}) = g(\mathbf{r}_i\sigma_i, \mathbf{r}_j\sigma_j) = \sum_{\mu, \nu} \lambda_{\mu, \nu} \phi_{\mu}(\mathbf{i}) \phi_{\nu}(\mathbf{j}) \quad (2.34)$$

where clearly the symmetry of  $g$  implies that  $\lambda_{\mu, \nu} = -\lambda_{\nu, \mu}$ . The coefficients  $\lambda_{\mu, \nu}$  define a  $2L \times 2L$  skew-symmetric matrix  $\lambda$ . If we define the  $2L$  dimensional vector  $\Phi_i = (\phi_1(\mathbf{i}), \dots, \phi_{2L}(\mathbf{i}))^T$ , Eq. (2.34) rewrites as  $g(\mathbf{i}, \mathbf{j}) = \Phi_i^T \lambda \Phi_j$ . Moreover, the overlaps  $S_{\mu, \nu} \equiv \langle \phi_{\mu} | \phi_{\nu} \rangle$  between atomic orbitals define the overlap matrix  $S$ , that in the case of a spin-dependent basis is block diagonal:

$$S = \left[ \begin{array}{c|c} S_{uu} & 0 \\ \hline 0 & S_{dd} \end{array} \right] \quad (2.35)$$

with  $S_{uu}$  and  $S_{dd}$  positive definite  $L \times L$  square matrices ( $S_{uu} = S_{dd}$  when orbitals are the same for spin up and spin down). In TURBORVB, the overlap matrix  $S$  is computed on a suitable uniform mesh with an efficient and general parallel algorithm. Then an orthonormal basis is defined:

$$\tilde{\phi}_{\mu}(\mathbf{i}) = \sum_{\nu} S_{\mu, \nu}^{-1/2} \phi_{\nu}(\mathbf{i}) \quad (2.36)$$

where  $S^{-1/2}$  is well defined since  $S$  is strictly positive definite.<sup>1</sup> The matrix elements of  $g$  can be recasted in the new orthonormal basis  $\{\tilde{\phi}_\mu(\mathbf{i})\}$ , yielding  $g(\mathbf{i}, \mathbf{j}) = \sum_{\mu, \nu}^{2L} \tilde{\lambda}_{\mu, \nu} \tilde{\phi}_\mu(\mathbf{i}) \tilde{\phi}_\nu(\mathbf{j}) = \tilde{\Phi}_i^T \tilde{A} \tilde{\Phi}_j$ , with the matrix  $\tilde{\lambda} \equiv S^{1/2} \lambda S^{1/2}$  that is antisymmetric.

At this point, from the spectral theory of skew-symmetric matrix, as it will be shown in the section 2.2.6, it is possible to perform the Youla decomposition of  $\tilde{\lambda}$ [28], which can be written in the form  $\tilde{\lambda} = Q\Sigma Q^T$ , where  $Q$  is unitary (also real if  $\tilde{\lambda}$  is real), and the matrix  $\Sigma$  is block diagonal with  $\Sigma_{2k-1, 2k} = a_k = -\Sigma_{2k, 2k-1}$  for  $k = 1, \dots, L$ , and zero everywhere else, with  $a_k \geq 0$ . So, the pairing function  $g(\mathbf{i}, \mathbf{j})$  can be written as  $\Psi_i^T \Sigma \Psi_j$ , where  $\Psi_i = Q^T \tilde{\Phi}_i$  for each  $i$ . This defines a basis of  $L$  MOs  $\{\psi_k(i)\}$  and corresponding twinned ones  $\{\bar{\psi}_k(i)\}$ , forming together a basis of  $2L$  mutually orthonormal elements for which the original geminal function reads:

$$g(\mathbf{i}, \mathbf{j}) = \sum_{k=1}^L a_k [\psi_k(\mathbf{i}) \bar{\psi}_k(\mathbf{j}) - \bar{\psi}_k(\mathbf{i}) \psi_k(\mathbf{j})] \quad (2.37)$$

with  $a_k \geq 0$ . After these transformations, the MOs can be finally written in the chosen (hybrid) atomic basis:

$$\begin{aligned} \psi_k(\mathbf{i}) &= \sum_{\nu=1}^{2L} P_{\nu, k} \phi_\nu(\mathbf{i}) \\ \bar{\psi}_k(\mathbf{i}) &= \sum_{\nu=1}^{2L} \bar{P}_{\nu, k} \phi_\nu(\mathbf{i}) \end{aligned} \quad (2.38)$$

by appropriate  $p \times 2L$  rectangular matrices  $P$  and  $\bar{P}$ . Then, with no loss of generality, we can assume that the molecular orbitals  $\{\psi_k(\mathbf{i}), \bar{\psi}_k(\mathbf{i})\}$  are ranked such that  $a_1 \geq a_2 \geq \dots \geq a_L \geq 0$ . The above expression highlights

---

<sup>1</sup> $S^{-1/2}$  can be computed after a standard block diagonalization of the matrix  $S = UDU^\dagger$ , being  $U$  a unitary matrix and  $D = \text{diag}\{d_1, \dots, d_{2L}\}$  is a diagonal matrix, such that  $S^{-1/2} = UD^{-1/2}U^\dagger$ , where  $D^{-1/2}$  is the diagonal matrix obtained by taking the inverse square root of each diagonal element  $d_i$  of  $D$ . At this stage, we carefully remove from the basis the elements corresponding to the smallest eigenvalues  $d_i$  in order to work with a sufficiently large condition number that guarantees stable finite precision numerical calculations.

that the most important MOs are those corresponding to the larger values of  $a_k$ . Therefore, it is possible to constrain the variational freedom by neglecting all the orbitals with  $k > n$ , yielding the pairing function:

$$g_n(\mathbf{i}, \mathbf{j}) = \sum_{k=1}^n a_k [\psi_k(\mathbf{i})\bar{\psi}_k(\mathbf{j}) - \bar{\psi}_k(\mathbf{i})\psi_k(\mathbf{j})] , \quad (2.39)$$

where  $n$  is conveniently chosen and is  $\ll L$ . This yields the AGPn *ansatz* and the  $\Phi_{\text{AGPn}}$  WF, which can be useful to improve the stability of the wavefunction optimization. The corresponding algorithm, based on projection operators in the space of the  $n$  molecular orbitals considered, has been described extensively in Ref.[5]. Moreover, in the original paper[61] introducing the AGPn, a precise recipe was given to improve the evaluation of the binding energies. Indeed, despite a constraint on the variational parameters necessarily increases the variational energy expectation value, energy differences may actually improve by an appropriate choice of  $n$ . In the mentioned work[61], this promising approach was applied with an AGP containing only singlet correlations, but the binding energies were defined without using a rigorous size consistent criterium. This drawback can be now removed, by exploiting the full variational freedom of the Pf WF combined with a general spin-dependent Jastrow factor.

The variational optimization of an AGP with a fixed number  $n$  of molecular orbitals can be easily generalized to the Pf case, by exploiting that the constrained Pf WF, dubbed Pfn, can be written either in the canonical form with MOs as in Eq. (2.39) or in the localized basis set expansion, as in Eq. (2.34), with a corresponding matrix  $\lambda_{\mu,\nu}^n = \sum_{k=1}^n a_k [P_{\mu,k}\bar{P}_{\nu,k} - P_{\nu,k}\bar{P}_{\mu,k}]$ .

According to Eq. (2.39) a small but arbitrary variation  $\delta g_n$  of the con-

strained pairing function  $g_n$  reads:

$$\begin{aligned}
\delta g_n(\mathbf{i}, \mathbf{j}) &= \sum_{k=1}^n \delta a_k [\psi_k(\mathbf{i})\bar{\psi}_k(\mathbf{j}) - \bar{\psi}_k(\mathbf{i})\psi_k(\mathbf{j})] \\
&+ \sum_{k=1}^n a_k [\delta\psi_k(\mathbf{i})\bar{\psi}_k(\mathbf{j}) - \delta\bar{\psi}_k(\mathbf{i})\psi_k(\mathbf{j})] \\
&+ \sum_{k=1}^n a_k [\psi_k(\mathbf{i})\delta\bar{\psi}_k(\mathbf{j}) - \bar{\psi}_k(\mathbf{i})\delta\psi_k(\mathbf{j})] \quad (2.40)
\end{aligned}$$

and therefore satisfies the following property, as it will be shown later:

$$(\hat{I} - \mathbf{L})\delta g_n(\hat{I} - \mathbf{R}) = 0 \quad (2.41)$$

where  $\hat{I}$  is the identity operator,  $\mathbf{L}$  and  $\mathbf{R}$  are projection operators over the occupied MOs, *i.e.*,  $\mathbf{L}^2(\mathbf{i}, \mathbf{j}) = \int d\mathbf{k} \mathbf{L}(\mathbf{i}, \mathbf{k})\mathbf{L}(\mathbf{k}, \mathbf{j}) = \mathbf{L}(\mathbf{i}, \mathbf{j})$ , and similarly  $\mathbf{R}^2 = \mathbf{R}$ , where here and henceforth the shorthand integration symbol  $\int d\mathbf{k} = \sum_{\sigma_k} \int d\mathbf{r}_k$  contains implicitly also the spin summation. These operators are then defined as follows:

$$\begin{aligned}
\mathbf{L}(\mathbf{i}, \mathbf{j}) &= \sum_{k=1}^n [\psi_k(\mathbf{i})\psi_k^*(\mathbf{j}) + \bar{\psi}_k(\mathbf{i})\bar{\psi}_k^*(\mathbf{j})] \\
\mathbf{R}(\mathbf{i}, \mathbf{j}) &= \sum_{k=1}^n [\psi_k^*(\mathbf{i})\psi_k(\mathbf{j}) + \bar{\psi}_k^*(\mathbf{i})\bar{\psi}_k(\mathbf{j})]. \quad (2.42)
\end{aligned}$$

With the above definitions, Eq. (2.40) is easily verified because each term of Eq. (2.40) is annihilated either by the left  $(\hat{I} - \mathbf{L})$  or the right  $(\hat{I} - \mathbf{R})$  projection over the unoccupied MOs. Notice that  $\mathbf{L} = \mathbf{R}$  in the real case and  $\mathbf{L} = \mathbf{R}^*$  in the most general complex case. In this way, in order to implement a constrained variation  $\delta g_n$  of the Pfn WF, corresponding to an appropriate variation of its matrix  $\delta\lambda_{\mu,\nu}^n$ , it is useful to work with a small free variation  $\delta g$  (with corresponding  $\delta A_{\mu,\nu}$ ). This is then projected onto the chosen restricted *ansatz* by means of the following equation:

$$\delta g_n = \delta g - (\hat{I} - \mathbf{L})\delta g(\hat{I} - \mathbf{R}). \quad (2.43)$$

Indeed, it is easy to show that the right-hand side of the above equation vanishes if we apply  $\hat{I} - \mathbf{L}$  and  $\hat{I} - \mathbf{R}$  to its left and its right, respectively, just because  $\hat{I} - \mathbf{R}$  and  $\hat{I} - \mathbf{L}$  are projection operators, being such  $\mathbf{R}$  and  $\mathbf{L}$ , yielding  $(\hat{I} - \mathbf{R})^2 = (\hat{I} - \mathbf{R})$  and  $(\hat{I} - \mathbf{L})^2 = (\hat{I} - \mathbf{L})$ , from which Eq. (2.43) fulfills Eq. (2.41). Eq. (2.43) represents, therefore, a *linear relation* applied to the variational parameter matrix change  $\delta a_{\mu,\nu}$  corresponding to the unconstrained geminal  $g$  in Eq. (2.34), yielding the new constrained variation  $\delta a_{\mu,\nu}^n$ . Indeed, by using the definitions of the projector operators in Eq. (2.42) and the expansion of the MOs in the atomic (hybrid) basis the implementation of Eq. (2.43) turns to a number of matrix-matrix operations acting on  $\lambda$ ,  $P$ ,  $\bar{P}$  and the overlap matrix  $S$  that can be easily and efficiently worked out[5].

This linear relation between  $\lambda$  and  $\lambda^n$  can be therefore easily implemented together with the corresponding derivatives necessary to the optimization of the energy <sup>2</sup> and allows the explicit calculation of the new matrix  $\lambda_{\mu,\nu}^n$ , yielding the new constrained geminal  $g_n + \delta g_n$ . Then the new geminal can be recasted in the form of Eq. (2.39) by the mentioned diagonalization of skew-symmetric matrices, presented in the subsection 2.2.7, in this way implicitly neglecting nonlinear contributions that are irrelevant close to convergence, when  $\delta g_n \rightarrow 0$ . After employing several iterations of this type, the lowest energy *ansatz* of the Pfn type can be obtained in a relatively simple and very efficient way.

It is also important to emphasize that this constrained optimization algorithm allows a further reduction of the number of parameters, by efficiently exploiting locality, namely that variational parameters  $\lambda_{\mu,\nu}$  corresponding to atoms at a distance larger than a reasonable cutoff can be safely disregarded with negligible error[5].

Finally, we can notice that if  $n$  is equal to half the number of electrons, the

---

<sup>2</sup>The output of AAD are matrices  $D_{\mu,\nu} = \frac{\partial F}{\partial \lambda_{\mu,\nu}}$  where  $F$  is either the log of the WF or the corresponding local energy computed on a given configuration. Then the projected derivatives corresponding to  $\delta \lambda_{\mu,\nu}^n$  easily follows from Eq. (2.43), by applying the chain rule.

AGPn and the Pfn are equivalent to a SD.

## 2.2.6 Conversion of the Wave Functions

The choice of the proper *ansatz* depends on the target system, considering the computational cost of a chosen *ansatz* and the relevant physical and chemical properties of a target material. During the simulation, it can be convenient (or necessary) to go back and forth between the *ansatze*, with/without losing the information of an optimized *ansatz* (Fig. 2.2). The first case is to add molecular orbitals to an *ansatz*, *i.e.*, AGP  $\Rightarrow$  SD, AGP  $\Rightarrow$  AGPn, or Pf  $\Rightarrow$  Pfn. In TURBORVB, this is obtained by rewriting the expansion of the geminal in terms of molecular orbitals, as shown in the previous section. This operation requires to solve the eigenvalues and eigenvectors problem of the parameters matrix. For the AGP this means that one has to diagonalize a symmetric matrix (using the standard LAPACK routines[63]). For the Pf WF, instead, one has to deal with an antisymmetric matrix. In the section 2.2.7, I will describe a fast and general procedure to transform a generic complex antisymmetric matrix into a canonical Youla’s form that represents the equivalent of the standard diagonalization of Hermitian matrices.

The second important case is to convert an *ansatz* among the available ones, *i.e.*, SD, AGP, or AGPn  $\Rightarrow$  AGP or Pf, Pfn  $\Rightarrow$  Pf. This is achieved by maximizing the overlap between the two WFs. In more details, in TURBORVB, the following overlap between two geminals is maximized:

$$\max Q = \frac{\langle g^{\text{new}} | g^{\text{ori}} \rangle^2}{\langle g^{\text{new}} | g^{\text{new}} \rangle \langle g^{\text{ori}} | g^{\text{ori}} \rangle}, \quad (2.44)$$

in order to obtain new geminal matrix coefficients  $\lambda_{\mu,\nu}^{\text{new}}$ , defining the new pairing function as:

$$g^{\text{new}}(\mathbf{i}, \mathbf{j}) = \sum_{\mu,\nu} \lambda_{\mu,\nu}^{\text{new}} \phi_{\mu}^{\text{new}}(\mathbf{i}) \phi_{\nu}^{\text{new}}(\mathbf{j}), \quad (2.45)$$

while the original geminal was given in terms of the parameter matrix  $\lambda_{\mu,\nu}^{\text{ori}}$ :

$$g^{\text{ori}}(\mathbf{i}, \mathbf{j}) = \sum_{\mu,\nu} \lambda_{\mu,\nu}^{\text{ori}} \phi_{\mu}^{\text{ori}}(\mathbf{i}) \phi_{\nu}^{\text{ori}}(\mathbf{j}). \quad (2.46)$$

Notice that  $0 \leq Q \leq 1$ ; therefore the larger is  $Q$ , the better is the conversion, and  $Q$  approaches the unit value if the conversion is perfect.

The final case is to convert a AGP *ansatz* to Pf. Since the AGP *ansatz* is a special case of the Pf one, where only  $W_{\uparrow\downarrow}$  and  $W_{\downarrow\uparrow}$  terms are defined as described in section 2.2.4, the conversion can be realized just by direct substitution. Therefore, the main challenge is to find a reasonable initialization for the two spin-triplet sectors  $W_{\uparrow\uparrow}$  and  $W_{\downarrow\downarrow}$  that are not described in the AGP and that otherwise have to be set to 0. There are two possible approaches[28]. For polarized systems, we can build the  $W_{\uparrow\uparrow}$  block of the matrix by using an even number of unpaired orbitals  $\{\Theta_i\}$  and build an antisymmetric  $W_{\uparrow\uparrow}$  by means of Eq. (2.39), where the eigenvalues  $a_k$  are chosen to be large enough to occupy certainly these unpaired states, as in the standard Slater determinant used for the initialization. This works only for polarized systems. The second approach that also works in a spin-unpolarized case is to determine a standard broken symmetry single determinant *ansatz* (*e.g.*, DFT within the LSDA) and modify it with a global spin rotation. Indeed, in the presence of finite local magnetic moments, it is often convenient to rotate the spin moments of the WF in a direction perpendicular to the spin quantization axis chosen for our spin-dependent Jastrow factor, *i.e.*, the  $z$  quantization axis. In this way one can obtain reasonable initializations for  $W_{\uparrow\uparrow}$  and  $W_{\downarrow\downarrow}$ . The corresponding tool developed in TURBORVB allows every possible rotation, including an arbitrary small one close to the identity. A particularly important case is when a rotation of  $\pi/2$  is applied around the  $y$  direction. This operation maps

$$|\uparrow\rangle \rightarrow \frac{1}{\sqrt{2}} (|\uparrow\rangle + |\downarrow\rangle) \quad \text{and} \quad |\downarrow\rangle \rightarrow \frac{1}{\sqrt{2}} (|\uparrow\rangle - |\downarrow\rangle). \quad (2.47)$$

One can convert from an AGP the pairing function that is obtained from a

VMC optimization:

$$g_{AGP}(\mathbf{i}, \mathbf{j}) = g_+(\mathbf{r}_i, \mathbf{r}_j) \frac{|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle}{\sqrt{2}} + g_-(\mathbf{r}_i, \mathbf{r}_j) \frac{|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle}{\sqrt{2}} \quad (2.48)$$

to a Pf one:

$$g_{AGP}(\mathbf{i}, \mathbf{j}) \rightarrow g(\mathbf{i}, \mathbf{j}) = g_+(\mathbf{r}_i, \mathbf{r}_j) \frac{|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle}{\sqrt{2}} + g_-(\mathbf{r}_i, \mathbf{r}_j) (|\uparrow\uparrow\rangle - |\downarrow\downarrow\rangle). \quad (2.49)$$

This transformation provides a meaningful initialization to the Pf WF that can be then optimized for reaching the best possible description of the ground state within this *ansatz*.

### 2.2.7 Diagonalization of a antisymmetric generally complex matrix $\lambda$

In the following we will discuss a general procedure to transform a generic complex antisymmetric matrix into a canonical Youla's form that represents the equivalent of the standard diagonalization of Hermitian matrices. This is obtained by means of an appropriate unitary matrix  $U$  defined by an orthonormal set of states that we will call in the following MOs.

Given a  $\bar{N} \times \bar{N}$  antisymmetric matrix  $\lambda$ , our goal is to identify a set of  $p$  paired states  $\{(\phi_j^1, \phi_j^2)\}$  of orthonormal MOs, such that  $p \leq \bar{N}$  is even and

$$\lambda\phi_j^1 = a_j\phi_j^2 \quad (2.50)$$

$$\lambda\phi_j^2 = -a_j\phi_j^1, \quad (2.51)$$

where the left-hand side of the above equations indicates standard matrix vector products, with shorthand notations adopted also in the remaining part of this appendix. In this basis we can write any skew-symmetric matrix  $\lambda$  in

the canonical Youla's form:

$$\lambda_{MO} = \begin{pmatrix} 0 & a_1 & 0 & 0 & \cdots & 0 \\ -a_1 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & a_2 & \cdots & 0 \\ 0 & 0 & -a_2 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \end{pmatrix}, \quad (2.52)$$

using only  $p/2$  strictly positive parameters  $a_j$ . These ones play the same role of the eigenvalues for an ordinary Hermitian matrix and henceforth we will use this name for them, even if the matrix  $\lambda_{MO}$  is not diagonal but represents the simplest non-vanishing skew-symmetric matrix.

The transformation of the original matrix  $\lambda$  to the corresponding canonical Youla's form by means of an appropriate unitary transformation  $\lambda = U^* \lambda_{MO} U^\dagger$  provides us also a very simple way to regularize the matrix  $\lambda$  by replacing the too small eigenvalues with reasonable lower bounds.

In the case of odd  $\bar{N}$  it will be shown later that there exists always an eigenvector of  $\lambda$  with vanishing eigenvalue, but the decomposition remains possible, as  $\lambda_{MO}$  will contain at least one vanishing row and corresponding column. In the following, we define that an eigenvector is singular if it corresponds to a vanishing eigenvalue, as in the odd  $\bar{N}$  case.

It would be ideal for this calculation to use a very robust and stable diagonalization routine to maintain machine accuracy for the MOs. Unfortunately, these routines are not commonly available for antisymmetric matrices and thus several mathematical transformations are necessary to map our task to a sequence of more commonly used or at least easily available algorithms.

A generic  $\bar{N} \times \bar{N}$  antisymmetric matrix is written in the following way:

$$\lambda = \begin{pmatrix} 0 & a_{1,2} & a_{1,3} & \cdots & a_{1,\bar{N}} \\ -a_{1,2} & 0 & a_{2,3} & \cdots & a_{2,\bar{N}} \\ -a_{1,3} & -a_{2,3} & 0 & \cdots & a_{3,\bar{N}} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -a_{1,\bar{N}} & -a_{2,\bar{N}} & -a_{3,\bar{N}} & \cdots & 0 \end{pmatrix}. \quad (2.53)$$

The first step is to transform  $\lambda$  in a tridiagonal antisymmetric real matrix. This operation is implemented in the subroutine `zsktrd` (`dkstrd`) contained in the PFAPACK library [64]. The use of the Householder algorithm allows us to decompose the generic matrix  $\lambda$  as

$$\lambda = U_1^* \lambda_{Tr} U_1^\dagger, \quad (2.54)$$

where  $U_1$  is the transformation matrix output of the algorithm, while  $\lambda_{Tr}$  is a tridiagonal real antisymmetric matrix written in the standard tridiagonal form

$$\lambda_{Tr} = \begin{pmatrix} 0 & b_1 & 0 & \cdots & 0 \\ -b_1 & 0 & b_2 & \cdots & 0 \\ 0 & -b_2 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \end{pmatrix}. \quad (2.55)$$

Thus we can multiply the matrix  $\lambda_{Tr}$  for the imaginary unit  $i$ , yielding a more conventional tridiagonal hermitian matrix  $\lambda_{iH}$ , defined by purely imaginary matrix elements.

We highlight that it is possible to map the matrix  $\lambda_{iH}$  into a real hermitian matrix via a unitary transformation and use the appropriate LAPACK routine for its fast diagonalization. This procedure is well known and will be discussed later.

At this point, we can use the spectral theorem for Hermitian matrices to decompose the matrix  $\lambda_{iH} = \psi \lambda_{diag} \psi^\dagger$ , where  $\lambda_{diag}$  is a diagonal matrix containing in its diagonal part the real eigenvalues  $a_i$  of  $\lambda_{iH}$  and  $\psi$  is the unitary matrix, where each column is given by the eigenvector, in principle complex, corresponding to each eigenvalue, in the chosen order. This decomposition implies:

$$\lambda = -i U_1^* \psi \lambda_{diag} \psi^\dagger U_1^\dagger. \quad (2.56)$$

However, since the matrix  $\psi$  is generally complex and  $\psi^\dagger \neq \psi^T$ , some manip-

ulation is necessary if we want to satisfy the skew-symmetry property of  $\lambda$ , in an easy and transparent way.

If we consider one eigenvector  $\bar{\psi}_j$  associated to an eigenvalue  $a_j > 0$  we have that

$$\lambda_{iH}\bar{\psi}_j = i\lambda_{Tr}\bar{\psi}_j = a_j\bar{\psi}_j, \quad (2.57)$$

the complex conjugate of this expression is

$$-i\lambda_{Tr}\bar{\psi}_j^* = -a_j\bar{\psi}_j^*, \quad (2.58)$$

where we have used that both  $\lambda_{Tr}$  and the eigenvalues  $a_j$  are real. This means that if  $\bar{\psi}_j$  is an eigenvector of  $\lambda_{iH}$  relative to the eigenvalue  $a_j$ , then  $\bar{\psi}_j^*$  is an eigenvector corresponding to the eigenvalue  $-a_j$  and thus orthogonal to  $\bar{\psi}_j$  because of the orthogonality between eigenvectors of a Hermitian matrix corresponding to different eigenvalues  $\pm a_j$ . We can thus easily verify, by using the relations given in Eq. (2.57) and Eq. (2.58), the following simple equations<sup>3</sup>:

$$\lambda_{iH}(\bar{\psi}_j + \bar{\psi}_j^*) = a_j(\bar{\psi}_j - \bar{\psi}_j^*) \quad (2.59)$$

$$\lambda_{iH}(\bar{\psi}_j - \bar{\psi}_j^*) = a_j(\bar{\psi}_j + \bar{\psi}_j^*) \quad (2.60)$$

. In this way we can define pairs of *real* orthogonal vectors  $\bar{\psi}_j^1 = \sqrt{2} \operatorname{Re}(\bar{\psi}_j)$  and  $\bar{\psi}_j^2 = \sqrt{2} \operatorname{Im}(\bar{\psi}_j)$  such that

$$\lambda_{iH}\bar{\psi}_j^1 = ia_j\bar{\psi}_j^2 \quad (2.61)$$

$$\lambda_{iH}\bar{\psi}_j^2 = -ia_j\bar{\psi}_j^1. \quad (2.62)$$

Once we have identified all the pairs corresponding to all positive eigenvalues  $a_j > 0$  we can write the unitary matrix  $\bar{\psi}$  that is now real, by adding the remaining eigenvectors (that can be also chosen real as shown in the following)

---

<sup>3</sup>The same argument holds if the eigenvalue  $a_j$  corresponds to  $p > 1$  degenerate eigenvectors. The mentioned orthogonality property of Hermitian matrix eigenvectors leads to the straightforward definition of  $p$  pairs of mutually orthonormal real ones used for the present decomposition with a block diagonal matrix, where each  $2 \times 2$  block corresponds to one of the  $p$  degenerate eigenvectors.

with vanishing eigenvalues in the remaining rightmost columns. In this way we can finally define a unitary real matrix  $\bar{\psi}$  yielding  $\lambda_{Tr} = -i\lambda_{iH} = \bar{\psi}\lambda_{MO}\bar{\psi}^T$  where  $\lambda_{MO}$  is defined in Eq. (2.52) and therefore by using Eq. (2.54)

$$\lambda = U_1^* \bar{\psi} \lambda_{MO} \bar{\psi}^T U_1^\dagger. \quad (2.63)$$

which represents the desired decomposition because the product of two unitary matrices  $U^* = U_1^* \bar{\psi}$  remains a unitary matrix and its transpose  $U^\dagger$  coincides with  $\bar{\psi}^T U_1^\dagger$ , yielding  $\lambda = U^* \lambda_{MO} U^\dagger$ .

### Triangular hermitian matrices: a mapping from imaginary to real

In order to use the LAPACK routines for the diagonalization we have to map the tridiagonal fully imaginary hermitian matrix  $\lambda_{iH}$ , defined only (the diagonal elements are zero to fulfill hermitianity) by its upper diagonal elements  $ib_j$  with  $b_j$  real for  $j = 1, 2, \dots, \bar{N} - 1$ , into a tridiagonal real symmetric matrix  $\lambda_R$ . We can implement this mapping by applying a unitary transformation to the matrix  $\lambda_{iH}$ . For this purpose we introduce the following transformation described by the matrix  $U_2$

$$\lambda_R = U_2^\dagger \lambda_{iH} U_2. \quad (2.64)$$

The matrix  $U_2$  is a complex diagonal matrix defined as

$$U_2 = \begin{pmatrix} e^{i\phi_1} & 0 & \dots & 0 \\ 0 & e^{i\phi_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & e^{i\phi_{\bar{N}-1}} \end{pmatrix}. \quad (2.65)$$

The explicit calculation of the right-hand side of the Eq. (2.64) gives

$$\lambda_R = \begin{pmatrix} 0 & ib_1 e^{i(\phi_2 - \phi_1)} & 0 & \dots & 0 \\ -ib_1 e^{i(\phi_1 - \phi_2)} & 0 & ib_2 e^{i(\phi_3 - \phi_2)} & \dots & 0 \\ 0 & -ib_2 e^{i(\phi_2 - \phi_3)} & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \end{pmatrix}. \quad (2.66)$$

By setting the imaginary units  $\pm i = \exp(\pm i\frac{\pi}{2})$  (when not exponentiated in the previous equation), we can easily impose that all the phase factors cancel in all the corresponding matrix elements of  $\lambda_R$  with the choice:

$$\phi_j = -\frac{\pi}{2}(j-1), \quad (2.67)$$

that, therefore, implies that  $\lambda_R$ , with the above definition, is a real symmetric matrix. At this point we can diagonalize the matrix  $\lambda_R$  by means of a real unitary matrix  $U_R$ , that is the output of a standard LAPACK diagonalization routine of tridiagonal real matrices (e.g. `dstevx` for double-precision arithmetic). In this way  $\lambda_{iH}$  can be diagonalized as  $\lambda_{iH} = U_2 U_R \lambda_{diag} U_R^T U_2^\dagger$  where  $\lambda_{diag}$  is a diagonal matrix containing the corresponding eigenvalues of the LAPACK diagonalization.

### Singular eigenvectors

Within this formulation it is also particularly easy to compute all the real singular eigenvectors of  $\lambda_{iH}$  corresponding to the possible vanishing eigenvalues. They were used in this appendix to complete the columns of the unitary real matrix  $\bar{\psi}$ . From the outcome of the previous subsection any eigenvector  $\phi_k^j$  of  $\lambda_{iH}$  can be obtained by applying the diagonal matrix  $U_2$  to a real eigenvector  $\bar{\phi}_k^j$  of  $\lambda_R$ , namely  $\phi_k^j = \bar{\phi}_k^j \exp(-i\frac{\pi}{2}(k-1))$ , implying that even  $k$ -components are purely imaginary and odd  $k$ -components are purely real. Then it is immediate to realize that if  $\bar{\phi}_k^j$  corresponds to a singular eigenvector of  $\lambda_R$  also  $\text{Re}(\phi_k^j)$  and  $\text{Im}(\phi_k^j)$  ( and therefore also  $\text{Re}(\phi_k^j) + \text{Im}(\phi_k^j)$ ) correspond to singular eigenvectors or at most null vectors (not both) of  $\lambda_{iH}$  because this matrix is purely imaginary and the complex conjugation of a singular eigenvector is again a singular eigenvector by Eq. (2.57) and Eq. (2.58) with  $a_j = 0$ .

Then it follows that all the orthogonal eigenvectors  $\bar{\phi}^j$  (output of `dstevx`) corresponding to the zero eigenvalues of the matrix  $\lambda_R$  can be used to define the corresponding real singular eigenvectors corresponding to the matrix  $\lambda_{iH}$ ,

i.e.:

$$\tilde{\phi}_k^j = \text{Re}(\phi_k^j) + \text{Im}(\phi_k^j) = \begin{cases} (-1)^{(k-1)/2} \bar{\phi}_k^j & k \text{ odd} \\ (-1)^{k/2} \bar{\phi}_k^j & k \text{ even} \end{cases} \quad (2.68)$$

that are explicitly real, and orthogonal each other because  $\sum_k \tilde{\phi}_k^j \tilde{\phi}_k^l = \sum_k \bar{\phi}_k^j \bar{\phi}_k^l = \delta_{l,j}$ . They are also orthogonal to all the other pairs of non-singular eigenvectors because of the orthogonality property of eigenvectors of a Hermitian matrix  $\lambda_{iH}$ , that we have already used in the previous section.

## 2.3 Jastrow Factor

Within QMC, it is easy to improve the quality of the WF by multiplying the WF with an exponential JF. This last one enriches the description of the ground state (GS) by encoding explicitly the electronic correlation, while speeding up the convergence to the complete basis set limit[11]. Indeed, with an appropriate choice, the JF can satisfy exactly the electron-electron and electron-ion cusp conditions of the many-body WF, consequences of the Coulomb  $1/r$  singularity at short distance. In Ref. [28] we introduced a new kind of JF that contains a richer dependence on the spin and that plays a fundamental role when used in combination with the Pf WF.

The JF is defined as

$$J(\mathbf{X}) = e^{U_{ei} + U_{ee}}, \quad (2.69)$$

where  $U_{ei}$  is a single body term that deals explicitly with the electron-ion interaction and  $U_{ee}$  is a many-body term that properly accounts for the electronic correlation. The single body term is

$$U_{ei} = \sum_{i=1}^N u_{ei}(\mathbf{r}_i), \quad (2.70)$$

with  $u_{ei}$  being

$$u_{ei}(\mathbf{r}_i) = \sum_{I=1}^{\#ions} -Z_I \frac{1 - \exp(b_{eI} |\mathbf{r}_i - \mathbf{R}_I|)}{b_{eI}} + g_I(\mathbf{r}_i). \quad (2.71)$$

In Eq. (2.71)  $Z_I$  is the atomic number of the atom  $I$  and  $b_{ei}$  is a variational parameter defined for each atomic species, while  $g_I(\mathbf{r}_i)$  encodes the most general non-homogeneous electron-ion one-body term, i.e. depending explicitly on all nuclear and electron coordinates and not only on their relative distances, that is defined as

$$g_I(\mathbf{r}_i) = \sum_{\nu} \xi_{I,\nu} \phi_{I,\nu}(\mathbf{r}_i), \quad (2.72)$$

where the summation is extended over all the gaussian orbitals in the JF basis set centered on the  $I$ -th atom. The electron-electron term instead is written as

$$U_{ee} = \sum_{i < j} u_{ee}(\mathbf{r}_i \sigma_i, \mathbf{r}_j \sigma_j), \quad (2.73)$$

where the sum is extended over the pairs of different electrons and where

$$u_{ee}(\mathbf{r}_i \sigma_i, \mathbf{r}_j \sigma_j) = k_{\sigma_i, \sigma_j} \frac{|\mathbf{r}_i - \mathbf{r}_j|}{1 + b_{\sigma_i, \sigma_j}^{ee} |\mathbf{r}_i - \mathbf{r}_j|} + g_{ee}(\mathbf{r}_i \sigma_i, \mathbf{r}_j \sigma_j), \quad (2.74)$$

with the  $2 \times 2$  matrix  $b_{\sigma, \sigma'}^{ee}$  described by one  $b_{\sigma, \sigma'}^{ee} = b^{ee}$  or two variational parameters for  $\sigma_i = \sigma_j$  when  $k_{\sigma_i, \sigma_j} = 1/4$  and  $b_{\sigma, \sigma'}^{ee} = b_{\parallel}^{ee}$  and for  $\sigma_i \neq \sigma_j$  when  $k_{\sigma_i, \sigma_j} = 1/2$  and  $b_{\sigma, \sigma'}^{ee} = b_{\perp}^{ee}$ . The conventional expression for the JF can be obtained by removing all spin dependency in the previous expressions and remaining only with the variational parameters corresponding to the opposite spin case  $k_{\sigma_i, \sigma_j} = 1/2$  and  $b_{\sigma, \sigma'}^{ee} = b^{ee}$ .

In our expression the first term in Eq. (2.74), named two-body Jastrow, deals explicitly with the electron-electron cusp conditions, the second term in Eq. (2.74) instead is a bosonic pairing function in the form

$$g_{ee}(\mathbf{r}_1 \sigma_1, \mathbf{r}_2 \sigma_2) = \sum_{k, l} \zeta_{k, l}^{\sigma_1 \sigma_2} \phi_{k \sigma_1}(\mathbf{r}_1) \phi_{l \sigma_2}(\mathbf{r}_2), \quad (2.75)$$

with the elements of the matrix  $\zeta$  defining further variational parameters. Notice that both  $g_I$  and  $g_{ee}$  do not affect the cusp conditions because they are expanded over cusplless gaussian orbitals. The  $g_{ee}$  term has the same form of Eq. (2.14), but, since the fermionic behavior is already encoded in

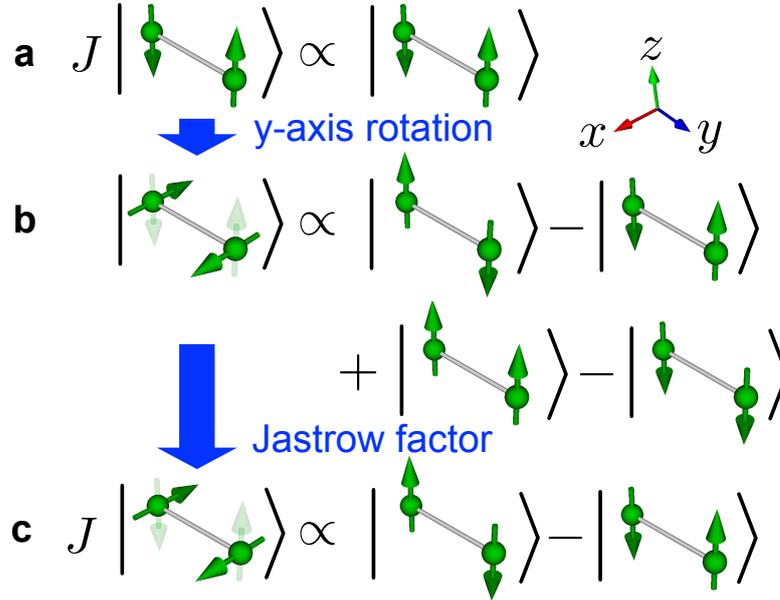


Figure 2.4: Restoring the singlet state for the Jastrow correlated broken symmetry *ansatz*. The JF cannot change the broken symmetry *ansatz* if it is oriented in the same quantization axis ( $z$ -axis) of the electron basis. If we rotate the spins of the broken symmetry *ansatz* by 90 degrees around the  $y$ -axis, the state becomes a quite general linear combination of spin configurations in the original basis. By carefully tuning the weights of each configuration with an appropriate spin-dependent Jastrow factor, we can recover the exact expansion of the singlet state in this basis.

the fermionic part of the WF, this term is symmetric under particle exchange. The use of a pairing function in the JF enriches the description of the charge and spin correlations of the system, by noticeably improving the quality of the global WF. It is a common practice to adopt a simplified or even absent spin dependency in the function  $u$  of Eq. (2.74). This is often accurate for systems where the magnetic properties are not relevant. We will refer to it in the following with the prefix Js in the WF, in order to distinguish it from the prefix J used for the full spin dependent JF.

A perfect singlet remains such after the multiplication of a spin independent JF, and so our spin dependent JF is not appropriate if we do not want to break the spin symmetry. It is, instead, necessary if we want to recover, at least

approximately, the singlet from a spin contaminated broken symmetry uncorrelated fermionic *ansatz*. A general spin dependent  $u$ , as defined in Eq. (2.74), is therefore of fundamental importance for the AGPu or the Pf *ansatze*. Let me start with a simple example. We consider two atoms with opposite spins and break the spin symmetry by orienting the spins of the atoms along the  $z$ -quantization axis. In this case the JF is not able to change the classical antiferromagnetic spin state because it acts as an irrelevant constant when applied to it. It is instead more physical to orient the spin moment of the atoms in a direction perpendicular to the quantization axis chosen for the JF. In this way the JF can act on the electrons and the spins while the magnetic moment is free to fluctuate and recover its genuine quantum character. As previously mentioned with the Pf it is possible to rotate the spin of the WF in every direction and orient the magnetic moment in any direction of the space. This works particularly well in combination with our Jastrow that can suppress the unfavored triplet configurations with parallel spins generated by the rotation as shown in Fig. 2.4. This optimal spin-orientation of the atoms, i.e. perpendicular to the JF one, is rigorously valid within the well known spin-wave theory of a quantum antiferromagnet[30]. In this case the JF defined with a spin-quantization axis perpendicular to the magnetic moment of the atoms allows the description of the quantum fluctuations and the corresponding zero point energy, even for a finite (as is our case) number of atoms[30].

### 2.3.1 The $S^2$ operator

The standard QMC algorithms rely on the possibility to sample the real space configurations of a general electronic system. All the observables can be indeed calculated in the basis where the electron positions and their spins are defined. In particular, for the systems considered, it is interesting to estimate the spin observables in order to understand their magnetic properties and the quality of the corresponding WFs. Though the total  $S_z$  is fixed during the simulation,

the value of the  $S^2$  depends on the coupling between the JF and the Pf or the AGPu that is particularly important when the spin symmetry is broken. The efficient computation of the expectation value of the  $S^2$  operator has already been described in [65] for the JsAGPu and therefore we will derive in the following an analogous expression valid for the JPf.

In particular, I will show how to evaluate  $S^2$  in a region of the space with a fast and computationally cheap approach based on the fast-update algebra of the Pf and the spin dependent JF.

Let us consider the expectation value of the  $S^2$  operator over a generic WF  $\Psi$  by direct application of its definition. In the following we use the completeness of the spatial configurations:

$$\sum_{\mathbf{x}} |\mathbf{x}\rangle\langle\mathbf{x}| = \mathbb{I} \quad (2.76)$$

where the summation symbol implies here also a  $3N$ -multidimensional integral over the electron coordinates. Assuming a fixed polarization  $S_z$  we can write the explicit expression of the total spin square as

$$\begin{aligned} \langle S^2 \rangle_{\Psi} &= \frac{\langle \psi | (S^z)^2 | \psi \rangle}{\langle \psi | \psi \rangle} + \frac{1}{2} \sum_{\mathbf{x}} \sum_{i,j}^N \frac{\langle \psi | (S_i^+ S_j^- + S_i^- S_j^+) | \mathbf{x} \rangle \langle \mathbf{x} | \psi \rangle}{\langle \psi | \psi \rangle} \\ &= \sum_{i=1}^{N_{\uparrow}} \sum_{j=N_{\uparrow}+1}^N \sum_{\mathbf{x}} \frac{\langle \psi | \mathbf{x} \rangle \langle \mathbf{x} | S_i^+ S_j^- | \psi \rangle}{\langle \psi | \psi \rangle} \\ &\quad + \frac{1}{4} (N_{\uparrow} - N_{\downarrow})^2 + \frac{1}{2} (N_{\uparrow} + N_{\downarrow}) \\ &= \sum_{i=1}^{N_{\uparrow}} \sum_{j=N_{\uparrow}+1}^N \sum_{\mathbf{x}} p(\mathbf{x}) \frac{\langle \mathbf{x} | S_i^+ S_j^- | \psi \rangle}{\langle \mathbf{x} | \psi \rangle} \\ &\quad + \frac{1}{4} (N_{\uparrow} - N_{\downarrow})^2 + \frac{1}{2} N, \end{aligned} \quad (2.77)$$

where the operator  $\vec{S}_i$  in the above equation acts on the spin component corresponding to the electron position  $\mathbf{r}_i$  of the configuration  $\mathbf{x}$ . In the above expression

$$p(\mathbf{x}) = \frac{|\langle \psi | \mathbf{x} \rangle|^2}{\sum_{\mathbf{x}} |\langle \psi | \mathbf{x} \rangle|^2}, \quad (2.78)$$

therefore, by using QMC sampling, we generate configurations according to the probability density  $p(\mathbf{x})$ . Thus we can evaluate the above multidimensional integral by directly sampling the estimator  $S^2(\mathbf{x})$  that multiplies  $p(\mathbf{x})$  in Eq. (2.77), namely

$$S^2(\mathbf{x}) = \sum_{i=1}^{N_{\uparrow}} \sum_{j=N_{\uparrow}+1}^N \frac{\langle \mathbf{x} | S_i^+ S_j^- | \psi \rangle}{\langle \mathbf{x} | \psi \rangle} + \frac{1}{4}(N_{\uparrow} - N_{\downarrow})^2 + \frac{1}{2}N. \quad (2.79)$$

The content of the former equation can be evaluated efficiently as I will explain in the following. Indeed, the application of the operator  $S_i^+ S_j^-$  to the configuration  $\mathbf{x}$  generates only a single configuration  $\mathbf{x}_{ij} = \{(\mathbf{r}_1 \uparrow), \dots, (\mathbf{r}_i \downarrow), \dots, (\mathbf{r}_j \uparrow), \dots, (\mathbf{r}_N \downarrow)\}$ . Considering  $\mathbf{x}$  our sampled configuration and using the previously given definition of  $\mathbf{x}_{ij}$ , we can recast Eq. (2.79) as

$$S^2(\mathbf{x}) = \frac{1}{4}(N_{\uparrow} - N_{\downarrow})^2 + \frac{1}{2}N + \sum_{i=1}^{N_{\uparrow}} \sum_{j=N_{\uparrow}+1}^N \frac{\langle \mathbf{x}_{ij} | \psi \rangle}{\langle \mathbf{x} | \psi \rangle}. \quad (2.80)$$

The only hard challenge of Eq. (2.80) is the calculation of the  $N_{\uparrow} \times N_{\downarrow}$  ratios

$$r_{ij} = \frac{\langle \mathbf{x}_{ij} | \psi \rangle}{\langle \mathbf{x} | \psi \rangle} \quad (2.81)$$

for  $i = 1, 2, \dots, N_{\uparrow}$  and  $j = N_{\uparrow} + 1, N_{\uparrow} + 2, \dots, N$ , that in our case reads

$$r_{ij} = \frac{J(\mathbf{x}_{ij}) \Psi_{Pf}(\mathbf{x}_{ij})}{J(\mathbf{x}) \Psi_{Pf}(\mathbf{x})} = r_{ij}^{Pf} r_{ij}^{JF}. \quad (2.82)$$

The configurations  $\mathbf{x}$  and  $\mathbf{x}_{ij}$  differ for a spin flip of the electrons  $i$  and  $j$ , but we can also consider  $\mathbf{x}_{ij}$  as the configuration in which the electron  $i$  evolved to the position previously occupied by  $j$  and vice versa. We can then calculate the ratios in Eq. (2.82) using a fast algebra to update two positions for the Pf based on the Sherman-Morrison algebra and some simple manipulations and for the JF with a direct evaluation, as discussed in detail later on.

It is also possible to calculate the value  $S^2(\Lambda)$  of the  $S^2$  operator in a sub-region of the space  $\Lambda$ . For this quantity, we can obtain an expression similar to Eq. (2.80):

$$S_{\Lambda}^2(\mathbf{x}) = \frac{1}{4}(N_{\uparrow}^{\Lambda} - N_{\downarrow}^{\Lambda})^2 + \frac{1}{2}N^{\Lambda} + \sum_{i=\{\Lambda,\uparrow\}} \sum_{j=\{\Lambda,\downarrow\}} \frac{\langle \mathbf{x}_{ij} | \psi \rangle}{\langle \mathbf{x} | \psi \rangle}, \quad (2.83)$$

where  $N_{\sigma}^{\Lambda}$  ( $\sigma = \uparrow, \downarrow$ ) is the number of  $\sigma$ -electrons in the region  $\Lambda$ ,  $N^{\Lambda} = N_{\uparrow}^{\Lambda} + N_{\downarrow}^{\Lambda}$ . The summation symbol over  $i \in \{\Lambda, \sigma\}$  indicates the sum for all  $\sigma$ -electron whose coordinate is in the region  $\Lambda$ . Therefore also this quantity can be easily evaluated in terms of the ratios  $r_{ij}$  of Eq. (2.82), that will be derived in the following subsections.

### The Pfaffian contribution

In order to calculate the Pf contribution to  $r_{ij}$  we used a slim and fast algebra by making extensive use of the Pfaffian properties[66] with a computational cost  $O(N^3)$ , mostly BLAS3 operations. This was extremely important because otherwise this computation could easily become the bottleneck of the whole procedure. In this way we could ensure the evaluation cost of  $S^2$  to be comparable with the one of a typical QMC cycle over all the  $N$  electrons that is at most  $O(N^3)$ . Before describing the fast-updating rules for the position of two electrons with a single move, we need to introduce some quantities fundamental for the calculation.

Let us denote as  $W^{-1}$  the inverse of  $W$ . This inverse  $W^{-1}$  can be computed from scratch for each configuration used to sample the spin square. The electron coordinates  $\mathbf{r}_i$  are given for  $i = 1, \dots, N$ , but since the corresponding spin can change with respect to the original choice ( $\uparrow$  for  $i \leq N_{\uparrow}$ , and  $\downarrow$  for  $i > N_{\uparrow}$ ) due to the spin flips mentioned in the previous subsection, we will consider explicitly the values of the spin here.

We then define the matrix  $\theta$  as

$$\theta_{ij} = g(\mathbf{r}_i \uparrow, \mathbf{r}_j \downarrow) + g(\mathbf{r}_i \downarrow, \mathbf{r}_j \uparrow) - g(\mathbf{r}_i \uparrow, \mathbf{r}_j \uparrow) - g(\mathbf{r}_i \downarrow, \mathbf{r}_j \downarrow). \quad (2.84)$$

For each spin  $\uparrow$  electron labeled by  $k \leq N_\uparrow$  we can define the column vectors

$$v_k^\uparrow = \begin{pmatrix} g(\mathbf{r}_1 \uparrow, \mathbf{r}_k \uparrow) - g(\mathbf{r}_1 \uparrow, \mathbf{r}_k \downarrow) \\ g(\mathbf{r}_2 \uparrow, \mathbf{r}_k \uparrow) - g(\mathbf{r}_2 \uparrow, \mathbf{r}_k \downarrow) \\ \vdots \\ g(\mathbf{r}_N \downarrow, \mathbf{r}_k \uparrow) - g(\mathbf{r}_N \downarrow, \mathbf{r}_k \downarrow) \end{pmatrix}, \quad (2.85)$$

while for the spin  $\downarrow$  we have instead  $k \geq N_\uparrow$  and

$$v_k^\downarrow = \begin{pmatrix} g(\mathbf{r}_1 \uparrow, \mathbf{r}_k \downarrow) - g(\mathbf{r}_1 \uparrow, \mathbf{r}_k \uparrow) \\ g(\mathbf{r}_2 \uparrow, \mathbf{r}_k \downarrow) - g(\mathbf{r}_2 \uparrow, \mathbf{r}_k \uparrow) \\ \vdots \\ g(\mathbf{r}_N \downarrow, \mathbf{r}_k \downarrow) - g(\mathbf{r}_N \downarrow, \mathbf{r}_k \uparrow) \end{pmatrix}. \quad (2.86)$$

We can use these vectors to build the  $N \times N$  matrix

$$V = \left( v_1^\uparrow v_2^\uparrow \cdots v_{N_\uparrow}^\uparrow v_{N_\uparrow+1}^\downarrow \cdots v_N^\downarrow \right) = (V^\uparrow V^\downarrow), \quad (2.87)$$

that allows us to define

$$U = (U^\uparrow U^\downarrow) = W^{-1}V = (W^{-1}V^\uparrow W^{-1}V^\downarrow), \quad (2.88)$$

and finally

$$D = (V^\uparrow)^T U^\downarrow. \quad (2.89)$$

Now we have all the ingredients that we need for our fast-updating algebra, and upon application of Sherman-Morrison algebra, we arrive at the ratio

$$\begin{aligned} r_{ij}^{Pf} &= \text{Pf}[W(\mathbf{X}_{ij})]/\text{Pf}[W(\mathbf{X})] \\ &= (1 + U_{ii})(1 + U_{jj}) - U_{ij}U_{ji} - (\theta_{ij} + D_{ij})W_{ij}^{-1}. \end{aligned} \quad (2.90)$$

We can notice that the preliminary calculation of the auxiliary matrices  $\theta$ ,  $V$ ,  $U$ , and  $D$ , including the inversion of  $W$ , amounts to a total of  $O(N^3)$  operations, while the calculation of the ratios is  $O(N^2)$  once the matrices have been computed.

### The JF contribution

In the JF that we introduced in the previous section only the electron-electron term of Eq. (2.73) has a spin dependence and thus only this part gives a contribution to the ratio. By simple substitution it is easy to prove that

$$\begin{aligned} r_{ij}^{JF} &= \exp(D_i - D_j + u_{ee}(\mathbf{r}_i \uparrow, \mathbf{r}_j \downarrow) + u_{ee}(\mathbf{r}_i \downarrow, \mathbf{r}_j \uparrow) \\ &\quad - u_{ee}(\mathbf{r}_i \uparrow, \mathbf{r}_j \uparrow) - u_{ee}(\mathbf{r}_i \downarrow, \mathbf{r}_j \downarrow)), \end{aligned} \quad (2.91)$$

where we have defined

$$D_k = \sum_l u_{ee}(\mathbf{r}_l \sigma_l, \mathbf{r}_k \downarrow) - u_{ee}(\mathbf{r}_l \sigma_l, \mathbf{r}_k \uparrow). \quad (2.92)$$

The whole operation has a  $O(N^2)$  computational cost and so does not limit the calculation in terms of performances.

## Chapter 3

# Hydrogen Chain

In this chapter I will discuss the ground state properties of the hydrogen chain, a project successfully completed within the "Many Electron Collaboration" of the Simons Foundation. I will report the results of a combined application of cutting-edge computational methods to determine the properties of the hydrogen chain in its quantum-mechanical ground state, outlining the results contained in Ref.[12].

A linear chain of hydrogen atoms ( $N$  protons equispaced along a line, with  $N$  electrons) [67, 68, 69, 70, 11] embodies many central themes of modern condensed matter physics while retaining a certain degree of computational tractability. It features a periodic atomic potential, incorporates realistic Coulomb interactions, requires the construction of accurate and compact basis sets, and yet maintains a connection with the fundamental Hubbard model which has been a hallmark of the theory of interacting fermions. Varying the separation between the nuclei mimics applying pressure to a crystal, which we found leads to a rich phase diagram, including an antiferromagnetic Mott phase, electron density dimerization with power-law correlations, an insulator-to-metal transition and an intricate set of intertwined magnetic orders. There have also been several previous studies of this system [67, 68, 69, 70]. However, they were restricted either to small basis sets or finite system sizes, which prevented a realistic description of the H chain, or by their accuracy and capa-

bilities, which prevented reliable resolution of the delicate scales and discovery of all the phases. The synergistic application of complementary methods distinguishes the work presented here and allowed us to make robust predictions with a multi-messenger approach in this challenging problem. We accessed the ground-state properties using multiple first-principles many-body methods, including standard and sliced-basis density-matrix renormalization group (DMRG, sb-DMRG) [71, 72, 73, 74], auxiliary-field (AFQMC) [75], and VMC and DMC (our main contributions). For reference, independent-electron calculations have also been performed, including restricted (RHF) and unrestricted (UHF) Hartree-Fock [76] and DFT [77]. While it is not practically possible with any single method to simultaneously converge all aspects of the electronic structure calculations (such as basis set limit, many-electron correlation, and thermodynamic limit) across different regimes of the GS phase diagram, we draw our conclusions based on the convergence of multiple approaches to a consistent physical picture. For what concerns the VMC and DMC, in this study we used JsSD WF *ansatz*, averaging over twisted boundary condition, as described in section 2.1. The descriptions of the other methods adopted can be found in the supplemental information (SI) of Ref. [12].

In the following sections I will describe the different phases of the ground state as a function of the lattice spacing.

## 3.1 Insulating phase

### 3.1.1 Antiferromagnetism.

At large proton-proton separation to a first approximation the system is a collection of isolated H atoms, each with a single electron in the atomic  $1s$  orbital. This is very similar to the half-filled Hubbard model in the large coupling ( $U/t$ ) limit. However the weakly bound nature of  $H^-$  and its very diffuse orbitals can create excitons with strong binding in the H chain.

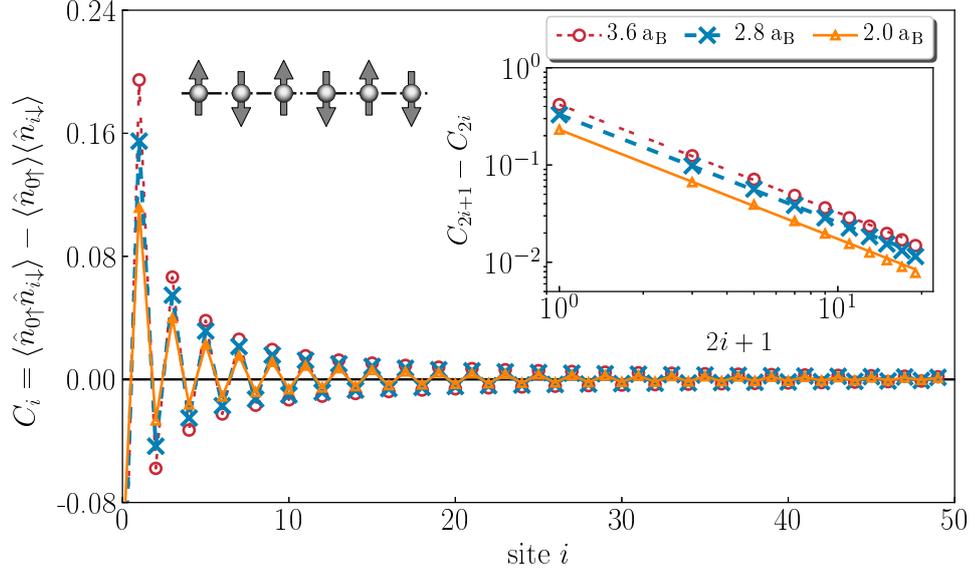


Figure 3.1: Insulating phase, antiferromagnetism. Main: correlation function  $C_i$  at  $R = 3.6$ ,  $2.8$ , and  $2.0 \text{ a}_B$  (red circles, blue crosses, and orange triangles) computed with DMRG for a chain of  $N = 50$  with OBC. The oscillations, with wavelength  $\lambda = 2R$ , show AFM correlations. Inset: oscillation amplitudes  $C_{2i+1} - C_{2i}$  versus  $i$  (log-log scale). Lines show the results of a linear fit. From Ref. [12].

At large  $R$ , the correlations in the H chain can be characterized in terms of a spin- $\frac{1}{2}$  Heisenberg chain. The Heisenberg chain is a critical system, with power-law decay of antiferromagnetic (AFM) spin-spin correlations,  $\langle \hat{\mathbf{S}}_0 \cdot \hat{\mathbf{S}}_i \rangle$ . In response to a local perturbation (such as a local magnetic field), one observes local ordering (such as local Néel order) which decays as a power law in the distance from the perturbation. To probe AFM correlations in the H chain, in Fig. 3.1, the quantity  $C_i = \langle \hat{n}_{0\uparrow} \hat{n}_{i\downarrow} \rangle - \langle \hat{n}_{0\uparrow} \rangle \langle \hat{n}_{i\downarrow} \rangle$  has been computed with DMRG in the minimal (STO-6G) basis, where  $\hat{n}_{i\sigma}$  denotes the number of electrons occupying the (orthogonalized) atomic orbital  $i$  with spin polarization  $\sigma$  along  $z$ .  $C_i$  oscillates with wavelength  $\lambda = 2R$ , corresponding to the Néel vector of two sublattice antiferromagnetism; the wavevector may also be thought of as twice the Fermi wave-vector  $q = \frac{2\pi}{\lambda} = 2k_F^0$  of a paramagnetic 1D ideal Fermi gas of density  $\rho = \frac{1}{R}$ . In Fig. 3.1, it is possible to observe that

the oscillations in  $C_i$  decay with a power-law envelope,  $C_i = C_0 i^{-\eta} (-1)^i$ . The decrease of  $C_0$  with  $R$  (see inset) indicates the weakening of AFM as the chain is compressed. The power-law decay of the AFM correlation is consistent with quasi long-range order in 1D. The fitted exponent of  $\eta \simeq 1.11(1)$ , likely affected by finite-size effects, is slightly higher than the prediction from conformal field theory, which gives  $\langle \mathbf{S}_i \cdot \mathbf{S}_0 \rangle \propto (-1)^i \sqrt{\ln i} / i^\eta$ , with  $\eta = 1$  for systems within the same universality class as the 1D Heisenberg chain [78, 79].

### 3.1.2 Dimerization

As  $R$  is reduced in the large  $R$  regime, a charge dimerization is observed. With PBC, dimerization can be probed by density-density correlations. With OBC, or with a local perturbation, density dimerization can be measured by the electronic density, for example integrated along transverse slabs (i.e., over  $x$  and  $y$  for a finite  $\delta z$ ),  $n(z)$ . The upper portion of Fig. 3.2a shows  $n(z)$  versus  $z$ , for a segment at the center of the chain under OBC, computed with AFQMC. The density has maxima at the proton sites and minima halfway between. The dimerization measure for an  $N$ -atom chain,  $\Delta_N$ , is defined by the difference between the two adjacent local minima of  $n(z)$  in the center of the chain as illustrated in the lower portion of Fig. 3.2a. When dimerization is present, we find that its amplitude increases as  $R$  is decreased.

In Fig. 3.2b the dimerization in the H chain at  $R = 2.0 a_B$  has been investigated using AFQMC, sb-DMRG, DMC, and VMC. We found that, in the middle of the chain, dimerization decays with chain length as  $\Delta_N \propto N^{-d}$ , with exponents  $d = 0.57(3)$ ,  $0.563(4)$ ,  $0.58(11)$  and  $0.90(26)$  from AFQMC, sb-DMRG, DMC and VMC, respectively. The subtle power-law physics of 1D correlated systems is not easy to capture accurately with numerical methods. It is encouraging that a quantitative agreement is seen between our very different many-body methods. HF is qualitatively incorrect here, giving either no order or true long-range order; DFT results are similar.

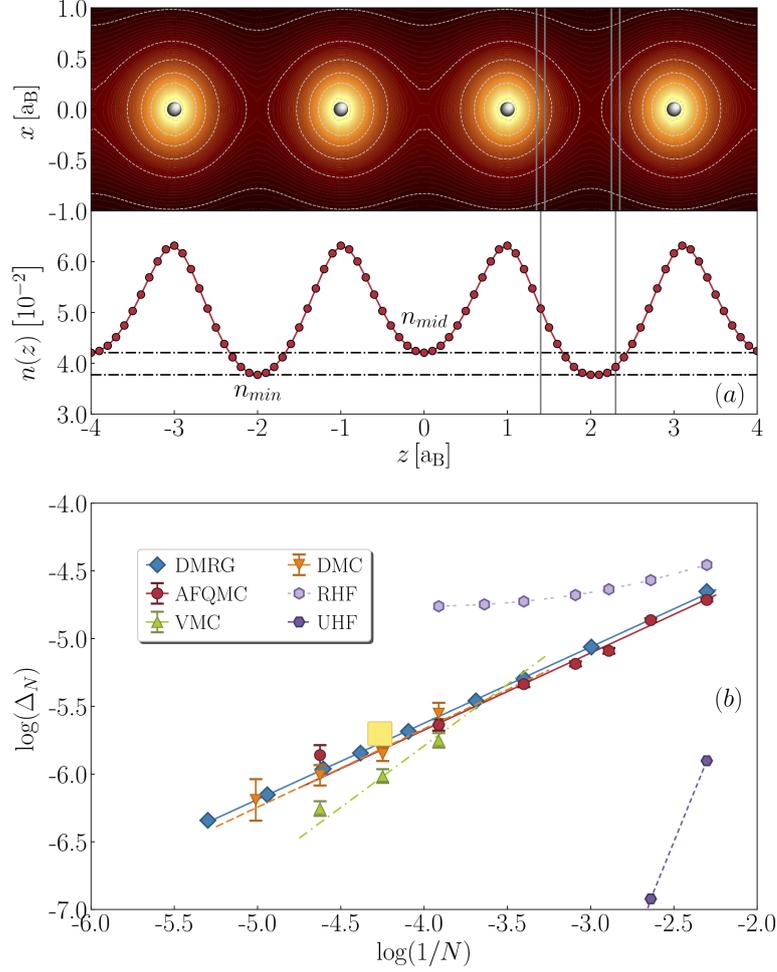


Figure 3.2: Insulating phase, density dimerization. (a) Top: electronic density  $n(z)$  for  $R = 2.0 a_B$ , with OBC. Bottom: definition of the dimerization measure  $\Delta_N$  for a system under OBC. The gray vertical lines indicate how  $n(z)$  is integrated along thin (width  $\delta z = 0.1 a_B$ ) slabs perpendicular to the chain. (b) Dependence of dimerization measure  $\Delta_N$  on the number of atoms in the open chain,  $N$ , from AFQMC, sb-DMRG, VMC, and DMC.  $\Delta_N$  decays as a power-law,  $\Delta_N \propto N^{-d}$  with exponent  $d \simeq 0.5$  from correlated methods. Results from HF are also shown for reference. From Ref. [12].

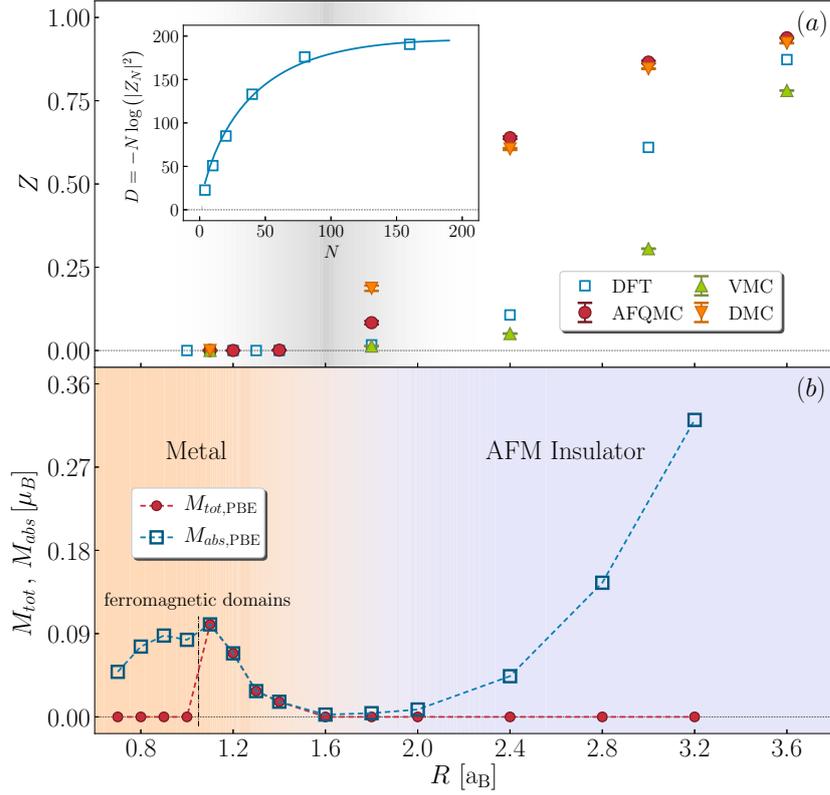


Figure 3.3: Metal-insulator transition. (a) Complex polarization measure as a function of  $R$  from multiple methods, identifying a MIT separating a region at smaller  $R$  with  $|Z| = 0$  from  $|Z| > 0$  at larger  $R$ . (b) Magnetic phases from DFT-PBE, as indicated by  $M_{tot} = \int m(r)d^3r/N$  and  $M_{abs} = \int |m(r)|d^3r/N$ , where  $m(r)$  is the magnetization of the simulation cell. From Ref. [12].

## 3.2 Metallic phase

### 3.2.1 Insulator-to-metal transition

Correlated electron materials often exhibit metal-insulator transitions as parameters such as temperature, pressure or crystal structure are varied [80, 81]. From the perspective of the 1D one-band Hubbard model which, as we have seen, captures the universal aspects of the physics at large  $R$  — and more generally from the perspective of one-band models with second-order Umklapp processes, no MIT should occur here[82]. With multiple methods and multiple probes, we have shown conclusive evidence that a MIT occurs in the H chain, and provided a characterization of the physical origin and properties of the transition.

The concept of macroscopic localization [83, 84, 70] provides a direct wavefunction-based characterization of system properties. For periodic systems, one defines the complex polarization  $Z_N = \langle \Psi_N | e^{i\frac{2\pi}{L} \sum_i \hat{z}_i} | \Psi_N \rangle$ , where  $|\Psi_N\rangle$  is the ground state of the  $N$  electrons in a supercell of size  $L = NR$  along the chain direction. The electron localization length  $\Lambda = \frac{\sqrt{D}}{2\pi\rho}$  is related to the complex polarization by  $D = -\lim_{N \rightarrow \infty} N \log |Z_N|^2$ . In localized systems,  $\lim_{N \rightarrow \infty} |Z_N|^2 = 1$ , and  $\Lambda$  is finite; in metallic systems  $\lim_{N \rightarrow \infty} |Z_N|^2 = 0$ , and  $\Lambda \rightarrow \infty$ .

In Fig. 3.3a, we established the MIT by computing  $Z_N$  with AFQMC, DMC and VMC. For small  $R$ , all methods give  $Z_N$  equal to or statistically compatible with 0 across a wide range of system sizes  $N$ , indicating a metallic many-body ground state. For large  $R$ , all methods yield a non-zero  $Z_N$ . The many-body methods point to a transition point located approximately at  $R_{\text{MIT}} \sim 1.70(5) a_B$ . While there is some uncertainty in the critical value, because of computational limitations, it is remarkable that two methods working with completely different basis sets and projection algorithms yield results in excellent quantitative agreement.

In particular the values of  $|Z_N|$  in the insulating phase, which fall further

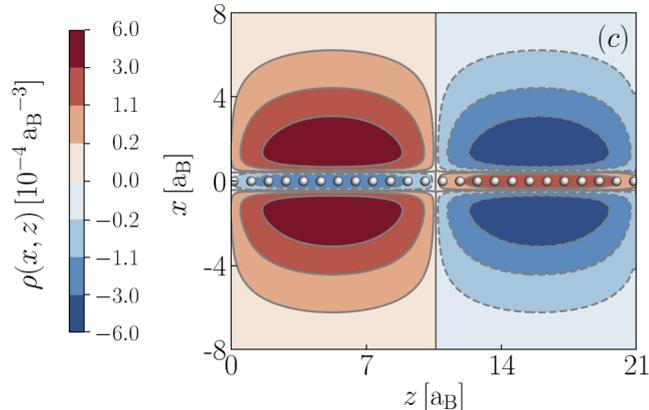


Figure 3.4: Spin-density computed from DFT-PBE, shown along a plane containing the chain, illustrating long-wavelength ferromagnetic domains at  $R = 0.9 a_B$ . From Ref. [12].

from unity at smaller  $R$ , are sensitive to finite-size effects. At large  $N$ , we expect  $|Z_N| = 1 - \tilde{g}(\xi/L)$  in the insulating phase, where  $\tilde{g}$  is a scaling function and  $\xi$  the MIT correlation length. The smooth decrease of  $|Z|$  as  $R_{\text{MIT}}$  is approached suggests a second-order transition (related to gap closure).

Before exploring the origin of the MIT, we briefly discuss magnetic correlations in the metallic phase. The evolution of the magnetic moments within DFT is shown in Fig. 3.3b, and the spin density at  $R = 0.9 a_B$  is plotted in Fig. 3.4 as a point of reference. To account for spin correlations, the DFT solution breaks translational symmetry to create antiferromagnetic domains of varying periods, often associated with very diffuse orbitals as seen in Fig. 3.4. Of course, DFT observations are independent-electron in nature. In the many-body solution, in particular, translational symmetry is restored, and two-body correlation functions are needed to probe magnetic correlations.

### 3.2.2 Origin of the MIT and properties of the metallic phase

It is theoretically established that the ground state of a one-band model with commensurate filling is insulating. Our calculations revealed that the MIT

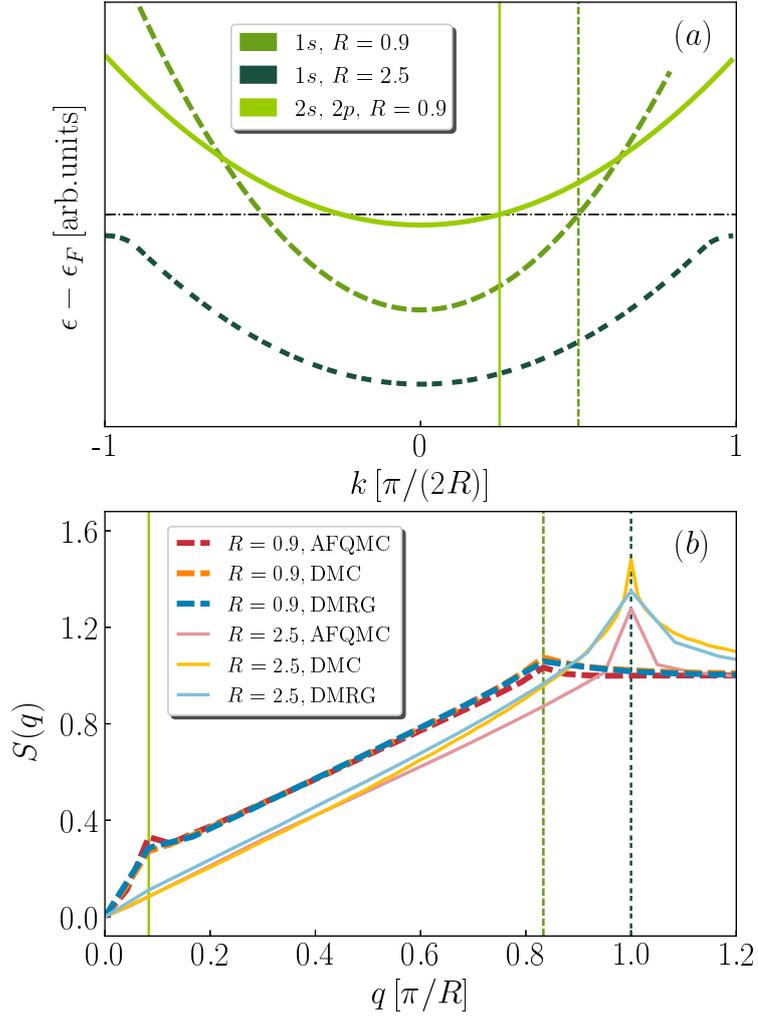


Figure 3.5: Mechanism of the insulator-to-metal transition. (a) Schematic illustration of the self-doping mechanism inducing a MIT in the H chain. Black line: large  $R$ , no band overlap. Green lines: small  $R$ , multiple bands crossing  $\epsilon_F$ . (b) Structure factor of the spin-spin correlation function at  $R = 0.9 a_B$  and at  $R = 2.5 a_B$ , computed from AFQMC, DMC and DMRG. Wave-vectors  $q$  are along the chain. Vertical lines, as an aid to the eye, mark the kinks in  $S(q)$  associated with the Fermi surfaces. From Ref. [12].

arises from a self-doping mechanism in which the one-band picture breaks down. The basic idea is illustrated in Fig. 3.5a, using a band-theory based cartoon of the electronic structure. The isolated H atom has multiple states, including the occupied  $1s$  and excited states  $2s, 2p$ , etc. At large  $R$  the band-

widths are small compared to the energy gaps, and a one-band approximation is reasonable. As  $R$  is decreased, the bands broaden, band overlap occurs, and metallic behavior results. Quantitative calculations involve a correlation problem that requires treating interactions in a multi-band situation, effects that are entirely absent in the standard Hubbard model.

We computed the spin structure factor, defined as  $S(q) = \frac{1}{N} \langle \Psi | \hat{\rho}_q^\dagger \hat{\rho}_q | \Psi \rangle$ , where  $\hat{\rho}_q = \sum_i \hat{S}_{z,i} e^{i\mathbf{q}\cdot\hat{\mathbf{r}}_i}$  is the Fourier transform of the spin density at  $\mathbf{q} = (0, 0, q)$ , with  $\hat{S}_{z,i}$  and  $\hat{\mathbf{r}}_i = (\hat{x}_i, \hat{y}_i, \hat{z}_i)$  denoting the spin- $z$  and position operators of the electron  $i$ , respectively. The result is shown in Fig. 3.5b. At  $R = 2.5$  only one peak is seen at  $q = 2k_F^0$ , signaling power-law AFM order as discussed earlier. To probe the nature of the metallic phase, we focus on a representative case of  $R = 0.9 a_B$ , away from the vicinity of the MIT transition.  $S(q)$  is shown from two different QMC calculations, in supercells of  $N = 48$  atoms averaging over 11 twist angles. In addition, DMRG calculations in a supercell of  $N = 24$  atoms are also reported. In a metallic system we expect peaks at  $q = 2k_F$ , where  $k_F$  is one of the Fermi wavevectors of the system. Two cusps are seen in each result at locations,  $q_1$  and  $q_2$ , in precise agreement among the different calculations. We interpreted the larger wavevector as arising from the  $2k_F$  process in the  $1s$ -dominated lower band. The position  $q_2 = 2(1-x)k_F^0$  then gives the doping  $x$  of this band. In a simple two-band picture, the lower wave vector is given by  $q_1 = 2xk_F^0/g$ , where  $g$  gives the degeneracy of the ‘‘upper band’’ which is occupied. At  $R = 0.9 a_B$ , the locations of  $q_1$  and  $q_2$  satisfy  $gq_1 + q_2 = \pi/R$  in all the results, with  $g = 2$ . This is consistent with a doubly degenerate upper band (e.g.  $2p_{x,y}$ ).

# Chapter 4

## The $H_4$ model system

Even relatively simple systems can hide pitfalls that can be very difficult to solve. The case of the  $(H_2)_2$ , a system of two diatomic molecules of hydrogen at equilibrium distance, first introduced in the literature by Anderson [85], is emblematic from this point of view: as recently shown by Gasperich et al. [14], a single SD can only give a very poor description of this system when it approaches the square geometry. This is due to the HOMO-LUMO degeneracy in the square limit that a single SD is not able to reproduce. Remarkably in Ref. [13] we show that the JsAGP allows a perfect description of this highly entangled ground state.

The simplicity of this model system allows us to study the role of the optimization in determining an accurate nodal surface, because, by repeating several times the optimization, we can be safely confident that the absolute minimum energy WF is obtained. On the other hand we can also verify that our stochastic optimization[5] works also when we remove the Jastrow from our *ansatz*, providing the lowest energy AGP, clearly with much larger computational effort compared to deterministic methods, that are not easily available for the AGP. We are able to show that the use of a pure AGP determinant (without any JF) can give rise not only to a poor description of the electronic correlation but also to a qualitatively wrong picture of the chemical bond.

Remarkably, the DMC energies obtained with the JsAGP trial WF in

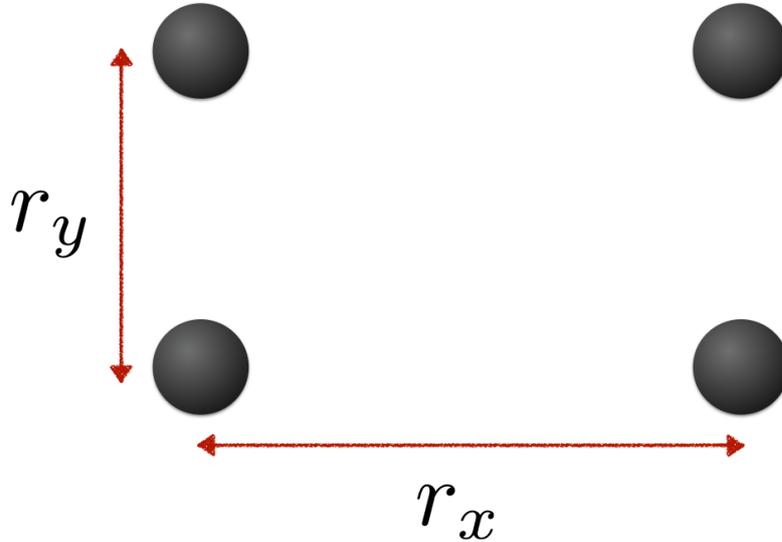


Figure 4.1: Stylized picture of the system. While  $r_y$  is kept constant for all the calculations at a distance equal to 2.4 a.u., the distance  $r_x$  is varied for different system shapes. From Ref.[13].

ccpVDZ basis are better than the ones obtained with the complete active space (CAS) (2,2) and CAS(4,4) and also with the full configuration interaction (FCI) all calculated with a quadruple zeta basis [14].

The geometry of the system studied has a fixed bond distance along the  $y$  direction  $r_y = 2.4$  a.u.. This value gives the lowest energy result for the square geometry [86]. As sketched in Fig. (4.1), we study the system as a function of the distance  $r_x$  between the two vertical molecules, comparing the dispersion curves obtained with the JsAGP and the JsSD and analyzing the effect of the JF.

For the optimization of the JsAGP and JsSD WFs we used the same procedure. We consider two types of initializations that we denote in the following by OPT  $r_x > r_y$  or OPT  $r_x < r_y$ , to indicate that the tetragonal symmetry is broken. In the first (second) case we take  $r_x = 4$ a.u. ( $r_x = 1.8$ a.u.) and perform a DFT calculation for the initial SD. We initially optimize only the

JF and we proceed with the full optimization of the AGP or SD with the JF. Then we move the atoms to a new position close to the original one maintaining the values of the variational parameters. If the new solution is reasonably close to the previous one, the stochastic optimization drives the WF to its new minimum. We iterated this procedure to obtain the WFs at all the  $r_x$  distances for the JsAGP and JsSD. As we will discuss more extensively later, the JsAGP optimization does not depend much on the starting WF, that is instead crucial for the JsSD. In this latter case the optimization procedure determines completely different results depending on the initial geometry when we get close to the symmetric square case.

To optimize the AGP without the JF we followed two different procedures yielding the same results. In one case we started for every geometry from the corresponding optimized JsAGP WF: we set the JF to 1 and we optimized the AGP from there. In the second case we used the same procedure adopted for the JsAGP and JsSD cases and obtained consistent energy values, validating the optimization procedure even in this difficult case without the JF.

## 4.1 Wave functions comparison: the failure of the Slater Determinant

The variational energies for the considered WFs are visible in Fig. (4.2)a and reported in table (4.1). As shown in Fig. (4.2)a, the JsSD values are reasonably accurate when the system is far from the square geometry, but very poor when  $r_x \approx r_y$ . We notice that for the JsSD the starting point is fundamental and the optimization result can significantly differ depending on the two different initializations. A particularly evident effect is the crossing of the JsSD energy dispersions in Fig. (4.2)a.

As expected this problem does not affect the JsAGP WF that shows the correct profile because, close to the square geometry, it contains implicitly the

Table 4.1: Variational energies for different optimized WFs. The basis set used for the JsAGP and JsSD is indicated between round parenthesis. We show one point for each case:  $r_x = r_y$ ,  $r_x < r_y$  and  $r_x > r_y$ . All the energies are expressed in Hartree.

$r_x$	JsSD(ccpVDZ)	JsAGP(ccpVDZ)	JsAGP(ccpVTZ)
1.80	$-2.1909 \pm 0.0003$	$-2.1957 \pm 0.0004$	$-2.1953 \pm 0.0003$
2.40	$-2.0694 \pm 0.0004$	$-2.1075 \pm 0.0004$	$-2.1084 \pm 0.0003$
3.00	$-2.1435 \pm 0.0003$	$-2.1491 \pm 0.0003$	$-2.1504 \pm 0.0003$

two important SDs with strong bonds either in the  $x$  or in the  $y$  direction. The optimizations of the JsAGP both from  $\text{OPT}r_x > r_y$  and  $\text{OPT}r_x < r_y$  lead exactly to the same result. The qualitative difference between the two *ansatze* is clearly shown in Fig. 4.3. The MOs try to localize the charge between two pairs of atoms to form two  $\text{H}_2$  molecules. In particular the JsSD binds the atoms that are at smaller distances in the initial geometry: if we consider the  $\text{OPT}r_x > r_y$  case we obtain a higher charge density along the  $y$  direction, while if we start from the  $\text{OPT}r_x < r_y$  case a higher charge along the  $x$  direction shows up. The JsAGP, instead, can resonate between these two configurations and catch the resonance valence bond (RVB)[87] behavior expected for the ground state of the square geometry.

The JsAGP result is not only good at the variational level, but it provides also particularly accurate nodal surfaces for the DMC calculations. Indeed, as we can notice from Fig. (4.2b) and from table (4.2), the DMC energies calculated using the nodes of the JsAGP (cc-pVDZ) are lower than the ones calculated with the multi-determinant WF CAS(4,4), and FCI with the quadruple zeta basis [14]. This shows that, even with a small basis set, the JsAGP leads, in this controlled case, to almost optimal nodes and, by consequence, very accurate DMC energies. This is indeed remarkable, considering also that other more standard methods suffer not only for poor accuracy but also for the too

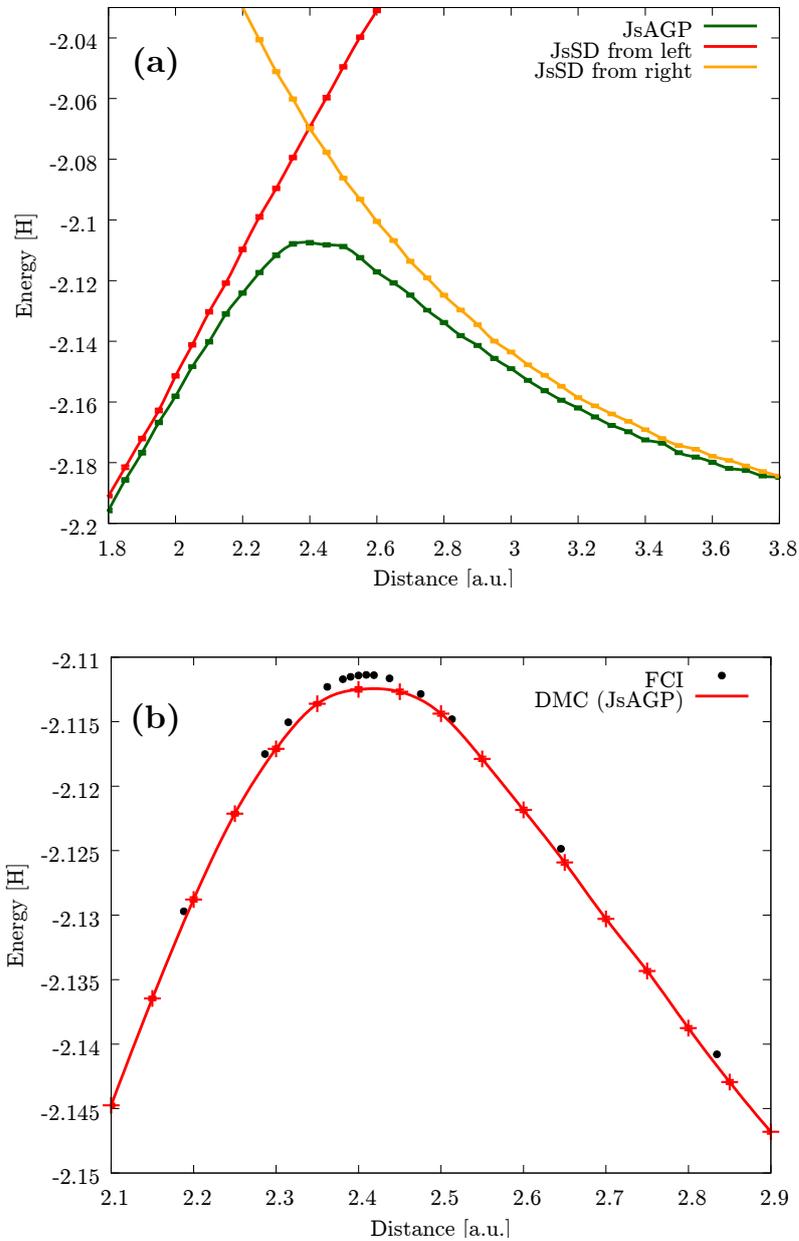


Figure 4.2: Energy comparisons between different methods. (a) VMC energies with different *ansatzes*: in orange the energies of the JsSD starting from the calculation at large  $r_x$ , in red the ones starting from the small  $r_x$ , while in green the JsAGP variational energies are reported. (b) Comparison between the DMC energies calculated using the nodes of the JsAGP (ccpVDZ) and the FCI (ccpVQZ) calculation. On this scale the error bars of the DMC calculations are not visible. From Ref.[13].

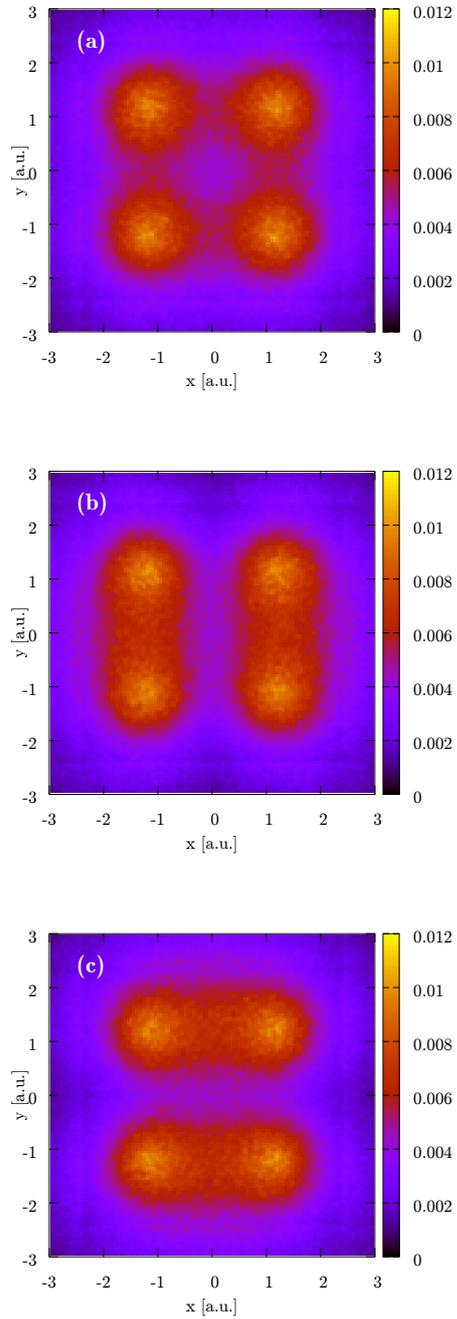


Figure 4.3: Charge density on the  $xy$  plane of the systems with square geometry.. In panel (a) the density obtained with the JsAGP WF, in the panel (b) the density of the JsSD optimized from  $r_x > r_y$ , while in the panel (c) the one of the JsSD from  $r_x < r_y$ . From Ref.[13].

Table 4.2: Difference between the energies calculated with the DMC performed using the nodes of the JsAGP, the ones of the CAS(4,4) and the FCI [14]. All the energies are expressed in Hartree.

$r_x$	JsAGP	CAS(4,4)	FCI
2.188	$-2.1307 \pm 0.0001$	$-2.13033 \pm 0.00010$	-2.1297
2.400	$-2.1125 \pm 0.0002$	$-2.11193 \pm 0.00005$	-2.1114
2.646	$-2.1257 \pm 0.0001$	$-2.12558 \pm 0.00003$	-2.1248

large extension of the basis set. It is also worth noticing that we obtain a higher gain in the region  $r_x \approx r_y$  where the RVB picture is more relevant.

## 4.2 The role of the Jastrow Factor

Thanks to the simplicity of the  $H_4$  molecule, and the limited number of the WFs variational parameters, we used this model to study the genuine AGP without any JF. This case is particularly difficult with our stochastic optimization method because the statistical fluctuations of the energy are much larger compared to the JsAGP case. In principle the AGP should be able to describe the static correlation of this molecule also without JF, with the two main contributions in the WF as in Fig. 4.3a. At the variational level a much worse energy for the AGP WF is expected because the correlation described by the JF is very important. However, it is very interesting to observe that the DMC results are significantly different with (JsAGP) or without (AGP) JF, even considering that the  $JF > 0$  cannot change the signs of the WF, and only the optimization of the AGP in presence of the JF leads to a very accurate nodal surface. In Fig. 4.4a we can see that the variational energies of the AGP WF are indeed considerably higher compared to the JsAGP ones. The smoothness of the curve and the reproducibility of the results indicate that the optimization is not stuck in spurious local minima. Instead the DMC

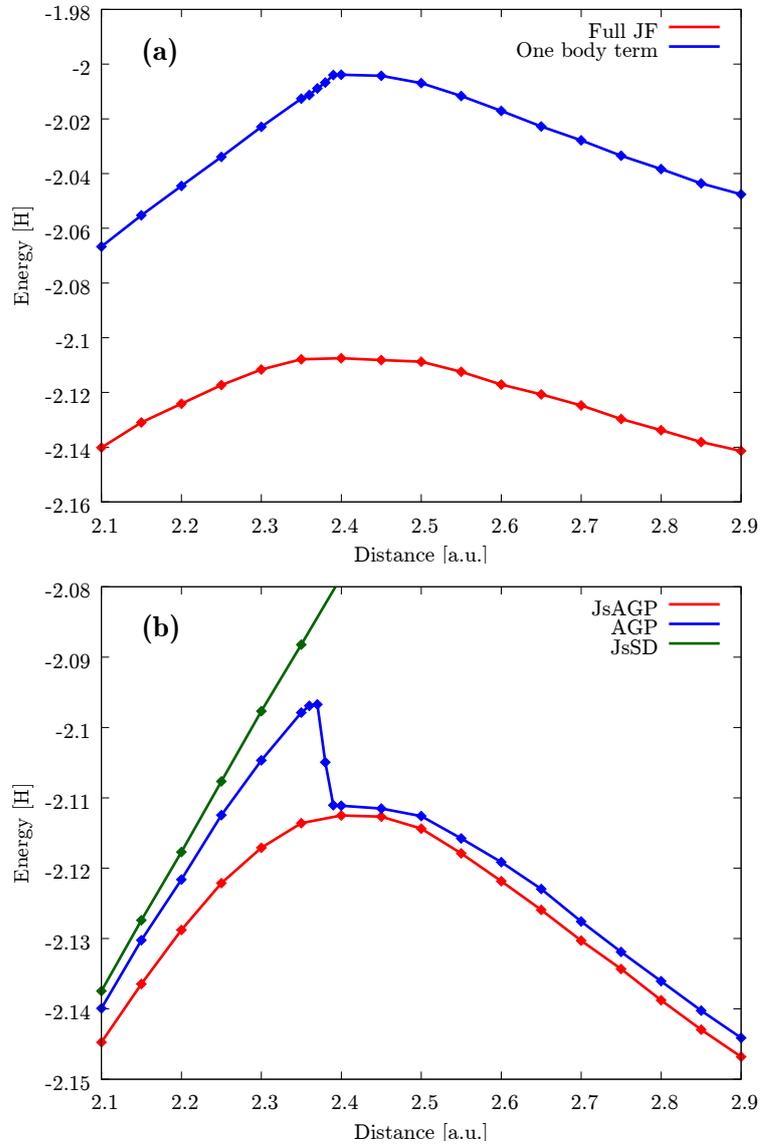


Figure 4.4: AGP and JsAGP energies. In panel (a) we compare the values at the VMC level, while in panel (b) the corresponding DMC energies, within fixed node approximation, are shown. From Ref.[13].

results shown in Fig. 4.4b indicate an unphysical jump of the energy between two different phases. When  $r_x \geq r_y$  the AGP is able to give very good energies that differ only a few mH from the JsAGP ones. Instead, when  $r_x < r_y$  we can see a clear jump in the energy indicating that the nodal surface of the

WF is not correctly described by the AGP. However, also in this regime the nodes are still better than the ones provided by the JsSD WF, with energy values between the ones of the JsSD and the JsAGP. In order to check that this transition was not due to some optimization error we have calculated the WFs for  $r_x < r_y$  starting from the one previously obtained for  $r_x = r_y$ , yielding exactly the same VMC and DMC results. Qualitatively speaking, when the AGP is optimized in presence of the JF, it can resonate between the correct configurations by avoiding double occupancies of singlet electron pairs [87, 88], that are energetically unfavorable. In some sense the Jastrow correlation drives the optimization of the AGP toward the correct ground state energy and the corresponding nodal surface.

Finally, as we can see from table (4.1), the JsAGP is almost converged to the complete basis set limit with only the double zeta cc-pVDZ basis. The differences in energy with the cc-pVTZ are much below one mH per atom. This fast convergence is due to the term in the Eq. (2.70) that fulfills the electron-ion cusp conditions and allows us to use a very small basis set to describe the system. In the AGP the number of variational parameters scales with the square of the number of elements of the basis. It is therefore very important to reach a very accurate description with the smallest possible basis set. This can have a very dramatic impact for large systems where the dimension of the basis set is one of the most important bottlenecks of our JsAGP calculations.

# Chapter 5

## Atoms, Dimers and Molecules

In this chapter I will discuss and compare the accuracy of the different *ansatze* on the atoms and dimers of the first row of the periodic table (Lithium, Beryllium, Boron, Carbon, Nitrogen, Oxygen and Fluorine) and on the benzene molecule and dimers. The results in this chapter are based on the content of Ref.[28, 33, 29, 34]. In every numerical technique designed to deal with a reasonably large number of electrons (QMC included), the exact value of the total energy cannot be reached in practice. However, one can hope that the approximations that are involved do not have any impact on the chemical properties of the system, as their effects are consistent for the atoms and the molecules, yielding an error cancellation that allows the accurate description of the chemical bonds.

In table 5.1 and Fig. 5.1 it is possible to observe the binding energies calculated using the DMC with different guiding function *ansatze*. In particular we compare our JsAGP and JPf to the JSD and the JFVCAS. There is a clear general trend with the JPf improving upon the JsAGP and the JSD results, showing an accuracy comparable with the multi-determinant JFVCAS WF. Beryllium and Boron are two exceptions. For the Beryllium dimer it is possible to guess that the JSD accuracy is due to a "lucky" error cancellation since the JSD is even more accurate than the JFVCAS. For the Boron case we suspect a similar scenario even if the large uncertainty on the exact value (see Tab. 5.1)

Table 5.1: Binding energy of the first-row dimers calculated with DMC and different types of *ansatze* as a guiding function, at the experimental bond length[16].

Binding energy (eV)	Li <sub>2</sub>	Be <sub>2</sub>	B <sub>2</sub>	C <sub>2</sub>	N <sub>2</sub>	O <sub>2</sub>	F <sub>2</sub>
JSD	0.976(6)	0.144(7)	2.83(1)	5.74(2)	9.67(1)	4.94(3)	1.27(1)
JsAGP	0.9812(12)	-0.0270(12)	2.66(1)	6.01(1)	9.88(1)	5.06(2)	1.56(1)
JPf	1.0580(12)	0.0304(22)	2.75(1)	6.31(1)	9.933(6)	5.127(5)	1.64(1)
JFVCAS	1.0465(6) <sup>a</sup>	0.0767(8) <sup>a</sup>	2.906(3) <sup>a</sup>	6.482(3) <sup>a</sup>	10.037(3) <sup>a</sup>	5.187(5) <sup>a</sup>	1.693(5) <sup>a</sup>
Est. exact	1.06(4) <sup>b</sup>	0.1153(3) <sup>c</sup>	2.91(6) <sup>d</sup>	6.44(2) <sup>e,f</sup>	9.902(3) <sup>e</sup>	5.233(3) <sup>e</sup>	1.693(5) <sup>e</sup>

<sup>a</sup> Reference [16].

<sup>b</sup> Reference [89].

<sup>c</sup> Reference [90].

<sup>d</sup> Reference [91].

<sup>e</sup> Reference [92].

<sup>f</sup> A more recent estimate yields 6.39eV (Cyrus Umrigar, private communication).

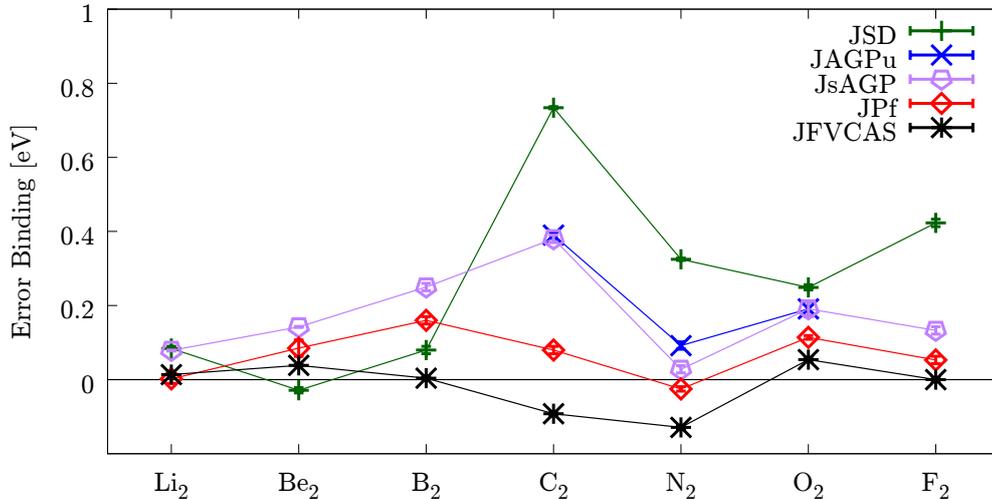


Figure 5.1: Comparison between the different DMC binding energies obtained with different WFs. The JFVCAS results are taken from literature[16].

does not allow a definite conclusion on the scale of the energy differences in play. In any case it is remarkable that all the JPf calculations yield error always below 0.2 eV.

In the following section we will examine in better detail three of these dimers that have particular magnetic properties. In the carbon and nitrogen dimers two atoms of spin 1 and 3/2 respectively combine into a singlet, while two spin 1 oxygen atoms combine themselves into a dimer that has also spin 1. We will see the importance of the spin fluctuations described by the JPf WF in all these cases. Moreover, due to the magnetic interaction, the JsAGP and the JsSD *ansatz*, that are constrained to orient the magnetic moments along the same direction of the  $z$  quantization axis, are not size consistent. This means that they do not recover the spin and the energy of two independent atoms when the dimer is stretched and the atoms are at large distance. The JPf will also provide some insight into the nature of the carbon dimer chemical bond. In this dimer any type of mean-field approach, such as HF or DFT, is completely off with errors of the order of the eV. Also highly correlated methods, such

Table 5.2: Spin measures with different WFs for the Carbon atom and dimer at VMC level.

	$S^2$		$2\mu_B$
	Atom	Molecule	Moment $\parallel z$
JsAGP	2.00	0.00	0.0005(4)
JAGPu	2.00534(3)	0.1743(5)	0.5833(4)
JsAGP	2.00418(5)	0.2880(4)	0.7194(4)
JAGP	2.00542(1)	0.0327(1)	0.0013(5)
Exact	2.00	0.00	-

as coupled cluster [93, 94, 95], face severe difficulties in describing its ground state properties, so that highly involved multi-configuration expansions [95, 16, 96] are often adopted. Recently, Shaik *et al.* have proposed that a fourth bond is necessary[97] to explain the C<sub>2</sub> spectrum at low energy. This result was rather surprising, especially considering that quadruple bonds should very rarely occur [98, 99, 100, 101, 15].

In the last section of this chapter we will also discuss the case of the benzene molecule, a system that represents the prototypical example of the RVB theory and thus a fundamental test case for our approach. Finally, we will discuss the benzene dimer in the two parallel displaced (PD) and T shape (T) configurations.

## 5.1 Magnetic Dimers

### 5.1.1 Carbon

Carbon dimer is probably the most interesting example among these three dimers. A full understanding of the behavior of the carbon-carbon interaction is indeed still missing and the bond order of this molecule is still under debate, with Shaik *et al.* that have recently proposed the existence of a quadruple bond

Table 5.3: Carbon Energies. The JsAGP, JAGPu and JPf results are calculated with an optimized ccpVTZ basis set.

<b>Carbon</b>			
	Atom	Molecule	Binding
Source	Energy[ $H$ ]	Energy[ $H$ ]	Energy[ $eV$ ]
JSD	-37.81705(6) <sup>a</sup>	-75.8088(5) <sup>a</sup>	4.75(1) <sup>a</sup>
JFVCAS	-37.82607(5) <sup>a</sup>	-75.8862(2) <sup>a</sup>	6.369(6) <sup>a</sup>
JsAGP	-37.8243(1)	-75.8611(2)	5.78(1)
JAGPu	-37.8263(1)	-75.8706(2)	5.93(1)
JPf	-37.827965(3)	-75.88650(4)	6.274(3)
JSD (DMC)	-37.82966(4) <sup>a</sup>	-75.8672(1) <sup>a</sup>	5.656(3) <sup>a</sup>
JFVCAS (DMC)	-37.83620(1) <sup>a</sup>	-75.9106(1) <sup>a</sup>	6.482(3) <sup>a</sup>
JsAGP (DMC)	-37.8364(1)	-75.8938(2)	6.01(1)
JAGPu (DMC)	-37.8364(1)	-75.8935(2)	6.00(1)
JPf (DMC)	-37.8363(1)	-75.9045(2)	6.31(1)
Estimated Exact	-37.8450 <sup>b</sup>	-75.9265 <sup>c</sup>	6.44(2) <sup>c,d</sup>

<sup>a</sup> Reference [16].

<sup>b</sup> Reference [102].

<sup>c</sup> Reference [92].

<sup>d</sup> A more recent estimate yields  $6.39eV$   
(Cyrus Umrigar, private communication).

for this dimer[97]. Within a correlated RVB approach, they have found that the  $2\sigma$  and  $3\sigma$  molecular orbitals, after s-p hybridization, change their nature as compared to standard molecular orbital theory and show a corresponding bonding character. By taking into account the remaining two  $\pi$  orbitals, obviously bonding, they argued that an unexpected quadruple bond should be a more appropriate description of the  $C_2$  molecule.

The carbon atoms have spin triplet electronic configurations, and their mutual interaction leads to a singlet molecule. As we can see from Fig. 5.1 and table 5.3, the JPf not only improves the results of the JSD WF, but

remarkably also the description given by the JsAGP and JAGPu. The huge difference between the multi-determinant expansion JFVCAS and the JSD binding energies helps to quantify the effect of the multi determinantal nature of this molecule, and this makes even more surprising the quality of the results obtained with a single JPf WF that, with a computational cost comparable to a SD, is already very close to the exact value.

The explanation for the impressive improvement of the binding energy from JsAGP and JAGPu to JPf resides on the description of the strong spin fluctuations in this molecule. The JPf gives a very accurate picture of its magnetic properties as we can see from table 5.2, giving results very close to  $S^2 = 2$  for the atom and  $S^2 = 0$  for the molecule. Conversely, by using the JsPf (the Pf without spin dependent JF) and the JAGPu, we cannot recover the singlet from the broken symmetry initialization. Interestingly, as expected, the molecule does not have any magnetic moment on the  $z$  direction, because it is an almost perfect singlet. The atomic spins, localized around each atom, point in opposite directions in order to form the singlet molecular state. Since there is no magnetic moment along  $z$  we can measure its magnetic moment only by separately evaluating the  $S^2$  in the two semi-infinite regions, each one containing a single atom, separated by a plane perpendicular to the molecular axis and at the same distance from the two atoms. In Fig. 5.3 we show that even at bond distance there is a very strong magnetic moment around the atoms and, in this way, we can explain the strong effect of the zero point energy of the spin fluctuations described by the JPf. This confirms the picture that the  $C_2$  molecule can be considered as the smallest antiferromagnet made of two atoms with opposite spins.

It is particularly instructive to see the role of correlation in modifying the molecular orbitals, by taking as a reference the ones corresponding to a simple DFT double bond picture. The Pf part of the JPf WF, after full optimization of the energy in presence of the JF, can be also recasted in terms of MOs

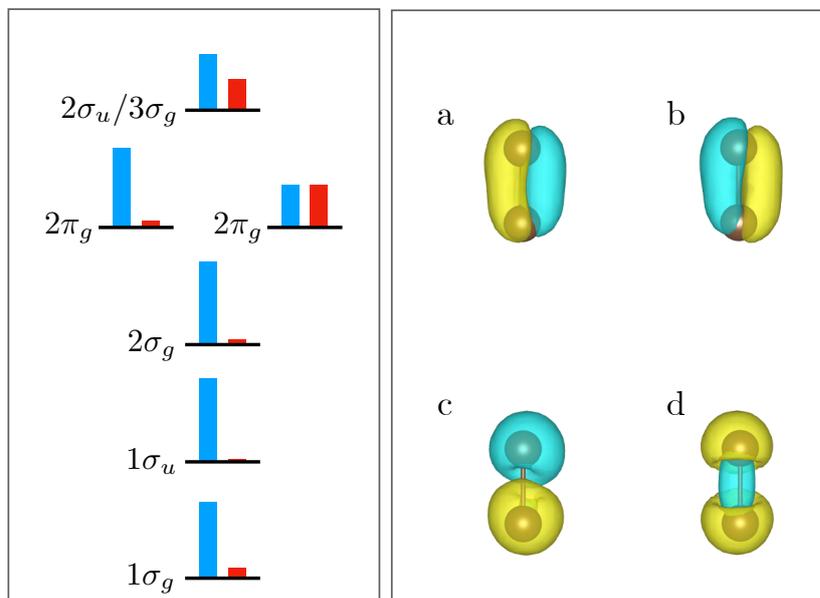


Figure 5.2: Left panel: after breaking the spin symmetry, each spin-independent orbital of an unrestricted SD *ansatz* splits into a pair of single occupied ones with no definite spin projection. The histograms represent the corresponding spin component weights: the height of the blue (red) rectangle indicates the percentage of the majority (minority) spin. Notice that the occupation order is different in DFT calculations where the order is  $1\sigma_g/1\sigma_u/2\sigma_g/2\sigma_u/2 \times (2\pi_g)$  and where the orbitals have a single spin component. Right panel: majority and minority spin orbitals with the most relevant spin contamination. The orbitals (a) and (b) come from the  $2\pi_g$  and have the same weight, the orbitals (c) and (d) are the last occupied ones (indicated as  $2\sigma_u/3\sigma_g$ ), in this case the (c) orbital has a 65% and the (d) one a 35% weight. From Ref.[33].

as explained in section 2.2.7. The spin character of the resulting orbitals is displayed in Fig. 5.2, where it is shown that i) at variance of the DFT mean-field *ansatz*, the  $3\sigma_g$  bonding orbital is eventually present as the minority spin component of the highest occupied molecular orbital (HOMO), in partial agreement with the quadruple bond picture. ii) in our *ansatz* however the main effect that mostly determines the chemical bond is the spin contamination of the most important occupied orbitals  $2\sigma_u, 3\sigma_g$  and  $2\pi'_g$ s. In this way they

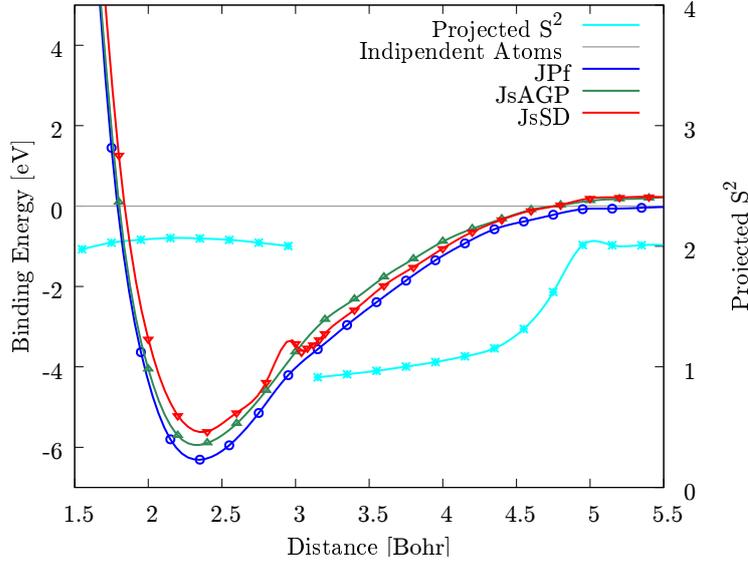


Figure 5.3: DMC energy dispersion of the carbon dimer: only the JPf allows the system to be size consistent at large distance, namely it is able to recover the energy and the expectation value of the  $S^2$  operator of two isolated atoms. At bond distance however the carbon atoms maintain a large value of  $S^2$ . The sharp change of the projected  $S^2$  value at around 3 *a.u.* is probably due to an avoided crossing of two energy levels belonging to the same irreducible representation, in agreement with DMRG[103]. Within LSDA this effect is reproduced by a discontinuous change in the occupation of the  $\pi$  orbitals in the corresponding Slater determinant. Lines are guides to the eye. From Ref.[28].

can contribute to the bonding by means of the corresponding spin-fluctuation energy gain, that is instead vanishing for the inner core orbitals (this is because they have a definite spin in the same quantization axis chosen for the Jastrow). This shows therefore that the bonding in  $C_2$  cannot be explained with charge electrostatic, and instead the large atomic spin value confirms that the energy is intimately due to correlation, the same that allows, by means of the JF, the evaluation of the spin-wave zero point energy of a quantum antiferromagnet.

Moreover Fig. 5.3 shows that only with the JPf WF we have a size consistent solution with the molecule that recovers the energy of two independent atoms at large distance. This feature is fundamental if we want to use this WF

to describe chemical reactions and perform large scale simulations, with a size consistent behavior at large distances. The importance of the variational optimization of the wave function is particularly evident in this small molecule. With the standard approach, by applying DMC to a SD taken by DFT (here obtained with Purdue and Zunger LDA[104]), a level crossing in the occupation of the  $\pi$  molecular orbitals occurs at around 3 *Bohr* distance, above which the  $\pi$  bonding orbitals are only partially occupied. This implies clear artifacts in the DMC energies. We have verified that this level crossing is reproduced with a standard DFT-LDA calculation by Gaussian16 A.03 revision [105] and an almost converged basis set (the standard cc-pVQZ). The level crossing has also been observed in Ref. [106]. In our variational optimization instead, we have verified that it is important to start at large distance with the WF predicted by LSDA, otherwise a sizeably higher energy is obtained. This effect is reflected also by the sharp change of the projected  $S^2$  at around 3 *Bohr* distance (see Fig. 5.3), that could be compatible with an avoided crossing between two energy levels belonging to the same  $^1\Sigma_g^+$  representation[103].

We further compared the JPF carbon energy dispersion with unrestricted single reference coupled cluster (UCCSD-T), and DMRG, heat-bath configuration interaction (HCI) and FCI from literature[103, 107, 96]. U-CCSD(T) calculations were performed using Gaussian16 A.03 revision with the counterpoise correction, with the frozen-core approximation and the full-core correlation [105]. Table 5.4 and Fig. 5.4 show that there are significant discrepancies between different methods in the carbon dimer dispersion curve at large distances. However, even in a quadruple zeta basis the FCI binding energy  $D_e = 6.22eV$ [96] is about 6mH lower than the estimated exact one. Therefore, if we reference all the curves at the bond length minimum energy, as reported in the mentioned figure, a method that is supposed to be weakly dependent on the basis, as our DMC, should be slightly higher in energy at large distance, provided it remains close to the exact dispersion energy curve. Moreover there

Table 5.4: Carbon Energy Dispersion [Hartree]. The JPf results were obtained with the optimized ccpVDZ basis set, the DMRG results with the ccpVQZ basis, the HCl with ccpV5Z basis set, the FCI with ccpVQZ basis set, whereas the UCCSD-T ones, both full and frozen core, are shown for ccpV5Z basis sets.

Distance	Numerical Technique						
	JPf (DMC)	DMRG	HCl	UCCSD-T <sub>frozen</sub>	UCCSD-T <sub>full</sub>	FCI	
2.0787	-75.86652(3)	-75.76125 <sup>b</sup>	-75.76701 <sup>c</sup>	-75.76085	-75.78683	-75.7624 <sup>d</sup>	
2.2677	-75.90207(3)	-75.79924 <sup>b</sup>	-75.80461 <sup>c</sup>	-75.78450	-75.80878	-75.7987 <sup>d</sup>	
2.3480	-75.90456(3)	-75.80269 <sup>b</sup>	-75.80786 <sup>a,c</sup>	-75.78370	-75.80754	-75.8025 <sup>d</sup>	
2.4566	-75.90008(3)	-75.79937 <sup>b</sup>	-75.80444 <sup>c</sup>	-75.77928	-75.80247	-75.7993 <sup>d</sup>	
2.6456	-75.87825(4)	-75.77937 <sup>b</sup>	-75.78460 <sup>c</sup>	-75.76465	-75.78664	-75.7798 <sup>d</sup>	
3.0235	-75.81700(8)	-75.72405 <sup>b</sup>	-75.72895 <sup>c</sup>	-75.71762	-75.73765	-75.7243 <sup>d</sup>	
3.7794	-75.73649(8)	-75.64560 <sup>b</sup>	-75.65043 <sup>a,c</sup>	-75.62162	-75.63996	-75.6454 <sup>d</sup>	

<sup>a</sup> Interpolated.

<sup>b</sup> Reference [103].

<sup>c</sup> Reference [107].

<sup>d</sup> Reference [96].

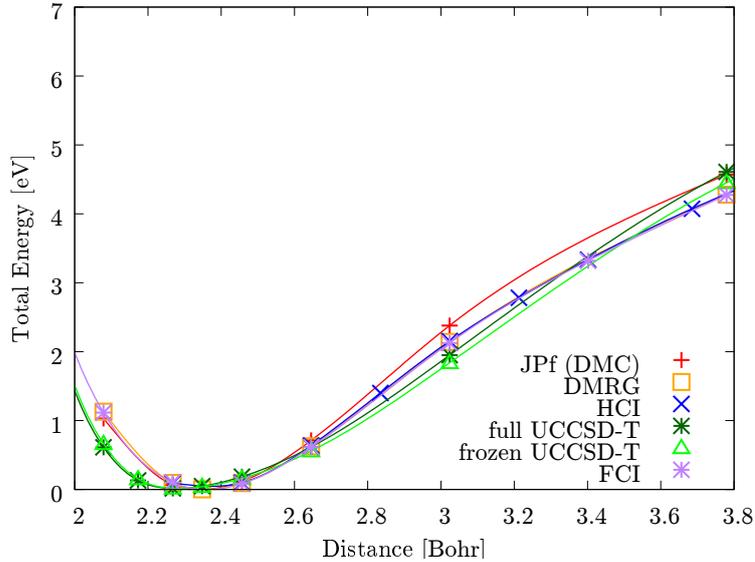


Figure 5.4: Energy dispersion of the carbon dimer calculated with JPf (DMC), UCCSD-T (ccpV5Z), DMRG[103], HCl[107], and FCI[96]. Lines are guides to the eye. From Ref.[28].

may be sizeable corrections due to the frozen core approximation employed by DMRG, HCl, and FCI. We have indeed verified that they are non negligible in the UCCSD-T calculation, implying that core-valence interaction can lead to a further non-parallelity error of about 3mH (see Fig. 5.4). Core-valence interaction is considered in DMC calculations simply because, within this technique, it is not possible to employ the frozen core approximation. Nevertheless, it is clear that our results may have some error, but it is remarkable that if we use the corresponding energy values for computing the zero point energy (ZPE) of the dimer we find excellent agreement with the experimental value, given by 0.1146 eV[108]. Indeed the ZPE calculated values, using a standard fit with a quartic polynomial close to the equilibrium distance, are 0.1153(6) eV, 0.108 eV, 0.106 eV, 0.112 eV, 0.114 eV and 1133(3) eV for DMC, UCCSD-T full core and frozen core, DMRG, HCl, and FCI, respectively. In summary, by taking into account all possible sources of error, we believe that our results are in reasonable agreement with the expected "exact result" converged in the com-

plete basis set limit and with full core-valence interaction taken into account. Indeed we believe that only a more direct comparison with experiments or a full core FCI/DMRG or HCI extrapolated to the complete basis set limit can further improve the accuracy of the dispersion curve.

### 5.1.2 Nitrogen

Nitrogen is in some sense similar to the carbon case: also its dimer is indeed a singlet formed by two large spin ( $3/2$ ) atoms.

As we can notice from Fig. 5.1 and table 5.5, at DMC level the JsAGP and JPf are both exact within chemical accuracy. All our calculations compare with the exact result better than the JFVCAS solution. Surprisingly, at VMC level the binding energies calculated with JPf, JsAGP, and JAGPu are also very good.

We remark that a very powerful method, like the recently proposed Fermi Net[109] (a neural network based WF), cannot reach the same precision in the binding energy even if the total energies of the molecule and atom are the best available ones. This clearly shows that all our *ansatze* allow a remarkable cancellation of errors when computing the total energy differences between the molecule and the two independent atoms.

In this case, however, the difference between JPf and JsAGP/JAGPu is much smaller than in the previous case. This should be related to a less important role of the spin fluctuations and also to a smaller magnetic moment of the atoms at equilibrium distance. By repeating the reasoning done for the carbon dimer, we can quantify the magnetic moment from the  $S^2$  value in the semi-infinite region separated by a plane perpendicular to the axis of the molecule and equidistant from the atoms. Looking at Fig. 5.5 we can see that, at bond distance, the  $S^2$  of the atom is much smaller than the one of an independent atom and therefore, even if the nitrogen atom has a large spin, when it is forming a dimer it does not give rise to strong antiferromagnetism.

Table 5.5: Nitrogen Energies. The JsAGP, JAGPu and JPf results are calculated with an optimized ccpVTZ basis set.

<b>Nitrogen</b>			
	Atom	Molecule	Binding
Source	Energy[ $H$ ]	Energy[ $H$ ]	Energy[ $eV$ ]
JSD	-54.5628(1) <sup>a</sup>	-109.4520(5) <sup>a</sup>	8.88(1) <sup>a</sup>
JFVCAS	-	-109.4851(3) <sup>a</sup>	9.78(1) <sup>a</sup>
JsAGP	-54.55794(6)	-109.4781(7)	9.856(3)
JAGPu	-54.55998(5)	-109.48155(7)	9.840(3)
JPf	-54.56633(5)	-109.49226(7)	9.785(3)
JSD (DMC)	-54.57587(4) <sup>a</sup>	-109.5039(1) <sup>a</sup>	9.583(3) <sup>a</sup>
JFVCAS (DMC)	-	-109.5206(1) <sup>a</sup>	10.037(3) <sup>a</sup>
JsAGP (DMC)	-54.5765(1)	-109.5164(2)	9.88(1)
JAGPu (DMC)	-54.5767(3)	-109.5140(2)	9.81(1)
JPf (DMC)	-54.57709(9)	-109.5192(1)	9.933(6)
Fermi Net	-54.58882(6) <sup>b</sup>	-109.5388(1) <sup>b</sup>	9.828(5) <sup>b</sup>
Estimated Exact	-54.5892 <sup>c</sup>	-109.5427 <sup>d</sup>	9.908(3) <sup>d</sup>

<sup>a</sup> Reference [16].

<sup>b</sup> Reference [109].

<sup>c</sup> Reference [102].

<sup>d</sup> Reference [92].

Also in this case it is important to notice that the JPf solution is size consistent both in energy and spin. Despite the very good description at bond distance provided by the JsAGP, we notice from Fig. 5.5 that it is not perfectly size consistent. Within our approach a fully consistent picture and a very accurate dispersion are possible only by means of the JPf *ansatz*, that is able to work properly also in the strong correlation regime, namely at large interatomic distance.

We also compared in table 5.6 the JPf nitrogen energy dispersion with

Table 5.6: Nitrogen Energy Dispersion [Hartree]. The JPf results were obtained with the optimized ccpVDZ basis set, the DMRG and MRCC results with the ccpVDZ basis, whereas the corresponding UCCSD-T ones are shown also for a much larger basis (ccpV5Z), resulting in much better agreement with the present DMC results.

Distance	Numerical Technique					
	JAGP (DMC)	DMRG	MRCC	UCCSD-T (DZ)	UCCSD-T (5Z)	
2.118	-109.51694(5)	-109.27833 <sup>b</sup>	-109.27683 <sup>b</sup>	-109.27652	-109.41303 <sup>a</sup>	
2.4	-109.46459(6)	-109.23838 <sup>b</sup>	-109.23687 <sup>b</sup>	-109.23202	-109.35926	
2.7	-109.37935(6)	-109.16029 <sup>b</sup>	-109.15895 <sup>b</sup>	-109.14731	-109.26936	
3.0	-109.29961(6)	-109.08619 <sup>b</sup>	-109.08442 <sup>b</sup>	-109.06570	-109.18331	
3.6	-109.19745(6)	-108.99489 <sup>b</sup>	-108.99272 <sup>b</sup>	-108.97982	-109.08833	
4.2	-109.16376(7)	-	-108.96471 <sup>b</sup>	-108.96002	-109.06204	

<sup>a</sup> Interpolated.

<sup>b</sup> Reference [110].

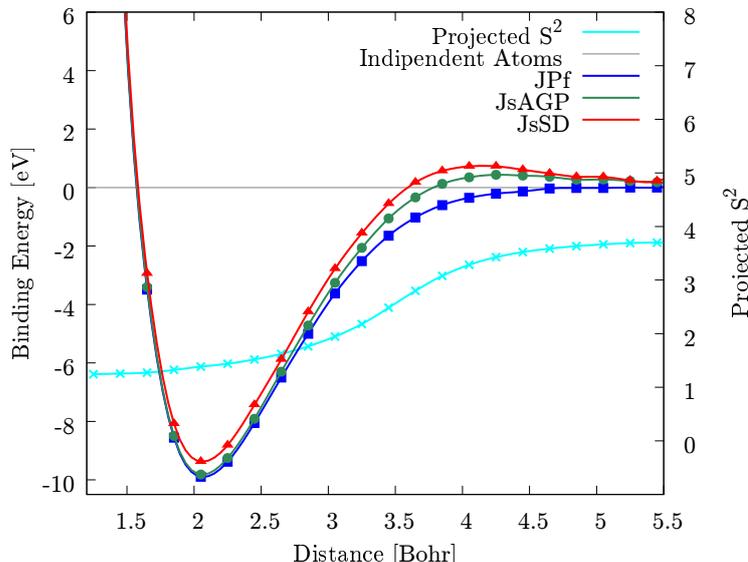


Figure 5.5: DMC energy dispersion of the nitrogen dimer: only the JPf appears to be perfectly size consistent, thus recovering at large interatomic distance the energy and the expectation value of the  $S^2$  operator of two isolated atoms. At bond distance however the nitrogen atoms have a smaller value of  $S^2$ , in contrast to what is observed for the carbon dimer. Lines are guides to the eye. From Ref.[28].

unrestricted single reference coupled cluster (UCCSD-T) with ccpVDZ and ccpV5Z basis sets, multi-reference coupled cluster (MRCC), and DMRG[110]. U-CCSD(T) calculations were performed using Gaussian16 A.03 revision with the counterpoise correction, with the frozen-core approximation and the full-core correlation [105]. The DMC dispersion is in excellent agreement with UCCSD-T calculations in ccpV5Z basis set, while DMRG and MRCC suffer from a large basis set error compatible with the UCCSD-T one in ccpVDZ.

### 5.1.3 Oxygen

The oxygen is very different from the previous cases but nevertheless very interesting for different reasons. The oxygen dimer consists of two triplet atoms, but this time the molecule is a triplet. There are small atomic magnetic moments in the GS of the oxygen molecule, but the role of the magnetic

Table 5.7: Oxygen Energies. The JsAGP, JAGPu and JPf results are calculated with an optimized ccpVTZ basis set.

<b>Oxygen</b>			
	Atom	Molecule	Binding
Source	Energy[ $H$ ]	Energy[ $H$ ]	Energy[ $eV$ ]
JSD	-75.0352(1) <sup>a</sup>	-150.2248(5) <sup>a</sup>	4.20(1) <sup>a</sup>
JFVCAS	-	-150.2436(2) <sup>a</sup>	4.713(8) <sup>a</sup>
JsAGP	-75.0268(3)	-150.2372(6)	5.00(3)
JAGPu	-75.0339(3)	-150.2503(5)	4.97(3)
JPf	-75.0346(2)	-150.2572(4)	5.11(2)
JSD (DMC)	-75.05187(7) <sup>a</sup>	-150.2872(2) <sup>a</sup>	4.992(7) <sup>a</sup>
JFVCAS (DMC)	-	-150.29437(9) <sup>a</sup>	5.187(5) <sup>a</sup>
JsAGP (DMC)	-75.0518(3)	-150.2894(3)	5.06(2)
JAGPu (DMC)	-75.0519(3)	-150.2902(4)	5.06(2)
JPf (DMC)	-75.05289(7)	-150.2942(1)	5.127(5)
Estimated Exact	-75.0673 <sup>b</sup>	-150.3724 <sup>c</sup>	5.241 <sup>c</sup>

<sup>a</sup> Reference [16].

<sup>b</sup> Reference [102].

<sup>c</sup> Reference [92].

interaction remains important, as shown by the application of the JPf *ansatz*. In this case it looks that the interaction of parallel spins electrons is particularly important, and this can be described by the JPf *ansatz* more accurately than the corresponding JsAGP and JAGPu ones. Thus we expect to recover with the JPf some correlation that we miss when we simplify the *ansatz* by using the unpaired orbitals in the JsAGP and in the JAGPu WFs.

By looking at Fig. 5.1 and table 5.7 we can see that, at DMC level, the energies obtained with the JPf WF are extremely good even for the oxygen dimer. In this case the correct description of the triplet pairing correlations, possible within the JPf *ansatz*, appears to be fundamental. Indeed the final

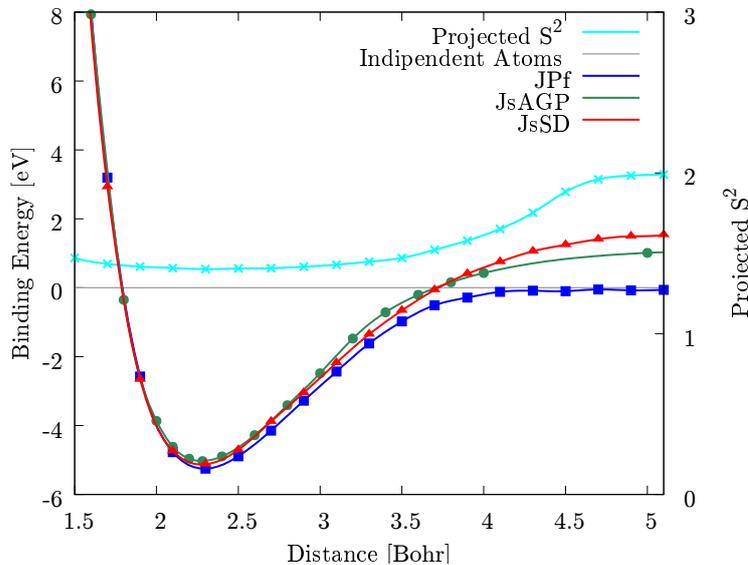


Figure 5.6: DMC energy dispersion of the oxygen dimer with the JPf, JsAGP [111] and JSD (with the SD obtained from DFT calculations): at large distance only the JPf WF is size consistent. In the plot also the expectation value of the projected  $S^2$  operator on the atoms for the JPf that recovers at large distance the value of two isolated atoms. Lines are guides to the eye. From Ref.[28].

result is so accurate that the binding energy is comparable to the one obtained with the multi-determinant JFVCAS WF. It is even more surprising that the absolute energies of the atom and molecule are very close where not even better than the ones provided by the multi-determinant expansion both at VMC and DMC level. We have to point out, however, that within JFVCAS method it is not possible to improve the JSD atom [16] and that the binding energy slightly better than the JPf one derives from the poorer quality of the atom rather than a better description of the molecule.

The problem of the size consistency for the oxygen dimer is absolutely non trivial and even more complicated than the previous cases. Starting from bond distance we have a molecule of spin one and fixed projection  $S_z = 1$  but we have to recover the behavior of two independent atoms. At large distance this means that, by keeping the projection  $S_z = 1$  constant, while separating the

atoms far apart, we have to recover the correct atoms of spin one and thus we need to have one atom with  $S_z = 0$ . This is impossible for the JsAGP and the JAGPu but allowed by the JPf, a remarkable and absolutely non trivial feature of this WF. As we can see from Fig. 5.6, at large distance only with the JPf the system recovers the energy and the spins of the independent atoms, showing that, by means of our advanced optimization tools, it is possible to dramatically change the WF up to the point of rotating completely the spin of an atom.

## 5.2 Benzene

The benzene molecule represents one of the most successful example of the RVB theory with the carbon-carbon bonds resonating among several valence bond configurations. QMC methods are able to provide a very good description of this important molecule [112, 40], and thus it is interesting to check whether, with our new approach, we can obtain a very accurate result. In particular in table 5.8 we compare the results obtained by JsSD, JsAGP, JAGPu and JPf WFs, showing that all the results obtained with a pairing function (from JsAGP to JPf) provide a very good estimate of the absolute energies, noticeably improving the results of the JsSD. Moreover the corresponding atomization energies are extremely accurate at the DMC level, whereas the JsSD largely overestimates it. It is finally interesting to notice that, even if there is a sizeable gain in terms of absolute energy with the JPf, the atomization energy does not change, suggesting that this could be almost converged to the exact value, therefore in slightly disagreement with Ref. [113]. This might be in principle explained because, at present, the accuracy of the state of the art "estimated exact" calculation is probably not enough to establish an energy difference  $\ll 0.1eV$ . For instance the zero point energy (ZPE) has been estimated by DFT[113] and some work is certainly necessary to clarify this issue,

Table 5.8: Benzene Energies.

<b>Benzene</b>			
	C atom <sup>a</sup>	Molecule	Atomization Energy
Source	Energy[ <i>H</i> ]	Energy[ <i>H</i> ]	Energy[ <i>eV</i> ]
JsSD	-37.8074(1)	-232.0261(3)	59.37(1)
JsAGP	-37.82383(4)	-232.0805(3)	58.166(8)
JAGPu	-37.82651(5)	-232.0900(3)	57.986(8)
JPf	-37.82921(4)	-232.1060(2)	57.982(7)
JsSD(DMC)	-37.8299(1)	-232.1879(6)	60.09(2)
JsAGP(DMC)	-37.8368(1)	-232.1947(6)	59.16(2)
JAGPu(DMC)	-37.8367(1)	-232.1943(6)	59.16(2)
JPf(DMC)	-37.83751(9)	-232.1998(5)	59.18(2)
Estimated Exact	-37.8450 <sup>b</sup>	-232.250(1)	59.32(2) <sup>c</sup>

<sup>a</sup> Calculated with the same basis set used for the benzene molecule.

<sup>b</sup> Reference [102].

<sup>c</sup> Reference [113].

e.g. by calculating the ZPE directly with QMC.

We remark here that the JsAGP description of the benzene molecule is already very accurate and it is not improved by the JPf. This is probably due to the lack of any sizeable spin moment around any atom composing this molecule. Indeed the  $S^2$  value calculated for the JPf and JAGPu solutions is 0.032(1) and 0.0123(7), respectively, proving that any local magnetic moment is almost completely melted during the optimization, despite its non zero initialization. We conclude therefore that in the benzene molecule the spin fluctuations are not relevant and the use of the Pfaffian leads only to a marginal improvement of the total energy while the molecule is correctly described by a perfect singlet RVB *ansatz* given by the JsAGP, in agreement with the classical RVB picture by L. Pauling[114].

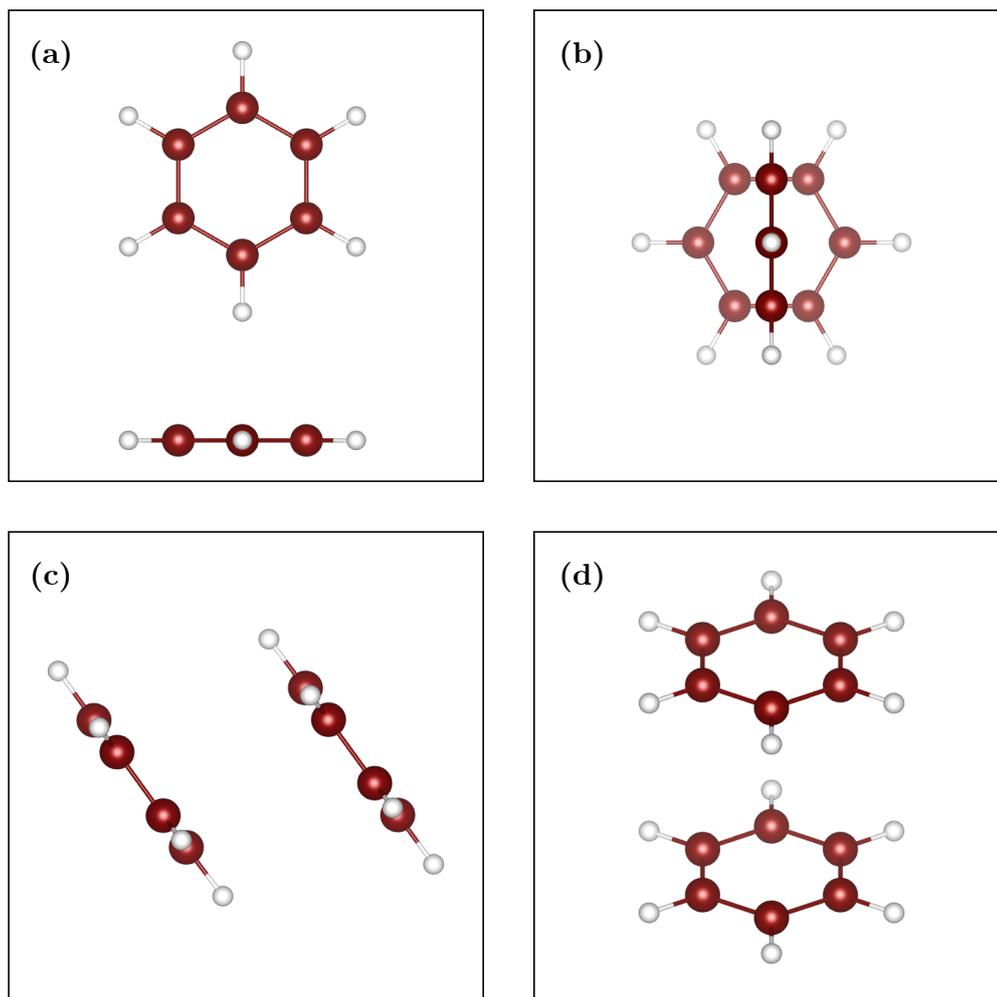


Figure 5.7: Orthogonal projections of the benzene dimers in the T (a and b) and PD (c and d) configurations.

### 5.2.1 The benzene dimer

Motivated by the success in the description of the benzene molecule we studied its dimer interaction. Of course there are many different configurations for the benzene dimer, so we decided to focus on the parallel displaced (PD) and the T shaped (T) configurations, that we can see in Fig. 5.7.

The dimers are bonded by non-covalent interaction and their interaction is weak, in the order of few  $Kcal/mol$ . Unfortunately, it is not possible to

Table 5.9: Benzene dimer.

<b>Benzene dimer binding energies</b>		
	Parallel Displaced	T-shape
Source	Binding Energy [ <i>Kcal/mol</i> ]	Binding Energy [ <i>Kcal/mol</i> ]
JsSD (DMC)	2.2(3)	3.0(3)
JsSD (DMC)	2.0(1) <sup>a</sup>	-
JsAGP (DMC)	3.3(2)	3.0(2)
JsSD (DMC+BF)	2.7(3) <sup>b</sup>	-
CCSD(T)	2.65 <sup>c</sup>	2.72 <sup>c</sup>

<sup>a</sup> Reference [117].

<sup>b</sup> Reference [112].

<sup>c</sup> Reference [115].

estimate experimentally the binding energy due to the many different possible configurations of the dimer and the absence of control on the geometry on this scale. Moreover, compared to the cases of study presented so far, these systems contain a much larger number of electrons and therefore they cannot be studied with multi-determinant methods. For this reason there is not a clear reference for an exact value. The most accurate references for these systems are the CCSD(T)[115] and the DMC with JsSD as a guiding function with the backflow transformation (BF)[112]. Relying on the results on the benzene molecule presented before, we decided to use the JsAGP for this calculation instead of the JPf since we were not expecting significant differences among the two *ansatze*. Moreover, the calculations on these systems (with correlation consistent effective core potential (ccECP)[116] for the carbon atoms and ccpVTZ basis set both for carbon and hydrogen) requires the simultaneous optimization of more than 20000 variational parameters with the JsAGP and more than 40000 with the JPf.

Table 5.9 shows the results from literature and our calculations, including the JsSD results in the same basis set used for the JsAGP. We can see that

the JsAGP coincides with the CCSD(T) results within chemical accuracy. The slightly higher binding energies may actually be the result of a better description of the Van Der Waals interaction. Unfortunately without an exact reference it is impossible to understand which of the two methods is more accurate in this case. The JsSD *ansatz*, instead, clearly underestimates the interaction between two benzene dimers in the PD configuration, yielding slightly wrong nodal surfaces for the DMC calculations and wrong binding energies. This is surprising considering that for non-covalent interaction the DMC with JsSD is supposed to provide a good accuracy in estimating the binding energies due to error cancellation within fixed node approximation[118]. However, it was already known in literature[117] that JsSD underestimates the binding energy for the PD configuration. It is evident that in this case the use of an optimized pairing function *ansatz*, or a more complicated one, is necessary to obtain a correct guiding function and consequently a good evaluation of the binding energies even for this non-covalent interaction.

# Conclusions

In this thesis we have seen and analyzed different possibilities for QMC calculations, with a particular focus on Jastrow correlated pairing function *ansatze*.

Going through the chapters, we have seen that VMC and DMC, even with the use of the standard Jastrow Slater *ansatz* (JsSD), can give insight on non trivial systems like the hydrogen chain, unveiling the metal-insulator phase transition and the dimerization in an excellent qualitative and quantitative agreement with the other state of the art techniques.

We have seen, however, that the use of a JsSD can hide some pitfalls and even in simple systems like the  $(H_2)_2$  it is possible to observe catastrophic failures. This system has been particularly instructive, allowing us to show that the full optimization of the JsAGP *ansatz* guarantees a very fast convergence in the basis set, so that no kind of extrapolation is necessary for almost converged results in the complete basis set limit. This property has also been verified for the atomic dimers where other methods like DMRG, HCI, FCI, and UCCSD-T have shown a critical dependency on the basis set, requiring the use of a much larger basis compared to the ones we have adopted.

Moreover, with the use of the  $(H_2)_2$  model system we were also able to prove that the AGP alone, without the use of a JF, miserably fails, leading not only to inaccurate DMC energies but qualitatively wrong results, as a discontinuity of the energy landscape as a function of the atomic positions was reported. In this case the wrong nodal surface determined by the AGP, was not detectable at the VMC level, because the optimized VMC energy was

indeed a smooth and continuous function of the ion positions, as it should be from general grounds.

The use of a powerful spin and charge dependent JF, together with the advanced optimization techniques, is probably the key to explain the remarkable improvement obtained with the JF *ansatz*, compared to previous attempts[26, 27]. With a computational cost comparable to a SD we were able to improve not only the results achieved with a simple JsSD but also with JsAGP and JAGPu, reaching a level of accuracy comparable to the multi-determinant JFVCAS WF one. In particular we have seen that the JF *ansatz* provides a very accurate description of high spin atoms and their dimers and that it is size consistent. This should increase the number of possible applications, providing a reasonably accurate and computationally feasible tool for studying chemical reactions. By considering the triplet correlations, we can now correctly take into account the zero point energy associated with the well-known spin fluctuations of a quantum magnetic system. Within our Pf WF they are correctly and efficiently taken into account by orienting the atomic magnetic moments in the direction perpendicular to the spin quantization axis chosen for the JF. For this reason we have obtained a very good description of the carbon and nitrogen dimers, remarkably, even when the first molecule was found to be very poorly described by the JsAGP and the JAGPu. Moreover, it is only thanks to the presence of the triplet correlations that we were able to improve the description of the oxygen dimer, a strongly correlated triplet molecule with a highly entangled spin interaction among the atoms. We also demonstrated that for the benzene dimer the JF, JsAGP, and JAGPu can provide very accurate atomization energies, in contrast to the commonly used JsSD (overbinding by more than 0.2 eV). In this case the triplet correlations do not seem to play a crucial role.

The case of the benzene dimer has been also particularly instructive. Through the use of an optimized pairing function it was possible to improve the descrip-

tion of the DMC binding energy for the PD configuration. With the JsSD, even at DMC level, there is a relevant difference between the binding energies of the two considered configurations that is not observed within CCSD(T) or JsAGP calculations. This shines a sinister light on the widespread assumption that the JsSD is able to provide correct nodal surfaces for non-covalent interaction since it is evident that optimization matters even in this case.

It is important to highlight that for all the systems in exam the accuracy in the binding energy is always much better than the accuracy in the total energy and that therefore there exists always a remarkable cancellation of errors in the total energy differences. This feature indeed is fundamental for compact *ansatze* like the JPf and the JsAGP, that, as we have shown, provide accuracy in binding energies better than other very expensive highly correlated methods, like the Fermi Net in the nitrogen dimer, even when these are able to achieve almost exact total energies.

The relatively low computational cost of QMC combined with powerful optimization techniques, allowing a reasonably large number of variational parameters, makes this approach ideal for studying systems even much larger than the ones considered in this work. Indeed, the paradigm presented in this thesis could represent in the future a very powerful tool to investigate the electronic structure of interesting chemical compounds and physical systems where the spin interaction plays an important role. The spin interaction may, indeed, be much more important than previously believed, as we have shown here for the  $C_2$  molecule, the very first and remarkable example of an antiferromagnetic chemical bond.

# Acknowledgements

At the end of such a wonderful experience as this PhD, I feel it is fundamental to acknowledge the persons that made this experience so stimulating and enjoyable.

The first and most important person that I want to thank is obviously Sandro. Thank you for all the passion, the effort, the teaching, and the time you put in this project but overall for my growth process during the years. It is useless to say that the level of this work could not have been nearly the same without you. You have always been present for all the technical problems and to improve the quality of what we did together. It is impossible not to admire your dedication to science and student education. I am extremely thankful for having the possibility to learn directly from one of the best and most passionate scientists in the world.

I thank Kousuke, Natanael, Tomonori, Andrea Zen, Andrea Tirelli, Giacomo, and Federico for all the useful discussions and lunches together. It has been a pleasure to work surrounded by people with your technical and human skills. It will not be easy to find another team like this.

Thanks to Lucas, Matteo, Davide, and Mattia. It has been so funny and satisfying to work with you as a team and to study new techniques and concepts together.

Thanks to all the PhD students and post-docs in the Condensed Matter area. Having the possibility to discuss and confront with, and to learn from such a wonderful community has been an honour. In particular, a big thanks

goes to the students of my year, I am proud to have been part of this extraordinary group. Among everyone let me thank also Andrea Urru for his friendship, for being such an incredible flatmate, and for the time and the journey we shared during this PhD.

I want to thank the whole SISSA community for making SISSA a wonderful place to grow as a person and as a scientist.

The work that you do reflects a lot of the life that you have. For this reason, I dedicate this thesis to Sara. Thank you for making my days and, hopefully, my whole life joyful and to give me the serenity to face my work with passion. I want to thank also my family that has always supported me and has always been with me in this journey despite the distance. Thank also to my friends and all the people who made these four years interesting and satisfying.

# Bibliography

- [1] E. Schrödinger, “An undulatory theory of the mechanics of atoms and molecules,” *Phys. Rev.*, vol. 28, pp. 1049–1070, Dec 1926.
- [2] P. A. M. Dirac, “The quantum theory of the electron,” *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, vol. 117, no. 778, pp. 610–624, 1928.
- [3] P. A. M. Dirac and R. H. Fowler, “Quantum mechanics of many-electron systems,” *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, vol. 123, no. 792, pp. 714–733, 1929.
- [4] W. M. C. Foulkes, L. Mitas, R. J. Needs, and G. Rajagopal, “Quantum monte carlo simulations of solids,” *Rev. Mod. Phys.*, vol. 73, pp. 33–83, Jan 2001.
- [5] F. Becca and S. Sorella, *Quantum Monte Carlo Approaches for Correlated Systems*. Cambridge University Press, 2017.
- [6] B. L. Hammond, W. A. Lester, and P. J. Reynolds, *Monte Carlo Methods in Ab Initio Quantum Chemistry*. WORLD SCIENTIFIC, 1994.
- [7] P. J. Reynolds, J. Tobochnik, and H. Gould, “Diffusion quantum monte carlo,” *Computers in Physics*, vol. 4, no. 6, pp. 662–668, 1990.

- [8] J. B. Anderson, “A random-walk simulation of the schrödinger equation:  $H+3$ ,” *J. Chem. Phys.*, vol. 63, no. 4, pp. 1499–1503, 1975.
- [9] W. Pauli, “Über den zusammenhang des abschlusses der elektronengruppen im atom mit der komplexstruktur der spektren,” *Zeitschrift für Physik*, vol. 31, no. 1, pp. 765–783, 1925.
- [10] J. C. Slater, “The Theory of Complex Spectra,” *Physical Review*, vol. 34, pp. 1293–1322, Nov. 1929.
- [11] M. Motta, D. M. Ceperley, G. K.-L. Chan, J. A. Gomez, E. Gull, S. Guo, C. A. Jiménez-Hoyos, T. N. Lan, J. Li, F. Ma, A. J. Millis, N. V. Prokof’ev, U. Ray, G. E. Scuseria, S. Sorella, E. M. Stoudenmire, Q. Sun, I. S. Tupitsyn, S. R. White, D. Zgid, and S. Zhang, “Towards the solution of the many-electron problem in real materials: Equation of state of the hydrogen chain with state-of-the-art many-body methods,” *Phys. Rev. X*, vol. 7, p. 031059, Sep 2017.
- [12] M. Motta\*, C. Genovese\*, F. Ma\*, Z.-H. Cui\*, R. Sawaya\*, G. K.-L. Chan, N. Chepiga, P. Helms, C. Jimenez-Hoyos, A. J. Millis, U. Ray, E. Ronca, H. Shi, S. Sorella, E. M. Stoudenmire, S. R. White, and S. Zhang, “Ground-state properties of the hydrogen chain: insulator-to-metal transition, dimerization, and magnetic phases,” Accepted in *Physical Review X*, 2019.
- [13] C. Genovese, A. Meninno, and S. Sorella, “Assessing the accuracy of the jastrow antisymmetrized geminal power in the h4 model system,” *J. Chem. Phys.*, vol. 150, no. 8, p. 084102, 2019.
- [14] K. Gasperich, M. Deible, and K. D. Jordan, “H4: A model system for assessing the performance of diffusion monte carlo calculations using a

- single slater determinant trial function,” *J. Chem. Phys.*, vol. 147, no. 7, p. 074106, 2017.
- [15] B. Braïda, J. Toulouse, M. Caffarel, and C. J. Umrigar, “Quantum monte carlo with jastrow-valence-bond wave functions,” *J. Chem. Phys.*, vol. 134, no. 8, p. 084108, 2011.
- [16] J. Toulouse and C. J. Umrigar, “Full optimization of jastrow–slater wave functions with application to the first-row atoms and homonuclear diatomic molecules,” *J. Chem. Phys.*, vol. 128, no. 17, p. 174101, 2008.
- [17] P. Knowles and N. Handy, “A new determinant-based full configuration interaction method,” *Chem. Phys. Lett.*, vol. 111, no. 4, pp. 315 – 321, 1984.
- [18] J. Olsen, B. O. Roos, P. Jørgensen, and H. J. A. Jensen, “Determinant based configuration interaction algorithms for complete and restricted configuration interaction spaces,” *J. Chem. Phys.*, vol. 89, no. 4, pp. 2185–2192, 1988.
- [19] B. O. Roos, P. R. Taylor, and P. E. Sigbahn, “A complete active space scf method (casscf) using a density matrix formulated super-ci approach,” *Chem. Phys.*, vol. 48, no. 2, pp. 157 – 173, 1980.
- [20] C. Filippi and C. J. Umrigar, “Multiconfiguration wave functions for quantum monte carlo calculations of first-row diatomic molecules,” *J. Chem. Phys.*, vol. 105, no. 1, pp. 213–226, 1996.
- [21] M. A. Morales, J. McMinis, B. K. Clark, J. Kim, and G. E. Scuseria, “Multideterminant wave functions in quantum monte carlo,” *J. Chem. Theory Comput.*, vol. 8, no. 7, pp. 2181–2188, 2012.

- [22] J. Linderberg, “The antisymmetrized geminal power, a simple correlated wave function for chemical bonding,” *Isr. J. Chem.*, vol. 19, no. 1-4, pp. 93–98, 1980.
- [23] M. Casula and S. Sorella, “Geminal wave functions with jastrow correlation: A first application to atoms,” *J. Chem. Phys.*, vol. 119, no. 13, pp. 6500–6511, 2003.
- [24] M. Casula, C. Attaccalite, and S. Sorella, “Correlated geminal wave function for molecules: An efficient resonating valence bond approach,” *J. Chem. Phys.*, vol. 121, no. 15, pp. 7110–7126, 2004.
- [25] E. Neuscamman, “Improved optimization for the cluster jastrow antisymmetric geminal power and tests on triple-bond dissociations,” *J. Chem. Theory Comput.*, vol. 12, no. 7, pp. 3149–3159, 2016.
- [26] M. Bajdich, L. Mitas, G. Drobný, L. K. Wagner, and K. E. Schmidt, “Pfaffian pairing wave functions in electronic-structure quantum monte carlo simulations,” *Phys. Rev. Lett.*, vol. 96, p. 130201, Apr 2006.
- [27] 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*, vol. 77, p. 115112, Mar 2008.
- [28] C. Genovese, T. Shirakawa, K. Nakano, and S. Sorella, “General correlated geminal ansatz for electronic structure calculations: exploiting pfaffians in place of determinants,” Accepted in the Journal of Chemical Theory and Computation, 2020.
- [29] K. Nakano, C. Attaccalite, M. Barborini, L. Capriotti, M. Casula, E. Coccia, M. Dagrada, C. Genovese, Y. Luo, G. Mazzola, A. Zen, and S. Sorella, “Turborvb: A many-body toolkit for ab initio electronic sim-

- ulations by quantum monte carlo,” *The Journal of Chemical Physics*, vol. 152, no. 20, p. 204121, 2020.
- [30] F. Franjić and S. Sorella, “Spin-Wave Wave Function for Quantum Spin Models,” *Prog. Theor. Phys.*, vol. 97, pp. 399–406, 03 1997.
- [31] S. Sorella, K. Seki, O. O. Brovko, T. Shirakawa, S. Miyakoshi, S. Yunoki, and E. Tosatti, “Correlation-driven dimerization and topological gap opening in isotropically strained graphene,” *Phys. Rev. Lett.*, vol. 121, p. 066402, Aug 2018.
- [32] K. Nakano, R. Maezono, and S. Sorella, “All-electron quantum monte carlo with jastrow single determinant ansatz: Application to the sodium dimer,” *J. Chem. Theory Comput.*, vol. 15, no. 7, pp. 4044–4055, 2019.
- [33] C. Genovese and S. Sorella, “The nature of the quadruple chemical bond in the dicarbon molecule,” 2019.
- [34] C. Genovese, A. Zen, and S. Sorella, “The benzene dimer and the role of the optimization for non covalent interaction in qmc,” In preparation.
- [35] S. Sorella and L. Capriotti, “Algorithmic differentiation and the calculation of forces by quantum monte carlo,” *J. Chem. Phys.*, vol. 133, no. 23, p. 234111, 2010.
- [36] Wikipedia contributors, “Monte carlo method — Wikipedia, the free encyclopedia,” 2020. [Online; accessed 20-May-2020].
- [37] C. Umrigar, “Accelerated metropolis method,” *Phys. Rev. Lett.*, vol. 71, no. 3, p. 408, 1993.
- [38] M. Stedman, W. Foulkes, and M. Nekovee, “An accelerated metropolis method,” *J. Chem. Phys.*, vol. 109, no. 7, pp. 2630–2634, 1998.

- [39] S. Sorella, “Green function monte carlo with stochastic reconfiguration,” *Phys. Rev. Lett.*, vol. 80, no. 20, p. 4558, 1998.
- [40] S. Sorella, M. Casula, and D. Rocca, “Weak binding between two aromatic rings: Feeling the van der waals attraction by quantum monte carlo methods,” *J. Chem. Phys.*, vol. 127, no. 1, p. 014105, 2007.
- [41] S. Sorella, “Wave function optimization in the variational monte carlo method,” *Phys. Rev. B*, vol. 71, no. 24, p. 241103, 2005.
- [42] C. J. Umrigar, J. Toulouse, C. Filippi, S. Sorella, and R. G. Hennig, “Alleviation of the fermion-sign problem by optimization of many-body wave functions,” *Phys. Rev. Lett.*, vol. 98, p. 110201, Mar 2007.
- [43] J. Toulouse and C. J. Umrigar, “Optimization of quantum monte carlo wave functions by energy minimization,” *J. Chem. Phys.*, vol. 126, no. 8, p. 084102, 2007.
- [44] R. Assaraf, S. Moroni, and C. Filippi, “Optimizing the energy with quantum monte carlo: A lower numerical scaling for jastrow–slater expansions,” *J. Chem. Theory Comput.*, vol. 13, no. 11, pp. 5273–5281, 2017.
- [45] G. Mazzola, A. Zen, and S. Sorella, “Finite-temperature electronic simulations without the born-oppenheimer constraint,” *J. Chem. Phys.*, vol. 137, no. 13, p. 134112, 2012.
- [46] S.-I. Amari, “Natural gradient works efficiently in learning,” *Neural Comput.*, vol. 10, no. 2, pp. 251–276, 1998.
- [47] S. Sorella, M. Casula, and D. Rocca, “Weak binding between two aromatic rings: Feeling the van der waals attraction by quantum monte carlo methods,” *J. Chem. Phys.*, vol. 127, no. 1, p. 014105, 2007.

- [48] M. Casula, C. Filippi, and S. Sorella, “Diffusion Monte Carlo method with lattice regularization,” *Phys. Rev. Lett.*, vol. 95, no. 10, pp. 1–5, 2005.
- [49] D. F. Ten Haaf, H. J. Van Bommel, J. M. Van Leeuwen, W. Van Saarloos, and D. M. Ceperley, “Proof for an upper bound in fixed-node Monte Carlo for lattice fermions,” *Phys. Rev. B*, vol. 51, no. 19, pp. 13039–13045, 1995.
- [50] M. Calandra Buonauro and S. Sorella, “Numerical study of the two-dimensional Heisenberg model using a Green function Monte Carlo technique with a fixed number of walkers,” *Phys. Rev. B*, vol. 57, pp. 11446–11456, may 1998.
- [51] S. Sorella and L. Capriotti, “Green function Monte Carlo with stochastic reconfiguration: An effective remedy for the sign problem,” *Phys. Rev. B*, vol. 61, no. 4, pp. 2599–2612, 2000.
- [52] M. Casula, *New QMC approaches for the simulation of electronic systems: a first application to aromatic molecules and transition metal compounds*. PhD thesis, International School for Advanced Studies (SISSA), 2005.
- [53] M. Calandra Buonauro and S. Sorella, “Numerical study of the two-dimensional heisenberg model using a green function monte carlo technique with a fixed number of walkers,” *Phys. Rev. B*, vol. 57, pp. 11446–11456, May 1998.
- [54] K. Nakano, R. Maezono, and S. Sorella, “Speeding up ab initio diffusion monte carlo simulations by a smart lattice regularization,” *Phys. Rev. B*, vol. 101, p. 155106, Apr 2020.

- [55] P. López Ríos, A. Ma, N. D. Drummond, M. D. Towler, and R. J. Needs, “Inhomogeneous backflow transformations in quantum Monte Carlo calculations,” *Phys. Rev. E*, vol. 74, no. 6, 2006.
- [56] 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.*, vol. 73, pp. 1959–1962, Oct 1994.
- [57] G. Rajagopal, R. Needs, A. James, S. Kenny, and W. Foulkes, “Variational and diffusion quantum monte carlo calculations at nonzero wave vectors: Theory and application to diamond-structure germanium,” *Phys. Rev. B*, vol. 51, no. 16, p. 10591, 1995.
- [58] C. Pisani, R. Dovesi, and C. Roetti, “Lecture notes in chemistry,” in *Hartree-Fock ab initio treatment of crystalline systems*, vol. 48, Springer Berlin, 1988.
- [59] R. Dovesi, A. Erba, R. Orlando, C. M. Zicovich-Wilson, B. Civaleri, L. Maschio, M. Rérat, S. Casassa, J. Baima, S. Salustro, *et al.*, “Quantum-mechanical condensed matter simulations with crystal,” *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, vol. 8, no. 4, p. e1360, 2018.
- [60] S. Karakuzu, L. F. Tocchio, S. Sorella, and F. Becca, “Superconductivity, charge-density waves, antiferromagnetism, and phase separation in the hubbard-holstein model,” *Phys. Rev. B*, vol. 96, p. 205145, Nov 2017.
- [61] M. Marchi, S. Azadi, M. Casula, and S. Sorella, “Resonating valence bond wave function with molecular orbitals: Application to first-row molecules,” *J. Chem. Phys.*, vol. 131, no. 15, p. 154116, 2009.
- [62] P. W. Kasteleyn, “Dimer statistics and phase transitions,” *J. Math. Phys.*, vol. 4, no. 2, pp. 287–293, 1963.

- [63] 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*. Philadelphia, PA: Society for Industrial and Applied Mathematics, third ed., 1999.
- [64] M. Wimmer, "Algorithm 923: Efficient numerical computation of the pfaffian for dense and banded skew-symmetric matrices," *ACM Trans. Math. Softw.*, vol. 38, pp. 30:1–30:17, Aug. 2012.
- [65] A. Zen, E. Coccia, Y. Luo, S. Sorella, and L. Guidoni, "Static and dynamical correlation in diradical molecules by quantum monte carlo using the jastrow antisymmetrized geminal power ansatz," *J. Chem. Theory Comput.*, vol. 10, no. 3, pp. 1048–1061, 2014.
- [66] S. Caracciolo, A. D. Sokal, and A. Sportiello, "Algebraic/combinatorial proofs of cayley-type identities for derivatives of determinants and pfaffians," *Adv. Appl. Math.*, vol. 50, no. 4, pp. 474 – 594, 2013.
- [67] J. Hachmann, W. Cardoen, and G. K.-L. Chan, "Multireference correlation in long molecules with the quadratic scaling density matrix renormalization group," *The Journal of Chemical Physics*, vol. 125, no. 14, p. 144101, 2006.
- [68] W. A. Al-Saidi, S. Zhang, and H. Krakauer, "Bond breaking with auxiliary-field quantum Monte Carlo," *The Journal of Chemical Physics*, vol. 127, no. 14, p. 144101, 2007.
- [69] A. V. Sinitskiy, L. Greenman, and D. A. Mazziotti, "Strong correlation in Hydrogen chains and lattices using the variational two-electron reduced density matrix method," *The Journal of Chemical Physics*, vol. 133, no. 1, p. 014104, 2010.

- [70] L. Stella, C. Attaccalite, S. Sorella, and A. Rubio, “Strong electronic correlation in the Hydrogen chain: A variational Monte Carlo study,” *Physical Review B*, vol. 84, p. 245117, Dec 2011.
- [71] S. R. White, “Density matrix formulation for quantum renormalization groups,” *Phys. Rev. Lett.*, vol. 69, pp. 2863–2866, Nov 1992.
- [72] S. R. White and R. L. Martin, “Ab initio quantum Chemistry using the density matrix renormalization group,” *The Journal of Chemical Physics*, vol. 110, no. 9, pp. 4127–4130, 1999.
- [73] S. Sharma and G. K.-L. Chan, “Spin-adapted density matrix renormalization group algorithms for quantum chemistry,” *The Journal of chemical physics*, vol. 136, no. 12, p. 124121, 2012.
- [74] E. M. Stoudenmire and S. R. White, “Sliced basis density matrix renormalization group for electronic structure,” *Physical Review Letters*, vol. 119, p. 046401, Jul 2017.
- [75] S. Zhang and H. Krakauer, “Quantum Monte Carlo method using phase-free random walks with Slater determinants,” *Physical Review Letters*, vol. 90, p. 136401, Apr 2003.
- [76] A. Szabo and N. Ostlund, *Modern Quantum Chemistry*. McGraw-Hill, 1989.
- [77] R. M. Martin, *Electronic Structure*. Cambridge University Press, 2004.
- [78] I. Affleck, “Exact correlation amplitude for the heisenberg antiferromagnetic chain,” *Journal of Physics A: Mathematical and General*, vol. 31, pp. 4573–4581, may 1998.
- [79] I. Affleck, D. Gepner, H. J. Schulz, and T. Ziman, “Critical behaviour of spin-s heisenberg antiferromagnetic chains: analytic and numerical re-

- sults,” *Journal of Physics A: Mathematical and General*, vol. 22, pp. 511–529, mar 1989.
- [80] N. F. Mott, “Metal-insulator transition,” *Reviews of Modern Physics*, vol. 40, pp. 677–683, Oct 1968.
- [81] M. Imada, A. Fujimori, and Y. Tokura, “Metal-insulator transitions,” *Reviews of Modern Physics*, vol. 70, pp. 1039–1263, Oct 1998.
- [82] E. H. Lieb and F. Y. Wu, “Absence of Mott transition in an exact solution of the short-range, one-band model in one dimension,” *Physical Review Letters*, vol. 20, pp. 1445–1448, Jun 1968.
- [83] R. Resta and S. Sorella, “Electron localization in the insulating state,” *Physical Review Letters*, vol. 82, pp. 370–373, Jan 1999.
- [84] W. Kohn, “Theory of the insulating state,” *Physical Review*, vol. 133, pp. A171–A181, Jan 1964.
- [85] J. B. Anderson, “Quantum chemistry by random walk: H4 square,” *International Journal of Quantum Chemistry*, vol. 15, no. 1, pp. 109–120, 1979.
- [86] D. M. Silver and R. M. Stevens, “Reaction paths on the h4 potential energy surface,” *The Journal of Chemical Physics*, vol. 59, no. 6, pp. 3378–3394, 1973.
- [87] L. Pauling, *The nature of the chemical bond*. Cornell University Press, 1960.
- [88] P. W. Anderson, “The resonating valence bond state in  $La_2CuO_4$  and superconductivity,” *Science*, vol. 235, p. 1196, 1987.
- [89] R. A. Aziz and M. Slaman, “The ne-ne interatomic potential revisited,” *Chem. Phys.*, vol. 130, no. 1, pp. 187 – 194, 1989.

- [90] J. M. Merritt, V. E. Bondybey, and M. C. Heaven, "Beryllium dimer—caught in the act of bonding," *Science*, vol. 324, no. 5934, pp. 1548–1551, 2009.
- [91] S. R. Langhoff and C. W. Bauschlicher, "Theoretical study of the spectroscopy of  $b_2$ ," *J. Chem. Phys.*, vol. 95, no. 8, pp. 5882–5888, 1991.
- [92] L. Bytautas and K. Ruedenberg, "Correlation energy extrapolation by intrinsic scaling. iv. accurate binding energies of the homonuclear diatomic molecules carbon, nitrogen, oxygen, and fluorine," *J. Chem. Phys.*, vol. 122, no. 15, p. 154110, 2005.
- [93] J. D. Watts and R. J. Bartlett, "Coupled-cluster calculations on the  $c_2$  molecule and the  $c+2$  and  $c-2$  molecular ions," *J. Chem. Phys.*, vol. 96, no. 8, pp. 6073–6084, 1992.
- [94] A. Karton, A. Tarnopolsky, and J. M. L. Martin, "Atomization energies of the carbon clusters  $c_n$  ( $n = 2-10$ ) revisited by means of  $w_4$  theory as well as density functional,  $gn$ , and  $cbs$  methods," *Mol. Phys.*, vol. 107, pp. 977–990, 2008.
- [95] W. Zou and D. Cremer, "C<sub>2</sub> in a box: Determining its intrinsic bond strength for the  $x1\sigma_g+$  ground state," *Chemistry – A European Journal*, vol. 22, no. 12, pp. 4087–4099, 2016.
- [96] G. H. Booth, D. Cleland, A. J. W. Thom, and A. Alavi, "Breaking the carbon dimer: The challenges of multiple bond dissociation with full configuration interaction quantum monte carlo methods," *J. Chem. Phys.*, vol. 135, no. 8, p. 084104, 2011.
- [97] S. Shaik, D. Danovich, W. Wu, P. Su, H. S. Rzepa, and P. C. Hiberty, "Quadruple bonding in  $c_2$  and analogous eight-valence electron species," *Nature chemistry*, vol. 4 3, pp. 195–200, 2012.

- [98] R. Zhong, M. Zhang, H. Xu, and Z. Su, “Latent harmony in dicarbon between vb and mo theories through orthogonal hybridization of  $3\sigma$  g and  $2\sigma$  u,” *Chemical science*, vol. 7, no. 2, pp. 1028–1032, 2016.
- [99] G. Frenking and M. Hermann, “Comment on “the quadruple bonding in c2 reproduces the properties of the molecule”,” *Chemistry – A European Journal*, vol. 22, no. 52, pp. 18975–18976, 2016.
- [100] S. Shaik, D. Danovich, B. Braida, and P. C. Hiberty, “The quadruple bonding in c2 reproduces the properties of the molecule,” *Chemistry – A European Journal*, vol. 22, no. 12, pp. 4116–4128, 2016.
- [101] D. W. O. de Sousa and M. A. C. Nascimento, “Is there a quadruple bond in  $c_2$ ?,” *J. Chem. Theory Comput.*, vol. 12, no. 5, pp. 2234–2241, 2016.
- [102] S. J. Chakravorty, S. R. Gwaltney, E. R. Davidson, F. A. Parpia, and C. F. p Fischer, “Ground-state correlation energies for atomic ions with 3 to 18 electrons,” *Phys. Rev. A*, vol. 47, pp. 3649–3670, May 1993.
- [103] S. Sharma, “A general non-abelian density matrix renormalization group algorithm with application to the c2 dimer,” *The Journal of Chemical Physics*, vol. 142, no. 2, p. 024107, 2015.
- [104] J. P. Perdew and A. Zunger, “Self-interaction correction to density-functional approximations for many-electron systems,” *Physical Review B*, vol. 23, no. 10, p. 5048, 1981.
- [105] M. Frisch, G. Trucks, H. Schlegel, G. Scuseria, M. Robb, J. Cheeseman, G. Scalmani, V. Barone, G. Petersson, H. Nakatsuji, *et al.*, “Gaussian 16,” 2016.
- [106] S. Sharma and A. Alavi, “Multireference linearized coupled cluster theory for strongly correlated systems using matrix product states,” *J. Chem. Phys.*, vol. 143, no. 10, p. 102815, 2015.

- [107] A. A. Holmes, C. J. Umrigar, and S. Sharma, “Excited states using semistochastic heat-bath configuration interaction,” *The Journal of Chemical Physics*, vol. 147, no. 16, p. 164111, 2017.
- [108] K. K. Irikura, “Experimental vibrational zero-point energies: Diatomic molecules,” *Journal of Physical and Chemical Reference Data*, vol. 36, no. 2, pp. 389–397, 2007.
- [109] D. Pfau, J. S. Spencer, A. G. de G. Matthews, and W. M. C. Foulkes, “Ab-initio solution of the many-electron schrödinger equation with deep neural networks,” 2019.
- [110] G. K.-L. Chan, M. Kállay, and J. Gauss, “State-of-the-art density matrix renormalization group and coupled cluster theory studies of the nitrogen binding curve,” *The Journal of Chemical Physics*, vol. 121, no. 13, pp. 6110–6116, 2004.
- [111] A. Zen, B. L. Trout, and L. Guidoni, “Properties of reactive oxygen species by quantum monte carlo,” *The Journal of Chemical Physics*, vol. 141, no. 1, p. 014305, 2014.
- [112] S. Azadi and R. E. Cohen, “Chemical accuracy from quantum monte carlo for the benzene dimer,” *J. Chem. Phys.*, vol. 143, no. 10, p. 104301, 2015.
- [113] E. U. Franck, “J. d. cox, d. d. wagman, v. a. medvedev: Codata — key values for thermodynamics, aus der reihe: Codata, series on thermodynamic properties. hemisphere publishing corporation, new york, washington, philadelphia, london 1989. 271 seiten, preis: £ 28.00,” *Ber. Bunsenges. Phys. Chem.*, vol. 94, no. 1, pp. 93–93, 1990.
- [114] L. Pauling, *The nature of the Chemical Bond*. 3rd ed. Cornell University Press, Itaca, New York, 1960.

- [115] M. S. Marshall, L. A. Burns, and C. D. Sherrill, “Basis set convergence of the coupled-cluster correction, dmp2ccsd(t): Best practices for benchmarking non-covalent interactions and the attendant revision of the s22, nbc10, hbc6, and hsg databases,” *The Journal of Chemical Physics*, vol. 135, no. 19, p. 194102, 2011.
- [116] M. C. Bennett, G. Wang, A. Annaberdiyev, C. A. Melton, L. Shulenburg, and L. Mitas, “A new generation of effective core potentials from correlated calculations: 2nd row elements,” *The Journal of Chemical Physics*, vol. 149, no. 10, p. 104108, 2018.
- [117] K. E. Gasperich, *Calculation of Intermolecular Interactions via Diffusion Monte Carlo*. PhD thesis, University of Pittsburgh, 2020.
- [118] M. Dubecký, L. Mitas, and P. Jurečka, “Noncovalent interactions by quantum monte carlo,” *Chemical Reviews*, vol. 116, pp. 5188–5215, May 2016.